Working Paper 24-003

Beyond the Hype: Unveiling the Marginal Benefits of 3D Virtual Tours in Real Estate

Mengxia Zhang
Isamar Troncoso

Harvard Business School

Beyond the Hype: Unveiling the Marginal Benefits of 3D Virtual Tours in Real Estate

Mengxia Zhang
Western University

Isamar Troncoso
Harvard Business School

Working Paper 24-003

Funding for this research was provided in part by Harvard Business School.

Beyond the Hype: Unveiling the Marginal Benefits of 3D Virtual Tours in Real Estate

Mengxia Zhang and Isamar Troncoso*†

July 18, 2023

*Authors' names are listed in reversed alphabetical order. The two authors equally contributed to this work.

†Mengxia Zhang (mezhang@ivey.ca) is an Assistant Professor in Marketing at the Ivey Business School, Western University. Isamar Troncoso (itroncoso@hbs.edu) is an Assistant Professor of Business Administration at Harvard Business School, Harvard University. We thank participants at the 2022 Conference on Artificial Intelligence, Machine Learning, and Business Analytics, the 2023 Workshop in Management Science, the 45th ISMS Marketing Science Conference, and seminar attendants at Boston University and Santa Clara University for their helpful comments.

Abstract

3D virtual tours (VTs) have become a popular digital tool in real estate platforms, enabling potential buyers to virtually walk through the houses they search for online. In this paper, we study home sellers' adoption of VTs and the VTs' relative benefits compared to other marketing efforts. Using data on 75,178 houses sold in the greater Los Angeles area from March 2019 to March 2021, employing computer vision, natural language processing, and causal machine learning, we present four key findings. First, sellers using VTs provide longer descriptions and more photos, suggesting VTs supplement traditional marketing efforts rather than replace them. Second, after controlling for photo content and quality and listing description sentiment and topics, VTs do not significantly increase sale prices and slightly increase time on the market, suggesting their marginal benefits are smaller than previously thought. Third, smaller firms/agents and neighborhoods with more racial minorities and lower-income households benefit most from VT adoption but do so less frequently, presenting opportunities for platforms and policymakers to encourage VT adoption among these groups. Lastly, although VTs do not significantly improve final sale outcomes, they may offer benefits such as helping consumers screen options and assisting sellers in portfolio growth and management.

1 Introduction

Virtual reality (VR) has gained considerable momentum in numerous industries, including gaming, healthcare, retail, and real estate. In real estate, 3D virtual tours (VTs) have emerged as a popular non-immersive VR tool sellers employ when listing their properties online. By providing a navigable digital representation of a house accessible through computers or smartphones, these VTs allow potential home-buyers searching for properties online to "virtually walk through" a home's rooms and examine its various features and layouts from multiple angles. And while available for over 20 years, recent technological advances in cameras and imaging software have made VTs more accessible and affordable.[1]

The prevailing belief among real estate platforms and other industry players is that VTs can significantly improve house sale outcomes. For instance, Matterport, a 3D VT creator company, estimates that houses that include VT in their listings can sell at up to 4-9% higher prices and stay 20-30% fewer days on the market.[2] However, the scarce empirical research on this topic has found more moderate effects of a 2-3% increase in sale prices and an unclear effect on days on market (Carrillo 2008, Yan et al. 2021, Yu et al. 2020). Moreover, little is known about the relationship between VT adoption and other marketing efforts in a house listing/advertisement, such as the photos and descriptions sellers provide.

In this paper, we aim to fill these gaps by empirically exploring answers to three research questions. First, what factors explain 3D virtual tour adoption (e.g., seller, property, neighborhood characteristics, marketing strategies, etc.)? Second, what is the marginal effect of VT on sale outcomes (i.e., sale price and time on market)? And third, when and how can VT adoption be more beneficial for sellers?

Our empirical study uses data from 75,178 houses sold in the greater Los Angeles area (California, US) between March 2019 and March 2021. The choice of the region allows us to benchmark our findings to prior literature on the topic while exploiting the diversity of its

[1]As per 2023, 3D VT creators typically charge a flat fee (e.g., $300 for up to 2,000 square feet) plus a varying fee function of square footage (e.g., $10 to $50 per additional 1,000 square footage).

[2]See: `https://matterport.com/blog/3d-tours-properties-sell-31-faster-and-higher-price`.

housing stock in our analyses. The choice of the time window allows us to examine potential disruptions to in-person tours following the COVID-19 pandemic. The data is collected from Redfin, a leading online real estate platform that operates in the US and Canadian markets and provides detailed information for each listing, including its description, photos, whether it included a VT, the name of the seller agent and real estate firm, neighborhood characteristics, initial listing price, and final sale outcomes, among others.

A major empirical challenge in studying the effect of VTs on sale outcomes is that their adoption is not random but likely influenced by various factors that also affect outcomes, such as the quality of the house, the seller agent's industry experience, and the overall effort invested in selling the property. To account for such selection bias, we adopt the double machine learning (DML) approach (Chernozhukov et al. 2018), which allows us to directly model VT adoption, sale outcomes, and their interdependence on a potentially large set of confounders X in a highly flexible manner. In addition to accounting for confounders previously suggested in the literature, such as seller agent experience and firms' size/technology adoption capability (Benefield et al. 2019), we follow recent work that leverages unstructured data to uncover potential confounders in observational studies (Zeng et al. 2022). Specifically, we apply computer vision and natural language processing (NLP) to the photos and open-ended descriptions in the listings, respectively, and uncover proxies for several confounders that are intuitively relevant but typically hard to measure, such as the aesthetic quality of the photos to proxy for the visual appeal of the house, and recent renovations or investment opportunities discussed in the descriptions to proxy for the house's quality/condition.

In our first analysis, we focus on the selection equation of the DML approach to explore what factors explain VT adoption, using the XGBoost classifier to predict VT adoption and SHAP values to examine which variables are more important in explaining such predictions (Lundberg et al. 2020). We find that the three most important explaining factors of VT adoption are the aesthetic quality of the photos (7.73% importance), the number of photos (7.58% importance), and the description word count (6.60% importance). Moreover, several photo and text-related variables are more important explaining factors of VT adoption than standard house and neighborhood characteristics, and they are also among the most important explaining factors of sale prices and time on market (TOM), highlighting the value of leveraging unstructured data to uncover confounders such as a house's visual aesthetic or condition. We also find that the probability of adopting VT increases with the aesthetic quality of the photos, the number of photos, and the number of words included in the listings, suggesting that sellers tend to use VTs as a supplementary tool rather than a replacement for traditional information sources like photos and descriptions. Furthermore, we observe that VT adoption is more likely for houses in higher-value areas and listings with descriptions emphasizing topics such as space and design. Notably, VT adoption experienced a significant surge following the World Health Organization's March 2020 announcement about the COVID-19 pandemic, perhaps to compensate for reduced in-person tours due to government lockdown policies.

In our second analysis, we focus on the DML estimates of the average treatment effects of VT on sale prices and TOM. To explore the impact of adjusting for additional confounders uncovered from the unstructured data on each listing, we first create a baseline model specification that controls for house, neighborhood, seller agent's characteristics, time trends, and the number of photos in the listing only. Consistent with prior literature, we find that VTs are associated with a 1.1% increase in sale prices and have a non-significant effect on the time a house stays on the market. However, after controlling for the additional confounders uncovered from the photos and descriptions in the listings, the effect of VT on sale prices becomes insignificant, and the effect of VT on TOM increases slightly. These results can be explained by a pricing strategy sellers tend to use when using VTs: starting with a higher listing price and adjusting it downwards over time, which leads to longer TOM.

We acknowledge that, as in many observational studies, our estimates cannot be interpreted as *unbiased* as they might be affected by unmeasured confounders whose existence we cannot rule out. Nonetheless, they do suggest that *after removing meaningful sources* *of bias* previously overlooked in the literature, such as the visual appeal and quality of the house and the effort made by the seller in providing visual and textual information about the property, the average benefits of using VT to improve sale outcomes are smaller than previously thought.

Finally, we conduct three descriptive analyses to further investigate when and how VT adoption can be more beneficial for sellers. First, we examine whether the benefits of VTs on sale outcomes vary with listing, seller agent/firm, house, and neighborhood characteristics. Among others, we find that the benefits of using VTs for sale prices are above average for houses in neighborhoods with more racial minorities and lower income, smaller seller agents/firms, and listings that minimize other marketing efforts (i.e., listings that include fewer words and photos). Notably, these are also the groups with lower VT adoption. Second, we examine whether the effect of VT adoption on sale outcomes changed with the COVID-19 pandemic, an event that significantly increased VT adoption among sellers and potentially altered home buyers' search and purchase behavior. We find that VTs helped significantly reduce TOM in the first two months following the COVID-19 pandemic stay-at-home order, but the benefits returned to pre-pandemic levels afterward. Third, we examine the relationship between VT adoption and other sellers' and buyers' outcomes and find suggestive evidence of other potential benefits of these tools: for sellers, VT can help attract more attention to individual listings and expand their listing portfolios; and for buyers, VT can help screen out their alternatives more efficiently.

Our research makes the following contributions to the literature. First, it adds to the growing literature on the impact of using virtual reality in the marketplace (e.g., Gallino and Moreno 2018, Meißner et al. 2020, Tan et al. 2022). While prior research has shown that virtual reality can benefit sales of lower-involvement and low-risk products such as grocery and cosmetic products, we discover that its benefits are more moderate for higher-involvement and high-risk goods. Second, our findings add to the marketing and economics literature studying the effect of VTs on sales outcomes in the real estate market context (e.g., Carrillo 2008, Yu et al. 2020, Yan et al. 2021). Our research uncovers factors associated with VT adoption and leverages unstructured text and photo data to uncover interpretable confounders inspired by several recent observational studies (e.g., Keith et al. 2020, Zeng et al. 2022). Lastly, our research adds to the marketing literature on the role of visual information in different marketing outcomes (e.g., Hartmann et al. 2021, Zhang et al. 2022, Zhang and Luo 2023). In particular, it sheds light on the marginal value of newer visual technologies (e.g., interactive VTs) relative to more traditional ones (e.g., static photos) and other marketing strategies.

Our research offers sellers several guidelines on when and how to use VT to obtain higher benefits. First, when working on a tight budget/schedule, focusing on providing a variety of high-quality photos and rich textual information can be as effective as adding VTs. Second, using VTs can bring more substantial benefits for properties in neighborhoods with higher minority ethnic group densities or lower average household incomes. Such an insight might also have implications for platforms and policymakers, highlighting the potential of encouraging VT adoption in those neighborhoods that could benefit the most by, for example, raising awareness about more accessible do-it-yourself tools for creating VT (e.g., Zillow 3D house app, Klapty, and Matterport). Last, our findings suggest that even if VTs do not improve average sale outcomes, sellers might use them to screen out less interested buyers and grow their listing portfolio.

The remainder of the paper is organized as follows. We present a summary of the related literature in Section 2, describe our data in Section 3, introduce our methodology in Section 4, explore the drivers of VT adoption in Section 5, investigate the impact of VT adoption on sales outcomes and other potential benefits in Section 6, and conclude the paper in Section 7.

2 Related literature

Our research relates to the growing marketing literature studying the impact of virtual reality (VR) and augmented reality (AR) applications in retailing and selling. To date, most empirical research has focused on the impact of VR/AR on consumer choices and sale outcomes of low-involvement products. For example, Meißner et al. (2020) shows that high-immersive VR can increase consumers' variety seeking and reduce price sensitivity for grocery products, and Tan et al. (2022) show that AR-based "virtual try-on" can increase sales of cosmetic products, especially when consumers experience a high level of uncertainty (products from less popular brands, with narrower appeal and higher price). In this paper, we extend this stream of literature by exploring the impact of VR in a context involving high-involvement and heterogeneous products.

Our research also relates closely to the marketing and economics literature studying the impact of virtual tours on house sale outcomes. Theoretical predictions suggest a positive correlation between VTs and sale prices, as higher-quality properties are more likely to use VT (Carrillo 2008) and VTs might help reduce search cost and improve matching (Ford et al. 2005). Empirical research has found that VT is associated with a 2-3% increase in sale prices (Carrillo 2008, Yu et al. 2020). However, the theory is less conclusive regarding the effect of VTs adoption on TOM, as the overall effect might depend on the type of properties that adopt the tool. Empirical evidence is also mixed and ranges from a 20-52% decrease (Carrillo 2008, Yan et al. 2021) to a 3-7% increase (Yu et al. 2020) on the number of days a house stays on the market. In this paper, we extend this literature in two ways. First, our analysis examines the endogenous decision to use VT and sheds light on the factors (house, neighborhood, and seller characteristics) associated with VT adoption and its relationship with other selling strategies (number of photos, description length, listing price, etc.). Second, our methodology leverages computer vision and NLP to capture other factors that would otherwise be unobserved and confound the effect of VT on sale outcomes, such as recent renovations, design and aesthetic quality of the house, etc.

Finally, our research also relates to the broader marketing literature studying the role of visual information on different marketing outcomes. Prior literature has examined questions such as how user-generated pictures can affect brand engagement in social media (Hartmann et al. 2021), how to leverage product images to predict returns (Dzyabura et al. 2019), how consumer-generated pictures can predict restaurant survival (Zhang and Luo 2023), how house photos can affect demand in the sharing economy (Zhang et al. 2022), and how profile pictures can affect matching outcomes in online labor marketplaces (Troncoso and Luo 2022), among others. In this paper, we extend this stream of literature by studying the marginal impact of a newer type of visual information relative to more traditional ones, namely, interactive virtual tours vs. static photos.

3 Data

Our empirical analyses are based on data from Redfin, a major real estate platform that operates in 95 markets in the US and Canada and provides house information obtained from multiple listing services (MLS databases) and public records. Using the platform's API, we collected information on 75,178 houses sold in the greater Los Angeles area (CA) between March 2019 and March 2021. The choice of the region allows us to benchmark our findings to prior literature on the topic while exploiting the diversity of its housing stock in our analyses. The choice of the time window allows us to examine potential disruptions to in-person tours following the COVID-19 pandemic.

For each house in our sample, we use its sale history provided by Redfin to collect our two outcomes of interest, i.e., the sale price and the number of days the house stayed on the market (from list date to sale date), as well as the initial listing price and any price changes the seller made while the house was on the market. Our treatment variable, i.e., whether the listings include a 3D virtual tour, is identified by the presence of a link to this tool. Figure 1 showcases an example of a 3D virtual tour for a listing. Overall, 21.63% of houses adopted VT in our observation window.

To estimate the treatment effects of virtual tour adoption, we collect two sets of control variables. The first is a set of baseline control variables motivated by prior research (Section 3.1). Table 1a defines all these baseline control variables (except the photo count variable), and Table A1a of Web Appendix A presents their summary statistics. For the second set of control variables, we exploit computer vision and NLP methods to uncover additional variables from the photos and the open-ended descriptions on listings that can explain both VT adoption and sales outcomes, i.e., confounders that could bias the estimated effects of VT on sale outcomes. Table 1b defines all these additional photo and text-related variables[3], and Table A1b of Web Appendix A presents their summary statistics.

3.1 Extracting standard house, neighborhood, and agent/firm characteristics

Motivated by prior research, we collect a comprehensive set of baseline control variables relevant to our analyses, which we categorize into the following groups:

House characteristics The literature on real estate and housing markets has extensively acknowledged that factors such as size, number of rooms, and age are important determinants of prices (e.g., Goodman 1978, Palmquist 1984, Selim 2009).

To account for these in our analyses, we extract 20 standard house attributes available on each listing, including property type, square footage, number of bedrooms, number of bathrooms, and age since the year built, among others. Furthermore, we use the text descriptions in the listings to identify whether the house is move-in-ready, a signal that the house does not need any significant repairs that could affect sale prices, and whether the house is vacant, which could translate into a shorter TOM.

[3]The photo count variable belongs to the baseline control variables, but we define it in Table 1b for ease of exposition.

Neighborhood characteristics In addition to the characteristics of the building itself, the price of a house is also affected by the attributes of its location or surrounding area, factors like the proximity of parks or beaches, the quality of nearby schools, or demographics of the population (e.g., Archer et al. 1996, Kolbe et al. 2012, Carrillo 2008).

To account for these in our analyses, we collect variables from multiple sources. First, we extract location-related information available on each listing, including the number and rating of the schools in the area, the walk-ability and bike-ability scores of the area, and the avg. number of offers that houses in the area receive, which reflects the general attractiveness of the area for buyers. Second, we use SafeGraph's Open Census Data to collect census-block-level demographics that could also affect sale outcomes, including population size, medium household income, and proportion of Black and Hispanic populations.[4] Third, we use the Zillow House Value Index to proxy for the average value of houses at the zipcode-year-month level and account for potential time-variant factors that could also affect sale prices, such as upcoming neighborhoods or changes in local crime levels.

We further use house address and real estate firm information on the listing to construct two additional variables: zip code dummies, which can capture any additional time-invariant neighborhood characteristics that could affect sale outcomes, and the proportion of past listings from competitors in the same zip code that include VTs, which can capture competitive forces driving VT adoption.[5]

Agent and firm characteristics Real estate listing agents are key intermediaries in a house selling process, and different agent and firm characteristics like experience or gen-

[4]SafeGraph's Open Census Data includes all data from the American Community Survey 5-year Estimates by Census Block Group (CBG) for the release years 2016 - 2020. We match houses and the census data on latitude and longitude. The data is available at: `https://docs.safegraph.com/docs/open-census-data`.

[5]Prior research has used a similar variable as an instrument for VT adoption (e.g., Carrillo (2008) uses the average number of virtual tours displayed by broker firms within a local area as an instrument). Nevertheless, if VTs indeed affect sale outcomes, this instrument could violate the exclusion restrictions, as the sale outcomes of competitors can also affect the sale outcomes of the focal house (even if not using VT). As such, we use this variable as a control that could affect both VT adoption and sale outcomes in our analyses.

der can directly influence sale outcomes (Beck et al. 2013, Seagraves and Gallimore 2013, Gilbukh and Goldsmith-Pinkham 2019). Moreover, the matching between houses and seller agents/firms is likely selective, with more experienced players likely working with higher-quality houses and more motivated sellers.

To account for the role of seller agent characteristics, we use the name of the agent associated with each listing to proxy for his/her experience and gender.[6] We also capture the number of listings he/she is managing in the same month as a proxy for workload, which could affect his/her ability to close sales quickly. Similarly, we use the name of the real estate firm associated with each listing to proxy for a firm's workload, and the proportion of past listings that include VTs to proxy for familiarity with the technology, which could affect the use of VTs in the future.[7] Further, we use the name of the real estate firm to create firm fixed effects and capture additional time-invariant firm characteristics, such as size, reputation, or resources.

Time dummies. Based on each house sale date, we create year-month dummies to account for time-varying market conditions that could affect sale outcomes, e.g., seasonality effects in the housing market, and that could also affect the probability of adopting VTs, e.g., limited mobility after stay-at-home orders triggered by the COVID-19 pandemic.

3.2 Uncovering potential confounders from photos

The photos included in each house listing can be used to capture two types of confounders. The first type relates to the overall effort a seller agent spends on marketing the house, which can predict both the decision to use VT and sale outcomes. For example, a seller

[6] Following Proserpio et al. (2021), We adopt the name-gender matching method, using a dictionary of name-gender pairs from the R software's gender package. This dataset provides probabilities of name-gender associations based on historical birth databases.

[7] Prior research has used a similar variable as an instrument for VT adoption (e.g., Carrillo (2008) uses the average number of virtual tours displayed by other agents within the same firm as an instrument). Nevertheless, we cannot discard the possibility that firms using VTs are also learning and updating their marketing expertise which could directly impact house sale outcomes. As such, we use this variable as a control that could affect both VT adoption and sale outcomes in our analyses.

agent who spends time and money on VTs might also care about adding "good pictures" to his/her listings, and photo quality and content can directly impact sale outcomes. Indeed, prior research has shown that photo quality can significantly impact demand for short-term rentals (Zhang et al. 2022), and real estate practitioners suggest that photos of the interior and exterior of a house, as well as an image of its floor plan, can make the selling process more efficient. The second type corresponds to unmeasured house characteristics that can predict both the decision to use VT and sale outcomes. For example, properties with more favorable visual appeal, which usually sell at a higher price (Elam and Stigarll 2012), might also be more willing to use VT tours (Carrillo 2008).

To account for these factors in our analyses, we capture the *photo count* on each house listing as a proxy for the amount of visual information the seller agent discloses. Furthermore, we leverage computer vision methods to capture three additional photo-related variables. Web Appendix B presents illustrating examples.

The first variable is the *aesthetic quality* of the photos, which can also serve as a proxy for the visual appeal of a house. To create this variable, we use the Convolutional Neural Network Image Assessment (NIMA), a model that leverages deep object recognition networks to predict human perceptions of the quality of an image and capture semantic-level characteristics associated with emotions and beauty in images (Talebi and Milanfar 2018).

The second variable is the *content distribution* of each photo, which captures the scene type displayed in the photos. To create this variable, we employ the state-of-the-art model for scene recognition, Places-CNNs (Zhou et al. 2017), a Convolutional Neural Network model trained on more than 10 million photos of 400+ scene types (e.g., bedroom, kitchen, yard, swimming pool, etc.)[8] Given that the photos depicting houses on Redfin belong to the same domain of photos for training the model, we use the calibrated model (Zhou et al. 2017) to

[8]An earlier version of this model has been applied in prior marketing research for detecting scene types of Airbnb listing photos (Zhang et al. 2022).

extract the scene types of our photos.[9] Based on the extracted scene types, we then include the top 10 scene types and others.

The third variable is the *content diversity*, which allows us to distinguish between the number and the informativeness of the photos included in a listing. For example, two listings with 10 pictures each could have different degrees of informativeness if, say, the first house has only photos of the bedroom and the second one has photos of the bedroom, kitchen, and outside view of the house. We create this variable based on the Gini coefficient of the distribution of extracted scene types.

3.3 Uncovering potential confounders from open-ended descriptions

The open-ended descriptions seller agents write for a particular house can also capture the two types of confounders discussed above. For example, a seller willing to spend the time and cost to use VT might also put effort into writing detailed and appealing descriptions to attract potential buyers. Moreover, real estate platforms like Redfin and Zillow encourage sellers to use descriptions to convey relevant information that cannot be disclosed with pictures, such as recent upgrades or renovations, details about the design or unique features of the house, etc.

To account for these factors in our analyses, we measure the *word count* in the description, which can proxy for the amount of information the seller provides in the description (e.g., longer descriptions might contain more details about the house) and the effort he/she is putting into selling the house (e.g., writing longer description takes time).

We further leverage NLP methods to capture two additional text-related variables. The first variable is the *sentiment* of the description, which can proxy for the emotional tone in a seller's style that can affect the house-buying process (Levy et al. 2008). To create this

[9]The code to implement this deep CNN image scene recognition tool is available at `https://github.com/CSAILVision/places365/blob/master/run_placesCNN_basic.py`.

variable, we use the lexicon-based approach of the Valence Aware Dictionary and Sentiment Reasoner (VADER) (Hutto and Gilbert 2014).

The second variable is the *topic distribution* of each description, which captures the content that sellers emphasize. To create this variable, we use the Latent Dirichlet allocation (LDA) model (Blei et al. 2003) and uncover seven topics[10] which, based on their most relevant terms shown in Table 2, we interpret as: renovations, condo amenities, design, location, spaciousness, investment opportunity, and details on the interiors. Note that these topics allow us to capture more nuanced house characteristics than standard facts on the listings. For instance, two houses may have the same square footage, but the layout can create a different sense of *spaciousness*. Additionally, two houses in the same zip code might emphasize *location* differently based on their proximity to amenities such as convenience stores, restaurants, or public transportation. Furthermore, a description that mentions *investment opportunities* is likely indicating that the house is in a worse condition and requires significant repairs compared to other similar listings.

4 Methodology

A primary goal of our empirical analyses is to estimate the effect of including a virtual tour (VT) on a house listing on two key sale outcomes, namely (i) the sale price and (ii) the time the house stays on the market (TOM). A major empirical challenge is that the adoption of virtual tours is not random, and it is likely that certain listings may include virtual tours depending on various factors that also affect sale outcomes, such as the design or visual appeal of the house, the desirability of the neighborhood, the experience the seller agent has, or the effort put into selling the house. To address this challenge and account for such selection bias, we adopt an empirical approach that directly links sale outcomes with the decision to use VT through a set of confounders X that might simultaneously affect both.

[10] We determine the optimal number of topics by balancing several criteria, i.e., perplexity score, coherence score, inter-topic distance, and human interpretability.

More specifically, we use an interactive regression model in which the sale outcome for a house i (Y_i = Sale Price_i or Y_i = TOM_i) is a function of whether its listing includes a virtual tour (VT_i) and a set of confounders (X) that also affect the decision of adopting VT:

$$\begin{aligned} \mathrm{Y}_i &= g(\mathrm{VT}_i, \mathbf{X}_i) + U_i \quad , \quad \mathbb{E}(U_i|\mathrm{VT}_i, \mathbf{X}_i) = 0 \\ \mathrm{VT}_i &= m(\mathbf{X}_i) + V_i \quad , \quad \mathbb{E}(V_i|\mathbf{X}_i) = 0 \end{aligned} \tag{1}$$

In our setting, $\mathbf{X}_i$ is a rich set of 1,156 variables comprising baseline controls defined in Table 1a, which include several house, neighborhood, and seller agent's characteristics and time controls, as well as the additional photo and text-related variables defined in Table 1b, which include the additional confounders uncovered from the unstructured data on each listing. The functions $g(\cdot)$ and $m(\cdot)$ that govern the outcome model for Y_i and the selection model for VT_i, respectively, are a priori unknown and potentially complex (e.g., high-level interactions between house characteristics), and the relationship between VT adoption and sale outcomes is potentially heterogeneous on X_i.

To estimate the average treatment effect of VT on sale outcome Y, θ^Y, we leverage recent advances at the intersection of causal inference and machine learning. In particular, we adopt the Double Machine Learning (DML) approach (Chernozhukov et al. 2018), which allows us to approximate functions $g(\cdot)$ and $m(\cdot)$ with flexible and scalable machine learning methods. The DML estimator is a doubly robust estimator that combines both regression adjustment and propensity score weighting strategies, and it is robust to misspecifications in either the outcome or selection model:

$$\begin{aligned} \hat{\theta}^Y_{DML} &= \frac{1}{N}\sum_i^N \psi_i^Y(Y_i, VT_i, \mathbf{X}_i) \\ \psi_i^Y(Y_i, VT_i, \mathbf{X}_i) &= \hat{g}(1, \mathbf{X}_i) - \hat{g}(0, \mathbf{X}_i) + \frac{\mathrm{VT}_i(\mathrm{Y}_i - \hat{g}(1, \mathbf{X}_i))}{\hat{m}(\mathbf{X}_i)} - \frac{(1-\mathrm{VT}_i)(\mathrm{Y}_i - \hat{g}(0, \mathbf{X}_i))}{1-\hat{m}(\mathbf{X}_i)} \end{aligned} \tag{2}$$

In practice, the DML estimation procedure consists of two steps. The first step is to use machine learning methods to estimate $g(\cdot)$ and $m(\cdot)$ with cross-validation/sample-splitting. The second step is to compute the orthogonalized scores or individual treatment effect (ITE) signals (ψ_i^Y in Equation 2), which can be averaged across individuals to obtain $\hat{\theta}^Y_{DML}$. The two critical ingredients of the estimation procedure, cross-fitting and orthogonality, help eliminate regularization and over-fitting biases from the machine learning methods.

In our implementation of the DML approach, we use the XGBoost algorithm (Chen et al. 2015) to estimate $g(\cdot)$ and $m(\cdot)$ following standard practices in the literature. The choice of this algorithm is motivated by its flexibility in capturing potentially nonlinear relationships and its excellent performance in many predictive tasks.[11] In our analyses, we focus on the first step (i.e., selection equation $m(\cdot)$ in Equation 1) to explore the most important factors that explain VT adoption (see Section 5), the final estimates $\hat{\theta}^Y_{DML}$ to quantify the average effects of adopting VT on sale outcomes (see Section 6.1), and the individual treatment effects to explore potential sources of heterogeneity in the effects of VT on sale outcomes (see Section 6.2).

As with many other methods to adjust for confounders in observational studies (e.g., matching, propensity score weighting, regression adjustment), the DML estimator relies on two identification assumptions. The first assumption is unconfoundedness, which requires that all confounders be measured and accounted for in the model (i.e., $\mathbb{E}(U_i|\mathrm{VT}_i, \mathbf{X}_i) = 0$ and $\mathbb{E}(V_i|\mathbf{X}_i) = 0$ in Equation 1). Despite our efforts to leverage computer vision and NLP to uncover additional confounders from the text and photos in the listings, we cannot empirically discard the existence of other unmeasured confounders not captured by this approach. As such, as we explain after presenting our results, our estimates might not be interpreted as *unbiased* but rather as *after removing meaningful sources of biases* previously overlooked in the literature (e.g., visual appeal, condition of the house, and seller's effort).

[11] The superior predictive performance of this algorithm has been widely recognized across many machine learning challenges held by Kaggle and KDD cups (Chen and Guestrin 2016). Several prior marketing studies have also used XGBoost (e.g., Rafieian and Yoganarasimhan 2021, Zhang and Luo 2023) or similar gradient boosting trees methods (Yoganarasimhan 2020) for complex customer behavior prediction problems.

The second identification assumption is sufficient overlap, which requires that all units in the data have a positive probability of being assigned to each condition (i.e., $0 < m(\mathbf{X}_i) < 1$). In Web Appendix C, we present an empirical assessment supporting the validity of this assumption. Moreover, and following Ellickson et al. (2022), we trim the dataset at the 0.01 and 0.99 level of the propensity scores $\hat{m}(\mathbf{X}_i)$ to ensure that our results are robust to extreme values of obtained treatment propensities.

5 What factors explain VT adoption?

In this section, we focus on the selection equation (i.e., $m(\cdot)$ in Equation 1) to explore what factors can help explain VT adoption. In particular, we are interested in examining: (i) the relative importance of factors previously studied in the literature (e.g., house, neighborhood, and seller agent characteristics) and the potential confounders uncovered from the photos and descriptions on each listing; and (ii) how the use of VT relates to other marketing efforts in the listing, such as the quality and content of the photos and descriptions of the listings.

To this aim, we use SHAP values (Lundberg et al. 2020) to examine the importance of different features in explaining VT adoption. In Table 3, we summarize the aggregated importance for each of the different groups of variables. At the aggregated level, the most important group of explaining factors of VT adoption are the 14 photo-related variables, which add up to 26.60% importance. The second most important group are the 24-year-month fixed effects, which add up to 22.90% importance. The importance of this group is expected given the trend we observe in the data: the percentage of listings that included a VT almost tripled after the stay-at-home orders triggered by the World Health Organization announcement of the COVID-19 pandemic in March 2020 (see Figure 2). The third most important group are the eight text-related variables extracted from the descriptions of the listings, which add up to 14.34% importance. Notably, the photo and text-related variables uncovered from the unstructured data on each listing are more important in explaining VT adoption than the standard house and neighborhood characteristics typically used in the literature.

Next, we examine each feature's importance in explaining VT adoption. In Figure 3, we summarize our main results for the top 30 most important features that are not fixed effects variables, and use color code to indicate the correlation between each feature and VT adoption. At the individual level, the most important explaining factors of VT adoption are the average aesthetic quality score, the log of the number of photos on the listing, and the log number of words in the description. By looking at the sign of the correlations, we find that the probability of adopting VT increases with the aesthetic quality of the photos included in the listing, which suggests that sellers strategically decide to use VT for properties with a higher visual appeal or simply do a better job at providing more higher-quality photos. We also note that the probability of adopting VT increases with the number of photos and words in the description, which suggests that sellers who use VTs put significant effort into describing/promoting the house and use VTs to supplement rather than replace more traditional sources of information.

Figure 3 also allows us to explore how VT adoption relates to other house, neighborhood, and seller agent/firm characteristics. For example, we find that the probability of adopting VT increases with the proportion of listings from both the focal firm and competitors that featured VT in the past month, which suggests that sellers may have both internal and competitive incentives to adopt VT. Additionally, we find that the likelihood of adopting VT increases with the average house value estimate (as measured by the Zillow Home Value Index) and the number of schools in the area, which suggests that sellers might strategically opt for VT for houses in more "desirable" neighborhoods with the potential to sell at a higher price. Furthermore, we note that the probability of adopting VT increases with the seller agents' experience and the number of listings a seller agent manages in the same month, which could suggest that the technology might enable them to grow or manage their listing portfolio. We provide additional evidence supporting this claim in Section 6.3.

Finally, we assess the validity of leveraging unstructured data to uncover confounders to reduce selection bias. Thus, and following Zeng et al. (2022), we examine whether the photo and text-related variables from the unstructured data on each listing are important explaining factors of not only the treatment (VT adoption) but also the outcomes of interest (sale price and TOM). In Figure A3 and Figure A4 of Web Appendix D, we illustrate the top 30 most important variables in explaining sale price and TOM, respectively. Notably, and as highlighted with "*/**" in Figure 3, many of the photo and text-related variables that are important explaining factors of VT adoption are also important explaining factors of sale price and TOM. Moreover, the signs of correlations between these variables can provide some preliminary evidence of the direction of the selection bias. For instance, listings with a higher proportion of spaciousness and design topics, which are also more likely to adopt VT, tend to sell at a higher price. Ignoring such positive correlations would overestimate the effect of VT on sale prices. Additionally, listings with a higher proportion of living room photos and a higher proportion of spaciousness topic, which are more likely to adopt VT, tend to have a shorter TOM. Neglecting such negative correlations would underestimate the effect of VT on TOM.

In sum, our analyses provide three insights. First, sellers use VTs to supplement more traditional sources of information in the listings (photos and descriptions). Second, sellers might strategically opt for VT for houses in more "desirable" neighborhoods or when managing multiple listings. Third, the unstructured data on each listing allow us to identify confounders that are hard to measure and have been overlooked in previous studies, such as photo aesthetic quality, content, and description topics.

6 How does VT adoption relate to sale outcomes?

In this section, we investigate the relationship between VT adoption and sale outcomes. We first present the estimates of the average treatment effects of VT adoption on sale outcomes and discuss the potential limitations and robustness of our findings (Section 6.1). Then, we dig deeper into when and how VT adoption can be more beneficial for sellers by exploring potential sources of heterogeneity in the effects of VT adoption on sale outcomes (Section 6.2) and by examining the other potential benefits of using VTs (Section 6.3).

6.1 The average effect of VT adoption on sale outcomes

In Table 4, we present the estimated average treatment effect of VT on sale prices and TOM, i.e., the DML estimator $\hat{\theta}^{Y}_{DML}$ in Equation 2. In column 1 and column 2, we start with a baseline model specification that, similar to prior literature, only controls for house characteristics, neighborhood characteristics, agent/firm characteristics, time dummies, and the number of photos included in the listing. Under this baseline specification, we find that using VT is associated with a 1.1% increase in the sale price (column 1) and a non-significant change in the time the house stays on the market (column 2). The former estimate is consistent in sign but smaller in magnitude than the 2%-3% previously found in the literature (Carrillo 2008, Yu et al. 2020), a difference arguably driven by the richer agent and firm characteristics we control for. The latter estimate is also consistent with the null or positive effect found in studies of markets close to ours (Yu et al. 2020).

Next, in columns 3 and 4 of Table 4, we consider a full model specification that incorporates the additional confounders uncovered from the photos and descriptions on each listing. Column 3 indicates that after controlling for factors such as the photo aesthetic quality and proportion of spaciousness topic, which are positively correlated with sale prices and VT adoption, the effect of VT on sale prices is negligible. Column 4 indicates that after controlling for factors such as spaciousness (or location) discussed in the descriptions, which are negatively (positively) associated with TOM but positively (negatively) associated with VT adoption, the effect of VT on TOM turns positive and significant.[12]

[12]Our definition of TOM is based on the number of days between the listing date and the sale date. In Web Appendix E, we show that our main findings are consistent when using an alternative definition used in prior literature and the real estate industry, i.e., the number is the number of days between the listing date

In summary, our findings challenge the prevailing belief that using VTs can substantially improve house sale outcomes. However, like many observational studies, our findings have limitations. Despite our efforts to capture confounders previously overlooked in the literature from the unstructured data on each listing, we cannot entirely rule out the possibility of other unmeasured confounders biasing our estimates. Therefore, our estimates should not be interpreted as *unbiased* but rather as *after removing meaningful sources of biases* that were overlooked in previous literature, such as the visual appeal and quality of the house and the seller's effort in providing visual and textual information about the property. Below, we summarize two sets of analyses that assess the robustness of our findings to confounders such as reference prices or different seller's pricing strategies, and the sensitivity of our findings to other unmeasured confounders.

Robustness checks with additional measured confounders Major online real estate platforms like Redfin and Zillow provide algorithmic estimates of each house's market value, which help reduce sellers' and buyers' uncertainty about house value (Fu et al. 2023). By serving as a reference point, these estimates might affect both sale prices and the decision to use VT. For example, because VTs are costly, sellers may be more likely to use them for houses they are predicted to sell at a higher price. To account for this in our analysis, we further control for the Redfin estimate of each house value, the most accurate algorithm-based estimate of home value available online.[13] Our results, reported in Table A7 of Web Appendix F are consistent with our main results in Table 4, i.e., they suggest that controlling for additional photo and text variables reduces the claimed benefits of VT on sale outcomes.[14]

and pending contract date, which tends to be significantly shorter but ignores the non-trivial percentage of pending contracts that fail to result in a completed transaction.

[13]Redfin estimates are based on more than 500 data points about the market, the neighborhood, and the home itself, and have a median error rate of 2.16% among for-sale homes and 6.80% among off-market homes. For more details, see: `https://www.redfin.com/redfin-estimate`.

[14]One concern with these home-value estimates is that they were collected after the properties in our sample were sold. Recent research suggests that these estimates change in response to recent market activity, list prices, and sale prices (Fu et al. 2022, Malik and Manzoor 2023). As such, these estimates not only reflect home value but might also pick up any potential effect of VTs on sellers' list price decisions or sale prices. However, we hope that our findings regarding the small effect of VT on list prices (see Table A8 in Web Appendix F) and the insignificant effect of VT on sale prices (see Table 4) help mitigate this concern.

Another type of confounder that could threaten the interpretation of our findings relates to pricing strategies consistent with the results we observe. For example, a seller who uses VT might also start with a higher listing price if she/he wants to increase the probability of recouping their costs, and higher listing prices reduce the pool of potential buyers and extend the time a house stays on the market (Knight 2002). To explore this possibility, we use the same DML framework described above but replace the outcome variable in Equation 1 by (i) the initial list price, (ii) the number of times the seller decreased the list price, and (ii) the number of times the seller increased the list price. Our results, reported in Table A8 of Web Appendix F, indicate that VT adoption is associated with higher list prices, more frequent subsequent price cuts, and fewer subsequent price rises, thus, pricing strategies that are more likely to result in longer TOM.[15] Moreover, in Web Appendix F, we use a different approach to show that such pricing strategies fully mediate the increase in TOM associated with VT adoption. These results suggest that accounting for these confounders does not change our main finding that the benefits of VT on sale outcomes are smaller than previously thought.

Sensitivity to other unmeasured confounders Although we cannot precisely quantify how other unmeasured confounders would impact our findings, theoretical knowledge and market intuition can help us speculate on the direction of the bias they would introduce. For instance, theoretical "unraveling predictions" suggest that sellers may be more willing to pay the cost of using VTs for properties they anticipate will benefit from disclosing additional visual information, e.g., higher quality houses that will sell at a higher price (Carrillo 2008). As a result, unmeasured confounders are likely to positively bias the effect of VT on sale prices, which is still in line with our key finding that such an effect is smaller than previously thought.

[15] A higher list price followed by more frequent price cuts is a pricing strategy commonly employed by sellers when uncertain about a house's value (Lazear 1986), which might suggest that sellers may use VTs for unique/atypical houses with uncertain value. Nonetheless, our estimates indicate that VT adoption is associated with a 0.3% increase in listing prices, equivalent to a \$2,500 increase in the average list price. Such a moderate increase might be less likely to reflect significant differences in house characteristics and more likely to reflect sellers trying to recoup the cost of creating a VT (typically \$300-\$1000) or overestimating the benefits of VT on sale price.

However, the impact of unmeasured confounders on our estimated effect of VT on TOM is less evident, as it might depend on whether sellers use it to accelerate the selling process of already "popular houses" that they anticipate selling fast, which could negatively bias our estimates, or to boost interest for "niche houses" that they anticipate taking longer to sell, which could positively bias our estimates. Since buyers tend to perceive longer than average TOM as a bad signal of the house quality (Tucker et al. 2013, Fu et al. 2022), the benefits of the latter are arguably more prominent than the former. As such, the effect of VT on TOM is more likely to be positively than negatively biased, which could change our finding on the positive effect of VT on TOM.

To further examine the sensitivity of the estimated effect of VT on TOM, we follow prior literature (Manchanda et al. 2015, Zhang et al. 2022) and adopt the Rosenbaum Bounds Approach (Rosenbaum 2002). As shown in Table A11 of Web Appendix G, we find that unmeasured confounders of small magnitude can explain away the positive and significant effect of VT on TOM. Despite changing our conclusion on the direction of the effect of VT on TOM from positive to insignificant, these results still challenge the prevailing belief that VTs can help substantially reduce TOM.

6.2 Heterogeneity of the effect of VT adoption on sale outcomes

Thus far, our analyses suggest that, on average, VT adoption does not substantially benefit house sale outcomes. In this section, we leverage the DML framework to explore potential sources of heterogeneity of the effect along different dimensions that can shed light on when and how sellers can use VTs more effectively, including other listing marketing efforts (i.e., amount and quality of visual and textual information), seller agent and firm characteristics, and house and neighborhood characteristics. We also examine differences before and after the COVID-19 pandemic, which accelerated VT adoption and potentially altered consumer search and purchase behavior.

Following prior literature that has studied treatment effect heterogeneity in marketing settings (e.g., Ascarza 2018, Guo et al. 2020, Ellickson et al. 2022), we compare the estimated individual treatment effect (ITE) signals or orthogonalized scores ψ_i^Y in Equation 2 among different groups. To simplify our analysis, we focus on one potential source of heterogeneity (Z) at a time and compare the ITE signals of groups with high (above median) and low (below median) values of Z. Table 5 summarizes the averages per group, the p-value of differences between groups, and highlights in bold the group that benefits significantly more from using VT (i.e., obtains higher sale price or shorter TOM). Below, we highlight some of the main significant differences on each dimension.

Differences by other listing efforts Rows 1 and 2 of Table 5 indicate that the estimated treatment effect of VT on sale prices is smaller for listings with a high photo count or a high word count. Rows 3, 4, and 6 of Table 5 indicate that the estimated treatment effect of VT on sale prices is also smaller for listings with high aesthetic quality photos and descriptions with a higher proportion of design or spaciousness topics. These results suggest that the benefits of using VT on sale price diminish when sellers have already provided ample and high-quality information about the house through photos and text, or when houses are already visually attractive due to their aesthetic, design, or spaciousness.

Differences by seller agent and firm characteristics Rows 7 and 8 of Table 5 indicate that the estimated treatment effect of VT on price is larger for agents and firms with a lower number of listings, and the estimated treatment effect of VT on TOM is smaller for firms with a lower number of listings. These results suggest that the benefits of using VTs are higher for smaller seller agents and firms.

Differences by house and neighborhood characteristics Rows 9 and 10 of Table 5 indicate that the estimated treatment effect of VT on TOM is smaller for Townhouses than for Condos, and smaller for Single Family houses than for Condos. Since Condos are typically more standardized than Townhouses and Single Family houses, these results suggest that the benefit of using VT on TOM is smaller when the properties are more homogeneous.

Rows 11 and 12 of Table 5 indicate that the estimated treatment effect of VT on sale prices is higher for houses in neighborhoods with a higher proportion of Black and Hispanic populations or lower median household income, and the estimated treatment effect of VT on TOM is smaller for these same groups. These results suggest that the benefits of using VT on both sale price and TOM are higher for houses in neighborhoods with more racial minorities and lower-income populations. Although our empirical setting does not allow us to identify the precise mechanism behind this finding, one possible explanation is that 3D virtual tours (VTs) may be particularly valuable in mitigating the greater uncertainty surrounding property values in these neighborhoods. Additionally, we note that listings in these neighborhoods typically feature fewer photos with lower aesthetic quality and are more likely to be managed by smaller seller agents or firms, which based on our findings above, also tend to benefit more from VT adoption.

Differences before and after COVID-19 pandemic Row 14 of Table 5 indicates the effect of VT on TOM was larger in the period before than after the COVID-19 pandemic, suggesting that using VTs might have been more beneficial for sellers when mobility was restricted and in-person tours were not allowed. While the pandemic generated significant disruptions that permanently changed consumption in low-involvement categories such as restaurant delivery (Oblander and McCarthy 2022) or music (Sim et al. 2022), it is not immediately clear whether changes in high-involvement categories such as the housing market will be permanent too. To further investigate this, we regress the individual treatment effect (ITE) signals ψ_i^Y in Equation 2 against dummies indicating the number of months after the pandemic (April 2020 is the first month after the pandemic). Our results, reported in Table 6, suggest that the effect of using VT on sale price is not significant before the stay-at-home orders and does not change significantly afterward, and that the effect of using VT on TOM was positive before the stay-at-home orders but decreases significantly and turns negative for houses sold two or three months afterward. Therefore, using VT helped to reduce TOM for properties sold in May or June 2020, but the benefits were short-term and returned to pre-pandemic levels after that period.

In sum, our findings uncover conditions when using VTs could be more or less beneficial for sale outcomes. In particular, VTs can provide higher benefits in terms of sale prices and TOM for sellers who need to minimize other listing marketing efforts (i.e., fewer pictures or shorter descriptions), for houses in neighborhoods with a higher proportion of racial minorities or lower-household income, and seller agents/firms of smaller sizes. It is worth noting that even in these most beneficial cases, the effect of VTs on sale outcomes is modest and reaches a maximum of a 0.6% increase in sale price. One exception is the more sizable 13.5% (17.8%-4.3%) reduction in TOM during the COVID-19 pandemic stay-at-home order, but this benefit was short-term and returned to pre-pandemic levels after a few months.

6.3 Other potential benefits of VT adoption

Thus far, our findings challenge the prevailing belief that VTs can significantly improve sales outcomes and suggest that the benefits of VTs are modest, even in the best-case scenarios. These results raise the question of why this costly tool has gained so much popularity in the real estate industry. In this section, we explore other incentives sellers might have to use VTs as a marketing tool. For instance, if VTs help home buyers evaluate options more efficiently, these tools might help seller agents or firms filter out less-serious buyers and potentially make their selling process more efficient. Or, if homeowners perceive VT as a signal of tech-savviness or experience, using VT might allow seller agents or firms to attract more customers and increase their portfolio of listings.

The relationship between VTs and different stages of the home buying process
First, we examine the relationship between VTs and different stages of the home buying process and its potential implications for sellers. To this aim, we examine the relationship

between VT adoption and several metrics provided by Redfin that summarize how users interact with listings, including the number of users that have: (i) favorited a listing they want to track, (ii) crossed out a listing they are no longer interested in, and (iii) scheduled a tour to visit the property with a Redfin buyer agent. To examine how VT adoption relates to these metrics, we use the same DML framework described in Section 4, using these metrics as the outcome variable in Equation 1.

We present our results in Table 7. Column 1 indicates that listings with VTs have a higher favorite count, suggesting that VTs help sellers attract more interest from potential buyers. Such a finding is consistent with Yan et al. (2021), who find that listings with VT receive higher web traffic and likes on a Chinese real estate platform. Notably, column 2 indicates that listings with VTs also have a higher cross-out count, suggesting that VTs help consumers evaluate and screen out their alternatives. Lastly, column 3 indicates that listings with VTs have a higher Redfin tour request count, suggesting that buyers still prefer to tour a house in person as they continue their decision-making process, which could explain the limited benefits of using VTs on final sale outcomes.[16]

We acknowledge that the aggregated metrics used for this analysis have limitations in tracking individual home buyers' use of VTs at different stages of their decision-making process and only capture the interest a house receives through Redfin. Nevertheless, and to the best of our knowledge, these results are among the first suggestive evidence that: (i) buyers use VTs in the early stages of the search process to evaluate and screen out properties more efficiently but are likely to continue the rest of their journey as they typically would and (ii) sellers who use VTs might attract more interest to their listing and also screen out less interested buyers.

[16] One concern with these results is that these outcomes are a function of TOM, i.e., houses that are listed for longer should naturally get more views and interactions from users, and at the same time, TOM might be affected by VT adoption. Nonetheless, in Web Appendix H we show that these results are consistent after normalizing these outcomes by the number of days the house has been on the market.

The relationship between VT adoption and listing portfolio size Next, we examine the possibility that VTs help seller agents and firms expand their listing portfolio or manage a larger listing portfolio. To this aim, we run a regression where the dependent variable is the log-transformed number of properties a seller agent/firm listed in a given month, and the independent variable is the proportion of those listings that use VTs.

We present our results in Table 8. At the firm level, column 2 indicates that after controlling for firm and year-month fixed effects, using VT in all listings is associated with a 2.4% increase in the average number of listings a firm manages in the same month. Similarly, the estimates in column 4 suggest that after controlling for firm and year-month fixed effects, using VT in all listings is associated with a 1.8% increase in the number of listings an individual agent manages in a given month.

We acknowledge that these results can only be interpreted in a correlational fashion. Nevertheless, and to the best of our knowledge, these results are among the first suggestive evidence that sellers can use VTs to expand or manage their listing portfolios.

7 Conclusions

Virtual reality (VR) has become increasingly prevalent in many sectors. One type of VR tool that has gained significant popularity in the real estate industry is 3D virtual tours (VT), which allow buyers searching for properties online to virtually walk through properties online and promise sellers boosted sale prices and reduced time on the market. Nonetheless, empirical research on the topic has been limited and has paid little attention to what factors explain VTs adoption and VTs' marginal impact on sale outcomes, especially relative to other marketing efforts. In this paper, we conduct an empirical study to fill in some of these gaps and arrive at four main findings. First, we find that sellers who use VT are also more likely to include longer descriptions, more photos, and higher aesthetic photos on a house listing, suggesting that sellers use VTs to supplement rather than replace more traditional sources of information. We also note that VTs are more likely to be adopted for houses located in areas of higher value and that the adoption of VT increased rapidly in the months following the COVID-19 lockdown policies. Second, we find that after controlling for interpretable confounders uncovered from the text and photos on the listing, which can proxy for the visual appeal and quality of the house and the effort made by the seller in providing visual and textual information about the property, VTs do not improve sale prices and TOM on average. Third, we find that VTs can better improve sale outcomes for sellers who need to minimize other listing marketing efforts (i.e., fewer pictures or shorter descriptions), for houses in neighborhoods with a higher proportion of Black and Hispanic populations or lower average household income, and sellers agents/firms of smaller size. Lastly, we provide suggestive evidence that VT adoption may provide benefits other than improving sale outcomes, such as helping consumers screen their options efficiently and helping sellers grow and manage their portfolios effectively.

Our research contributes to different streams of literature. First, it contributes to the growing literature on the impact of using virtual reality in the marketplace (e.g., Gallino and Moreno 2018, Meißner et al. 2020, Tan et al. 2022). Specifically, our research highlights that the benefits of using VT tools might be more moderate for high involvement and high-risk products/contexts. Second, our findings contribute to the marketing and economics literature about the impacts of VTs on sale outcomes in the real estate industry (e.g., Carrillo 2008, Yan et al. 2021). Inspired by recent literature (e.g., Keith et al. 2020, Zeng et al. 2022) that utilizes unstructured text data to uncover otherwise unobserved confounders, we leverage both unstructured text and photo data to investigate the factors associated with VT adoption as well as the marginal value of VTs. Lastly, our research adds to the marketing literature on the role of visual information in different marketing contexts (e.g., Hartmann et al. 2021, Zhang et al. 2022, Zhang and Luo 2023). More specifically, it examines the marginal value of newer visual technologies relative to more traditional ones and other marketing strategies.

Our research offers several guidelines on when and how to use VT to obtain higher benefits. First, we find that providing a variety of high-quality photos and rich textual information can be equally or more effective than adding VTs; hence, sellers working on a tight budget/schedule might consider focusing on the former. Second, we find that VTs can better benefit sale outcomes for properties in neighborhoods with higher Black and Hispanic populations and lower average household incomes; hence, sellers and policymakers might consider adopting/encouraging the adoption of VT in those neighborhoods. Last, VT creators might consider adjusting their clients' expectations regarding sale outcome improvements and highlight other benefits these tools can provide, such as screening out less interested buyers and more efficiently managing and growing their listing portfolio.

Our research also provides some fruitful directions for future work. First, despite our best efforts to leverage the unstructured data on each listing to adjust for as many confounders as possible, our findings cannot be interpreted as *unbiased* but rather as *after removing meaningful sources of biases* that were overlooked in previous literature (e.g., the visual appeal and quality of the house and the effort made by the seller in providing visual and textual information about the property). Future research could collaborate with real estate platforms or firms to run field experiments that randomize VT adoption to obtain unbiased estimates of their effect on sale outcomes. Second, while we study the role of VTs in house sales, future research could explore the role of VTs in the context of rentals, as VTs are already available on long-term rental platforms like Apartments.com and could be potentially adopted for short-term rental platforms like Airbnb. In particular, it would be interesting to examine whether the benefits of using VTs increase in these lower-risk settings. Finally, while we provide some initial suggestive evidence of the benefits other than sale outcome improvements that VT can bring to sellers, more research is needed to understand how these tools can help the process of buying and selling a house. For example, it might be interesting to explore whether web traffic and engagement information from the VT links can help sellers segment and retarget interested buyers, etc.

References

Archer, Wayne R, Dean H Gatzlaff, David C Ling. 1996. Measuring the importance of location in house price appreciation. *Journal of Urban Economics* **40**(3) 334–353.

Ascarza, Eva. 2018. Retention futility: Targeting high-risk customers might be ineffective. *Journal of Marketing Research* **55**(1) 80–98.

Austin, Peter C. 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine* **28**(25) 3083–3107.

Austin, Peter C. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research* **46**(3) 399–424.

Beck, Jason, Michael Toma, Anew Page. 2013. Real estate brokerage firm characteristics and home sales price outcomes. *International Journal of Business and Social Science* **4**(13).

Benefield, Justin D, William G Hardin. 2015. Does time-on-market measurement matter? *The Journal of Real Estate Finance and Economics* **50** 52–73.

Benefield, Justin D, C Stace Sirmans, G Stacy Sirmans. 2019. Observable agent effort and limits to innovation in residential real estate. *Journal of Real Estate Research* **41**(1) 1–36.

Blei, David M, Andrew Y Ng, Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* **3**(Jan) 993–1022.

Carrillo, Paul E. 2008. Information and real estate transactions: the effects of pictures and virtual tours on home sales. *Department of Economics, The George Washington University, Washington, DC. Luettu* **14** 2016.

Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, et al. 2015. Xgboost: extreme gradient boosting. *R package version 0.4-2* **1**(4) 1–4.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, James Robins. 2018. Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. *The Econometrics Journal* **21**(1).

DiPrete, Thomas A, Markus Gangl. 2004. 7. assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology* **34**(1) 271–310.

Dzyabura, Daria, Siham El Kihal, Marat Ibragimov. 2019. Leveraging the power of images in predicting product return rates. *Available at SSRN 3209307*.

Elam, Emmett, Andrea Stigarll. 2012. Landscape and house appearance impacts on the price of single-family houses. *Journal of Environmental Horticulture* **30**(4) 182–188.

Ellickson, Paul B, Wreetabrata Kar, James C Reeder III. 2022. Estimating marketing component effects: Double machine learning from targeted digital promotions. *Marketing Science* .

Ford, James Scott, Ronald C Rutherford, Abdullah Yavas. 2005. The effects of the internet on marketing residential real estate. *Journal of Housing Economics* **14**(2) 92–108.

Fu, Runshan, Yan Huang, Nitin Mehta, Param Vir Singh, Kannan Srinivasan. 2023. Unequal impact of Zestimate on the housing market. *Available at SSRN 4480469*.

Fu, Runshan, Ginger Zhe Jin, Meng Liu. 2022. Human-algorithm interactions: Evidence from zillow. com. *Working paper 29880. National Bureau of Economic Research*.

Gallino, Santiago, Antonio Moreno. 2018. The value of fit information in online retail: Evidence from a randomized field experiment. *Manufacturing & Service Operations Management* **20**(4) 767–787.

Gilbukh, Sonia, Paul S Goldsmith-Pinkham. 2019. Heterogeneous real estate agents and the housing cycle. *Available at SSRN 3436797*.

Goodman, Allen C. 1978. Hedonic prices, price indices and housing markets. *Journal of urban economics* **5**(4) 471–484.

Guo, Tong, Srinivasaraghavan Sriram, Puneet Manchanda. 2020. "Let the sunshine in:" The impact of industry payment disclosure on physician prescription behavior. *Marketing Science* **39**(3) 516–539.

Hartmann, Jochen, Mark Heitmann, Christina Schamp, Oded Netzer. 2021. The power of brand selfies. *Journal of Marketing Research* **58**(6) 1159–1177.

Hutto, Clayton, Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the international AAAI conference on web and social media*, vol. 8. 216–225.

Keith, Katherine A, David Jensen, Brendan O'Connor. 2020. Text and causal inference: A review of using text to remove confounding from causal estimates. *arXiv preprint arXiv:2005.00649*.

Knight, John R. 2002. Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real estate economics* **30**(2) 213–237.

Kolbe, Jens, Rainer Schulz, Martin Wersing, Axel Werwatz. 2012. Location, location, location: Extracting location value from house prices. Tech. rep., SFB 649 discussion paper.

Lazear, Edward P. 1986. Retail pricing and clearance sales. *American Economic Review* **76**(1) 14–32.

Levy, Deborah, Laurence Murphy, Christina KC Lee. 2008. Influences and emotions: exploring family decision-making processes when buying a house. *Housing Studies* **23**(2) 271–289.

Lundberg, Scott M, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, Su-In Lee. 2020. From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence* **2**(1) 56–67.

Malik, Nikhil, Emaad Manzoor. 2023. Does machine learning amplify pricing errors in the housing market? The economics of machine learning feedback loops. *arXiv preprint arXiv:2302.09438*.

Manchanda, Puneet, Grant Packard, Adithya Pattabhiramaiah. 2015. Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Science* **34**(3) 367–387.

McCaffrey, Daniel F, Beth Ann Griffin, Daniel Almirall, Mary Ellen Slaughter, Rajeev Ramchand, Lane F Burgette. 2013. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in medicine* **32**(19) 3388–3414.

Meißner, Martin, Jella Pfeiffer, Christian Peukert, Holger Dietrich, Thies Pfeiffer. 2020. How virtual reality affects consumer choice. *Journal of Business Research* **117** 219–231.

Oblander, Shin, Daniel McCarthy. 2022. Frontiers: Estimating the long-term impact of major events on consumption patterns: Evidence from covid-19. *Marketing Science* .

Palmquist, Raymond B. 1984. Estimating the demand for the characteristics of housing. *The Review of Economics and Statistics* 394–404.

Park, Sungho, Sachin Gupta. 2012. Handling endogenous regressors by joint estimation using copulas. *Marketing Science* **31**(4) 567–586.

Proserpio, Davide, Isamar Troncoso, Francesca Valsesia. 2021. Does gender matter? the effect of management responses on reviewing behavior. *Marketing Science* **40**(6) 1199–1213.

Rafieian, Omid, Hema Yoganarasimhan. 2021. Targeting and privacy in mobile advertising. *Marketing Science* **40**(2) 193–218.

Rosenbaum, Paul R. 2002. *Overt bias in observational studies*. Springer.

Seagraves, Philip, Paul Gallimore. 2013. The gender gap in real estate sales: Negotiation skill or agent selection? *Real Estate Economics* **41**(3) 600–631.

Selim, Hasan. 2009. Determinants of house prices in turkey: Hedonic regression versus artificial neural network. *Expert systems with Applications* **36**(2) 2843–2852.

Sim, Jaeung, Daegon Cho, Youngdeok Hwang, Rahul Telang. 2022. Frontiers: virus shook the streaming star: estimating the covid-19 impact on music consumption. *Marketing Science* **41**(1) 19–32.

Talebi, Hossein, Peyman Milanfar. 2018. Nima: Neural image assessment. *IEEE transactions on image processing* **27**(8) 3998–4011.

Tan, Yong-Chin, Sandeep R Chandukala, Srinivas K Reddy. 2022. Augmented reality in retail and its impact on sales. *Journal of Marketing* **86**(1) 48–66.

Troncoso, Isamar, Lan Luo. 2022. Look the part? The role of profile pictures in online labor markets. *Marketing Science* .

Tucker, Catherine, Juanjuan Zhang, Ting Zhu. 2013. Days on market and home sales. *The RAND Journal of Economics* **44**(2) 337–360.

Yan, Zhenbin, Zixuan Meng, Yong Tan. 2021. How does virtual reality matter? Evidence from an online real estate platform. *Working paper*.

Yoganarasimhan, Hema. 2020. Search personalization using machine learning. *Management Science* **66**(3) 1045–1070.

Yu, Wei, Zhongming Ma, Gautam Pant, Jing Hu. 2020. The effect of virtual tours on house price and time on market. *Journal of Real Estate Literature* **28**(2) 133–149.

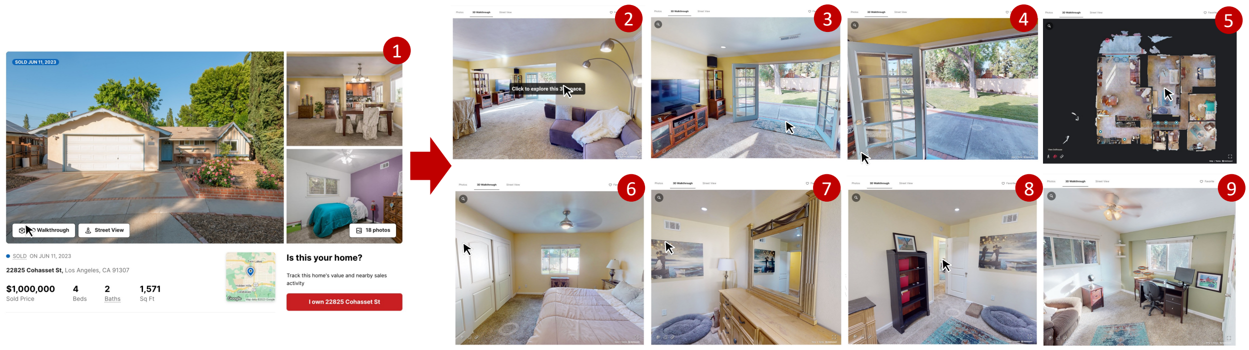
Zeng, Jiaming, Michael F Gensheimer, Daniel L Rubin, Susan Athey, Ross D Shachter. 2022. Uncovering interpretable potential confounders in electronic medical records. *Nature communications* **13**(1) 1–14.

Zhang, Mengxia, Lan Luo. 2023. Can consumer-posted photos serve as a leading indicator of restaurant survival? Evidence from yelp. *Management Science* **69**(1) 25–50.

Zhang, Shunyuan, Dokyun Lee, Param Vir Singh, Kannan Srinivasan. 2022. What makes a good image? airbnb demand analytics leveraging interpretable image features. *Management Science* **68**(8) 5644–5666.

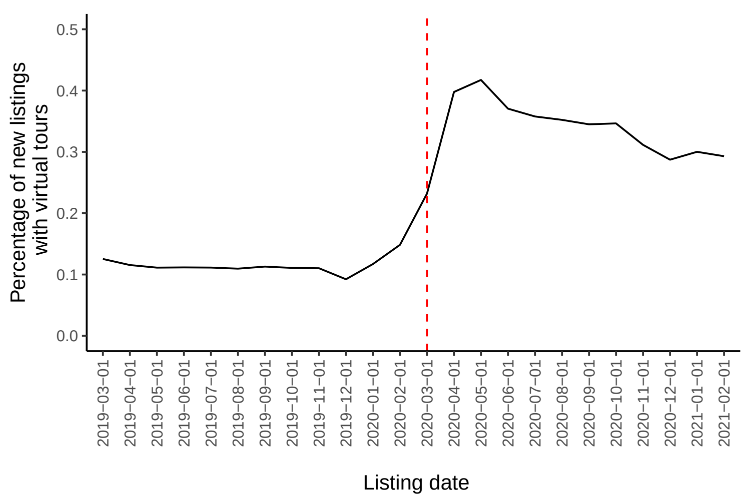
Zhou, Bolei, Agata Lapedriza, Aditya Khosla, Aude Oliva, Antonio Torralba. 2017. Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence* **40**(6) 1452–1464.

Figure 1: Illustration of a 3D Virtual Tour



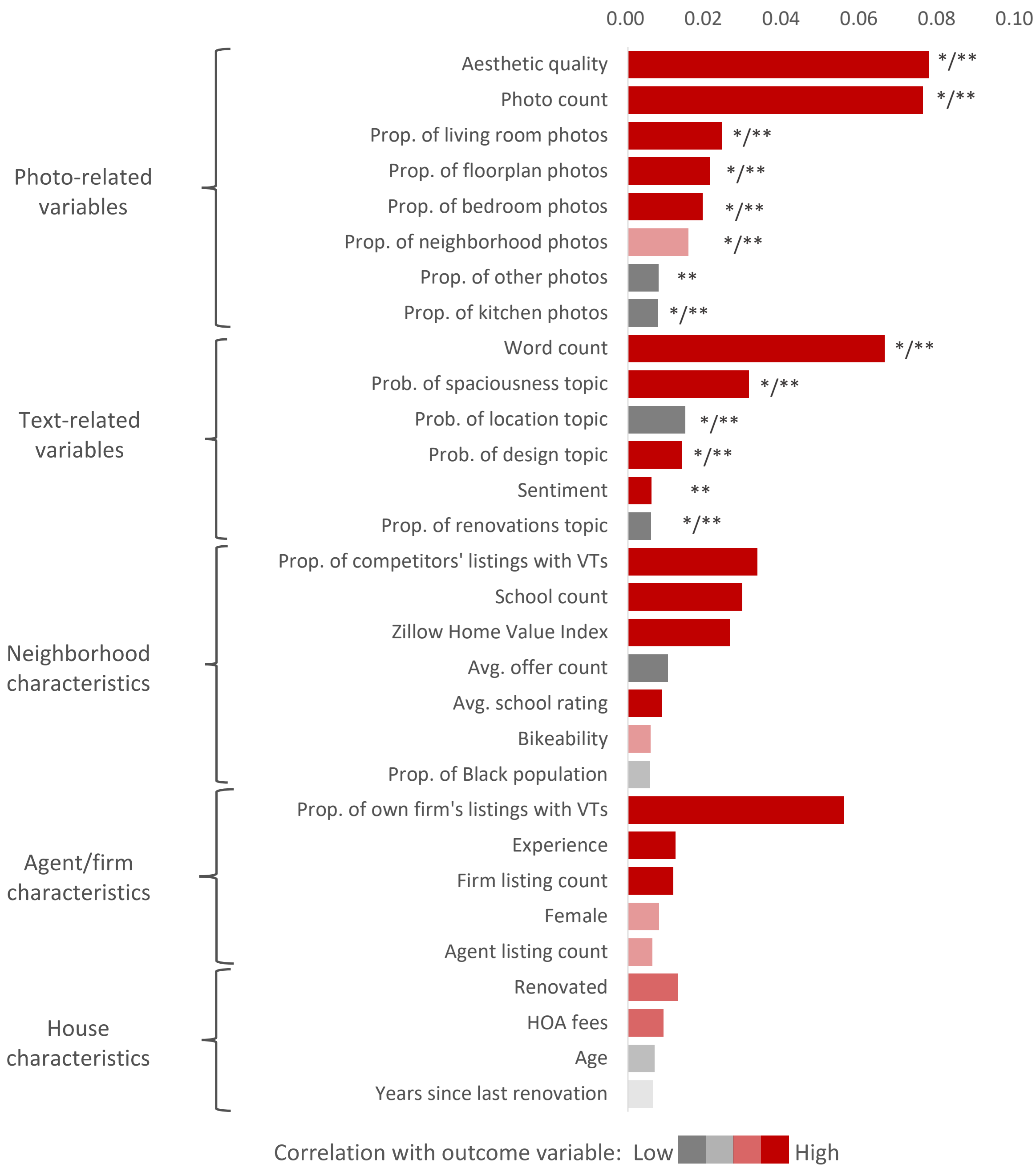
Note: Images 1 to 9 capture the sequence of interactions with the 3D virtual tour in this listing. In each image, the mouse cursor highlights the user's click, leading to the view of the house in the subsequent image. In images 1 and 2, the user clicks to access and start the 3D virtual tour, respectively. In image 3, the user clicks on the direction of the patio. In image 4, the user clicks on the floor plan. In image 5, the user clicks on the larger bedroom. In images 6 and 7, the user clicks to turn left and rotate the view of the room. In image 8, the user clicks toward the office room door.

Figure 2: Percentage of listings with VTs by listing dates



Note: The red dashed line indicates the date of the World Health Organization announcement of the COVID-19 pandemic in March 2020.

Figure 3: Top 30 most important explaining factors of VT adoption, excluding fixed effects



Note: The bar sizes represent each variable's importance, measured as the mean of the absolute SHAP values. The color codes represent the correlation between each variable and VT adoption. The red (gray) tones indicate a positive (negative) correlation, and the cutoff within each tone is based on the median correlations in the data. "*" indicates that the photo or text variable is also among the top 30 most important explaining factors of sale price, and "**" indicates that the photo or text variable is also among the top 30 most important explaining factors of TOM.

Table 1a: Definitions of control variables: house, neighborhood, agent/firm, time

Name	Definition
House characteristics:	
Property Type	Single family, townhouse, or condo
Square footage	The square footage of a house (log-transformed)
Bedroom count	The number of bedrooms in a house
Bathroom count	The number of bathrooms in a house
Age	The number of years since initial construction
Renovated	Whether the house has been renovated since initial construction
Years since renovation	The number of years since last renovation
Lot size	The square footage of the entire lot (log-transformed)
Garage space	The space available for housing a motor vehicle
Fireplace count	The number of fireplaces in a house
Pool	Whether there is a pool in a house
Story count	The number of floors of a house
Basement	Whether there is a basement in a house
Parking space	The space available for parking a motor vehicle
Laundry	Inside, community, other, or none
HOA fees	A monthly fee assessed by the homeowners association to pay for the services it provides (log-transformed)
Move-in ready	Whether the description mentions the house is “move-in ready”
Vacant	Whether the description mentions the house is vacant (i.e., not occupied by tenants or the owner)
Neighborhood characteristics:	
Population	The total population in the local census unit (log-transformed)
Prop. of Black population	Proportion of Black population in the local census unit
Prop. of Hispanic population	Proportion of Hispanic population in the local census unit
Average income	Median household income in the local census unit (log-transformed)
Zillow Home Value Index	Zillow’s estimates of the average home value in the local zip code area in the past month (log-transformed)
Avg. school rating	Schools’ average GreatSchhools rating
School count	The number of schools in the house’s service area
Walkability	The Walk Score (from 0 to 100) of an address. A higher score indicates better access to public transit, better commutes, and proximity to the people and places.
Bikeability	The Bike Score (from 0 to 100) of an address. A higher score indicates higher convenience of using bikes for daily errands.
Avg. offer count	The average number of offers received by houses listed in the local neighborhood (often smaller than a local zip code area) in the past month
Prop. of competitors’ listings with VTs	The proportion of listings with a virtual tour for competing firms in the past month in a local zip code area
Zip code dummies	There are 268 zip code dummies.

Note: Log transformation on a variable x means $log(1+x)$. “Past” and “current month” are with respect to the sale date.

Table 1a Continued: Definitions of control variables: house, neighborhood, agent/firm, time

Name	Definition
Agent/Firm characteristics:	
Experience	Measured by the number of properties sold by the same agent in the past year (log-transformed).
Female	The probability of the agent name being associated with the female gender
Agent listing count	The number of properties listed by the same agent in the current month.
Firm listing count	The number of properties listed by the same firm in the current month.
Prop. of own firm's listings with VTs	The proportion of listings with a virtual tour for the focal firm in the past month.
Firm dummies	The firm the agent belongs to. There are 806 firm dummies.
Time controls:	
Year-month dummies	The year-month of the sale date of the house. There are 25 year-month dummies (from 2019-03 to 2021-03).

Note: Log transformation on a variable x means $log(1 + x)$.

Table 1b: Definitions of control variables: photo and text description

Name	Definition
Photo-related variables:	
Photo count	The number of photos on a listed house (log-transformed)
Aesthetic quality	Photos' average aesthetic quality score of a house. The aesthetic scores are generated by Google's Neural Image Assessment tool.
Content distributions	The proportion of photos of each of the following 11 scene types: living room, kitchen, bathroom, house, bedroom, dining room, entrance hall, lawn, neighborhood, floorplan, and others. The scene types are generated by Places-CNNs.
Content diversity	Based on the Gini coefficient of the full content distributions.
Text-related variables:	
Word count	The number of words in the house text description provided by the selling agent (log-transformed)
Sentiment	The sentiment of the house text description. The text sentiment is extracted with VADER.
Topic distributions	The proportion of the text description of each of the following seven topics: renovations, condo amenities, design, location, spaciousness, investment opportunity, and interiors. The topics are extracted with LDA.

Note: Log transformation on a variable x means $log(1 + x)$.

Table 2: LDA topics

	Most relevant terms	Interpretation
Topic 1	New, remodel, brand, paint, roof, upgrade, window, throughout, tile, cabinet, door, fixture, water_heater, newer, bathroom, dual_pan, install, update	Renovations
Topic 2	Unit, condo, community, complex, balcony, park, townhouse, include, building, townhome, hoa, amenity, restaurant, private, space, two, storage	Condo amenities
Topic 3	Custom, build marble, cabinetry, high_end, glass, designer, finish, design, gourmet, luxurious, feature, suite, island, stun, appliance, modern	Design
Topic 4	School, great, freeway, family, close, shop, park, locate, nice, near, neighborhood, backyard, beautiful, location, perfect, plenty, lot	Location
Topic 5	Space, view, outdoor, deck, open, perfect, private, enjoy, entertain, office, expansive, living, din, provide, flow, fill, suite	Spaciousness
Topic 6	Property, lot, house, adu, potential, one, sell, back, original, need, primary, owner, year, make, dream, buyer, street, seller, investor	Investment opportunity
Topic 7	Master, upstairs, downstairs, fireplace, bathroom, family, shower, walk_closet, bath, separate, open, suite, feature, full, spacious, formal	Interiors

Table 3: Aggregated feature importance by groups

Group of features	N. of features	Aggregated importance
Photo-related variables	14	26.60%
Year-month fixed effects	24	22.90%
Text-related variables	8	14.34%
Neighborhood characteristics	11	13.42%
Agent/firm characteristics	5	9.34%
House characteristics	20	6.16%
Firm fixed effects	805	5.25%
Zip code fixed effects	269	1.98%
Total	1,156	100.00%

Table 4: Effect of VT adoption on house sale outcomes

	Baseline specification		Full specification	
	Sale price (1)	TOM (2)	Sale price (3)	TOM (4)
Has 3D virtual tour	0.011***	0.006	0.002	0.030***
Additional controls:				
Text-related variables	No	No	Yes	Yes
Other photo-related variables	No	No	Yes	Yes
Baseline controls:				
Photo count	Yes	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Observations	73,684	73,684	73,685	73,685
Selection model out-of-sample AUC	79.66%	79.66%	80.88%	80.88%
Outcome model out-of-sample R^2	0.947	0.114	0.952	0.127

Note: Dependent variables are log-transformed by log(1+y). We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Heterogeneity in the effects of VT on sale outcomes by different dimensions

	Effect of VT on sale price $\psi_i^{SalePrice}$			Effect of VT on TOM ψ_i^{TOM}		
	Low group	High group	p-value	Low group	High group	p-value
Photo and text-related variables:						
(1) Photo count	**0.004**	0.000	0.084	0.033	0.027	0.604
(2) Word count	**0.005**	-0.001	0.038	0.022	0.038	0.123
(3) Aesthetic quality	**0.005**	-0.001	0.009	0.028	0.032	0.684
(4) Prop. of design topic	**0.005**	-0.001	0.015	0.030	0.030	0.988
(5) Prop. of location topic	-0.002	**0.006**	0.002	0.040	**0.019**	0.048
(6) Prop. of spaciousness topic	**0.005**	-0.001	0.014	0.028	0.032	0.662
Agent/Firm characteristics:						
(7) Agent listing count	**0.003**	-0.001	0.087	0.029	0.034	0.655
(8) Firm listing count	**0.006**	-0.002	0.005	**0.017**	0.044	0.010
House characteristics:						
(9) Townhouse over Condo	0.003	-0.002	0.306	0.055	**0.015**	0.052
(10) Single Family Residential over Condo	0.003	0.002	0.696	0.055	**0.026**	0.035
Neighborhood characteristics:						
(11) Prop. of Black and Hispanic pop.	-0.002	**0.006**	0.001	0.052	**0.008**	0.000
(12) Average income	**0.006**	-0.002	0.003	**0.017**	0.042	0.018
(13) Zillow Home Value Index	0.004	0.000	0.114	**0.016**	0.044	0.011
After COVID-19:						
(14) After vs. before COVID-19	0.001	0.002	0.683	0.042	**0.019**	0.030

Note: Dependent variables are the estimated individual treatment effects for log-transformed sale price and TOM. Bold numbers are used to highlight the group that benefits significantly more from using VT (p-value < 0.10), i.e., obtains higher sale price or shorter TOM when using VT. We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Table 6: Heterogeneity in the effects of VT on sale outcomes by months after COVID-19 pandemic declaration

	Effect of VT on sale price $\psi_i^{SalePrice}$	Effect of VT on TOM ψ_i^{TOM}
Intercept	0.001	0.043***
After 1 month	0.005	−0.044
After 2 months	−0.011	−0.178***
After 3 months	0.005	−0.154***
After 4 months	0.003	−0.025
After 5 months	0.007	0.010
After 6 months	−0.005	0.005
After 7 months	0.002	−0.024
After 8 months	−0.002	−0.032
After 9 months	0.001	0.027
After 10 months	0.005	0.029
After 11 months	−0.007	0.006
After 12 months	0.008	−0.031
Observations	73,685	73,685
R^2	0.0001	0.001

Note: Dependent variables are the estimated individual treatment effects for log-transformed sale price and TOM. We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Effect of VT adoption on favorites, X-outs, and Redfin tours

	Favorite count	X-out count	Redfin tour count
	(1)	(2)	(3)
Has 3D virtual tour	0.100***	0.054***	0.016***
Additional controls:			
Text-related variables	Yes	Yes	Yes
Other photo-related variables	Yes	Yes	Yes
Baseline controls:			
Photo count	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Observations	73,685	73,684	73,684
Selection model out-of-sample AUC	80.88%	80.88%	80.88%
Outcome model out-of-sample R^2	0.509	0.424	0.113

Note: Dependent variables are log-transformed by log(1+y). We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Relationship Between VT usage and number of listings

	Firm listing count		Agent listing count	
	(1)	(2)	(3)	(4)
Prop. listings with VT	0.181***	0.024*	0.022***	0.018***
Controls:				
Year-month fixed effects	No	Yes	No	Yes
Firm fixed effects	No	Yes	No	Yes
Observations	13,098	13,098	41,213	41,213
R^2	0.006	0.767	0.002	0.125

Note: Dependent variables are log-transformed by log(1+y).

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Web appendices

A Summary statistics

Below, we provide summary statistics for all the variables used in our analyses.

Table A1a presents the summary statistics for house, neighborhood, and seller agent/firm characteristics extracted from house listings on Redfin, as well as additional sources such as Safegraph Open Census Data and Zillow Home Index Value (see Section 3.1 of the main manuscript). Table A1b presents the summary statistics for the text and photo-related variables uncovered from the unstructured data on the listings, i.e., the descriptions and the photos the seller includes in the listing (see Section 3.2 and Section 3.3 of the main manuscript). Last, Table A2 presents the summary statistics for the two main sale outcome variables, as well as different pricing strategy variables (see Section 6.1 of the main manuscript).

B Illustrating photo-related variables

Below, we provide examples to illustrate the photo-related variables described in Section 3.2 of the main manuscript.

Figure A1 illustrate two listings of houses of comparable characteristics (price range, location, and size) that have a similar number of photos but different *aesthetic quality scores*. The listings in panel (a) and panel (b) have below and above median photo aesthetic quality scores, respectively. Some notable differences between these two examples are the lighting of images (e.g., photos in panel (b) have much better lighting) and the angle at which the photos are taken and its impact on how much of the space is visible in the shots (e.g., larger zoom in panel (a) crops a lot of objects in the image and does not provide a full image of the rooms).

Table A1a: Summary statistics of control variables: house, neighborhood, agent/firm

	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
House characteristics:					
Single family	0.73	0.44	0	1	1
Townhouse	0.09	0.28	0	0	0
Condo	0.18	0.39	0	0	0
Square footage*	7.35	0.58	7.07	7.34	7.63
Bedroom count	3.09	1.00	2	3	4
Bathroom count	2.26	0.89	2.00	2.00	3.00
Age	51.51	26.38	32	53	69
Renovated	0.84	0.37	1	1	1
Years since renovation	40.88	28.42	15	42	63
Lot size*	9.15	1.74	8.64	8.90	9.62
Garage space	1.74	0.78	1.00	2.00	2.00
Fireplace count	0.69	0.51	0	1	1
Pool	0.34	0.47	0	0	1
Story count	0.77	0.77	0.00	1.00	1.00
Basement	0.02	0.14	0	0	0
Parking space	1.98	1.15	2.00	2.00	2.00
Laundry	2.11	1.02	1	3	3
HOA fees*	1.92	2.68	0.00	0.00	5.34
Move-in ready	0.03	0.18	0	0	0
Vacant	1.00	0.03	1	1	1
Neighborhood characteristics:					
Population*	7.44	0.55	7.07	7.43	7.77
Prop. of Black population	0.07	0.13	0.01	0.03	0.08
Prop. of Hispanic population	0.35	0.26	0.14	0.28	0.52
Average income*	11.37	0.43	11.09	11.39	11.65
Zillow Home Value Index*	13.44	0.49	13.11	13.37	13.71
Avg. school rating	5.98	1.73	4.67	6.00	7.33
School count	4.11	1.56	3	3	6
Walkability	52.93	27.22	32	58	75
Bikeability	46.31	22.32	31	50	62
Avg. offer count	7.46	7.21	2	5	10
Prop. of competitors' listings with VTs	0.12	0.15	0.00	0.07	0.20
Agent/firm characteristics:					
Experience*	1.32	1.14	0.00	1.10	2.08
Female	0.48	0.48	0.004	0.21	1.00
Agent listing count	1.62	1.73	1	1	2
Firm listing count	33.46	67.37	2	8	27
Prop. of own firm's listings with VTs	0.14	0.30	0.00	0.00	0.00

Note: * The averages of these variables before log-transformation are: 1,731 for square footage, 749,938 for lot size, 115 for HOA fees, 2,003 for population, 94,685 for average income, 791,387 for Zillow Home Value Index, and 7.158 for seller agent experience.

Table A1b: Summary statistics of photo and text-related variables

	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
Photo-related variables:					
Photo count*	3.26	0.56	3.00	3.33	3.61
Aesthetic quality	4.61	1.50	4.94	5.10	5.21
Prop. of living room photos	0.15	0.10	0.09	0.15	0.21
Prop. of kitchen photos	0.11	0.07	0.07	0.11	0.14
Prop. of bathroom photos	0.10	0.06	0.07	0.10	0.14
Prop. of house photos	0.10	0.12	0.04	0.08	0.13
Prop. of bedroom photos	0.08	0.07	0.00	0.07	0.12
Prop. of dining room photos	0.04	0.04	0.00	0.03	0.06
Prop. of entrance hall photos	0.02	0.04	0.00	0.00	0.03
Prop. of lawn photos	0.03	0.06	0.00	0.00	0.04
Prop. of neighborhood photos	0.03	0.06	0.00	0.00	0.04
Prop. of floorplan photos	0.01	0.02	0.00	0.00	0.00
Prop. of other photos	0.07	0.09	0.00	0.05	0.10
Content diversity	0.85	0.05	0.82	0.85	0.88
Text-related variables:					
Word count*	4.84	0.62	4.61	4.96	5.20
Sentiment	0.91	0.18	0.93	0.97	0.99
Prop. of renovations topic	0.14	0.20	0.002	0.05	0.23
Prop. of condo amenities topic	0.14	0.23	0.002	0.004	0.21
Prop. of design topic	0.09	0.15	0.002	0.004	0.13
Prop. of location topic	0.19	0.25	0.002	0.08	0.32
Prop. of spaciousness topic	0.14	0.20	0.003	0.01	0.23
Prop. of investment opportunity topic	0.10	0.18	0.002	0.004	0.14
Prop. of interiors topic	0.19	0.23	0.003	0.07	0.32

Note: * The average of these variables before log-transformation are: 28.6 for photo count and 141.7 for word count.

Table A2: Summary statistics of outcome variables

	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
Sale price*	13.48	0.54	13.12	13.40	13.76
Time on market*	4.15	0.53	3.76	4.08	4.49
List price*	13.49	0.55	13.12	13.40	13.76
Price decrease count*	0.18	0.38	0.00	0.00	0.00
Price increase count*	0.04	0.17	0.00	0.00	0.00

Note: * The average of these variables before log-transformation are: 844,050 for sale price, 72.66 for time on market, 855,877 for list price, 0.312 for price decrease count and 0.059 for price increase count.

Figure A1: Examples of listings with photos of different levels of aesthetic quality

(a) Listing with average aesthetic quality photos below median

(b) Listing with average aesthetic quality photos above median

Note: In each example, the first image corresponds to a screenshot of the photo layout shown on a property listing. The remaining photos are a sample of the album re-scaled to keep the original height-width ratio of each photo.

Figure A2 illustrates two listings that have a similar number of photos but different *content diversity*. In the listing in panel (a), the majority of the photos provide views of the exterior of the house (e.g., patio and lawn), and the photos of the interior do not show all the rooms in the house (e.g., the kitchen and bathroom are visible, but the living room and bedrooms are missing). In the listing in panel (b), each photo provides a view of different exterior and interior areas of the house.

C Assessing the overlap assumption

To assess whether our data meet the overlap or positivity assumption (i.e., $0 < m(\mathbf{X}_i) < 1$), we adopt two approaches. The first approach suggested by McCaffrey et al. (2013) and utilized by Ellickson et al. (2022) assesses the degree of covariate balance by calculating

Figure A2: Examples of listings with photos of different levels of content diversity

(a) Listing with photo content diversity below median

(b) Listing with photo content diversity above median

Note: In each example, the first image corresponds to a screenshot of the photo layout shown on a property listing. The remaining photos are a sample of the album re-scaled to keep the original height-width ratio of each photo.

each covariate's "population" standardized bias (PSB). For each pre-treatment covariate k ($k = 1, 2, ..., K$), a PSB is given by:

$$PSB_k = \frac{|\bar{X}_{kw} - \bar{X}_{kp}|}{\hat{\sigma}_{kp}} \tag{3}$$

where $\bar{X}_{kw} = \frac{\sum_i \frac{VT_i X_{ki}}{\hat{m}(\mathbf{X_i})}}{\sum_i \frac{VT_i}{\hat{m}(\mathbf{X_i})}}$ is the propensity-weighted mean of the covariate k, and $\hat{m}(\mathbf{X_i})$ is estimated propensity score. $\bar{X}_{kp}$ and $\hat{\sigma}_{kp}$ are the unweighted mean and standard deviation of covariate k for all observations. McCaffrey et al. (2013) note that generally PSBs "of less than 0.20 are considered small, 0.40 are considered moderate, and 0.6 are considered large." Table A3a and Table A3b report the PSB scores across all covariates (except for year-month, zip-code, and firm fixed-effect dummies). We find that most covariates have PSB smaller than 0.2. Only seven of the 58 covariates fall in the low category of 0.20-0.40. And no cases fell in the median (0.40-0.60) or large (>0.6) difference category. The results suggest that the overlap or positivity assumption is likely satisfied in our setting.

As a robustness check, we use a second approach suggested by Austin (2011) and utilized by Zhang and Luo (2023) to check the quality of estimated propensity score $\hat{m}(\mathbf{X_i})$ utilized in our DML framework. The intuition is that if the propensity scores' quality is good, we would see that the covariates adjusted by the propensity scores are balanced across the treated and control conditions. The covariate difference after propensity score adjustment between the treated and control is defined as:

$$d_{adjusted} = \frac{|\bar{X}^{adj}_{k,VT_i=1} - \bar{X}^{adj}_{k,VT_i=0}|}{\sqrt{\frac{{s^{adj}_{k,VT_i=1}}^2 + {s^{adj}_{k,VT_i=0}}^2}{2}}} \tag{4}$$

Where $\bar{X}^{adj}_{k,VT_i=v} = \frac{\sum_{i\in\{i|VT_i=v\}} w_i X_{ki}}{\sum_i w_i}$ is the weight adjusted average X_k for condition $VT_i = v$, ${s^{adj}_{k,VT_i=v}}^2 = \frac{\sum_{i\in\{i|VT_i=v\}} w_i}{(\sum_{i\in\{i|VT_i=v\}} w_i)^2 - \sum_{i\in\{i|VT_i=v\}} w_i^2} \sum_{i\in\{i|VT_i=v\}} w_i (X_{ki} - \bar{X}^{adj}_{k,VT_i=v})^2$ is the weight adjusted variance of X_k for condition $VT_i = v$, and $w_i = \frac{VT_i}{\hat{m}(\mathbf{X_i})} + \frac{1-VT_i}{1-\hat{m}(\mathbf{X_i})}$ is the weight. As per Austin (2009), $d_{adjusted} \leq 0.2$ indicates an adequate balance in covariates between treated and control conditions and thus a good quality of estimated propensity scores. Table A3a and Table A3b report the $d_{adjusted}$ scores across all covariates (except for year, zip-code, and firm fixed effect dummies). The $d_{adjusted}$ of all covariates are smaller than 0.2, indicating adequate balance in covariates between treated and control conditions after adjustment and providing further support that the overlap or positivity assumption is likely satisfied in our setting.

Table A3a: Balance checks for baseline control variables

Variable	PSB score	$d_{adjusted}$
House characteristics:		
Single family	0.04	0.02
Townhouse	0.03	0.02
Condo	0.03	0.00
Square footage	0.09	0.03
Bathroom count	0.12	0.03
Bedroom count	0.03	0.00
Age	0.01	0.00
Renovated	0.01	0.04
Years since renovation	0.02	0.01
Lot size	0.02	0.00
Garage space	0.02	0.01
Fireplace count	0.09	0.03
Pool	0.04	0.00
Story count	0.00	0.02
Basement	0.04	0.02
Parking space	0.08	0.05
Laundry	0.04	0.02
HOA fees	0.05	0.02
Move-in ready	0.03	0.02
Vacant	0.01	0.03
Neighborhood characteristics:		
Population	0.04	0.01
Prop. of Black population	0.02	0.00
Prop. of Hispanic population	0.15	0.06
Median household income	0.07	0.03
Avg. school rating	0.06	0.02
School count	0.13	0.05
Bikeability	0.05	0.00
Walkability	0.06	0.02
Zillow Home Value Index	0.22	0.08
Avg. offer count	0.06	0.03
Prop. of competitors' listings with VTs	0.35	0.08
Agent/firm characteristics:		
Experience	0.18	0.01
Female	0.05	0.01
Firm listing count	0.14	0.05
Agent listing count	0.09	0.03
Prop. of own firm's listings with VTs	0.37	0.07

Table A3b: Balance checks for text and photo related variables

Variable	PSB score	$d_{adjusted}$
Photo-related variables:		
Photo count	0.25	0.10
Aesthetic Score	0.09	0.01
Prop. of living room photos	0.14	0.05
Prop. of kitchen photos	0.01	0.01
Prop. of bathroom photos	0.02	0.00
Prop. of house photos	0.09	0.03
Prop. of bedroom photos	0.13	0.04
Prop. of dining room photos	0.10	0.03
Prop. of entrance hall photos	0.04	0.00
Prop. of lawn photos	0.06	0.02
Prop. of neighborhood photos	0.02	0.00
Prop. of floorplan photos	0.08	0.03
Prop. of other photos	0.08	0.03
Content diversity	0.10	0.03
Text-related variables:		
Word count	0.24	0.09
Sentiment	0.13	0.05
Prop. of renovations topic	0.05	0.03
Prop. of condo amenities topic	0.02	0.00
Prop. of design topic	0.18	0.06
Prop. of location topic	0.22	0.06
Prop. of spaciousness topic	0.25	0.06
Prop. of investment opportunity topic	0.07	0.01

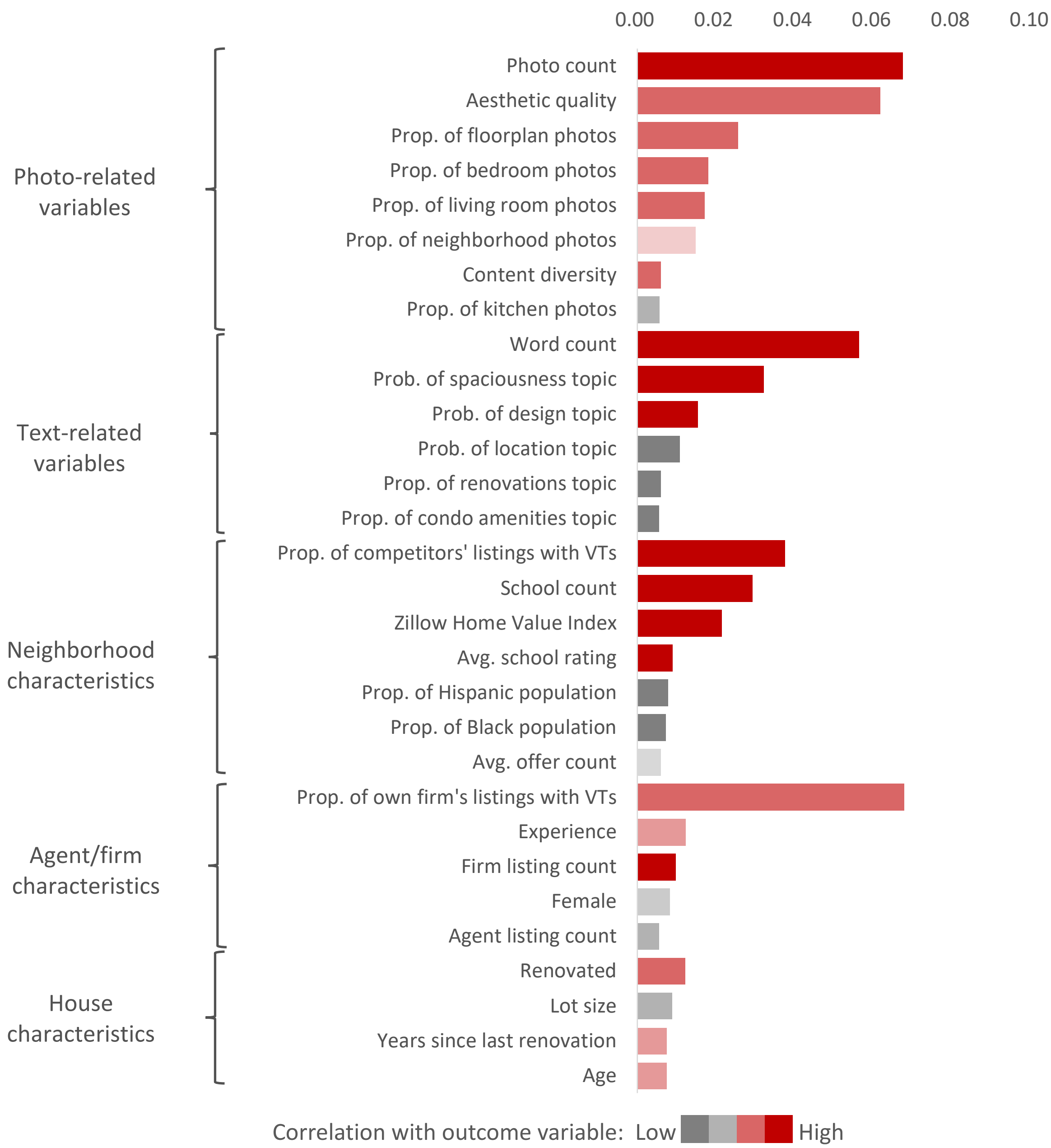
D Photo and text-related variables and sales outcomes

To assess the validity of leveraging unstructured data to uncover confounders to reduce selection bias, we follow Zeng et al. (2022) and examine whether the photo and text-related variables from the unstructured data in the listings are important explaining factors of not only the treatment (VT adoption) but also the outcomes of interest (sale price and TOM).

To this aim, and similar to Section 5, we predict outcomes using the Xgboost algorithm, this time for continuous rather than binary outcomes. For the sale price outcome, the algorithm achieves a 0.120 root-mean-square error in the test sample. In Table A4, we present the aggregated importance by the different groups of features. Relative to the results for VT adoption prediction, the importance of photo-related variables decreases by nearly 4 percentage points and the importance of real estate firm dummies increases by nearly 7 percentage points. The latter might be consistent with the idea that more experienced firms are more likely to work with higher quality houses and more motivated sellers (Gilbukh and Goldsmith-Pinkham 2019), which can directly relate to sale prices. Despite such changes in the absolute values, photo and text-related variables are still as important as many other readily available metrics, such as standard house characteristics.

In Figure A3, we summarize the individual importance of the 30 most important features in explaining sale prices (excluding fixed effects variables) and use color code to indicate the correlation between feature values and sale prices. A comparison with the results for VT adoption prediction highlights some correlations that, if ignored, could bias the estimated effect of VT on sale price. For instance, Figure A3 suggests that listings with more photos and higher aesthetic quality photos, which are also more likely to adopt VT (Figure 3), tend to sell at a higher price. Similarly, Figure A3 suggests that listings with a higher proportion of spaciousness and design topics, which are also more likely to adopt VT (Figure 3), tend to sell at a higher price. Ignoring these positive correlations between factors that are important in explaining VT adoption and sale price would overestimate the effect of VT on sale prices.

Figure A3: Top 30 most important explaining factors of sale price, excluding fixed effects



Note: The bar sizes represent each variable's importance, measured as the mean of the absolute SHAP values. The color codes represent the correlation between each variable and sale price. The red (gray) tones indicate a positive (negative) correlation, and the cutoff within each tone is based on the median correlations in the data.

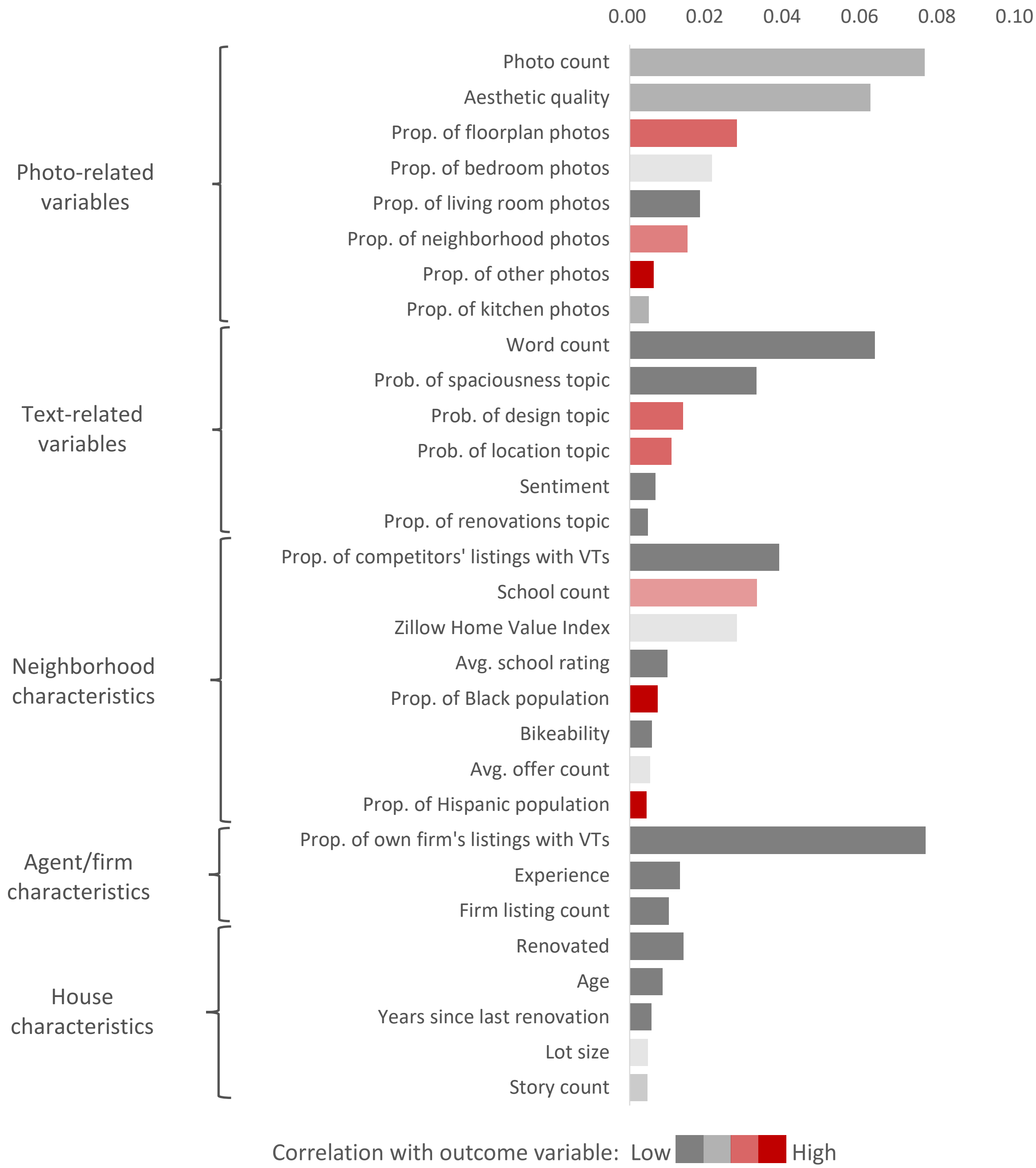
Table A4: Aggregated feature importance by groups for sale price prediction

Group of features	N. of features	Aggregated importance
Photo-related variables	14	23.03%
Year-month fixed effects	24	19.91%
Text-related variables	8	13.48%
Neighborhood characteristics	11	12.99%
Agent/firm characteristics	5	10.31%
House characteristics	20	5.52%
Firm fixed effects	805	11.94%
Zip code fixed effects	269	2.81%
Total	1,156	100.00%

For the TOM outcome, the algorithm achieves 0.498 root-square-mean error in the test sample. In Table A5, we present the aggregated importance by the different groups of features. Relative to the results for VT adoption prediction, the importance of photo-related variables decreases by nearly 3 percentage points and the importance of real estate firm dummies increases by 4 percentage points. Despite such changes in the absolute values, photo and text-related variables are still equally or more important than many other readily available metrics, such as standard house characteristics.

In Figure A4, we summarize the individual importance of the 30 most important features in explaining sale prices (excluding fixed effects variables) and use color code to indicate the correlation between feature values and TOM. A comparison with the results for VT adoption prediction highlights some correlations that, if ignored, could bias the estimated effect of VT on TOM. For instance, Figure A4 suggests that listings with a higher proportion of living room photos and a higher proportion of spaciousness topic, which are more to adopt VT (Figure 3), spend less TOM. Ignoring these factors that are important in explaining VT adoption and TOM would underestimate the effect of VT on TOM.

Figure A4: Top 30 most important explaining factors of TOM, excluding fixed effects



Note: The bar sizes represent each variable's importance, measured as the mean of the absolute SHAP values. The color codes represent the correlation between each variable and TOM. The red (gray) tones indicate a positive (negative) correlation, and the cutoff within each tone is based on the median correlations in the data.

Table A5: Aggregated feature importance by groups for time on market prediction

Group of features	N. of features	Aggregated importance
Photo-related variables	14	23.93%
Year-month fixed effects	24	22.08%
Text-related variables	8	13.64%
Neighborhood characteristics	11	13.63%
Agent/firm characteristics	5	10.69%
House characteristics	20	4.94%
Firm fixed effects	805	9.36%
Zip code fixed effects	269	1.73%
Total	1,156	100.00%

E Relation between VT adoption and TOM using pending date

Benefield and Hardin (2015) discuss two measures for TOM: listing to pending or listing to sale. They think that both measures are defensible. Specifically, one could argue that the pending contract date should signal the end of the marketing period, since agents are not likely to continue marketing a property at the same level once a contract has been signed. However, this ignores the non-trivial percentage of pending contracts that fail to result in a completed transaction. Our main analysis measures TOM using the duration between the list date and sale date. In Table A6, we show that our main result that the benefit of VT on TOM is smaller than previously found is consistent when defining TOM as the number of days between listing date and pending date.

Table A6: Effect of virtual tour adoption on TOM (listing to pending)

	Baseline specification	Full specification
	TOM	TOM
	(1)	(2)
Has 3D virtual tour	0.044***	0.072***
Additional controls:		
Text-related variables	No	Yes
Other photo-related variables	No	Yes
Baseline controls:		
Photo count	Yes	Yes
House characteristics	Yes	Yes
Neighborhood characteristics	Yes	Yes
Agent/firm characteristics	Yes	Yes
Year-month fixed effects	Yes	Yes
Zip code fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Observations	73,685	73,685
Selection model out-of-sample AUC	79.66%	80.88%
Outcome model out-of-sample R^2	0.108	0.108

Note: Dependent variables are log-transformed by log(1+y). We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

F Robustness checks with additional measured confounders

Below, we present additional details for the robustness checks to additional measured confounders discussed in Section 6.1.

Controlling for Redfin estimates of house market value As noted in Section 6.1, home sellers and buyers have access to publicly available estimates of a house market value provided by major real estate platforms like Redfin and Zillow. These estimates could serve as reference price points that consequently affect both sale prices and the decision to use VT. To examine how this could affect our estimates, we further control for the Redfin estimate of each house value in the model.

Table A7 presents the DML estimates of VT adoption on sale outcomes with Redfin Estimates controlled. Consistent with our main results, after the Redfin Estimate is controlled in the baseline, we observe that controlling additional photo and text variables reduces the claimed benefits of VT on sale outcomes.

Examining the link between VT adoption and pricing strategies As noted in Section 6.1, another type of confounder that could threaten the causal interpretation of our findings relate to unmeasured pricing strategies consistent with the results we observe. To explore this possibility, we use the DML framework described in Section 4 to examine the relationship between VT adoption and different pricing-related variables, namely: (i) the initial list price, (ii) the number of times the seller decreased the list price, and (ii) the number of times the seller increased the list price. Our results, presented in Table A8, suggest that VT adoption is associated with higher listing prices, more frequent subsequent price cuts, and less frequent subsequent price rises, a pricing strategy that is indeed consistent with the results we observe (i.e., longer TOM).

Table A7: Effect of VT adoption on house sale outcomes - Redfin estimates controlled

	Baseline specification		Full specification	
	Sale price (1)	TOM (2)	Sale price (3)	TOM (4)
Has 3D virtual tour	0.006***	0.008	0.002***	0.028***
Additional controls:				
Text-related variables	No	No	Yes	Yes
Other photo-related variables	No	No	Yes	Yes
Baseline controls:				
Redfin Estimate	Yes	Yes	Yes	Yes
Photo count	Yes	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Observations	73,889	73,889	73,890	73,890
Selection model out-of-sample AUC	79.71%	79.71%	80.72%	80.72%
Outcome model out-of-sample R^2	0.983	0.130	0.984	0.132

Note: Dependent variables are log-transformed by log(1+y). We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table A8: Effect of VT adoption on pricing strategies

	List price (1)	Price decrease count (2)	Price increase count (3)
Has 3D virtual tour	0.003***	0.011***	−0.007***
Controls:			
Text-related variables	Yes	Yes	Yes
Photo-related variables	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Observations	73,685	73,685	73,685
Selection model out-of-sample AUC	80.88%	80.88%	80.88%
Outcome model out-of-sample R^2	0.948	0.046	0.052

Note: Dependent variables are log-transformed by log(1+y). We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

To dig deeper into the extent to which the pricing strategies sellers are more likely to use when adopting VTs drive the results we observe, we turn into an OLS framework to conduct a mediation analysis. To this aim, we first estimate the effect of VT adoption on sale outcomes to examine whether our findings are consistent with those obtained from the DML framework. In particular, we estimate a fixed effects specification:

$$Y_i = \theta \cdot VT_i + \mathbf{X}_i \cdot \boldsymbol{\beta} + \alpha_{\text{zip-code}} + \mu_{\text{firm}} + \lambda_{\text{year-month}} + U_i \tag{5}$$

where Y_i is a linear function of VT adoption and other controls, fixed effects by zip code, seller agent's firm, and year-month, and the treatment effect θ here is the average treatment effect. The results, presented in Table A9, are consistent with those with results obtained with the DML framework in Table 4.

Next, we examine how pricing strategies mediate the effects of VT adoption on sale outcomes in Table A10. Columns 1 to 3 of Table A10 show that VT adoption is significantly

Table A9: Effects of VT adoption on house sale outcomes - fixed effects model

	Baseline specification		Full specification	
	Sale price (1)	TOM (2)	Sale price (3)	TOM (4)
Has 3D virtual tour	0.012***	0.010	0.002	0.017***
Additional controls:				
Text-related variables	No	No	Yes	Yes
Other photo-related variables	No	No	Yes	Yes
Baseline controls:				
Photo count	Yes	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Observations	69,160	69,160	62,652	62,652
R^2	0.914	0.116	0.923	0.125

Note: Dependent variables are log-transformed by log(1+y). Pricing strategies include: log-transformed list price, price decrease count, and price increase count.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

associated with a higher list price and more price downward adjustments afterward. Columns 4 and 5 of Table A10 show that higher list prices are associated with higher sale prices and longer TOM, price adjustments are associated with longer TOM, and intuitively, downward (upward) adjustments reduce (increase) sale prices. Moreover, columns 4 and 5 suggest that after controlling pricing strategies, VT adoption has a null effect on sale prices, and that the positive effect of VT on TOM in column 4 of Table A9 diminishes to null. Altogether, these results suggest that list prices might be higher for the houses with VTs not because of systematic differences from those without VTs, but because of the selling process itself, namely, seller agents aiming too high and then adjusting prices downward. Moreover, such a selling process might lead to a longer TOM for houses with VTs.

Table A10: Mediating role of pricing strategies - fixed effects model

	List price (1)	Price decrease count (2)	Price increase count (3)	Sale price (4)	TOM (5)
Has 3D virtual tour	0.003*	0.020***	−0.002	0.001	0.003
List price				0.871***	0.128***
Price decrease count				−0.076***	0.674***
Price increase count				0.033***	0.294***
Controls:					
Text-related variables	Yes	Yes	Yes	Yes	Yes
Photo-related variables	Yes	Yes	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	62,652	62,652	62,652	62,652	62,652
R^2	0.917	0.069	0.078	0.988	0.339

Note: Dependent variables are all log-transformed by log(1+y).
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

G Sensitivity to other unmeasured confounders

Following prior marketing literature (Manchanda et al. 2015, Zhang et al. 2022), we adopt the Rosenbaum Bounds Approach (Rosenbaum 2002) to assess the sensitivity of our estimates to unmeasured confounders. Given that the effect of VT on sale prices is no longer significant after controlling for the visual appeal and quality of the house and the effort made by the seller in providing visual and textual information about the property, we are mostly interested in the sensitivity of the estimated positive impact of VT on TOM.

We refer the reader to Rosenbaum (2002) for details on the Rosenbaum Bounds Approach. In short, it starts from the idea that the probability π_j that a unit j receives the treatment depends on both observed covariate x_j and an unmeasured covariate u_j: $\log(\pi_j/(1-\pi_j)) = f(x_j) + \gamma u_j$, with $0 \leq u_j \leq 1$, and $f(.)$ an unknown function and γ an unknown parameter. Then, for two units j and k with the same covariates x_j and x_k, the ratio between the odds of receiving the treatment:

$$\exp(-\gamma) \leq \frac{\pi_j(1-\pi_k)}{\pi_k(1-\pi_j)} \leq \exp(\gamma) \tag{6}$$

If $\Gamma = exp(\gamma) = 1$, both units would be equally likely to be treated and the study would be free of *hidden bias* (bias from the unmeasured confounder u). When $\Gamma > 1$, the odds of receiving the treatment change even when $x_j = x_k$. Thus, Γ can be seen as a measure of the degree of departure from a study free of *hidden bias*. A sensitivity analysis consists of conducting statistical hypothesis tests and estimates at different values of Γ. A study is sensitive if its inferences change for small deviations from Gamma equal to 1.

In Table A11, we present the results for the sensitivity of the estimated treatment effect of VT on TOM using three commonly used tests for matched-pair data.[1] Column 2 shows the p-values corresponding to the p-value from Wilcoxon's signed rank test to reject the null that the differences in outcomes of the paired treatment and control units are insignificant.[2] Columns 3 and 4 correspond to the Hogges-Lehman point estimates of an additive treatment effect. Columns 5 and 6 correspond to the Hogges-Lehman confidence intervals for the point estimates of an additive treatment effect.

The three tests in Table A11 suggest that the positive estimated effect of VT adoption on TOM is explained away by small magnitudes of *hidden biases*: the p-value to reject the null hypothesis of null treatment effect is larger than the $\alpha = 0.10$ threshold with $\Gamma = 1.1$, and similarly, both the Hogges-Lehman point estimates and confidence intervals contain zero with $\Gamma = 1.1$. As such, we conclude that the text and photo-related variables extracted from the unstructured data of the listings are not enough to account for all relevant confounders and that, after accounting for unmeasured confounders, the effect of VT on TOM is likely insignificant.

[1]We use a 1:1 nearest neighborhood matching based on the propensity score estimated with the full set of covariates used in our main analyses. Our findings are robust to alternative matching methods.

[2]Since this is a one-sided test, we focus on the relevant p-value according to the sign of the estimated treatment effect. We use the upper lower bound for the positive estimated effect of VT on TOM (DiPrete and Gangl 2004)

Table A11: Rosenbaum Bounds test for TOM outcome

Γ (1)	p-value (2)	t-hat+ (3)	t-hat- (4)	CI+ (5)	CI- (6)
1.0	0.001	0.018	0.018	0.007	0.030
1.1	0.958	−0.010	0.047	−0.022	0.059
1.2	1.000	−0.037	0.073	−0.048	0.085
1.3	1.000	−0.061	0.097	−0.073	0.109
1.4	1.000	−0.083	0.120	−0.095	0.131
1.5	1.000	−0.104	0.140	−0.116	0.152
1.6	1.000	−0.123	0.159	−0.135	0.171
1.7	1.000	−0.141	0.177	−0.153	0.189
1.8	1.000	−0.158	0.194	−0.171	0.206
1.9	1.000	−0.175	0.210	−0.187	0.222
2.0	1.000	−0.190	0.225	−0.202	0.238

Note: In Column (1), Gamma is the odds of differential assignment due to unobserved factors. In Column (2), the p-value is the Wilcoxon's signed rank statistics. In Columns (3) and (4), t-hat+ and t-hat- are the upper and lower bound Hodges-Lehmann point estimates, respectively. In Columns (5) and (6), CI+ and CI- are the upper and lower bound confidence interval at alpha = 0.95, respectively.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

We also conducted an additional sensitivity analysis based on the Gaussian Copulas approach proposed by Park and Gupta (2012), which directly models the joint distribution of an endogenous regressor and the error term in the outcome equation of interest. Despite the known limitations of this approach for binary endogenous variables like ours, its results indicate that the VT adoption has no significant effect on sale prices and TOM after controlling for photo and text-related variables, reassuring the validity of our conclusions above.[3]

H Effects of virtual adoption on favorites, x-outs, and Redfin tours normalized by TOM

In Table 7 we find that VT adoption has a positive impact on the favorite, x-out, and Redfin tour counts, respectively. One concern is that, as we also find that virtual adoption may have

[3]Please note that due to space constraints, these results are not included in the manuscript but can be provided upon request.

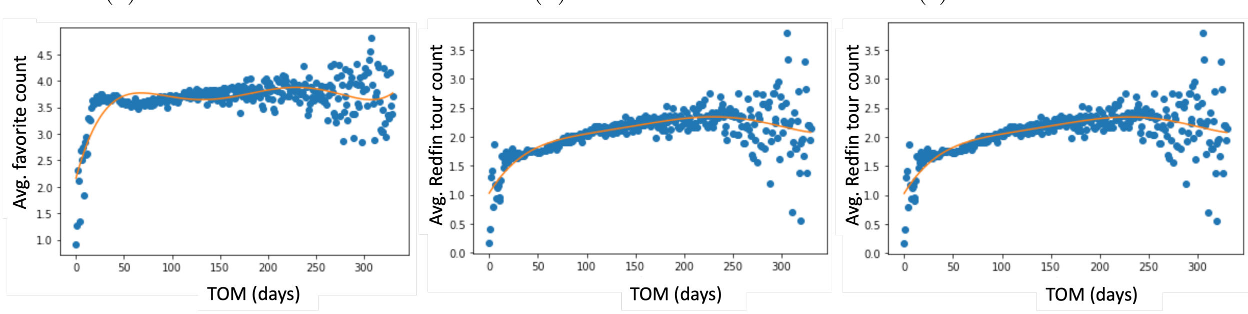
a positive impact on TOM, longer TOM may allow more views and thus more favorites, x-outs, and Redfin tours. Hence the positive effects on favorite, x-out, and Redfin tour counts could be due to two potential reasons: 1) VT adoption leads to longer TOM; 2) VT adoption increases the popularity of adopting houses in the early stages of buyers' search process.

To see whether the second reason (i.e., VT adoption increases houses' popularity in the search process) is at play, we normalized the favorite, x-out, and Redfin tour counts of a house based on its TOM. Specifically, we perform the following procedures:

- Plot the average log-transformed counts of favorites, x-outs, and Redfin tours against TOM as in Figure A5. We observe that the counts of favorites, x-outs, and Redfin tours indeed increase with TOM before the 50th day on market, but the pattern flattens after the 50th day on market.
- As the data becomes scarce and noisier when TOM is large, we fit a polynomial function with degree five of to smooth the average log-transformed counts with respect to days on the market.
- Then we normalize the log-transformed counts for a specific house by comparing it with the corresponding smoothed average log-transformed count as $\log(\text{count}+1)_t / \mathbb{E} \log(\text{count}+1)_t$ to capture how each house's count performs compared to other houses with the same TOM.

Figure A5: Average favorite, x-out, and Redfin tour counts by TOM

(a) Favorite count (b) X-out count (c) Redfin tour count



Note: Favorite, x-out, Redfin tour counts are log-transformed by log(1+y). The blue dots are the avg. count across houses for each day on the market. The orange line is fitted by a polynomial function of degree five.

Lastly, we use the normalized log-transformed count of favorites (x-outs, redfin tours) as new outcome variables and estimate the effect of VT adoption on them with the DML framework described in Section 4. The results are presented in Table A12. In general, compared to Table 7, after normalization, the effects on favorite and x-out counts become much smaller but still positive, and the effect on Redfin tour count is similar but less significant. The analysis here suggests that VT adoption can increase the popularity of houses in the early stages of the buyers' search process.

Table A12: Effect of virtual tour adoption on normalized favorite, x-out, and Redfin tour counts

	Favorite count (1)	X-out count (2)	Redfin tour count (3)
Has 3D virtual tour	0.027***	0.025***	0.025*
Controls:			
Text-related variables	Yes	Yes	Yes
Photo-related variables	Yes	Yes	Yes
House characteristics	Yes	Yes	Yes
Neighborhood characteristics	Yes	Yes	Yes
Agent/firm characteristics	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Observations	73,685	73,685	73,683
Selection model out-of-sample AUC	80.88%	80.88%	80.88%
Outcome model out-of-sample R^2	0.511	0.479	0.120

Note: Dependent variables are log-transformed by log(1+y). We trim the dataset at the 0.01 and 0.99 levels of the treatment propensities and the 0.0005 and 0.9995 quantiles of the individual treatment effects.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.