# Investigation of Alleged 'Algorithmic Collusion' In Rental Housing

## Impact and Implications of RealPage's Pricing Algorithms for Housing Affordability

April 14, 2025

Timothy Majidzadeh
Chelle Davies
Peter Benzoni
Patrick Yim
Ahmad Allaou

## Abstract

Recent investigative reports and lawsuits[1] allege that "Revenue Management" software operated by RealPage contributes to the sharply rental housing costs via price-fixing effects. We analyze penetration rate data disclosed by RealPage, aggregate rent indices calculated by Zillow, and building-level average rents acquired through web scraping to determine whether algorithmic pricing is associated with higher rent. Propensity score matching on the observable features of multifamily properties suggests that units managed by companies named as RealPage clients in recent lawsuits[2] are about 0.23 dollars per month per square foot more expensive on average. To test whether this is due to price-fixing effects, we apply difference-in-difference regression and synthetic control analysis to estimate the aggregate effect of RealPage's 2017 acquisition of Lease Rent Options (LRO), using MSAs with a high penetration rate as the 'treatment' group and MSAs with a low penetration rate as the 'control' group. No consistent statistically significant effect was detected in these analyses, and we cannot reject a null hypothesis that elevated prices for RealPage's associated units are not the result of price coordination. We debut our website which visualizes our findings and allows users to interact with our machine learning models, access our dataset, and understand housing market dynamics in the United States.

---

[1] See, e.g., Vogell, H., Coryne, H., & Little, R. (2022, October 15). *Rent Going Up? One Company's Algorithm Could Be Why.* ProPublica. https://www.propublica.org/article/yieldstar-rent-increase-realpage-rent.
[2] The Washington Post discloses a set of property management companies which were named in recent lawsuits against RealPage. See Flowers, A., Yu Chen, S., Rich, S., & Lerman, R. (2025, January 8). *Landlords are accused of colluding to raise rents. See where.* The Washington Post. https://www.washingtonpost.com/business/interactive/2025/realpage-lawsuit-rent-map.

## Introduction

In October 2022, ProPublica published an investigative report[3] alleging that YieldStar, a highly popular “Revenue Management” software operated by RealPage, contributes to the sharply rising cost of rent in the United States. This software uses an algorithm to recommend rental prices to landlords, who then maximize profit for housing units. ProPublica alleged that the usage of this software amounts to a tacit price-fixing scheme, which uses the algorithm instead of a person as an intermediary between competing landlords who implicitly agree to set prices at the algorithm’s supra-competitive level. Since 2022, RealPage has been sued by, at least: the Department of Justice, by independent individuals, and by the governments of Arizona and Washington, D.C.[4] These lawsuits allege that RealPage’s algorithms decrease competition among landlords, and that landlords use algorithms as a tool for collusion to raise prices.[5] RealPage denies the allegations,[6] saying that customers set their own rents, are not obligated to follow a pricing recommendation, are not given access to nonpublic data from other landlords except in aggregate form, and do not receive recommendations based on nonpublic competitor data.

Multiple researchers have investigated this issue in recent years. Calder-Wang and Kim (2023)[7] utilize a private panel dataset of rental properties over time to estimate the price elasticities of multifamily buildings; they find evidence of rent impact from pricing algorithms, but cannot rule out the possibility that these algorithms are simply more efficient and responsive without any collusion.[8] Conversely, Calvano, Calzolari et al. (2020, 2021)[9, 10] find that a group of forward-looking ‘asynchronous’ learning algorithms can learn to charge supra-competitive prices even if they are not communicating with each other; in this instance, price-coordination could be an inherent feature of the technology. The Council of Economic Advisers to the Biden-Harris administration attempted to estimate the rent impact under an assumption of price-coordination while relying on the elasticities calculated by Calder-Wang and Kim; they found a national average rent uplift of 4% under these assumptions.[11]

---

[3] Vogell, H., Coryne, H., & Little, R. (2022, October 15). *Rent Going Up? One Company’s Algorithm Could Be Why.* ProPublica. https://www.propublica.org/article/yieldstar-rent-increase-realpage-rent.
[4] See, e.g., Press release. (2024, August 23). *Justice Department Sues RealPage for Algorithmic Pricing Scheme that Harms Millions of American Renters*. U.S. Department of Justice.
[5] Flowers, A., Yu Chen, S., Rich, S., & Lerman, R. (2025, January 8). *Landlords are accused of colluding to raise rents. See where.* The Washington Post.
https://www.washingtonpost.com/business/interactive/2025/realpage-lawsuit-rent-map
[6] *The Real Story: RealPage’s Response to the False Allegations Concerning Its Revenue Management Software*. RealPage. https://www.realpagepublicpolicy.com/realpagestatement.
[7] Calder-Wang, S., & Kim, G. H. (2023). Coordinated vs efficient prices: the impact of algorithmic pricing on multifamily rental markets. Available at SSRN 4403058.
[8] According to the authors, a model of price coordination is only favored by their results if non-adopters without algorithms are assumed to price efficiently; they do not make this assumption and “consider non-adopters to be somewhat unsophisticated.”
[9] Calvano, E., Calzolari, G., Denicolo, V., & Pastorello, S. (2020). *Artificial intelligence, algorithmic pricing, and collusion.* American Economic Review, 110 (10), 3267-3297.
[10] Calzolari, G., Calvano, E., Denicolo, V., & Pastorello, S. (2021). *Algorithmic collusion, genuine and spurious* (No. 16393). CEPR Discussion Papers.
[11] This analysis was removed after the inauguration of the Trump-Vance administration, but is still accessible via web archives. See *The Cost of Anticompetitive Pricing Algorithms in Rental Housing.*

We rely on a twofold approach to investigate this issue using only publicly available sources. First, we acquire lists of multifamily rental buildings from RealPage's 'Explore' website[12] and from the property management companies named in relevant lawsuits.[13] We match addresses in these lists and identify the subset of properties which are likely to use a RealPage product; we use multiple machine learning approaches to estimate the relationship between RealPage usage and building rents. Second, we test whether RealPage's 2017 acquisition of its largest competitor, Lease Rent Options, had a greater effect on rent levels in markets where the merged company had higher market penetration rates and merger share gains. If such an effect were detected, this would be evidence that any observed price increase is due to algorithmic collusion and not a non-collusive factor. We continue in more detail below.

# Data and Methods

## I. Machine Learning Prediction for Rental Properties: Propensity Score, Random Forest, Gradient Boosting, and Feed-Forward Neural Network Approaches

### A. Propensity Score Modeling

Our initial steps towards model development started with an extended data retrieval stage, where we compiled our own dataset of properties that utilized RealPage versus properties that did not. The first source we used to understand which property management firms used RealPage was Washington Post reporting, which lists companies named in lawsuits against RealPage.[14] We then individually scraped each property management firm listed to create a consolidated dataset that had each property listed out, using a variety of manual, automatic scraping techniques including python scripts utilizing BeautifulSoup[15] and Selenium.[16] The second dataset was taken from the RealPage explore page, which contains approximately 42,000 multifamily properties representing the broader rental market. Finally, we used fuzzy matching techniques between the two datasets to identify the subset of these properties which are likely to use RealPage.

Fuzzy matching techniques required multiple stages, as there were inconsistent property identifiers, redundancy, and data quality issues, with different naming conventions, abbreviations, and address formatting making it difficult to directly match properties across

---

(2024, December 17). The White House. https://web.archive.org/web/20250116070912/https://www.whitehouse.gov/cea/written-materials/2024/12/17/the-cost-of-anticompetitive-pricing-algorithms-in-rental-housing/.

[12] *RealPage Explore: Multifamily Real Estate Analytics*. https://www.realpage.com/explore/main.

[13] See Appendix 1.

[14] See Appendix 1. See also Flowers, A., Yu Chen, S., Rich, S., & Lerman, R. (2025, January 8). *Landlords are accused of colluding to raise rents. See where.* The Washington Post. https://www.washingtonpost.com/business/interactive/2025/realpage-lawsuit-rent-map.

[15] Richardson, L. (2007, April). *Beautiful Soup Documentation*. https://www.crummy.com/software/BeautifulSoup/.

[16] Selenium Project. (2023). *Selenium WebDriver*. https://www.selenium.dev/.

datasets. Additionally, many fields contained missing values, redundant information, or lacked variability. First, we standardized property names by lowercasing text, removing special characters, and normalizing common abbreviations (e.g., “St” to “Street”, “Apt” to “Apartment”). We then computed multiple similarity metrics between records, including Levenshtein distance[17] for character-level edits, Jaccard similarity[18] based on shared words, and a custom address match score that compared structured address components. These scores were combined into a composite match score, which we used to classify matches into high-confidence (score ≥ 75), possible (50–75), and unmatched (<50). The final dataset of 5,175 Realpage matches was greater than 95% confidence in matching, with the resulting matched dataset forming the foundation for analyzing pricing behavior and potential price collusion among Property Management Firms that utilize RealPage.

After matching, data preprocessing included multiple steps.  We reduced feature count from 37 to 15 features,[19] we filtered out properties in Core-Based Statistical Areas (CBSAs) that lacked any RealPage-managed listing, we removed student housing properties due to their different pricing models, and we converted“Year Built” to “Building Age”. All numeric columns were explicitly cast, with non-numeric values transitioned to missing.

After preprocessing, we opted to use a Propensity Score Model[20] to match properties based on observable characteristics. This was determined to be relevant because before matching within the algorithm, RealPage properties were systematically different, as the property management firms that tended to use RealPage were usually larger, newer, and had different occupancy rates compared to firms that did not. The Propensity Score Matching Algorithm helped in balancing both the RealPage property with a similar non-RealPage property based on observable characteristics and reduced overall bias. We began model development by loading the cleaned dataset and removing any rows with missing values The treatment variable was defined using a binary indicator of RealPage usage. To model the likelihood of RealPage usage, we selected several covariates to influence both pricing strategy and software adoption:

- Square footage: Larger units may attract different renter profiles and are more likely to be associated with modern construction, where RealPage adoption is more common.
- Number of stories: This captures vertical size; taller buildings may differ operationally and logistically.
- Unit count: A proxy for property scale, with larger apartment complexes more likely to utilize RealPage.

---

[17] Levenshtein, V. I. (1966). *Binary codes capable of correcting deletions, insertions, and reversals.* Soviet Physics Doklady, 10(8), 707–710.
[18] Jaccard, P. (1912). *The distribution of the flora in the alpine zone.* New Phytologist, 11(2), 37–50.
[19] We identified features to remove by running an OLS regression against the dependent variable, Rent per Square Foot, and dropping features which lacked predictive power. We also dropped features with near-zero variance, more than 50% missing data, or collinearity over 0.9 with another feature.
[20] Rosenbaum, P. R., & Rubin, D. B. (1983). *The central role of the propensity score in observational studies for causal effects.* Biometrika, 70(1), 41–55. https://doi.org/10.1093/biomet/70.1.41.

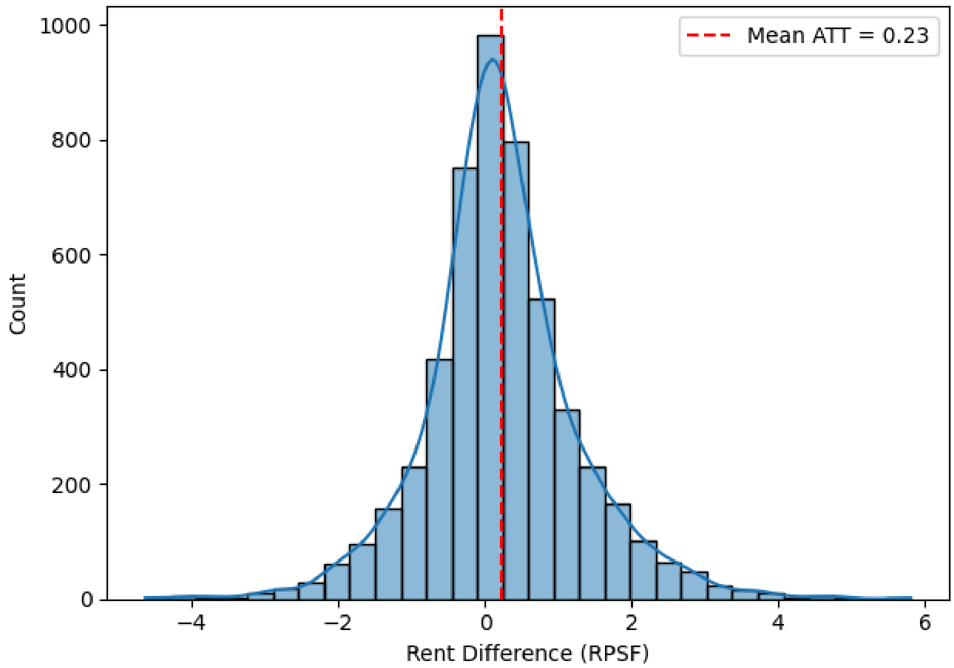
- Building age: Newer properties tend to command higher rents due to modern amenities and construction quality, and utilize RealPage.
- Occupancy rate: Reflects market demand and supply constraints, potentially influencing pricing behavior.

To account for regional variation in rental markets, we used each property's Core-Based Statistical Area (CBSA) code as a location proxy. Because CBSA is a categorical variable, we applied one-hot encoding[21] to generate binary indicator columns, allowing the model to account for regional heterogeneity without assuming a linear effect. We then constructed a covariate matrix comprising the selected numeric features and the one-hot encoded CBSA indicators. Prior to modeling, all covariates were standardized using a z-score transformation to provide comparability across features with different scales.

Next, we estimated each property's propensity to use RealPage via logistic regression. The model predicted the probability of treatment assignment conditional on covariates; these predicted probabilities, i.e., propensity scores were appended to the dataset. Using the estimated scores, we conducted 1:1 nearest-neighbor matching without replacement, pairing each treated property (RealPage user) with the closest control (non-user) based on propensity score. This matching approach was designed to replicate a randomized controlled trial by aligning each RealPage user with a statistically similar non-user, thereby reducing confounding from observable characteristics. The matching process resulted in 5,089 balanced treatment-control pairs. Following matching, we combined the treated and matched control units into a single dataset to estimate the Average Treatment Effect on the Treated (ATT). The outcome variable was the Market-Processed Rent per Square Foot (MPF-RPSF). We found that RealPage usage was associated with a $0.23 (shown in Figure 1) increase in average RPSF for the treated group, suggesting a statistically meaningful effect on rent pricing.[22]

[21] Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.
[22] Detailed results are presented in Appendix 2.

**Figure 1: Distribution of Rent Differences from the Propensity Algorithm**




## B. Additional Modeling, Simulations, and Statistical Testing

For robustness, we tested additional predictive models – including Linear Regression, Random Forest, Gradient Boosting, and Neural Networks – with the intention to methodically check for correlation between RealPage usage and increased rental pricing. These models relied on the same base dataset as in the Propensity Score model, but without the matching process, varying feature selection, and with the addition of derived market shares for most Metropolitan Statistical Areas (MSAs) which had missing market shares.[23] This was done by taking the total number of realpage users within a corresponding CBSA over the total number of properties in that CBSA in our data. Initial models, such as Linear Regression, provided baseline benchmarks, while Random Forest and Gradient Boosting improved on these by capturing complex interactions and nonlinear relationships among predictors.[24, 25, 26] Selected key variables included:

- Unit square footage

[23] This was done by calculating the share of properties in the available data which use RealPage. The average error rate using this method for non-missing MSAs was 7 percentage points.
[24] Virtanen, P., Gommers, R., et al. (2020). *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.* Nature Methods, 17(3), 261-272. *DOI: 10.1038/s41592-019-0686-2.*
[25] Buitinck, L., Louppe, G., et al. (2013). *API design for machine learning software: experiences from the scikit-learn project*. ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 108-122.
[26] Abadi, M., Agarwal, A., et al., (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* tensorflow.org.

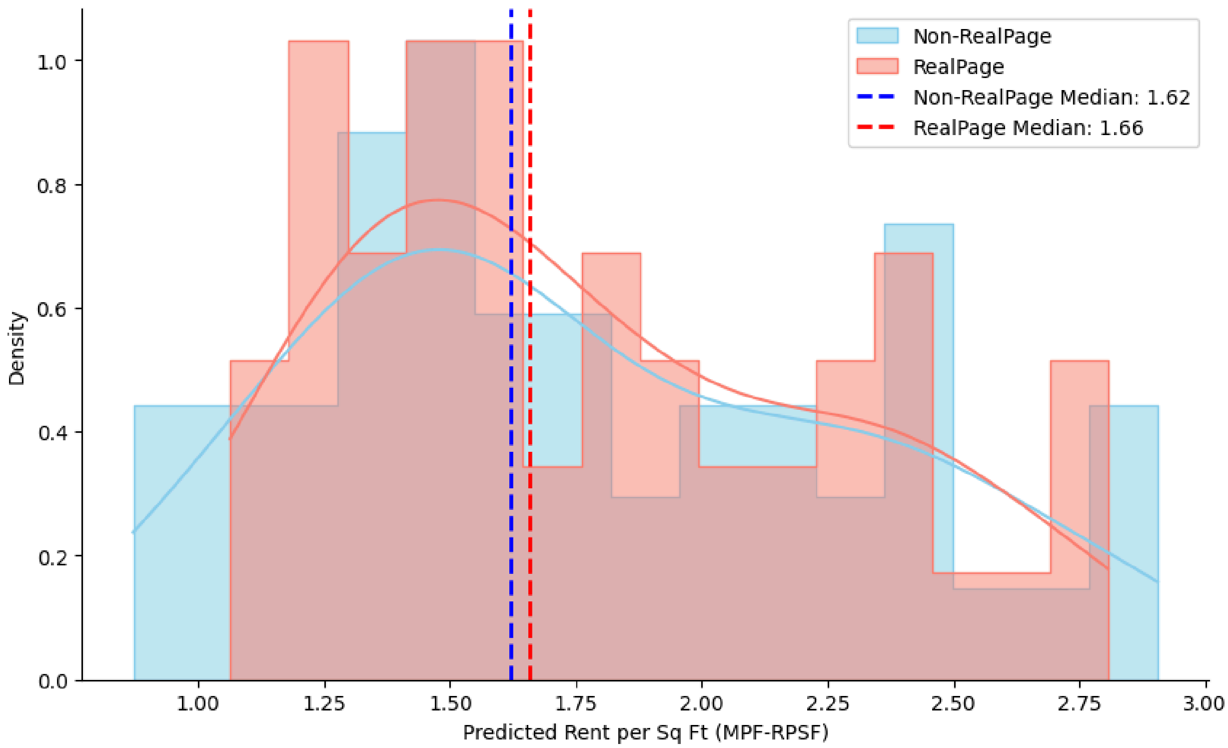
- Number of stories
- Building age (derived from year constructed)
- Unit count
- Occupancy rate
- Geographical market identifiers via CBSA codes
- Our proxy for Realpage Market share
- Quality of the property (class): a rating of A, B, or C is given
- A flag to indicate if the property was a realpage user or not
- Rent Per Square Foot (output variable)

We benchmark these models with metrics such as the Mean Squared Error (MSE) and $R^2$ ;the Neural Network was identified as the optimal model due to its superior scores in these metrics. The chosen architecture consisted of a multi-layer perceptron (MLP), implemented via TensorFlow and Keras frameworks, featuring one input layer, two hidden layers employing ReLU activation for capturing nonlinearity, and an output layer utilizing linear activation for continuous predictions.[27]

Using the FFNN, we performed predictions on simulated properties reflective of ones included within our dataset to quantify the impact of the Realpage algorithm's adoption on rental prices. Using the same properties, we provided predictions for the same properties with the model given that one of them was a realpage user and the other was not. This simulation consisted of 50 properties. Figure 2, below, represents the distributions of the predictions:

[27] Appendix 3 shows our benchmarking results and the neural network structure.

**Figure 2: Distributions of Predictions of RPSF for Realapge vs Non-RealPage Users**




To conclusively validate our findings, we employed a Mann-Whitney U test,[28] a robust non-parametric statistical method typically used in cases of non-normally distributed data. Our alternative hypothesis in this case was that the distribution of predicted prices for realpage users would be different from that of Non-RealPage users. Using a right-tailed test, the statistical test ended with the result of a p-value of *0.4465* which unfortunately does not show statistically significant results, therefore we cannot reject the null hypothesis using this method to show that there is a difference in price between the users.

## II. Test for Price-Coordination: RealPage's 2017 Merger with LRO

### A. Background

As demonstrated above, our predictive models show a relationship between RealPage usage and average rent per square foot at the level of individual properties in 2025. However, it remains to be seen whether this rent increase is the result of price-coordination. RealPage, in its official statement and its defense against nationwide lawsuits, objects to this allegation, asserting that customers set their own rents, are not obligated to follow a pricing recommendation, are not given access to nonpublic data from other landlords except in aggregate form, and do not receive recommendations based on nonpublic competitor data. In

[28] Shier, R. (2004). *Statistics: 2.3 The Mann-Whitney U Test.* Mathematics Learning Support Centre.

RealPage's view, algorithmic pricing products are pro-competitive tools which simply help landlords price their units efficiently under market rates.[29]

Price coordination can be demonstrated via a number of approaches. Some economists may test (using Ordinary Least Squares) whether a time series of market prices increases more than expected during an alleged period of collusion; others may test whether unit prices estimated from a model trained on a pre-collusion period substantially differ from reality in a post-collusion period.[30] Meanwhile, Calder-Wang and Kim (2023) tested for price coordination among rental algorithm users using a panel dataset of individual rental properties over time; they found that landlords using algorithms tend to be more responsive to economic shocks, and that markets with greater algorithmic penetration tend to have higher rents and lower occupancy. However, they caution that these findings could still be consistent with a model of efficient price responses without collusion.[31]

In our case, we turn our attention to an event which sharply increased the penetration rate of RealPage-owned products in the housing market. Specifically, in December 2017, RealPage completed an acquisition of Lease Rent Options (LRO).[32] LRO was the primary competitor of RealPage's products, YieldStar and AIRM. By 2024, the merged company commanded approximately 80% of the market for commercial revenue management.[33] We hypothesize that a significant increase in rent for highly affected markets as a result of the merger would be evidence of price coordination. This is because the merger did not involve any actual consolidation of competing landlords in the housing market; nor would it immediately change algorithmic pricing technology or increase the total penetration rate of the technology. The only change is to increase the number of competing landlords who license their revenue management software from the same ownership group. Those landlords should only gain market power directly from the merger if the merger facilitated increased price coordination.

## B. Zillow 'ZORI' Index Data and COVID Adjustment

---

[29] *The Real Story: RealPage's Response to the False Allegations Concerning Its Revenue Management Software*. RealPage. https://www.realpagepublicpolicy.com/realpagestatement.

[30] See summary by, e.g., Gilbert, S.D. (2019). Testing for Price-Fixing Effects: A Difference-in-Difference Approach.

[31] According to the authors, a model of price coordination is only favored by their results if non-adopters without algorithms are assumed to price efficiently; they do not make this assumption and "consider non-adopters to be somewhat unsophisticated." See Calder-Wang, S., & Kim, G. H. (2023). Coordinated vs efficient prices: the impact of algorithmic pricing on multifamily rental markets. Available at SSRN 4403058.

[32] Lane, B. (2017, December 5). *RealPage completes $300 million acquisition of Lease Rent Options* Housingwire, https://www.housingwire.com/articles/42017-realpage-completes-300-million-acquisition-of-lease-rent-options/.

[33] Press release. (2024, August 23). *Justice Department Sues RealPage for Algorithmic Pricing Scheme that Harms Millions of American Renters*. U.S. Department of Justice. https://www.justice.gov/archives/opa/pr/justice-department-sues-realpage-algorithmic-pricing-scheme-harms-millions-american-renters.

We analyze this merger using the Zillow Observed Rent Index (ZORI), a monthly panel dataset of rents for U.S. Metropolitan Statistical Areas (MSAs).[34] This index is calculated using the mean of listed rents between the 35th and 65th percentile for multifamily rental units in each region.[35] We limit this panel data to MSAs for which we can calculate YieldStar, AIRM, and LRO's penetration rates at a point in time based on RealPage's official public statement and American Community Survey data.[36] After calculating the penetration rates, we divide the cities into treatment and control groups based on the total penetration rate of the merged company, and the share added by acquiring LRO.[37]

Before testing these groups, however, we first adjust this dataset to account for the impact of the COVID-19 pandemic on the housing market. This adjustment is essential because the pandemic had a significant impact on rent indices in all cities, while also having *different*, *nonlinear* impacts on different cities at different times. For example, consider the difference in impact between Phoenix, AZ and Washington, DC in Figures 3 and 4 (below):

---

[34] *ZORI (Smoothed): Multi Family Residence Time Series ($)* [Data set]. Zillow Research. Retrieved March 8, 2025, from https://www.zillow.com/research/data/.

[35] This mean is weighted by the prevalence of units by structure type, decade built, and year rented triple. See Clark, J. (2022, September 19). *Methodology: Zillow Observed Rent Index*. Zillow Research. https://www.zillow.com/research/methodology-zori-repeat-rent-27092/.

[36] This is the set of MSAs for which RealPage disclosed its products' penetration rates in its 2023 statement. See *The Real Story: RealPage's Response to the False Allegations Concerning Its Revenue Management Software*. RealPage. https://www.realpagepublicpolicy.com/realpagestatement. RealPage presents these penetration rates as the share of all rental units, so we adjust the penetration rates to be the share of all *multifamily* rental units using 2023 American Community Survey and American Housing Survey data. See U.S. Census Bureau, U.S. Department of Commerce. (2023). *Units in Structure. American Community Survey, ACS 1-Year Estimates Detailed Tables, Table B25024* [Data set]. Retrieved March 8, 2025, from https://data.census.gov/table/ACSDT1Y2023.B25024?q=B25024&g=040XX00US20. See also U.S. Census Bureay, U.S. Department of Housing. (2023). *2023 National – General Housing Data – All Occupied Units, American Housing Survey (AHS) Table Creator* [Data set]. Retrieved March 8, 2025, from https://www.census.gov/programs-surveys/ahs/data/interactive/ahstablecreator.html?s_areas=00000&s_year=2023&s_tablename=TABLE1&s_bygroup1=2&s_bygroup2=1&s_filtergroup1=1&s_filtergroup2=1.

[37] The treatment group is the set of cities which has at least 35% penetration rate and at least 5% rate gain from the merger: Atlanta, Dallas, Phoenix, Denver, Tampa, and Washington DC. We test two versions of the control group: the first is the set of cities which have less than 5% share gain: Minneapolis, San Diego, Miami, San Francisco, Chicago, Detroit, Los Angeles, and New York (the remaining cities are dropped). The second control group includes every available non-treatment city as a control city instead of being dropped.

**Figure 3: Actual ZORI Index and COVID-Adjusted Index for Washington, DC**

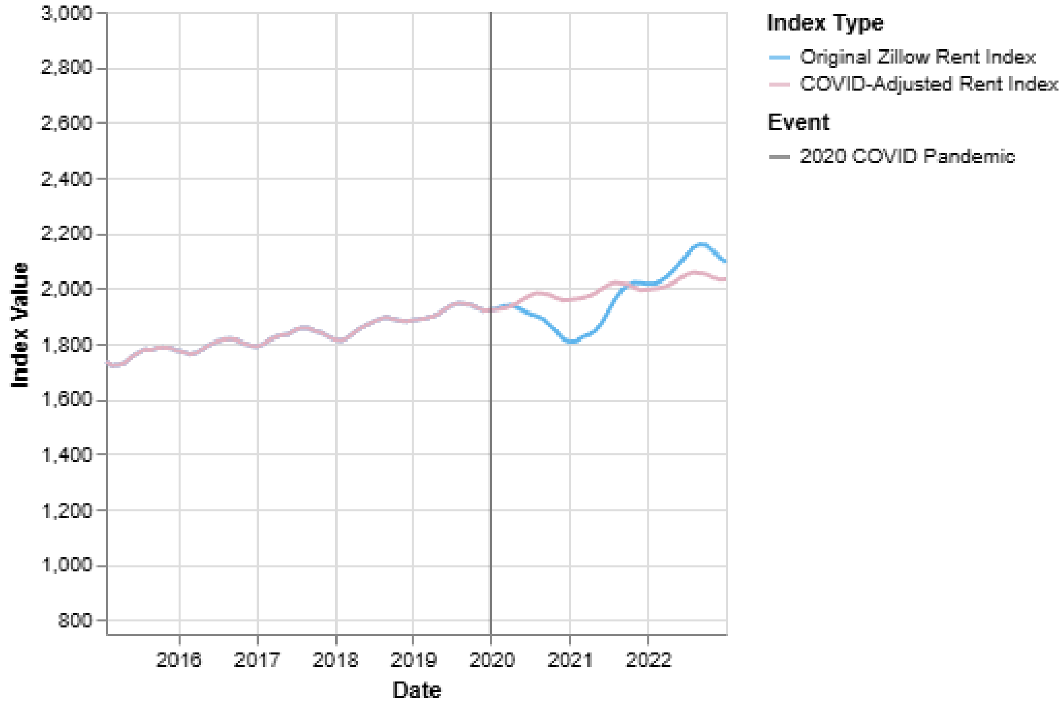
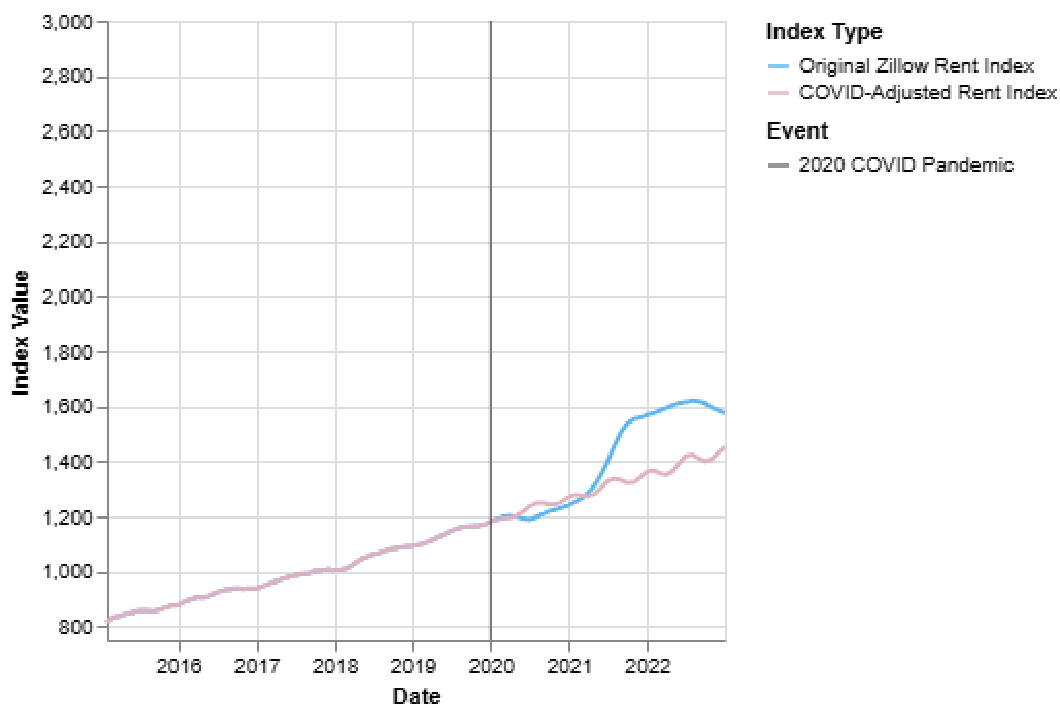



**Figure 4: Actual ZORI Index and COVID-Adjusted Index for Phoenix, AZ**



The nonlinear, highly variable impact of the pandemic is apparent in these visualizations. In 2020, both cities departed from a steady long-term trend of annual rent growth with a 12-month seasonal cycle, but in different ways. Washington, DC first saw a sharp drop in average rent levels, but later saw a sharp rebound to above-normal levels by the end of 2023. Meanwhile, Phoenix saw a smaller impact until the start of 2021, at which point rent sharply increased – perhaps because of migration trends as renters left high-cost cities to move to low-cost ones. Each city in our dataset experiences its own break from long-run trends as a result of COVID.

Clearly, it is essential to account for these impacts before attempting any analysis of the 2017 merger; otherwise, the impact of COVID may be improperly assigned to the merger. To do this, we use the statsmodels[38] package in Python to tune ARIMA models on the pre-COVID, post-merger period for each MSA. We use these models to forecast a counterfactual index value for each month post-COVID, up to the end of 2022. For robustness, we validate ARIMA by using models trained on 2018 data to predict 2019 values, and by using models trained on 2015-2016 data to predict 2017 values – the results of these tests are presented in Appendix 5. In all merger analysis including the post-merger period, we rely on the counterfactual ARIMA forecasts, which are presumed to include the effect of the merger but not the effect of COVID. Examples of the counterfactual forecasts for Washington, DC and Phoenix, AZ are presented in Figures 3 and 4, above.

## C. Difference-in-Difference and Synthetic Control Estimation

After creating these counterfactual forecasts for the post-COVID period, we analyze the impact of the December 2017 merger. We first consider difference-in-difference regression, drawing on Gilbert (2019) who points out that this method can be used to test for effects on a market during an alleged 'collusion' period.[39] The result of this approach is not statistically significant (see Appendix 6). Moreover, when testing for parallel trends in the pre-treatment period, we find that the treatment group on average has a lower rent growth trend than the control group, with a p-value well beyond a 1% significance threshold. The parallel trend assumption is critical for successful difference-in-difference analysis;[40] our tests reject this assumption, so we turn to an alternate method: the Synthetic Control method.

This method, pioneered by Abadie et al. (2003, 2010),[41] has grown in popularity over the past two decades, in part because it *controls for* unobserved differences between the treatment and control groups rather than *assuming* that these differences remain unchanged and

---

[38] Seabold, S., and Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with python*. 9th Python in Science Conference. https://www.statsmodels.org/stable/index.html.
[39] Gilbert, S.D. (2019). Testing for Price-Fixing Effects: A Difference-in-Difference Approach.
[40] See Appendix 6. See also, e.g., Angrist J., Pischke J.S. (2009). *Mostly Harmless Econometrics*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4j72.
[41] Abadie, A., and Gardeazabal, J. (2003, March). *The Economic Costs of Conflict: A Case Study of the Basque Country*. American Economic Review, *93*(1), 113-132. See also Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493-505.

parallel.[42] We implement this method using Ben-Michael et al. 's ‘augsynth’ R package.[43] For each treatment city, we calculate a separate weighted average of all control cities which closely fits the treatment city in the pre-merger period. This weighted average – the namesake ‘synthetic control’ – is assumed to control for the differences between the treatment city & control cities. We then observe whether the treatment city’s ZORI index and the ZORI of the synthetic control diverge in the post-merger period; any divergence is our estimated treatment effect of the merger. These results are presented in Figure 5, below.

**Figure 5: Synthetic Control Average Treatment Effects**

| | Specification 1 | | Specification 2 | | Specification 3 | |
|---|---|---|---|---|---|---|
| Control Group | Lowest Share Cities | | All Non-Treatment Cities | | Lowest Share Cities | |
| Data End Date | December 2022 | | December 2022 | | December 2019 | |
| | Estimate | P-Value | Estimate | P-Value | Estimate | P-Value |
| Atlanta, GA | -28.4 | 0.567 | -62.2** | 0.043 | 32.8*** | 0.000 |
| Dallas, TX | -69.6 | 0.166 | -65.7 | 0.103 | -26.5*** | 0.000 |
| Denver, CO | -7.6 | 0.934 | -7.3 | 0.376 | -1.9 | 0.733 |
| Phoenix, AZ | 103.7*** | 0.000 | 56.0** | 0.021 | 60.4*** | 0.000 |
| Tampa, FL | -1.5** | 0.017 | -34.0 | 0.390 | 14.9*** | 0.006 |
| Washingon, DC | -30.5*** | 0.000 | -35.1* | 0.080 | -5.7 | 0.135 |

*Notes: Cities are Metropolitan Statistical Areas (MSAs). The “Lowest Share Cities” are Minneapolis, San Diego, Miami, San Francisco, Chicago, Detroit, Los Angeles, and New York: See Footnote 37 and Appendix 4. Estimates are average monthly post-merger treatment effects on the ZORI multifamily rent index. Asterisks (*) indicate significance thresholds: (*) 10%, (**) 5%, (***) 1%.*

These results show that estimates of the merger’s effect on rent vary based on city, control group definition, and the date range considered. Of the six treatment group cities, only Phoenix is statistically significant at the 5% level with the same direction across all specifications. Atlanta and Tampa are significant for at least two specifications, but with different directions of the treatment effect, while the remaining three cities are only significant at the 5% level for one or zero specifications.

[42] Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., & Sutton, M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health economics*, *25*(12), 1514-1528.
[43] Ben-Michael, E., Feller, A., & Rothstein, J. (2021). The augmented synthetic control method. *Journal of the American Statistical Association*, *116*(536), 1789-1803.

In summary: no city except Phoenix has a significant result which is consistent across multiple specifications, no specification has a result which is significant and consistent across cities, and the direction of the treatment effect is not consistent across either specifications or cities. While there may be some evidence of a statistically significant post-merger rent increase in Phoenix, we caution against this interpretation. If the observed rent increase was because of price coordination based on market shares, we should expect to see a similar effect in Atlanta and Dallas, where RealPage's penetration rate is even higher than in Phoenix. Without such a result, there is no consistent causal interpretation of the merger's effects.

# Implications and Limitations

Our predictive models identify an average rent uplift for properties in publicly available data which are likely to use RealPage; in particular, our propensity score model is statistically significant with a magnitude of about $0.23 of rent increase per square foot. However, difference-in-difference and synthetic control analysis of the 2017 RealPage-LRO merger does not show any statistically significant effect on rent levels when the market for algorithmic pricing consolidated. These findings are consistent with previous authors such as Calder-Wang and Kim (2023),[44] who found that properties using algorithmic pricing tend to be more responsive to market shocks and set prices higher during periods of economic recovery; however, the authors could not rule out models where properties do not coordinate. Renters and policymakers nationwide may have an interest in the correlation between algorithmic pricing and rent levels; we communicate our findings on our interactive website, accessible at https://uc-berkeley-i-school.github.io/realpage-rent-impact/.[45]

It is important to understand the limitations of our findings. Specifically, we find that the 2017 merger did not lead to a statistically significant increase in rent levels. This means we do not find evidence of price coordination; however we also cannot *rule out* alternate models of price coordination. We show that the relative rent levels between highly-affected cities and less-affected cities did not change after the RealPage-LRO merger, but other possible explanations remain.

First, it is possible that RealPage's products and LRO's products independently have price-coordination effects, but this did not increase as a result of the merger. RealPage asserted in 2023 that its original products, YieldStar and AIRM, remain "completely separate" from LRO and do not have access to LRO's database (and vice versa).[46] If this is the case, then it would be unsurprising for the RealPage-LRO merger to show no effect; even if price coordination exists for YieldStar/AIRM and LRO separately, the merger would not create additional coordination between the different platforms.

[44] Calder-Wang, S., & Kim, G. H. (2023). Coordinated vs efficient prices: the impact of algorithmic pricing on multifamily rental markets. *Available at SSRN 4403058*.
[45] Davies, C., Benzoni, P., Yim, P., Allaou, A., and Majidzadeh, T. (2025, April 14). *Is Algorithmic Pricing Raising Your Rent? A Data-Driven Investigation*. https://uc-berkeley-i-school.github.io/realpage-rent-impact/.
[46] *The Real Story: RealPage's Response to the False Allegations Concerning Its Revenue Management Software*. RealPage. https://www.realpagepublicpolicy.com/realpagestatement.

Additionally, it could be the case that algorithms need not communicate with each other or share the same ownership to 'learn' to collude; this was the finding of Calvano, Calzolari et al. (2020, 2021),[47, 48] who found that some types of independent algorithms could still converge to a supra-competitive price even without explicitly working together. If this is the case, then the merger of RealPage with LRO might not be necessary to have implicit, 'learned' coordination between RealPage's algorithms and LRO's algorithms. Under this model, algorithmic coordination between YieldStar/AIRM users and LRO users already existed pre-merger and was not exacerbated by the merger.

Alternate methods and more detailed data may be necessary to test these explanations, which are neither supported nor rejected by our findings. These complex dynamics present a challenge for renters who may experience increased rents due to algorithmic factors, and for policymakers who must navigate these complexities to ensure fair housing practices. Further research and ongoing study of the rental housing market are crucial to understanding the full impact of algorithmic pricing and its potential effects on housing affordability and competition.

---

[47] Calvano, E., Calzolari, G., Denicolo, V., & Pastorello, S. (2020). *Artificial intelligence, algorithmic pricing, and collusion.* American Economic Review, 110 (10), 3267-3297.
[48] Calzolari, G., Calvano, E., Denicolo, V., & Pastorello, S. (2021). *Algorithmic collusion, genuine and spurious* (No. 16393). CEPR Discussion Papers.

# References

*Generative AI Acknowledgement*: We used ChatGPT and Gemini in a limited capacity as a creative writing aid to create introduction and concluding sentences in our presentation and report.

Abadie, A., and Gardeazabal, J. (2003, March). *The Economic Costs of Conflict: A Case Study of the Basque Country*. American Economic Review, *93*(1), 113-132.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). *Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program*. Journal of the American Statistical Association, *105*(490), 493-505.

Angrist J., Pischke J.S. (2009). *Mostly Harmless Econometrics*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4j72.

Asker, J., Chaim F., and Ariel, P., (2022, May). *Artificial intelligence, algorithm design, and pricing.* American Economic Association Papers and Proceedings. 112, 452-456.

Ben-Michael, E., Feller, A., & Rothstein, J. (2021). *The augmented synthetic control method.* Journal of the American Statistical Association, *116*(536), 1789-1803.

Buitinck, L., Louppe, G., et al. (2013). API design for machine learning software: experiences from the scikit-learn project. ECML PKDD Workshop: Languages for Data Mining and Machine Learning, 108-122.

Calder-Wang, S., & Kim, G. H. (2023). *Coordinated vs efficient prices: the impact of algorithmic pricing on multifamily rental markets.* Available at SSRN 4403058.

Calvano, E., Calzolari, G., Denicolo, V., & Pastorello, S. (2020). *Artificial intelligence, algorithmic pricing, and collusion.* American Economic Review, 110 (10), 3267-3297.

Calzolari, G., Calvano, E., Denicolo, V., & Pastorello, S. (2021). *Algorithmic collusion, genuine and spurious* (No. 16393). CEPR Discussion Papers.

Clark, J. (2022, September 19). *Methodology: Zillow Observed Rent Index*. Zillow Research. https://www.zillow.com/research/methodology-zori-repeat-rent-27092/.

Davies, C., Benzoni, P., Yim, P., Allaou, A., and Majidzadeh, T. (2025, April 14). *Is Algorithmic Pricing Raising Your Rent? A Data-Driven Investigation*. https://uc-berkeley-i-school.github.io/realpage-rent-impact/.

Flowers, A., Yu Chen, S., Rich, S., & Lerman, R. (2025, January 8). *Landlords are accused of colluding to raise rents. See where.* The Washington Post. https://www.washingtonpost.com/business/interactive/2025/realpage-lawsuit-rent-map.

Google (2025). Gemini [Large language model]. https://gemini.google.com/app.

Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media

Gilbert, S.D. (2019). *Testing for Price-Fixing Effects: A Difference-in-Difference Approach.*

Jaccard, P. (1912). *The distribution of the flora in the alpine zone.* New Phytologist, 11(2), 37–50.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., & Sutton, M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health economics*, *25*(12), 1514-1528.

Lane, B. (2017, December 5). *RealPage completes $300 million acquisition of Lease Rent Options*. Housingwire, https://www.housingwire.com/articles/42017-realpage-completes-300-million-acquisition-of-lease-rent-options/.

Levenshtein, V. I. (1966). *Binary codes capable of correcting deletions, insertions, and reversals.* Soviet Physics Doklady, 10(8), 707–710.

OpenAI. (2025). ChatGPT [Large language model]. https://chatgpt.com.

Press release. (2024, August 23). *Justice Department Sues RealPage for Algorithmic Pricing Scheme that Harms Millions of American Renters*. U.S. Department of Justice. https://www.justice.gov/archives/opa/pr/justice-department-sues-realpage-algorithmic-pricing-scheme-harms-millions-american-renters.

*RealPage Explore: Multifamily Real Estate Analytics*. https://www.realpage.com/explore/main.

Richardson, L. (2007, April). *Beautiful Soup Documentation*. https://www.crummy.com/software/BeautifulSoup/.

Rosenbaum, P. R., & Rubin, D. B. (1983). *The central role of the propensity score in observational studies for causal effects.* Biometrika, 70(1), 41–55. https://doi.org/10.1093/biomet/70.1.41.

Seabold, S., and Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with python*. 9th Python in Science Conference. https://www.statsmodels.org/stable/index.html.

Selenium Project. (2023). *Selenium WebDriver*. https://www.selenium.dev/.

Shier, R. (2004). *Statistics: 2.3 The Mann-Whitney U Test.* Mathematics Learning Support Centre.

*The Real Story: RealPage's Response to the False Allegations Concerning Its Revenue Management Software*. RealPage. https://www.realpagepublicpolicy.com/realpagestatement.

*The Cost of Anticompetitive Pricing Algorithms in Rental Housing.* (2024, December 17). The White House. https://web.archive.org/web/20250116070912/https://www.whitehouse.gov/cea/written-materials/2024/12/17/the-cost-of-anticompetitive-pricing-algorithms-in-rental-housing/.

U.S. Census Bureau, U.S. Department of Commerce. (2023). *Units in Structure. American Community Survey, ACS 1-Year Estimates Detailed Tables, Table B25024* [Data set]. Retrieved March 8, 2025, from https://data.census.gov/table/ACSDT1Y2023.B25024?q=B25024&g=040XX00US20.

U.S. Census Bureay, U.S. Department of Housing. (2023). *2023 National – General Housing Data – All Occupied Units, American Housing Survey (AHS) Table Creator* [Data set]. Retrieved March 8, 2025, from https://www.census.gov/programs-surveys/ahs/data/interactive/ahstablecreator.html?s_areas=00000&s_year=2023&s_tablename=TABLE1&s_bygroup1=2&s_bygroup2=1&s_filtergroup1=1&s_filtergroup2=1.

Virtanen, P., Gommers, R., et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17(3), 261-272. DOI: 10.1038/s41592-019-0686-2.

Vogell, H., Coryne, H., & Little, R. (2022, October 15). *Rent Going Up? One Company's Algorithm Could Be Why.* ProPublica. https://www.propublica.org/article/yieldstar-rent-increase-realpage-rent.

*ZORI (Smoothed): Multi Family Residence Time Series ($)* [Data set]. Zillow Research. Retrieved March 8, 2025, from https://www.zillow.com/research/data/.

# Appendices

## Appendix 1: List of Landlords Named in Lawsuits against RealPage

According to the Washington Post,[49] a list of property managers named in lawsuits against realpage includes: Allied Orion Group, Air Communities, Avenue 5 Residential, Bell Partners, BH Management Services LLC, Bozzuto Management Co., Brookfield Properties Multifamily, Camden Property Trust, Cortland Management Corp., CH Real Estate Services, CONAM Management Corp., CWS Apartment Homes, Dayrise Residential, ECI Group, Equity Residential Services, Essex Property Trust, FPI Management, Gables Residential Services, Greystar Management Services, Highmark Residential, HSL Properties, Independence Realty Trust, JBG Smith Properties, Kairoi Management, Knightvest Residential, Lantower Luxury Living, Mid-America Apartments, Mission Rock Residential, Morgan Properties Management Co., Paradigm Management, Pinnacle Property Management Services, Prometheus Real Estate Group, Rose Associates, RPM Living, Sares Regis Group, Security Properties Residential, Sherman Associates, Simpson Property Group, the Related Companies, Thrive Communities, UDR Inc., Weidner Property Management, William C. Smith & Co., Willow Bridge Property Co. (formerly Lincoln Property Co.), Windsor Property Management Co. and ZRS Management LLC.

Out of these, we were unable to access a list of managed properties for: Avenue 5 Residential, BH Management Services LLC, CH Real Estate Services, Cortland Management Corp., and Pinnacle Property Management Services.

[49] Flowers, A., Yu Chen, S., Rich, S., & Lerman, R. (2025, January 8). *Landlords are accused of colluding to raise rents. See where.* The Washington Post. https://www.washingtonpost.com/business/interactive/2025/realpage-lawsuit-rent-map.

## Appendix 2: Propensity Score Matching Algorithm Results

**Propensity Score Matching Regression Results**

| | Estimate | Std. Error | Z-Value | P-Value |
|---|---|---|---|---|
| Constant | -2.0034 | 0.0173 | -115.94 | <0.0001 |
| Avg. Square Footage | -0.0690 | 0.0172 | -4.01 | 0.0001 |
| Stories | 0.0914 | 0.0125 | 7.30 | <0.0001 |
| Unit Count | 0.3494 | 0.0144 | 24.21 | <0.0001 |
| Building Age | -0.5465 | 0.0199 | -27.41 | <0.0001 |
| Occupancy Rate | -0.0246 | 0.0136 | -1.82 | 0.0695 |
| Pseudo R-squared | 0.061 | | | |
| Observations | 37,294 | | | |

**Final OLS Regression Coefficients (MPF-RPSF)**

| | Coefficient | Std. Error | t | P-Value |
|---|---|---|---|---|
| Intercept | -2.0034 | 0.0173 | -115.94 | <0.0001 |
| Uses_realpage | -0.0690 | 0.0172 | -4.01 | 0.0001 |
| R-squared | 0.021 | | | |
| F-statistic | 219.4 | | | |
| Observations | 10,178 | | | |
| Model Type | Ordinary Least Squares (OLS) | | | |

## Appendix 3: Additional Predictive Models Specifications and Precision

**Predictive Models and Results**

| Model | Description | Mean Squared Error (MSE) | $R^2$ Score |
|---|---|---|---|
| Random Forest | Baseline Model | 0.040 | 0.921 |
| Gradient Boosting | Ensemble method that captures non-linear relationships. | 0.06 | 0.89 |
| Feed Forward Neural Network | 3 Hidden layers (64, 32, 16). ReLu activations and early stopping. | 0.038 | 0.925 |

**Appendix 4: RealPage Products' Penetration Rates for Multifamily Rental Units by City, May 2023**

| | YieldStar / AIRM Share | LRO Share | Total Share |
|---|---|---|---|
| Atlanta, GA | 40.0% | 22.3% | 62.3% |
| Dallas, TX | 33.2% | 15.6% | 48.8% |
| Phoenix, AZ | 30.9% | 14.4% | 45.4% |
| Denver, CO | 31.1% | 11.7% | 42.8% |
| Tampa, FL | 27.4% | 12.5% | 39.9% |
| Washington, DC | 20.5% | 16.4% | 36.9% |
| Houston, TX | 24.5% | 8.9% | 33.4% |
| Riverside, CA | 23.3% | 5.8% | 29.1% |
| Las Vegas, NV | 20.2% | 6.6% | 26.8% |
| Seattle, WA | 14.5% | 8.2% | 22.7% |
| Philadelphia, PA | 7.3% | 13.6% | 20.9% |
| Boston, MA | 12.1% | 7.7% | 19.8% |
| Minneapolis, MN | 11.8% | 4.9% | 16.7% |
| San Diego, CA | 12.1% | 3.3% | 15.5% |
| Miami, FL | 10.2% | 3.7% | 13.9% |
| San Francisco, CA | 8.0% | 5.3% | 13.3% |

| | | | |
|---|---|---|---|
| Chicago, IL | 8.4% | 3.6% | 12.0% |
| Detroit, MI | 5.0% | 4.5% | 9.5% |
| Los Angeles, CA | 4.8% | 2.9% | 7.7% |
| New York, NY | 1.4% | 2.2% | 3.6% |

*Note: Multifamily units are considered to be rental units in buildings with at least 5 units. RealPage disclosed its own penetration rates for all rental units; we adjust these to be the share of multifamily units using the 2023 American Community Survey and American Housing Survey.*

## Appendix 5: Validation of ARIMA Forecasting for COVID Adjustment

The ZORI monthly time series for each city is fit to an ARIMA model in a pre-period and evaluated on a post-period. The tuned ARIMA parameters are p=3, d=1, q=1, P=0, D=1, Q=1, S=12, except for Chicago, which uses p=1.[50]

**ARIMA Validation Results**

| Training Dates | January 2015 to December 2016 | December 2017 to December 2018 |
|---|---|---|
| Evaluation Dates | January 2017 to December 2017 | January 2019 to December 2019 |
| | Root Mean Squared Error of Predicted ZORI Values | |
| New York, NY | 14.295 | 12.223 |
| Los Angeles, CA | 10.906 | 14.539 |
| Chicago, IL | 17.887 | 7.745 |
| Dallas, TX | 20.673 | 6.712 |
| Houston, TX | 24.245 | 6.803 |
| Washington, DC | 22.057 | 19.074 |
| Philadelphia, PA | 7.880 | 6.977 |
| Miami, FL | 8.055 | 8.754 |
| Atlanta, GA | 13.148 | 23.040 |
| Boston, MA | 15.942 | 7.899 |
| Phoenix, AZ | 9.525 | 5.545 |
| San Francisco, CA | 87.706 | 22.189 |

[50] Chicago diverges when using a value of p=3 to predict post-COVID values.

| | | |
|---|---|---|
| Riverside, CA | 34.442 | 17.388 |
| Detroit, MI | 52.271 | 31.250 |
| Seattle, WA | 25.113 | 4.011 |
| Minneapolis, MN | 13.638 | 9.412 |
| San Diego, CA | 28.709 | 2.353 |
| Tampa, FL | 12.762 | 18.679 |
| Denver, CO | 8.304 | 5.781 |
| Las Vegas, NV | 27.912 | 20.090 |

## Appendix 6: Difference-in-Difference Regression, Parallel Trends and Placebo Tests

We divide our available cities into treatment and control groups,[51] and test the below equation:

$$ZORI_{it} = \beta_0 + \beta_1 AffectedCity_i + \beta_2 AffectedCity_i * PostMerger_t + \beta_3 MonthlyTimeTrend_t + MonthFixedEffects_t + u_{it}$$

**Difference-in-Difference Regression Results**

| | Base Model | | Placebo Test | |
|---|---|---|---|---|
| Control Group | All Non-Treatment Cities | | All Non-Treatment Cities | |
| Treatment Date | December 2017 | | December 2016 | |
| Data End Date | December 2019 | | November 2017 | |
| | Estimate | P-Value | Estimate | P-Value |
| Constant | 1446.74 (49.811)*** | 0.000 | 1440.442 (62.558)*** | 0.000 |
| Time Trend | 5.54 (0.930)*** | 0.000 | 5.95 (2.124)*** | 0.005 |
| Affected City | -325.37 (28.425)*** | 0.000 | -321.69 (35.800)*** | 0.000 |
| Affected City x Post-Treatment | -19.802 (40.528) | 0.625 | -11.76 (56.188) | 0.934 |
| R-squared | 0.142 | | 0.121 | |
| Adj. R-squared | 0.132 | | 0.103 | |
| Observations | 1200 | | 700 | |

*Notes: Includes month fixed effects. The "Treatment Cities" are Atlanta, Dallas, Denver, Phoenix, Tampa, and Washington DC: See Footnote 37 and Appendix 4. Estimates are average monthly post-merger treatment effects on the ZORI multifamily rent index. The RealPage-LRO Merger closed in December 2017; the placebo test arbitrarily uses a treatment date in December 2016. Standard errors indicated in parentheses. Asterisks (*) indicate significance thresholds: (*) 10%, (**) 5%, (***) 1%.*

[51] See Footnote 37. For the difference-in-difference test, we use control group definition 2, which includes all non-treatment cities as control cities.

We test the parallel trends assumption by checking whether the pre-merger time trend is statistically significantly different for treatment cities and control cities:

**Parallel Trends Test**

| | | |
|---|---|---|
| Control Group | All Non-Treatment Cities | |
| Data End Date | November 2017 | |
| | Estimate | P-Value |
| Constant | 1356.55<br>(32.565)*** | 0.000 |
| Time Trend | 10.200<br>(1.828)*** | 0.000 |
| Affected City x<br>Time Trend | -14.326<br>(1.521)*** | 0.000 |
| R-squared | 0.095 | |
| Adj. R-squared | 0.092 | |
| Observations | 700 | |

*Notes: The "Treatment Cities" are Atlanta, Dallas, Denver, Phoenix, Tampa, and Washington DC: See Footnote 37. The RealPage-LRO Merger closed in December 2017; this test on the pre-treatment period ends in November 2017. Standard errors indicated in parentheses. Asterisks (*) indicate significance thresholds: (*) 10%, (**) 5%, (***) 1%.*