04.14.2025

Investigation of Alleged ‘Algorithmic Collusion’ in Rental Housing

Impact and Implications of RealPage’s Pricing Algorithms for Housing Affordability

Timothy Majidzadeh
Chelle Davies
Peter Benzoni
Patrick Yim
Ahmad Allaou

Agenda

1. **Background**
2. Data Collection and Preparation
3. Modeling Realpage & Rent Levels
4. Interactive Website
5. Causal Inference Methods
6. Conclusion

Research Background

Reports and Lawsuits Draw National Attention

- ProPublica Report Spotlights Potential Issue (2022)
- Lawsuits follow within days
- Biden-Harris CEA Estimates 4% Rent Uplift (but assumes collusion)
- RealPage Rebuts Arguments & Discloses Nationwide Rent Data

Background

Technology

Rent Going Up? One Company's Algorithm Could Be Why.

by Heather Vogell, ProPublica, with data analysis by Haru Coryne, ProPublica, and Ryan Little

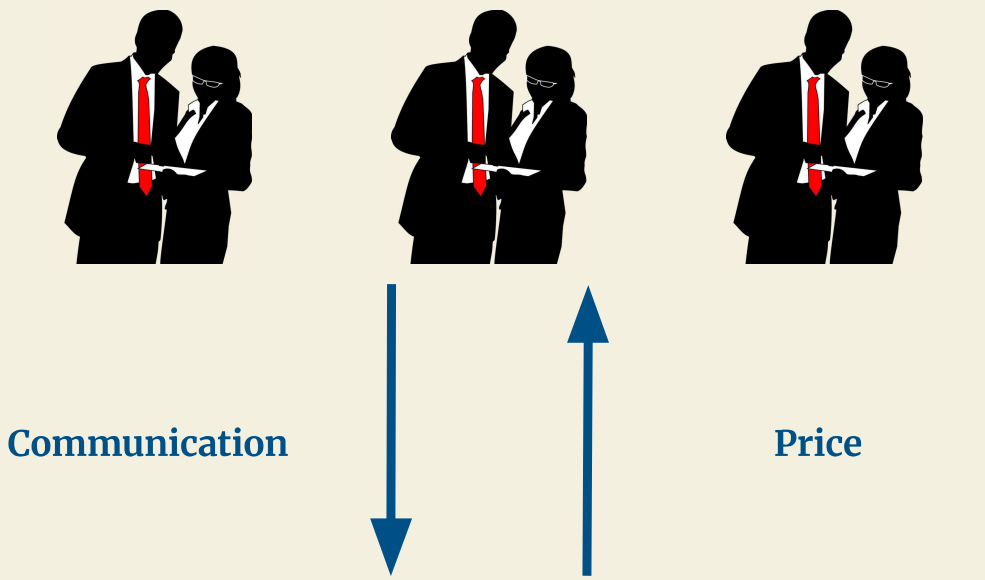
Oct. 15, 2022, 5 a.m. EDT

Animation by Lisa Larson-Walker/ProPublica

Background

Theory of Harm

- Housing Affordability is a Growing Crisis
- Landlords are increasingly using RealPage's Pricing Algorithms
- Using clients' data to train for other clients may be collusion
- Pricing Algorithms can 'learn' collusive behavior independently
  (see, e.g., Calvano, E., et al. (2020); Asker, J., et al. (2022))



Key Questions

Rent Impact

Does RealPage usage predict higher rent in public data?

If so, by how much?

Causal Inference

Is the rent impact we observe ***caused by*** price-coordination?

Is there statistical evidence to show this?

Communication

What's the best way to show this information?

How can renters and policymakers learn about the rent impact?

Agenda

1. Background
2. **Data Collection and Preparation**
3. Modeling Realpage & Rent Levels
4. Interactive Website
5. Causal Inference Methods
6. Conclusion

Tech Talk: Data Acquisition

RealPage	Scraped 42,000 properties from RealPage
Washington Post	Pulled Estimated Market Share; Clients
Property Managers	Scraped Properties from WaPo Named Clients
Zillow, HUD	Accessed Metro-Level Rent Indices
Census	Accessed Additional Metro-Level Data from Census

Data Tech Talk: 2025 Properties Snapshot

Matched Addresses on RealPage Explore (~42,000) with Addresses from Clients Named in Lawsuits (41 clients, 9,000 Properties, 5,173 Matched).

Utilized Open Source Maps & Fuzzy Matching

Data includes covariates such as:

- Year Built
- Market & Submarket
- Average Square Foot
- Average Rent & Rent Per Square Foot
- ... And more

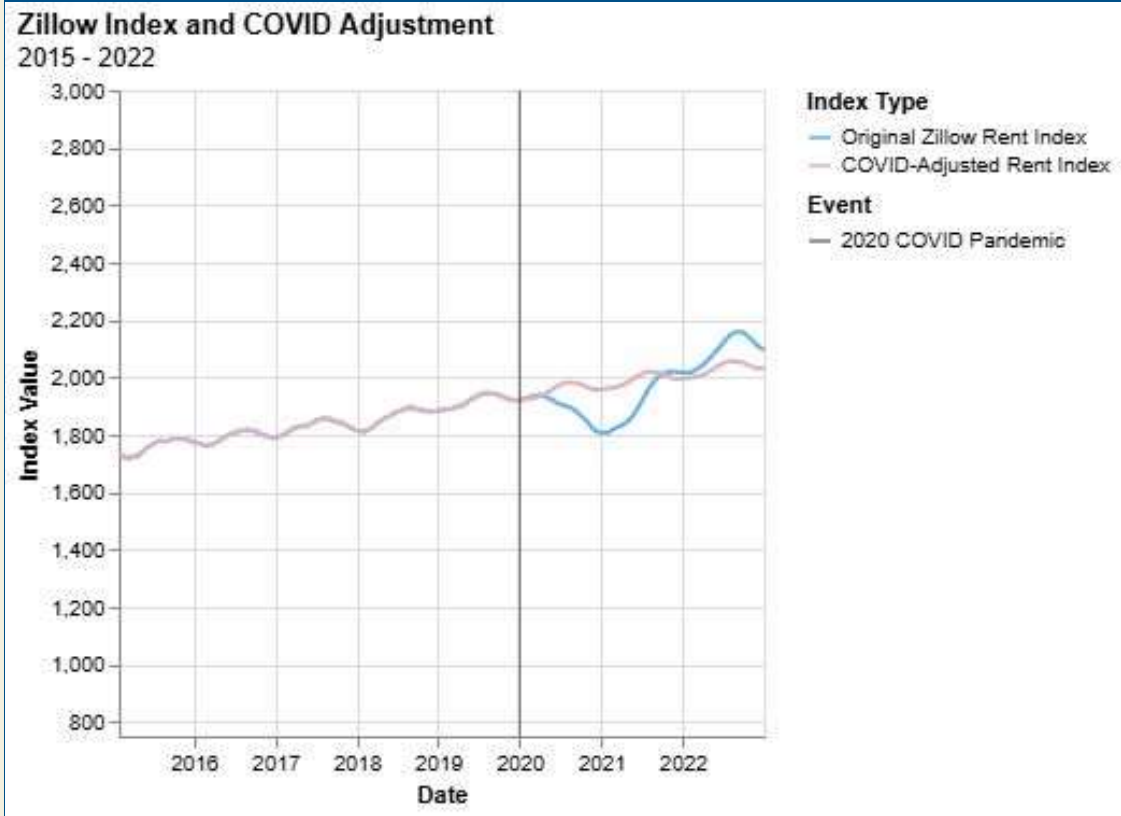
Data Tech Talk: Zillow Rent Index

Accessed Zillow panel data with the observed multifamily rent index.

Time: 2015 to Present

Geography: Metropolitan Statistical Areas (MSAs)

Utilized ARIMA forecasting to estimate a post-2020 counterfactual without COVID; necessary for our 2017 Merger Analysis.



Agenda

1. Background
2. Data Collection and Preparation
3. **Modeling Realpage & Rent Levels**
4. Interactive Website
5. Causal Inference Methods
6. Conclusion

Model 1: Propensity Model

Why Propensity?

- Differed systematically in observable characteristics (size, age, occupancy).
- Can bias any direct rent comparisons

Covariates Used to Predict RealPage Usage

Feature	Rationale
Square Footage	Larger, more modern buildings use RealPage
Stories	Vertical scale linked to RealPage use
Unit Count	Larger complexes often adopt RealPage and pricing algos
Building Age	Newer buildings charge more, often use RealPage
Occupancy Rate	Reflects demand/supply dynamics and pricing behavior
CBSA (One-Hot)	Controls for regional effects without linear assumptions

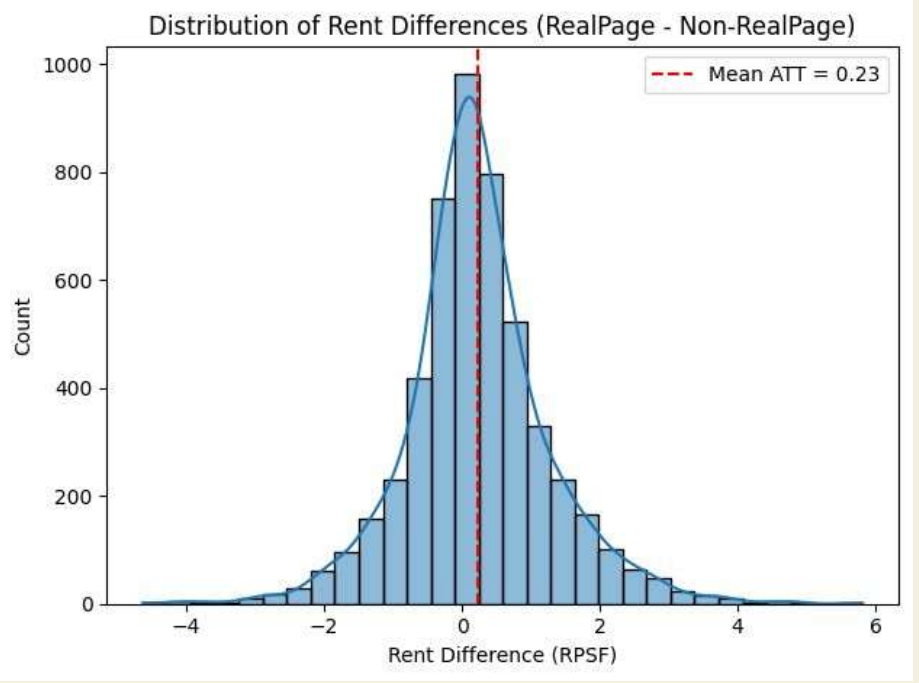
Model 1: Propensity Model Results

Matching Summary

- 1:1 Nearest Neighbor Matching without replacement
- 5,089 RealPage users matched to similar non-users
- All covariates achieved standardized mean difference < 0.1 post-match

Outcome Analysis

- Outcome Variable: MPF - RPSF (Rent per Square Foot)
- Average Treatment Effect on the Treated (ATT):
  + $0.23 increase in RPSF for RealPage properties



Additional Modeling: Predict RSPF

Model	Mean Squared Error (MSE)	R^2 Score	Description
Model 3 (Baseline RF)	0.04040	0.9214	Random Forest baseline using features including market_share
Gradient Boosting	0.06	0.89	Ensemble method that captures non-linear relationships. Feature importance shows market_share as influential.
Feed Forward Neural Network	0.03836	0.9254	Feed-forward NN with 3 hidden layers (64, 32, 16 neurons). Uses ReLU activations and early stopping; learns complex interactions with market_share.

Additional Modeling: Let's Simulate!

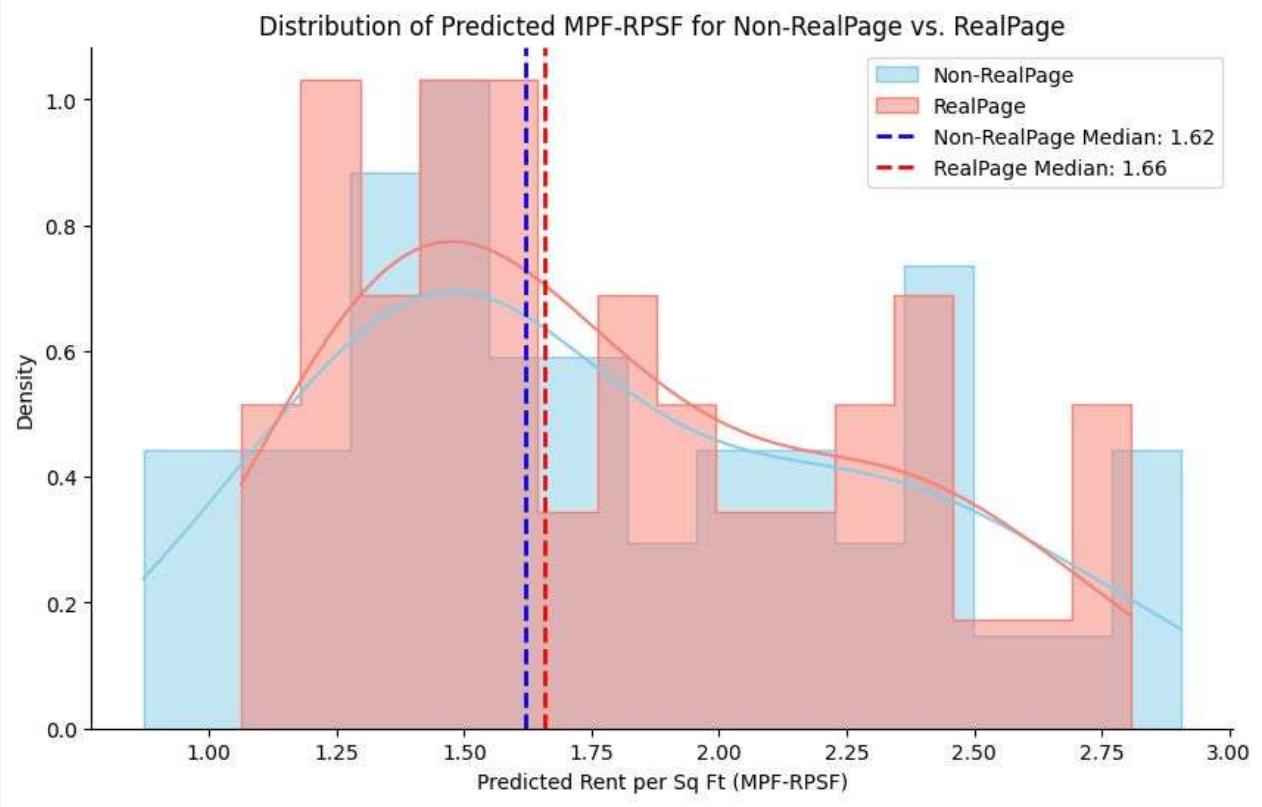
One sided Mann-Whitney U test:

Predictions made using the FFNN

Number of samples: 50

P-Value: 0.4465

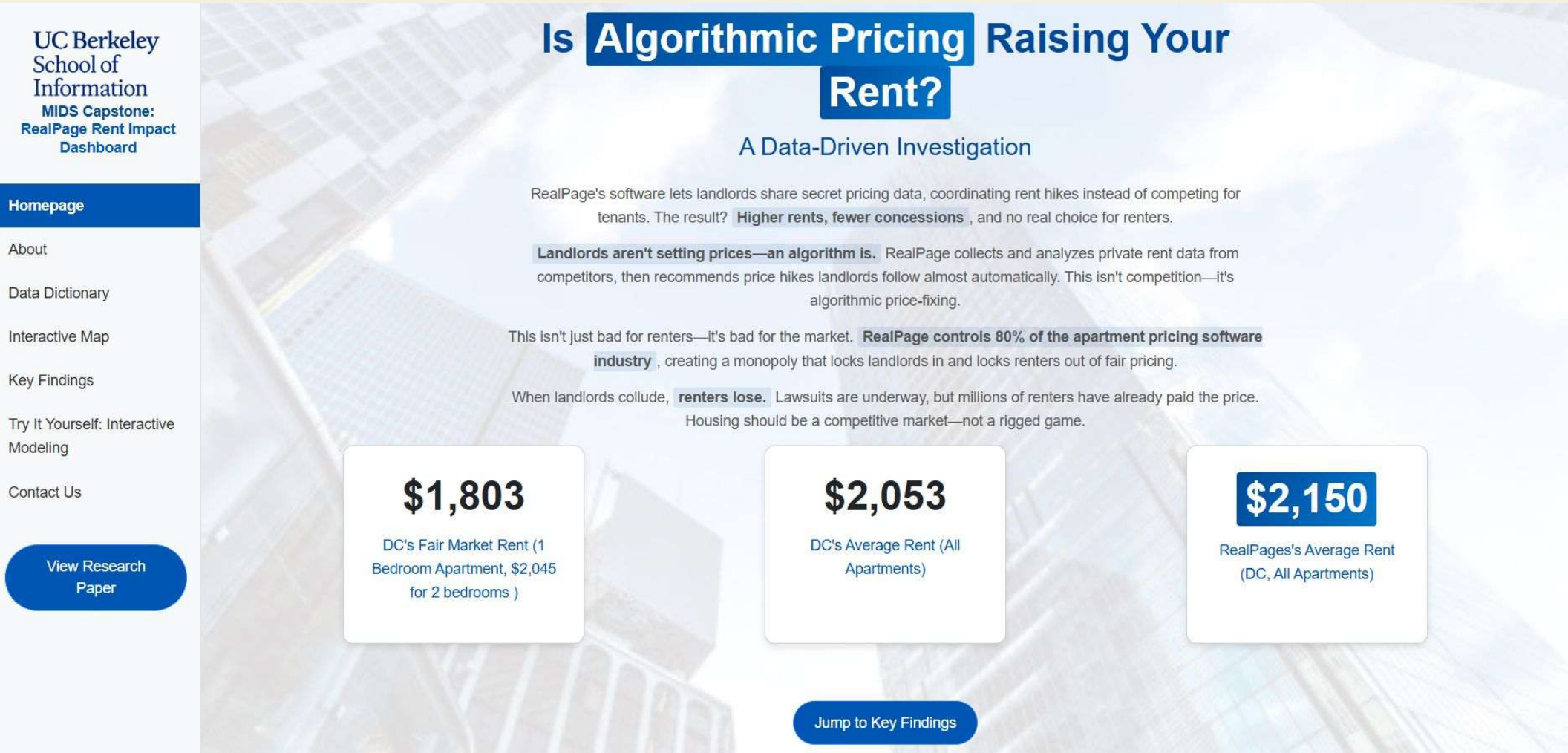
Not statistically significant



Agenda

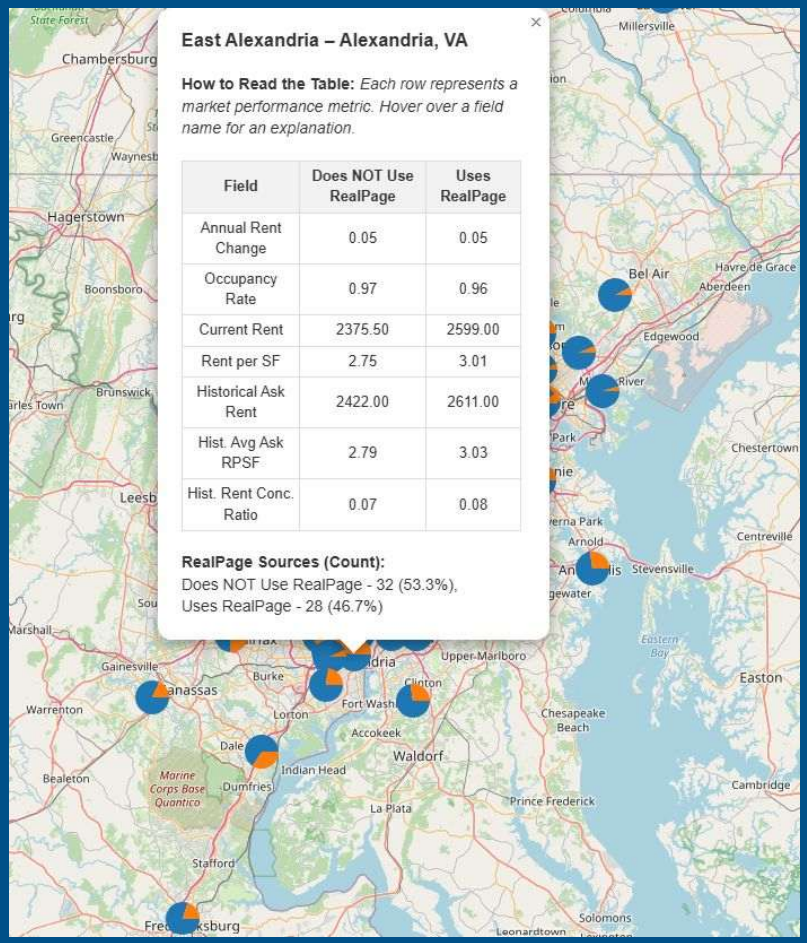
1. Background
2. Data Collection and Preparation
3. Modeling Realpage & Rent Levels
4. **Interactive Website**
5. Causal Inference Methods
6. Conclusion

Communicate with Interested Users



Interactive National Map

- Renters and policymakers can explore nationwide
- Uses snapshot of ~42,000 properties
- Quickly see estimate of Named Clients by submarket
- Drill deeper to see average differences in rent levels and property characteristics by submarket



‘Try It Yourself:’ Test Our Model

- Users can interact with model inputs directly
- Customize input with property details:
  - RealPage Usage
  - Year Built
  - Number of Stories
  - Submarket
  - Etc.
- Receive prediction of expected rent level

Agenda

1. Background
2. Data Collection and Preparation
3. Modeling Realpage & Rent Levels
4. Interactive Website
5. **Causal Inference Methods**
6. Conclusion

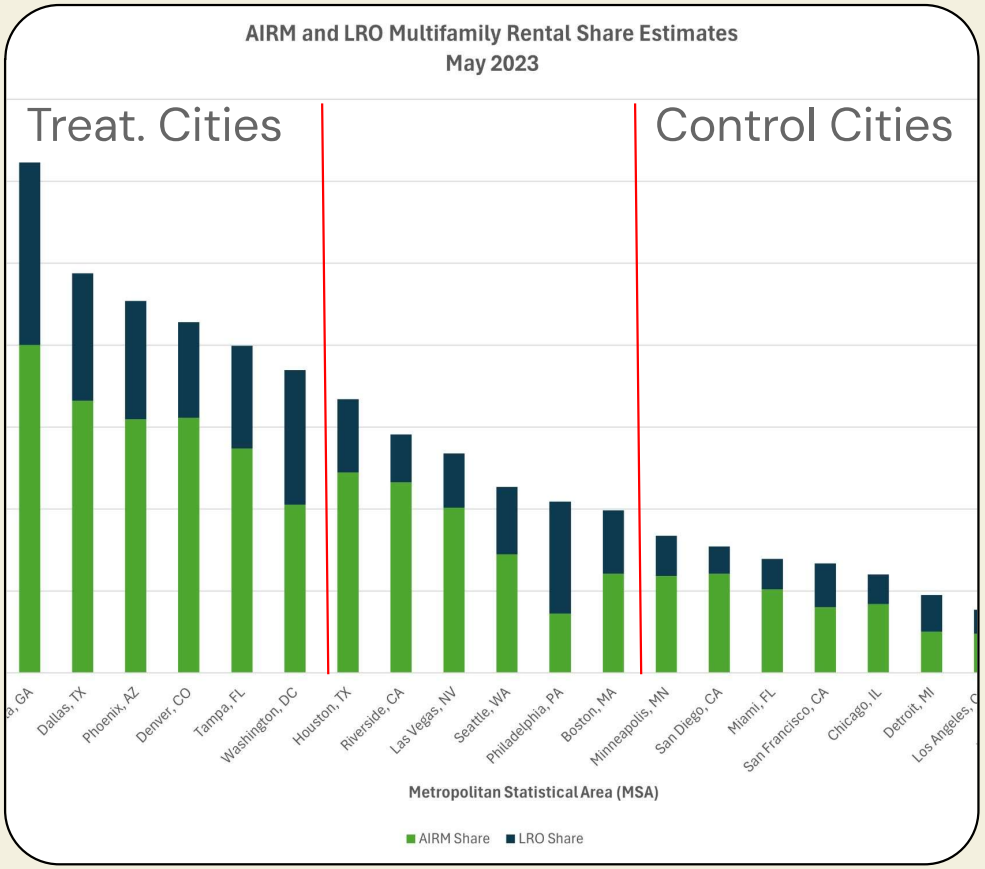
RealPage & Census Market Data

2017 RealPage-LRO Merger

RealPage acquired LRO, their largest competitor, in December 2017.

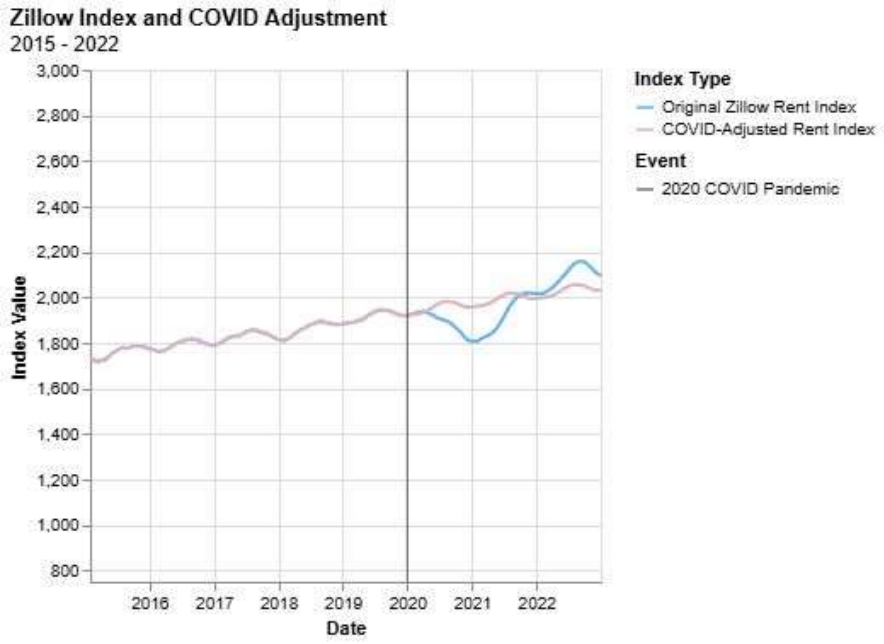
Did this impact rent? If so, that might causal evidence of price coordination.

City	Total Share	Share Gain
Atlanta, GA	62.3%	22.3%
Dallas, TX	48.8%	15.6%
...	...	...
Los Angeles, CA	7.7%	2.9%
New York, NY	3.6%	2.2%

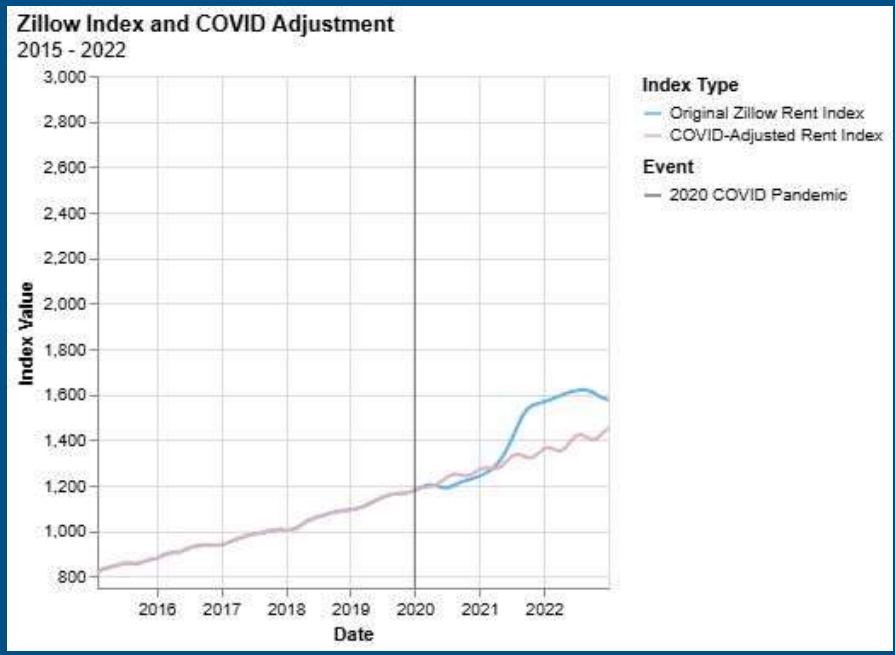


COVID Adjustment

Washington, DC

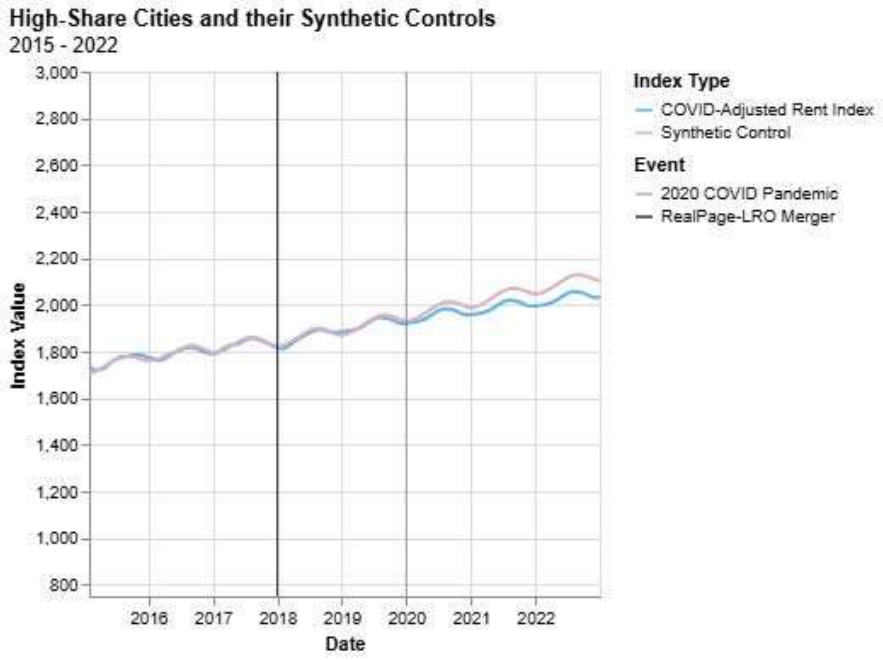


Phoenix, AZ

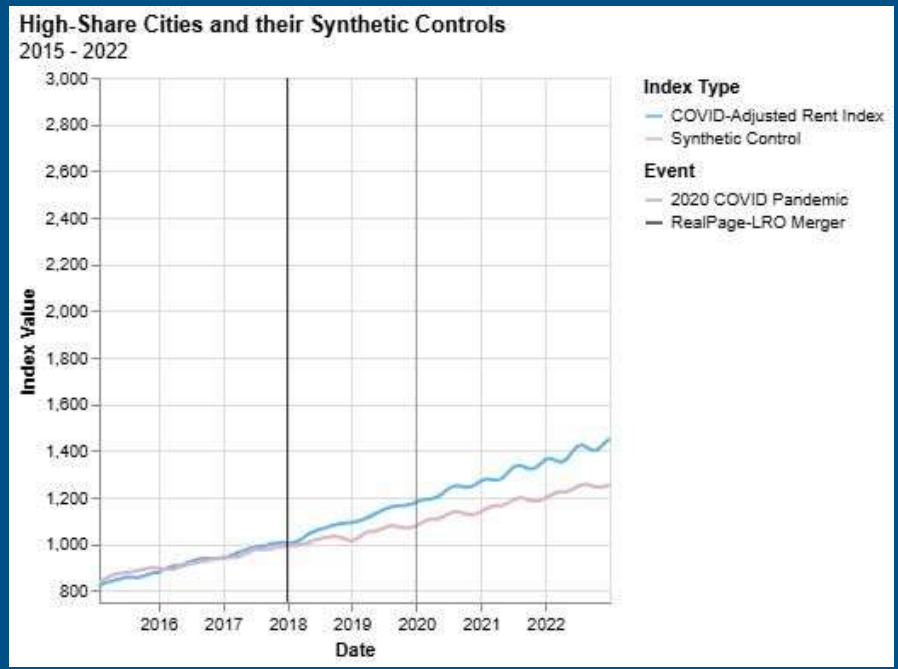


Synthetic Control

Washington, DC (Rent Decrease)



Phoenix, AZ (Rent Increase)



Zillow Data
COVID-Adjusted

Synthetic Control Results

Average Treatment Effect values for the 6 treatment cities are inconsistent.

- Rent Increases for Phoenix
- Rent Decreases for Washington
- Near zero for Tampa

Other cities are not statistically significant.

Robustness checks also lead to inconsistent results.

City	Average ATT	P-Value
Atlanta, GA	-28.40	0.567
Dallas, TX	-69.63	0.166
Denver, CO	-7.63	0.934
Phoenix, AZ	103.74	0.000***
Tampa, FL	-1.55	0.017***
Washington, DC	-30.46	0.000***

Agenda

1. Background
2. Data Collection and Preparation
3. Modeling Realpage & Rent Levels
4. Interactive Website
5. Causal Inference Methods
6. **Conclusion**

Summary of Findings

Models Show a Rent Impact

Propensity Score model shows + $0.23 increase in rent per square foot

Controls for unit size, building size, building age, occupancy rate, and geography

Impact Varies Based on Model

Not all models statistically significant. Simulation using Feed-Forward Neural Network has a p-value over 0.4.

Price Coordination Unclear

No consistent evidence that the 2017 merger raised rent across multiple cities.

Price coordination might still exist, but it would be unrelated to the merger.

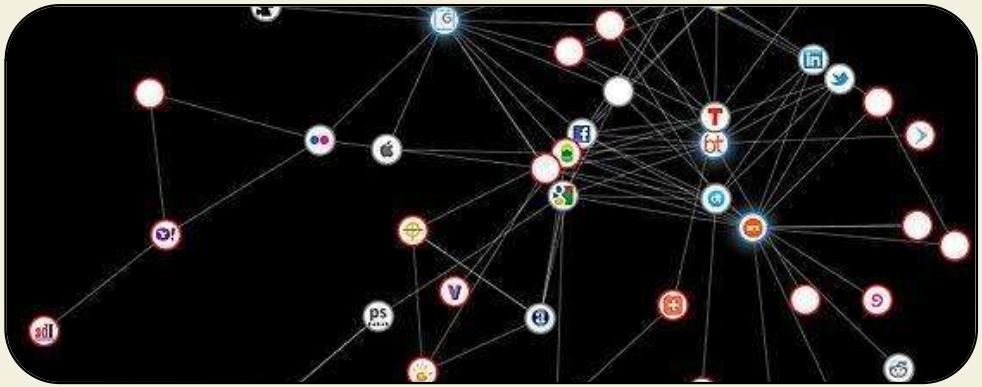
Our Website

https://uc-berkeley-i-school.github.io/realpage-rent-impact/

Our Paper

https://uc-berkeley-i-school.github.io/realpage-rent-impact/research-paper.pdf

THANK YOU



Tim Majidzadeh
Chelle Davies
Ahmad Allaou
Patrick Yim
Peter Benzoni

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/.

References

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/.

References

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.

References

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/.