The Market Impact of Algorithmic Pricing: Evidence from the Multifamily Rental Housing Market

Sophie Calder-Wang * Gi Kim†

Jan, 2023

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

We examine the impact of algorithmic pricing on the U.S. multifamily rental housing market. Adoption decisions of management companies are hand collected and merged with the universe of market-rate multifamily apartment data to construct a unique dataset for the analysis. Overall, we find that adoption of the software leads to 3.7% increase in average rents. This effect is highly dynamic, starting from 1% in the year of adoption and increasing to 5% four years into the adoption. The magnitude of the treatment effects is strongly and positively associated with the share of adopters in a local market-year, suggesting that the software could be facilitating price coordination among landlords using the software.

*Wharton School, University of Pennsylvania. sophiecw@wharton.upenn.edu
†Wharton School, University of Pennsylvania. robinkim@wharton.upenn.edu

1 Introduction

The recent advances in computing technology have shifted the paradigm of price setting practices of firms through software that dynamically and automatically set prices. These types of software often leverage high-frequency data collected across firms operating in the same industry or in the same market to suggest optimal prices for managers. The algorithms also make decisions based on the past behaviors of other firms and forecast market demands, which makes them especially prone to be the hub that facilitate price coordination across firms using similar technologies. Moreover, these programs are often powered by artificial intelligence (AI), raising concerns that these pricing agents might just learn to jointly play a collusive strategy rather than engaging in a price war. For these reasons, the technology has recently attracted lots of attentions from researchers, policy makers, and antitrust agencies alike (Fortin, 2020; Mcsweeny and ODea, 2017; OECD, 2017).

The industry that has recently came under heavy scrutiny of antitrust authorities for using such technology is the multifamily housing industry of the U.S. Following an article by Propublica with a title "*Rent Going Up? One Company's Algorithm Could Be Why.*"[1] accusing the pricing software of RealPage, a multifamily management solution IT company, of pushing "prices above competitive levels," a series of class action lawsuits were filed against both the software company and landlords using the software (*Yusupov v. RealPage, Inc. et al*, 2023; *Navarro v. RealPage, Inc. et al*, 2022; *Bason v. RealPage, Inc.*, 2022). It is also reported that the Department of Justice has opened an investigation to look for evidences of price coordination among landlords using the software[2].

Aside from the importance of the setting, there is a growing need for deeper understanding of pricing algorithms and more importantly, how they interact with varying market environments in practice. Many theoretical studies have been conducted under synthetic environments, yet there is no theoretical consensus on whether the algorithms end up facilitating collusion (Asker et al., 2022; Calvano et al., 2020; Mikls-Thal and Tucker, 2019; Brown and MacKay, 2021). To our knowledge, the only existing extensive empirical study is that of Assad et al. (2020), which examines the effect of algorithmic pricing on German retail gasoline market. Our study aims to reinforce the understanding how algorithmic pricing works in practice under a completely different empirical setting and its implications on an industry with serious

[1] https://www.propublica.org/article/yieldstar-rent-increase-realpage-rent
[2] https://www.propublica.org/article/yieldstar-realpage-rent-doj-investigation-antitrust

consequences for consumers.

First, we qualitatively highlight important patterns in pricing behavior of the algorithm in the rental housing market from a presentation that was intended to directly showcase the inner workings of the proprietary software. It shows the software's interface from the viewpoint of a property manager. From this evidence, we conclude that the software responds to *daily, unit-level* rent changes of competitors in a given market and heavily utilizes the data collected from competitors, who are actually other clients using the same software. The types of data the software uses are the necessary ingredients to facilitate price coordination among its clients.

With this in mind, we move onto the empirical analysis to assess the impact of pricing software adoption on rent prices. One of the main challenges in empirically studying algorithmic pricing is due to a lack of data on adoption decisions across firms over time. We tackle this problem by hand collecting the adoption data which we then merge with the universe of market-rate multifamily apartment pricing and vacancy data, constructing a unique dataset suited for this analysis. While the collected data is fairly underestimating the number of adopters, it covers top 20 management companies and has little false-positive error against buildings who are using the software as of 2022.

We then estimate the dynamic treatment effects of adopting the software on rent prices using the estimation method of Callaway and SantAnna (2020). We control for potential difference in building qualities by conditioning on the time trends of markets interacted with building's pre-period price position in its market along with other covariate-specific time trends. Moreover, since the adoption decisions are made by property management companies whose portfolio tends to span dozens of states, adoption decisions at the building-level is considered to be fairly exogenous. This logic is similar to the main identification argument of Assad et al. (2020). Under this identification strategy, we find significant dynamic treatment effects of adoption on rents. On average, the treatment effect starts out as moderate 1% increase in rents but culminates to 5% differential in rents four years after adoption compared to non-adopters. This effect is robust across specifications and estimation methods used.

Lastly, we examine heterogeneous treatment effects across geographical markets. We are interested in how the effect of adoption on rents varies with the degree of penetration of software in a given market. We find a positive and significant relationship between the magnitude of the treatment effects and the *share* of adopters in a local

market, conditioning on the *number* of adopters. There is no relationship between the effects and the number once we condition on the share. Moreover, this effect gets stronger with a narrower definition of a local market. Together, these findings suggest that the pricing algorithm raises rents from having more adopters relative to non-adopters in the market, but not through just being able to collect more market data that increases the precision of demand forecast. If this is a meaningful channel of optimization, we should see a relationship between the effect and the number of adopters after conditioning on the share.

Our findings are largely consistent with the theoretical studies of Asker et al. (2022) and Calvano et al. (2020) as well as the empirical findings of Assad et al. (2020). The algorithm exhibits strong dynamic effects and have heterogeneous effects on participants in markets with varying degrees of software penetration. While we do not make any claims about the conduct or explicit mechanisms, the reduced-form findings supplemented by the qualitative observations are consistent with the patterns of price coordination.

The remainder of the paper proceeds as follows. Section 2 provides background on the U.S. multifamily housing market and the pricing software used. Section 3 describes the data and shows stylized facts from the data. We then present our empirical strategy and the results in Section 4. Finally, Section 5 concludes with potential ways to further improve the study.

2 Multifamily Market and Pricing Software

2.1 Industry Background and Significance

The U.S. multifamily housing industry has experienced a fast-paced growth after the Great Recession, with 158% increase in value per square feet from 2010 to 2019[3]. While it has been an attractive investment opportunity for institutional investors with 80% increase in average nominal rents and 50% decrease in vacancy rates[4], multifamily units are primarily occupied by low-income demographics who do not have enough financial capability to own a property of their own and spend majority of their income on rents.

Institutional investors often outsource the day-to-day operation of buildings in

[3]https://www.nmhc.org/research-insight/quick-facts-figures/quick-facts-investment-returns-on-apartments/vstock

[4]https://www.nmhc.org/research-insight/quick-facts-figures/quick-facts-market-conditions

their portfolio to a management company. The management companies then "run" the buildings, setting monthly rents, managing lessees, running promotions, alongside with various maintenance activities. While there has been some noticeable consolidation of apartment *owners* that has caught the media's attention recently[5], more pronounced concentration has happened among the management companies. Greystar, the biggest management company in the U.S., has increased its number of units managing by 337% from 2010 to 2021[6], and other top 20 management companies as of 2022 have also shown a steady increase in number of units managing.

2.2 Algorithmic Pricing Software

The management companies rely on IT infrastructure to streamline their operation across thousands of units in dozens of buildings they manage across states. The management companies contract with enterprise solutions who develop proprietary property management software for implementation. These software companies provide a suite of services to construct a central database, manage bookkeeping, engage with tenants, automate lease renewals, monitor vacancies, assess market conditions, etc.

In addition to these services, three major companies started to offer rent optimization solutions starting early 2000s[7]. The "revenue management" solution is an automated pricing algorithm that suggests optimal rents in real time by unit type and lease lengths to property managers. It aims to take guess out of pricing rents in both new lease signing and renewal process. Investors often prioritized maintaining near full occupancy, which many property managers found as a suboptimal way of maximizing the return[8]. By 2011, around 15% of apartment units had adopted a version of pricing software[9], and 19 million units out of 45 million multifamily rental units were using a RealPage software in 2021. While it is unclear that all 19 million units use the company's pricing algorithm, it suggests that there has been great proliferation of algorithmic pricing across multifamily apartment management companies.

The details of how exactly these software compute optimal rents is not well-known. We were able to obtain a copy of presentation slides made in 2014 that showcases

[5]https://www.propublica.org/article/when-private-equity-becomes-your-landlord

[6]https://www.nmhc.org/research-insight/the-nmhc-50/top-50-lists/2022-top-managers-list/

[7]Yardi RentMAXImizer, RealPage Yieldstar, and Rainmaker LRO

[8]"If you're at 97% or 98% occupancy, your rents are probably too low." https://www.investors.com/news/landlords-consider-how-to-raise-rent/

[9]https://web.archive.org/web/20110824021635/http://www.multifamilyrevenue.com/revenue-management-users-multifamily/

the inner workings of one of the software, RealPage's Yieldstar. The most notable point is that the software estimates demand elasticity and forecasts dynamic demand at bedroom-level based on lease length and renewal probability by fully utilizing selected competitors' prices and vacancies. See Appendix Figure A1 for the exact wordings from the slide deck.

Figure 1. Manager's view of dynamic pricing by Yieldstar

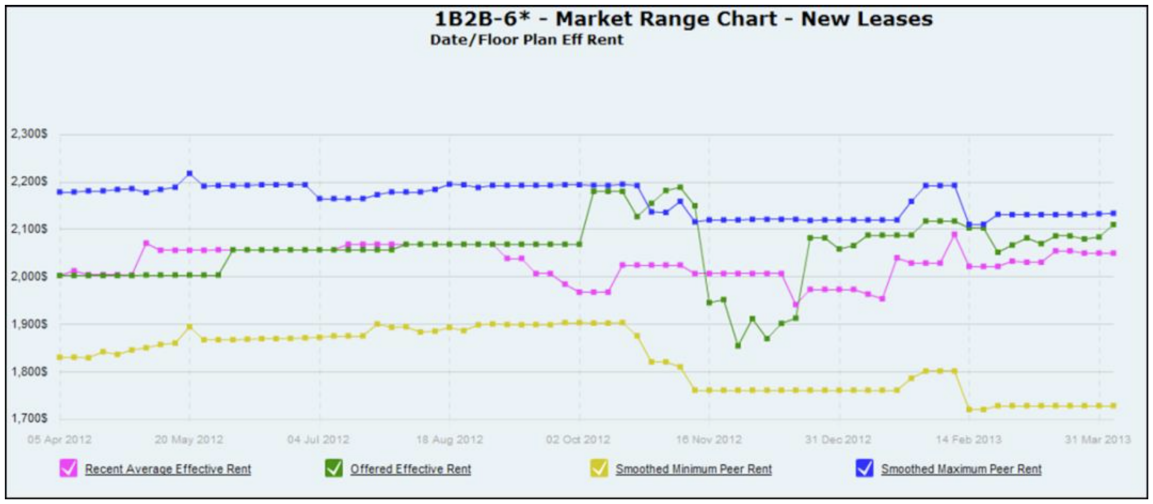


Figure 1 shows property manager's view of price optimization path with Yieldstar for a particular bedroom type in the showcased, but real, apartment property. The blue line and the yellow line represent competitors' maximum and minimum rent for similar bedroom units, respectively. The green line is the price optimizer's offered rent to new and renewing tenants, and the pink line is an average effective rent across tenants who recently accepted the offered rent. Although this is just one case study of a particular bedroom type in a particular market, there are many interesting signs of optimal pricing in this figure.

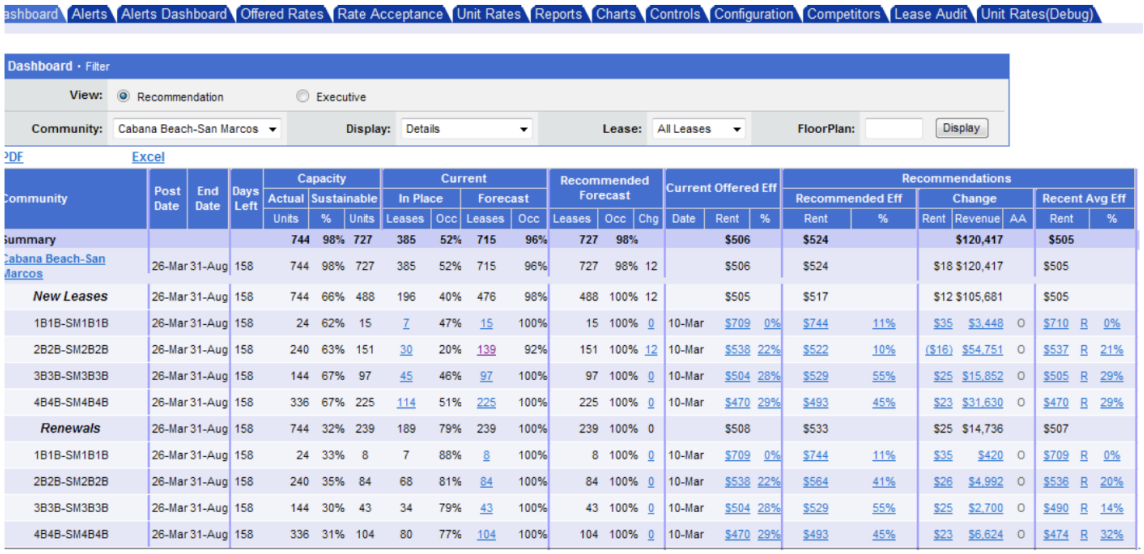
Up until Aug 18th., 2012, the market was somewhat stable and the minimum rent of the market (yellow line) has been increasing slowly. Then shortly after, the recent average effect rent (pink) has dropped for the particular unit to which the software responded by increasing the offered rent (green) to maintain the previous level of accepted average rent (pink).

Shortly after Oct 2nd, 2012, the market's minimum rent (yellow) has dropped significantly, which might have been associated with the previous decline of average effect rent of the unit (pink). In response to such market-wide price decline, the

software aggressively lowers its offered rent (green) from $2,200 to $1,850 over the span of couple weeks. From Nov. 16th to closer to Dec. 31st, the algorithm seems to be swinging widely to experiment with offered rents (green) to recover from the steep decline of the previous months. Finally after Dec. 31st, the rent stabilizes again, approaching the market's maximum rent (blue).

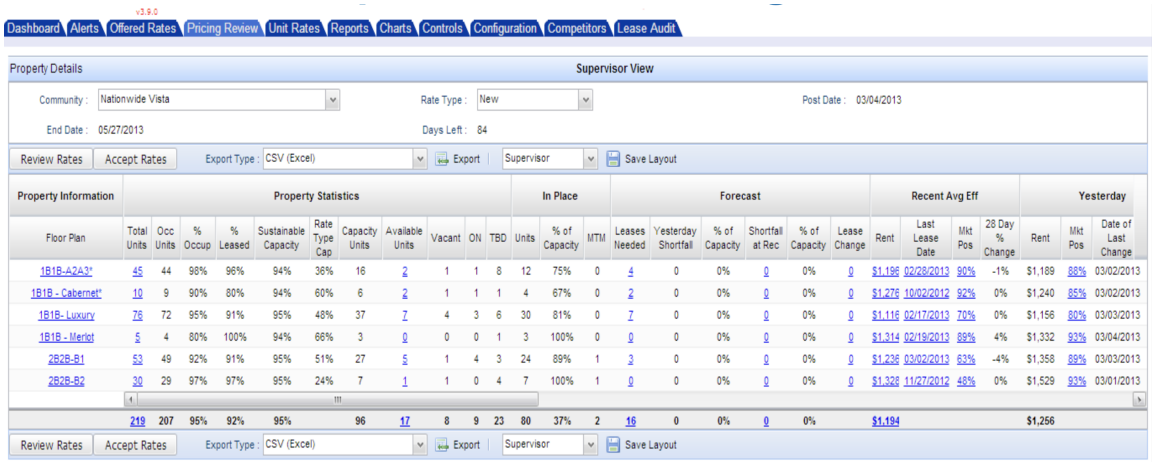
Figure 2. Manager's view of Yieldstar pricing dashboard

(a) Price recommendation made by Yieldstar



Community	Post Date	End Date	Days Left	Capacity Actual Units	Capacity Actual %	Capacity Sustainable Units	Current In Place Leases	Current In Place Occ	Current Forecast Leases	Current Forecast Occ	Recommended Forecast Leases	Recommended Forecast Occ	Recommended Forecast Chg	Current Offered Eff Date	Current Offered Eff Rent	Current Offered Eff %	Recommendations Recommended Eff Rent	Recommendations Recommended Eff %	Recommendations Change Rent	Recommendations Change Revenue	Recommendations Change AA	Recommendations Recent Avg Eff Rent	Recommendations Recent Avg Eff %
Summary				744	98%	727	385	52%	715	96%	727	98%			$506		$524			$120,417		$505	
Cabana Beach-San Marcos	26-Mar	31-Aug	158	744	98%	727	385	52%	715	96%	727	98%	12		$506		$524		$18	$120,417		$505	
New Leases	26-Mar	31-Aug	158	744	66%	488	196	40%	476	98%	488	100%	12		$505		$517		$12	$105,681		$505	
1B1B-SM1B1B	26-Mar	31-Aug	158	24	62%	15	7	47%	15	100%	15	100%	0	10-Mar	$709	0%	$744	11%	$35	$3,448	O	$710 R	0%
2B2B-SM2B2B	26-Mar	31-Aug	158	240	63%	151	30	20%	139	92%	151	100%	12	10-Mar	$538	22%	$522	10%	($16)	$54,751	O	$537 R	21%
3B3B-SM3B3B	26-Mar	31-Aug	158	144	67%	97	45	46%	97	100%	97	100%	0	10-Mar	$504	28%	$529	55%	$25	$15,852	O	$505 R	29%
4B4B-SM4B4B	26-Mar	31-Aug	158	336	67%	225	114	51%	225	100%	225	100%	0	10-Mar	$470	29%	$493	45%	$23	$31,630	O	$470 R	29%
Renewals	26-Mar	31-Aug	158	744	32%	239	189	79%	239	100%	239	100%	0		$508		$533		$25	$14,736		$507	
1B1B-SM1B1B	26-Mar	31-Aug	158	24	33%	8	7	88%	8	100%	8	100%	0	10-Mar	$709	0%	$744	11%	$35	$420	O	$709 R	0%
2B2B-SM2B2B	26-Mar	31-Aug	158	240	35%	84	68	81%	84	100%	84	100%	0	10-Mar	$538	22%	$564	41%	$26	$4,992	O	$536 R	20%
3B3B-SM3B3B	26-Mar	31-Aug	158	144	30%	43	34	79%	43	100%	43	100%	0	10-Mar	$504	28%	$529	55%	$25	$2,700	O	$490 R	14%
4B4B-SM4B4B	26-Mar	31-Aug	158	336	31%	104	80	77%	104	100%	104	100%	0	10-Mar	$470	29%	$493	45%	$23	$6,624	O	$474 R	32%

(b) Competitor data and recommendation acceptance



Property Information: Floor Plan	Property Statistics: Total Units	Occ Units	% Occup	% Leased	Sustainable Capacity	Rate Type Cap	Capacity Units	Available Units	Vacant	ON	TBD	In Place: Units	% of Capacity	MTM	Forecast: Leases Needed	Yesterday Shortfall	% of Capacity	Shortfall at Rec	% of Capacity	Lease Change	Recent Avg Eff: Rent	Last Lease Date	Mkt Pos	28 Day % Change	Yesterday: Rent	Mkt Pos	Date of Last Change
1B1B-A2A3*	45	44	98%	96%	94%	36%	16	2	1	1	8	12	75%	0	4	0	0%	0	0%	0	$1,196	02/28/2013	90%	-1%	$1,189	88%	03/02/2013
1B1B - Cabernet*	10	9	90%	80%	94%	60%	6	2	1	1	1	4	67%	0	2	0	0%	0	0%	0	$1,276	10/02/2012	92%	0%	$1,240	85%	03/02/2013
1B1B- Luxury	76	72	95%	91%	95%	48%	37	7	4	3	6	30	81%	0	7	0	0%	0	0%	0	$1,116	02/17/2013	70%	0%	$1,156	80%	03/03/2013
1B1B - Merlot	5	4	80%	100%	94%	66%	3	0	0	0	1	3	100%	0	0	0	0%	0	0%	0	$1,314	02/19/2013	89%	4%	$1,332	93%	03/04/2013
2B2B-B1	53	49	92%	91%	95%	51%	27	5	1	4	3	24	89%	1	3	0	0%	0	0%	0	$1,236	03/02/2013	63%	-4%	$1,358	89%	03/03/2013
2B2B-B2	30	29	97%	97%	95%	24%	7	1	1	0	4	7	100%	1	0	0	0%	0	0%	0	$1,328	11/27/2012	48%	0%	$1,529	93%	03/01/2013
	219	207	95%	92%	95%		96	17	8	9	23	80	37%	2	16	0	0%	0	0%		$1,194				$1,256		

Figure 2 shows the dashboard views for a property manager that displays price recommendations made by the software. The top figure shows current and forecasted demands and following rent price recommendations made by the software by bedroom

type. It also prices new leases and renewals differently. In the bottom figure, under "Recent Avg Eff" and "Yesterday" tabs, the column "Mkt Pos" indicates the position of the rent prices compared to competitors in the market, showing whether the units are being underpriced or overpriced at a *daily* frequency.

All of this information is purely for providing data points to manager and the cognitive burden for a manager to act on this myriad of information is minimal; a manager can simply click either the "Accept Rates" or the "Review Rates" (if something seems so out of place) button located on both top and bottom of the table. Propublica reports that managers accept recommended rents up to 90% of times.

Figure 1 and 2 show that the pricing software both reacts to changes in market conditions and heavily utilizes the detailed, high-frequency data of competitors down to the granularity of daily price for individual units. The question is then, where do they get the data from? In one of their promotional videos, Yieldstar claims that they

> "leverage the statistical analysis collected from the industry's largest lease transaction database, spanning over 11 million units and millions of transactions a year. No one else has this tremendous scope of real time data that determines daily acceptions and opportunities for maximizing rents and reducing vacancy with utmost accuracy."[10]

While it may be that property managers of individual buildings do not have access to detailed lease transaction data of their selected competitors, it is almost certain that Yieldstar utilizes its own clients' data to maximize other clients' profits. If this is the case, how would the software set prices across competing clients in the same market? Moreover, would the software have differential impact on rents across markets with varying degrees of penetration due to either price coordination or greater availability of market data?

3 Data and Descriptives

To answer these questions empirically, we use two main datasets. The first contains the universe of market-rate multifamily buildings in the U.S. from REIS by Moody's

[10] Yieldstar Revenue Management Overview Presentation Webinar, accessed by registration on Dec 1st, 2022.

Analytics. The second contains the adoption year of pricing software or "revenue management" software by management companies, hand collected from various sources. We discuss advantages and limitations of each data in following sections, then present stylized facts to motivate the main empirical analysis.

3.1 REIS

REIS by Moody's Analytics contains information on the universe of market-rate buildings in the US. This is a particularly attractive set of sample, since market-rate buildings are not subject to special subsidies or legal compensation. In the fourth quarter of each year, Moody's surveys owners and managers of these buildings and collect information on asking rents, occupancy, concessions, various amenities. The dataset contains building-year level observations spanning 14 years from 2005 to 2019 with the identity of management companies. REIS also provides its own definitions of markets. They assign each building to one of 625 "submarkets" in one of 50 "metros." Submarkets completely partition a metro without overlaps.

The limitations come from two sources. The first source is that the management companies were backfilled by Moody's from 2019, making the management company field for each building time-invariant. The decision to adopt pricing software is made on management company-level, and as mentioned in Section 2.1, the property management industry had experienced a fair bit of consolidation over the past decade. These two facts can potentially lead to over-counting of adopters in earlier periods. Because we are misclassifying non-adopted buildings to be adopters (measurement error in the independent variable), this would put an attenuation bias towards zero for the estimates.

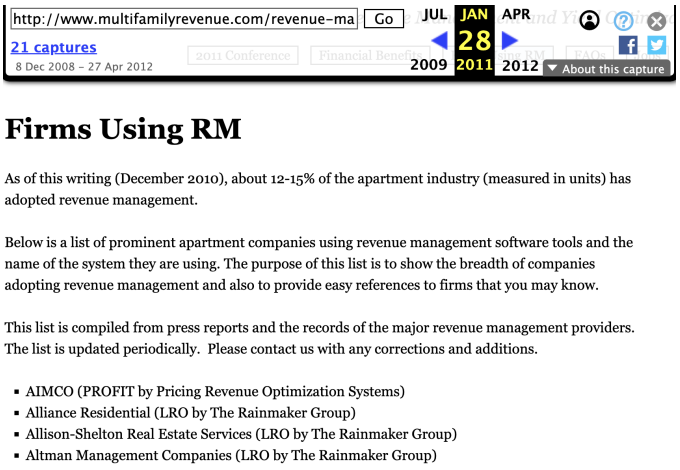
The second source of limitation is lack of high-frequency price data. The greatest advantages of high-frequency pricing data is that it can be used to detect structural breaks to infer adoption in absence of accurate adoption data as done in Assad et al. (2020). It can also highlight interesting price dynamics and responses to changes in competitors' prices as illustrated in Figure 1. This is not a major concern, however, since we were able to a collect fairly confident set of management companies who adopted the software along with *when* each of them had adopted.

3.2 Software Adoption Data

We collected the adoption data from three major sources. First source is yearly snapshots of a website who maintained and updated the list of management companies

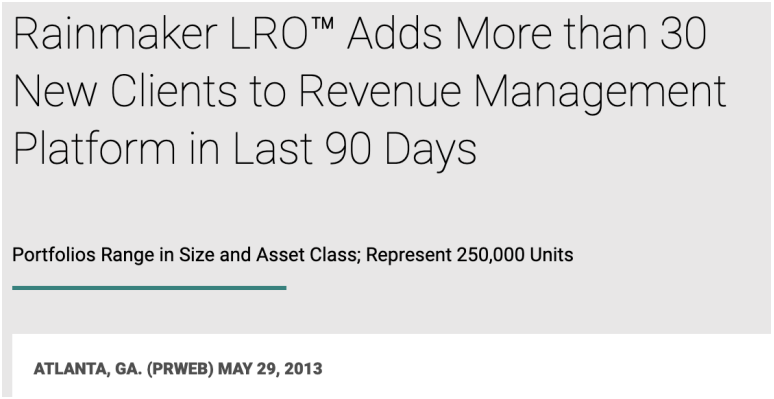
and owners who had adopted a pricing software. The owner of the website solicited a survey to participants at a major multifamily housing conference from 2008 to 2011 and updated the list every year. Companies appeared in snapshots are then assigned the year of adoption equal to the year of first appearance on the list. See Figure 3 for an example of the website snapshots.

Figure 3. Snapshot of surveyed list of adopters



Second source is from various media outlets. Rainmaker LRO, which was the major player in revenue management industry before it was acquired by RealPage in 2017, had an active media presence announcing their major customer acquisition. Through their main news outlet, PrWeb.com, they not only announced customer acquisitions but also major updates (or "patches") to their price optimization software. See Figure 4 for an example of an article.

Figure 4. Example article containing client acquisition made by Rainmaker LRO



Lastly, we supplement the data using the list from AppsRunTheWorld.com. This company collects data on adoption of enterprise IT applications and sells insights to salesforces of IT companies for better leads targeting. We use the list of companies who uses Yieldstar software. While they survey the companies and record adoption dates, it can be a bit noisy since some of the adoption dates are inferred from web-scraped data. Fortunately, this data contains only a small fraction of adopters compared to the other two more credible sources.

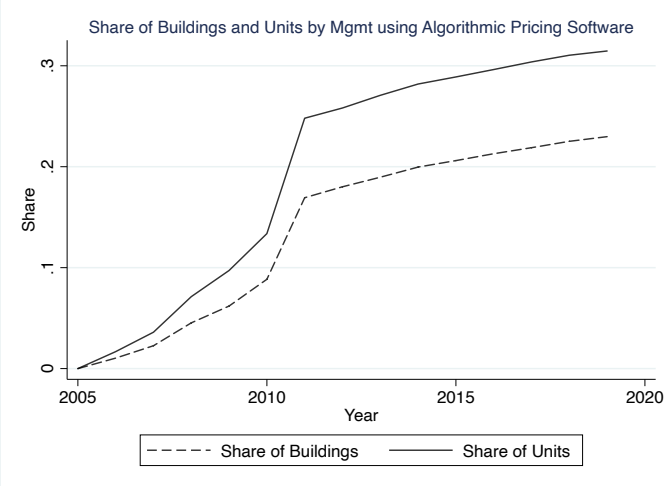
The main limitation is, of course, the measurement error. We expect that we vastly underestimate the number of adopters since the collected lists only cover major management companies. We conducted validation exercise against the current list of buildings that use a RealPage product[11]. Of 641 *buildings*, 274 buildings (42%) were using RealPage products as of 2022 but was not flagged as adopters in 2019 and 206 buildings (38%) and 146 buildings (22%) were rightly flagged as adopters and non-adopters in 2019, respectively. There are minimal false-positives; only 15 (2.3%) buildings were flagged as adopters in 2019 but not using RealPage in 2022. The majority of these buildings were said to be using Rainmaker LRO, and it is possible that they decided not to switch over to RealPage following its acquisition in 2017. We expect that any significant effect net of selection bias is reassuring due to the presence of attenuation bias from the measurement error.

3.3 Descriptives

In this section, we present stylized facts by combining the REIS data with the hand-collected data on software adoption and discuss sample selection for the main analysis. Figure 5 shows penetration trend of pricing software across buildings in the states.

[11] https://www.realpage.com/explore

Figure 5. Share of adoption over time

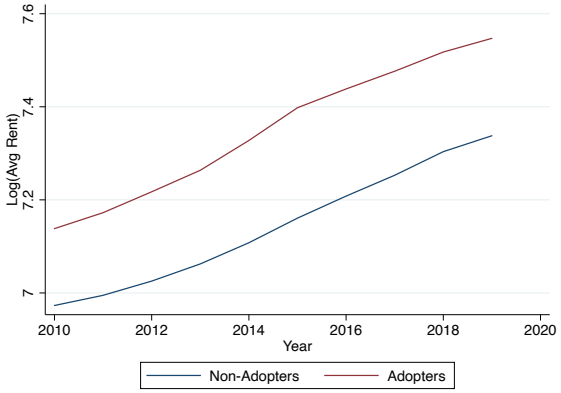


Since we are conditioning the sample on market-rate buildings, we expect these buildings to adopt pricing software more aggressively than other types of multifamily apartments. Compared to the guess of 15% penetration in 2012 made by the surveyor mentioned in Section 3.2, more than 20% of units in the REIS data had adopted the software. At the end of data period in 2019, about 30% of units and 20% of buildings had adopted the software.

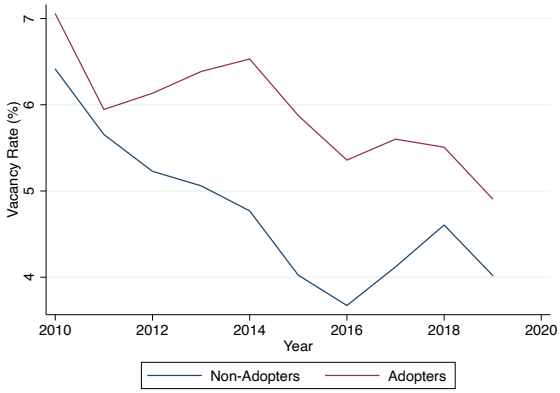
For the remainder of this paper, we restrict our sample period to 2010-2019. We do so based on a couple of reasons. First, dealing with various and potentially highly heterogeneous effects of the Great recession does not add much to the analysis given the long panel we have after the recession. Second, when management companies and owners have started to adopt the pricing software near its inception, they tested the software by rolling it out on selected properties for 6 months to a year before deciding to fully roll out across all properties. Earlier periods may be contaminated by the presence of "partial" adopters as opposed to "full" adopters. Lastly, the top management companies started to consolidate in 2010, thus conditioning on the sample period may help mitigate the attenuation bias caused by the backfill of management companies.

Figure 6. Time trend between adopters vs. non-adopters

(a) Log(Avg.Rent)



(b) Vacancy Rate



Next, we compare key variables along with building characteristics between adopters and non-adopters. Figure 6 plots evolution of log(avg.rent)[12] and vacancy rates. First, we see a systematic *level* difference in log(avg.rent) between adopters and non-adopters with no particularly noticeable sign of deviating trend in earlier years. Vacancy rates, on the other hand, is quite different. Adopters and non-adopters start out with small difference in vacancy rates in both 2010 and 2011. Then adopters tend to maintain high vacancy rate above 5% while non-adopters goes down to 4%. This echoes the findings mentioned in Section 2.2 that the pricing software may have achieved the optimal balance between the vacancy rate and asking rents.

[12]Average rents are computed by the average asking rents across different bedrooms (0,1,2,3,4) weighted by the number of units of each type.

Table 1. Building characteristics comparison between adopters vs. non-adopters

	Non-Adopters	Adopters
Avg.Rent($)	1355.00	1701.32
	(810.60)	(886.87)
Vacancy Rate(%)	5.78	7.66
	(7.22)	(8.71)
Free Rent(Month)	0.03	0.03
	(0.02)	(0.02)
Year Built	1979.91	1995.20
	(23.90)	(19.55)
Floors	3.87	5.05
	(4.36)	(6.24)
Pool	0.64	0.83
	(0.48)	(0.37)
Doorman	0.03	0.05
	(0.18)	(0.21)
Tennis Court	0.00	0.01
	(0.07)	(0.08)
Parking Garage	0.05	0.09
	(0.21)	(0.29)
Clubhouse	0.36	0.65
	(0.48)	(0.48)
$N_{building}$	28,632	8,584

Table 1 compares key variables as well as some available characteristics of buildings between adopters and non-adopters. Adopters tend to have higher rents and vacancy rates than non-adopters but they seem to comprised of buildings with higher quality. They are more recently built, tend to be taller in number of floors, and tend to have more amenities that can be found in "luxury" apartments. Combined with the trend shown in Figure 6, it suggests that carefully controlling for building-specific covariates and covariate-specific time trend is necessary to measure the impact of adopting the pricing software on rents.

4 Estimation and Results

4.1 Estimation Procedure

We are ultimately interested in recovering the dynamic treatment effect of adoption of algorithmic pricing software on average monthly rent. There are two identifi-

cation challenges to overcome. First is the classical violation of the parallel trend assumption. If unobserved characteristics of buildings are correlated with the adoption decision, this would result in biased inference. Since adoption decisions are made at the management company-level, it is unlikely that building-specific unobservable is correlated with company-level decision, similar to the case of gas stations in Assad et al. (2020). All of top 20 management companies in 2022 operates across dozens of states.[13] To further mitigate any remaining potential selection, we estimate the effects using the doubly robust estimator of SantAnna and Zhao (2020), controlling for covariate-specific time trend between the treated and never treated units.

The second source of challenge involves heterogeneity in treatment effects. From both theoretical and empirical literature, it has shown that adoption effects of pricing algorithm can vary by local market structure, such as share of other adopters (Asker et al., 2022; Assad et al., 2020; Calvano et al., 2020). Moreover, the treatment effects may vary temporally due to the algorithm's learning curve specific to local markets or software updates made by the software developers. This issue is a more nuanced one that has recently received lots of attention in the difference-in-differences literature. It is often required for researchers to carefully select the comparison units in order to avoid contamination of treatment effects by already-treated units.

The estimation procedure of Callaway and SantAnna (2020) provides a rich, yet transparent way of dealing with both issues. Let $Y_{ic(i)m(i)t}$ be the outcome of interest of building i managed by company c in local market m at year t. If $c(i)$ has not adopted a software, $C_i = 1$. If $c(i)$ has adopted a software in time $t = g$, then $G_{ig} = 1$. Under the potential outcome framework without conditioning on pre-period covariates, the average treatment effect on the treated (ATT) of treatment cohort $g \in \mathcal{G}$ in year $t \in \mathcal{T}$ is:

$$ATT(g,t) = E[Y_{ic(i)m(i)t} - Y_{ic(i)m(i)g-1} \mid G_{ig} = 1] - E[Y_{ic(i)m(i)t} - Y_{ic(i)m(i)g-1} \mid C_i = 1].$$

Both g and t are in calendar year. The estimator has an intuitive explanation; suppose we are interested in the instantaneous effect of the adoption. Then $ATT(g, t = g)$ is the instantaneous treatment effect of the buildings that adopted a software in year g (i.e. "cohort" g):

$$ATT(g,t=g) = E[Y_{ic(i)m(i)g} - Y_{ic(i)m(i)g-1} \mid G_{ig} = 1] - E[Y_{ic(i)m(i)t} - Y_{ic(i)m(i)g-1} \mid C_i = 1],$$

[13] https://www.nmhc.org/research-insight/the-nmhc-50/top-50-lists/2022-top-managers-list/

which is simply comparing year-over-year change of the outcome of interest between the treated group, $G_{ig} = 1$, and the never-treated group, $C_i = 1$. If we assume that the units treated in g and never-treated units would have evolved similarly over the course of one year between $g-1$ and g (i.e., the parallel trend assumption), the difference of the means yields a consistent estimate of $ATT(g, g = t)$. This can be extended to any period before or after the treatment for any given cohort.

Incorporating covariates to achieve the *conditional* parallel trend assumption resembles matching comparison groups on their pre-period covariates and removing covariate-specific temporal variation from the outcome variable. Specifically, given a matrix of pre-period covariates X, the "doubly-robust"(SantAnna and Zhao, 2020) $ATT^{dr}(g,t)$ is:

$$ATT^{dr}(g,t) = E\left[\left(\frac{G_{ig}}{E[G_{ig}]} - \frac{\frac{p_{ig}(X)C}{1-p_{ig}(X)}}{E[\frac{p_{ig}(X)C}{1-p_{ig}(X)}]}\right)(Y_{it} - Y_{ig-1} - \phi_{gt}(X))\right],$$

where $p_{ig}(X)$ is an inverse probability weighting of never-treated units by pre-period covariates following Abadie (2005), and

$$\phi_{gt}(X) = E[Y_{it} - Y_{ig-1} \mid X, C_i = 1],$$

which accounts for outcome evolution of never-treated units (Heckman et al., 1997, 1998). In practice, the propensity score, $p_{ig}(X)$, is estimated using a logit model, and the outcome regression, $\phi_{igt}(X)$, is estimated using a linear regression model. For comparison with the "usual" two-way fixed effects (TWFE) models, see Appendix B.

4.2 Dynamic Effects of Software Adoption

Once $\widehat{ATT}(g,t;X)$ is recovered for all $g \in \mathcal{G}$ and $t \in \mathcal{T}$, we can take weighted averages across dimensions and/or across subsets to construct appropriate treatment effects of interest. In doing so, we closely follow the procedure of Callaway and SantAnna (2020). Our main outcome of interest in $\log\left(Avg.Rent_{ic(i)m(i)t}\right)$. We first examine the dynamic effect of software adoption on average rents from five years prior to adoption to seven years post adoption in Table 2 with different sets of covariates.

Table 2. Dynamic effects of pricing software adoption

	(1)	(2)	(3)	(4)
Pre Avg	-0.003	-0.001	0.000	0.010
ATT	(0.002)	(0.002)	(0.002)	(0.007)
Post Avg	0.027***	0.019***	0.035***	0.037***
ATT	(0.007)	(0.007)	(0.007)	(0.007)
T-5	-0.007	-0.004	-0.004	-0.005
	(0.005)	(0.005)	(0.005)	(0.024)
T-4	0.003	0.006	0.006	0.042**
	(0.005)	(0.005)	(0.005)	(0.021)
T-3	-0.005	-0.004	-0.003	-0.002
	(0.005)	(0.005)	(0.004)	(0.023)
T-2	-0.004	-0.003	-0.000	-0.003
	(0.004)	(0.003)	(0.003)	(0.015)
T-1	0.001	0.001	0.002	0.020
	(0.003)	(0.003)	(0.003)	(0.017)
T+0	0.005**	0.005**	0.010***	0.013***
	(0.002)	(0.002)	(0.003)	(0.004)
T+1	0.012***	0.009***	0.017***	0.023***
	(0.004)	(0.003)	(0.004)	(0.007)
T+2	0.017***	0.012**	0.023***	0.032***
	(0.005)	(0.005)	(0.005)	(0.011)
T+3	0.029***	0.019***	0.034***	0.043***
	(0.007)	(0.007)	(0.007)	(0.011)
T+4	0.040***	0.026***	0.045***	0.051***
	(0.009)	(0.009)	(0.009)	(0.013)
T+5	0.042***	0.028***	0.051***	0.050***
	(0.010)	(0.011)	(0.010)	(0.009)
T+6	0.037***	0.026**	0.051***	0.050***
	(0.011)	(0.011)	(0.011)	(0.009)
T+7	0.034**	0.025*	0.052***	0.038***
	(0.013)	(0.013)	(0.012)	(0.006)
$N_{building}$	33,392	33,392	33,332	33,332
Controls	No Control	Building Chars	RentQ + Chars	Metro#RentQ + Chars

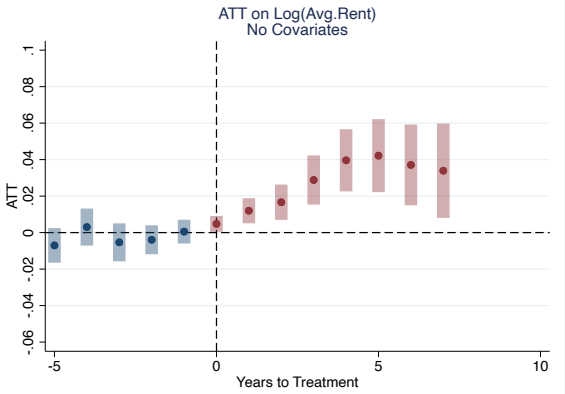
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Always-treated units and singletons are dropped from the analysis. Building characteristics include vacancy rate, avg. concession of the submarket and # of free rents offered. The set of amenities proxies for "high end" apartments and includes indicators for doorman, parking garage, pool, clubhouse, and tennis court. The rent quantiles are computed in each building's first year of apparence in the data within its metro area, representing its rent level relative to other buildings in the metro at its first year of operation. The standard errors are clustered at management company-level.

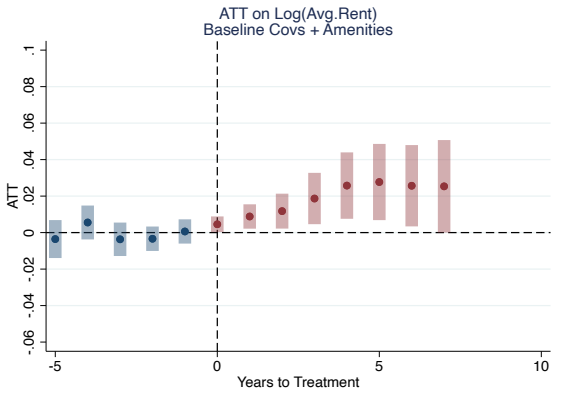
The first two rows show pre-period treatment and post-period average treatment effects. The pre-period *average* effect is close to zero across all specifications and post-period average effect is significant and positive across all specifications. It is common to assess pre-period relative time effects to look for signs of potential violation of the parallel trend assumption. We add controls incrementally and find no sign of potential violation of the assumption as shown in column (1)-(3). Because of this, we suspect that there is no systematic difference between the treated and never-treated buildings despite of our main specification showing statistically significant deviation in four years prior to adoption. Figure 7 shows a visual representation of Table 2. The post-period dynamic effects are also robust across specifications and we focus further analysis using the ATTs estimated from the specification of column (4).

Figure 7. Dynamic Effect of Software Adoption

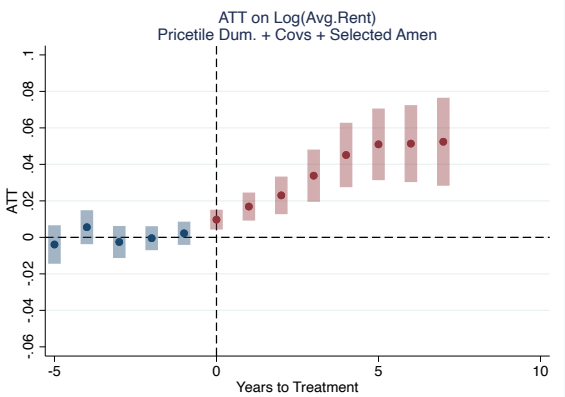
(a) No Control



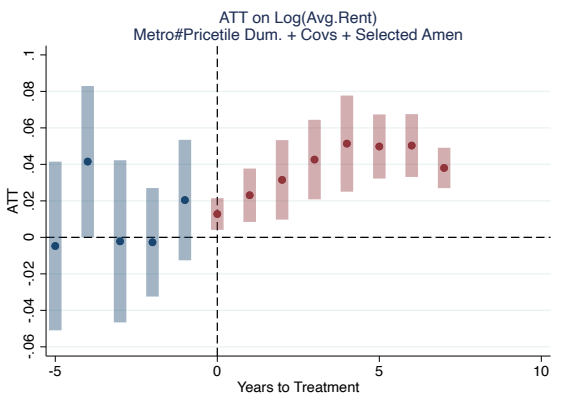
(b) Building Chars



(c) Rent Quantile + Chars



(d) Metro-Rent Quantile + Chars



Indeed, we find a convincing evidence suggesting the presence of dynamic effects.

On average, the adopters increased their rents by 3.7% points, but the effects during the first three years of adoption is less than the average. It then culminates to the "full effect" four years after the adoption at around 5% points more in average monthly rents compared to non-adopters, and stabilizes thereafter. This finding is consistent with the prior literature on algorithmic pricing software. Theoretically, Calvano et al. (2020) and Asker et al. (2022) attribute this to algorithm's "learning curve." In practice, there could be various mechanisms behind the observed dynamics, such as general technological advances, software updates made by developers, synergies from software companies acquiring others. While we do not attempt to fully address and disentangle the potential channels, we focus on one specific moderating factor: the degree of software penetration in a geographical market in a given year.

4.3 Heterogeneous Effects across Geographical Markets

How algorithmic pricing software can potentially facilitates price coordination among its users and shapes competitive landscape of a market is the main interest for both economists and policy makers. As briefly discussed in Section 2.2, the pricing software takes detailed past and current rents and vacancies of other buildings, which could potentially be competing buildings, into consideration when suggesting optimal rent prices. Since the software collects detailed data mainly through other buildings using the same software in a given geographical market, the treatment effect of adoption may vary with the level of penetration of the software in the market. Intuitively, if more buildings adopt the software, the more precise projection of future prices should be possible, leading to better optimization across all of its customers. While we do not make any definitive claim for possible price coordination, we carefully look for the signs by examining the heterogeneous treatment effects across markets.

The recovered ATT is for each group $g \in \mathcal{G}$ in year t, but buildings within each group are located across geographical markets. The treatment effect of interest is the average effect of adopting a software in geographical market m in year t across buildings that have adopted by t. Specifically,

$$ATT_{m,t} = \sum_{g \in \mathcal{G}} \mathbf{1}\{t \geq g\} P\left(g(i) = g | g(i) \leq t\right) ATT(g,t) \quad \forall i \text{ s.t } m(i) = m.$$

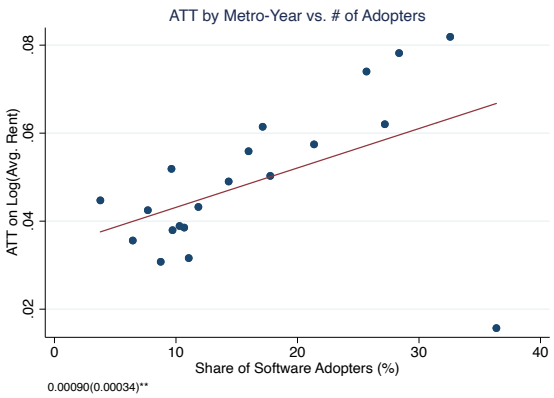
In words, we are taking the weighted average of ATT of year t across buildings within market m and adopted the software by year t.

We first filter market-year pairs with the p-value of $\widehat{ATT}_{m,t}$ less than or equal to

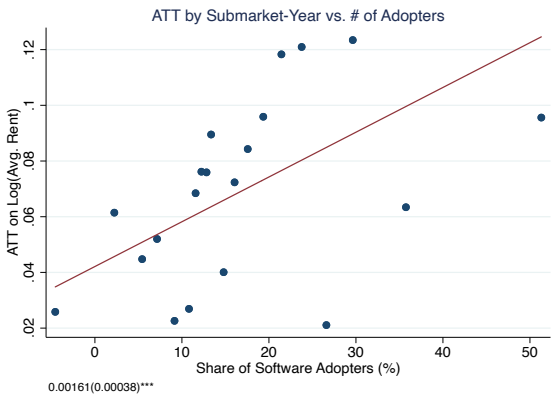
0.1. We then regress $\widehat{ATT}_{m,t}$ on share of buildings adopted the software in market m in year t, $AlgoShare_{m,t} = \frac{\#\text{ Adopted Buildings by } t \text{ in } m}{\#\text{ Buildings in } m,t}$, weighted by # of buildings in each m, t. For the definition of geographical market, we use the two pre-defined market definitions by REIS: metros and submarkets as described in Section 3. Figure 8 shows binned scatterplots of market-year level relationship between $\widehat{ATT}_{m,t}$ and $AlgoShare_{m,t}$.

Figure 8. Heterogeneous Treatment Effects by Degree of Penetration of Software

(a) Metro-Year



(b) Submarket-Year



The coefficient in the footnote of Figure 8 is the $\hat{\beta}_1$ from the following regression:

$$\widehat{ATT}_{mt} = \alpha + \beta_1 AlgoShare_{mt} + \beta_2 Num.Adopters_{mt} + \theta_t + \epsilon_{mt},$$

where θ_t is year fixed effect. Neither of $\hat{\beta}_2$ is significant. This is noteworthy because if the software optimizes rent setting through the scale effect of collecting data for more "precise" optimization, we should also see a greater treatment effect associated with the number of adopters conditioning on share of adopters. That is, collecting data across 80 adopters out of 100 buildings should have differential effect on ATT on prices than collecting data across 4 out of 5 buildings. This is not the case, however. The main channel of different magnitudes of treatment effect is through the variation of share of adopters, which suggests that higher treatment effects are achieved through coordination among adopters in presence of less non-adopters.

Figure 8 shows that there is a significant correlation between the magnitude of the treatment effects and the degree of penetration both visually and statistically. Conditioning on a year, geographical markets with higher penetration of the software

tend to be associated with greater price effects compared to non-adopters. In metro areas, 1% point increase in share of adopters is associated with 0.09% point increase in treatment effects compared to 0.16% point in submarkets. It is intuitive that this correlation is stronger both in magnitude and significance in submarkets than in metro areas, since the average number of buildings in submarkets are substantially smaller (59.3) than that of metro areas (744.3). Consistent with the hypothesis at the beginning of this section as well as the empirical findings of Assad et al. (2020), we find a strong evidence towards potential price coordination among the software adopters.

5 Conclusion

In this paper, we examined the impact of algorithmic pricing software adoption on average rent using unique datasets. We first document detailed inner workings of a pricing software. We find that the software collects high-frequency data from clients of its own, then fully utilizes the collected data to respond to changes in a local market. We then empirically find a robust effect of 3.7% point increase in rents after the adoption of a software. The adoption effect is highly dynamic, starting around 1% increase in rents then culminating to 5% increase in rents. Moreover, we find highly heterogeneous treatment effects across geographical markets. These effects vary by local market structure; the higher the share of software adopters of a market is, the higher the treatment effect of adoption. One percentage point increase in the share of adopters is associated with 0.09% point higher impact on rents in larger markets and 0.16% point in smaller sub-markets. The treatment effects do not vary with the number of adopters conditional on the share of adoption, suggesting that the increase in treatment effects are being achieved through price coordination among adopters, rather than better optimization of rent through collecting more data of the market.

This research contributes to the thin empirical literature on algorithmic pricing, and is surprisingly consistent with the previous findings from the literature. The findings are especially meaningful given the importance of the U.S. multifamily apartment industry. It could have significant distributional implications since multifamily apartment dwellers are often from the socioeconomically disadvantageous background. For future analysis, we hope to resolve the potential measurement errors due to the limitations of the datasets and conduct a structural analysis for evidence of price coordination among the adopters and its welfare implications.

References

Abadie, Alberto, "Semiparametric Difference-in-Differences Estimators," *The Review of Economic Studies*, 2005, *72* (1), 1–19. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].

Asker, John, Chaim Fershtman, and Ariel Pakes, "Artificial Intelligence, Algorithm Design, and Pricing," *AEA Papers and Proceedings*, May 2022, *112*, 452–456.

Assad, Stephanie, Robert Clark, Daniel Ershov, and Lei Xu, "Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market," *SSRN Electronic Journal*, 2020.
Bason v. RealPage, Inc.

Bason v. RealPage, Inc., *October 2022.*

Brown, Zach Y and Alexander MacKay, *"Competition in Pricing Algorithms," 2021, p. 65.*

Callaway, Brantly and Pedro H.C. SantAnna, *"Difference-in-Differences with multiple time periods," Journal of Econometrics, December 2020, p. S0304407620303948.*

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicol, and Sergio Pastorello, *"Artificial Intelligence, Algorithmic Pricing, and Collusion," American Economic Review, October 2020, 110 (10), 3267–3297.*

Fortin, John, *"Algorithms and Conscious Parallelism: Why Current Antitrust Doctrine is Prepared for the Twenty-First Century Challenges Posed by Dynamic Pricing," SSRN Electronic Journal, 2020.*

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd, *"Characterizing Selection Bias Using Experimental Data," Econometrica, 1998, 66 (5), 1017–1098. Publisher: [Wiley, Econometric Society].*

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd, *"Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," The Review of Economic Studies, 1997, 64 (4), 605–654. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].*

Mcsweeny, Terrell and Brian ODea, *"The Implications of Algorithmic Pricing for Coordinated Effects Analysis and Price Discrimination Markets in Antitrust Enforcement," 2017.*

Mikls-Thal, Jeanine and Catherine Tucker, *"Collusion by Algorithm: Does Better Demand Prediction Facilitate Coordination Between Sellers?," Management Science, April 2019, 65 (4), 1552–1561.*
Navarro v. RealPage, Inc. et al

Navarro v. RealPage, Inc. et al, November 2022.

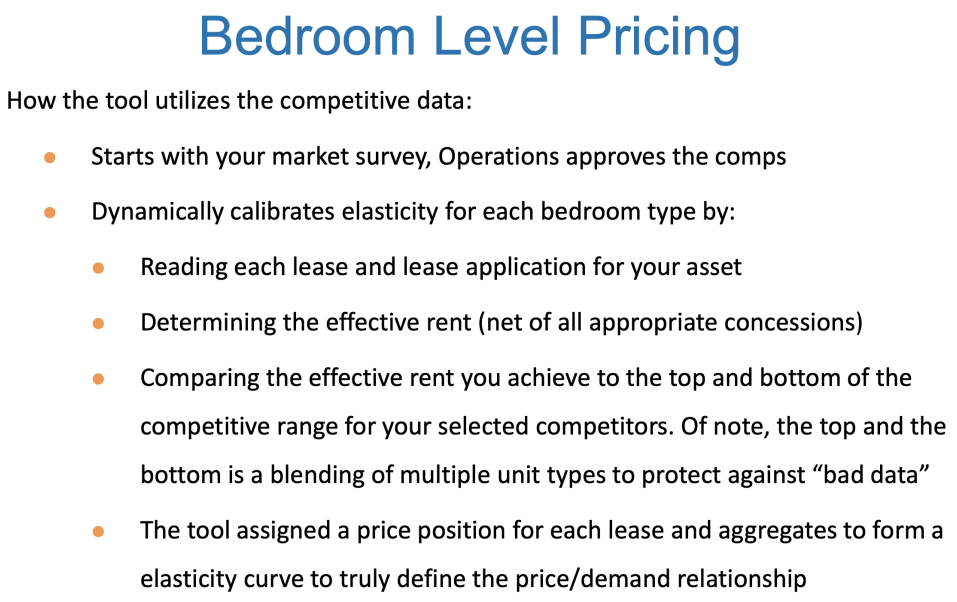
OECD, "Algorithms and Collusion - Note by the United States," 2017.

SantAnna, Pedro H.C. and Jun Zhao, "Doubly robust difference-in-differences estimators," *Journal of Econometrics*, November 2020, *219* (1), 101–122.
Yusupov v. RealPage, Inc. et al

Yusupov v. RealPage, Inc. et al*, January 2023.*

A Additional Slides from Yieldstar's Showcase

Appendix Figure A1. How Yieldstar optimizes rents



B TWFE Specification

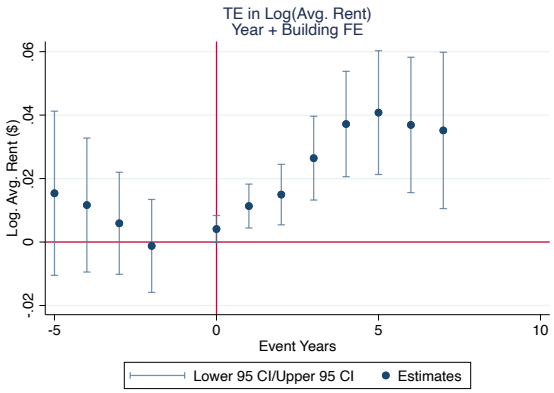
We supplement our analysis with the "usual" TWFE regression model. Specifically, we run building(i)-year(t) level TWFE regression:

$$\log(Avg.Rent_{it}) = \alpha + \sum_{\tau=-5}^{-2} \delta_\tau \mathbf{1}\{g(i) - t = \tau\} + \sum_{\tau=0}^{7} \delta_\tau \mathbf{1}\{g(i) + t = \tau\} + X_{it}\beta + \theta_i + \theta_t + \epsilon_{it},$$

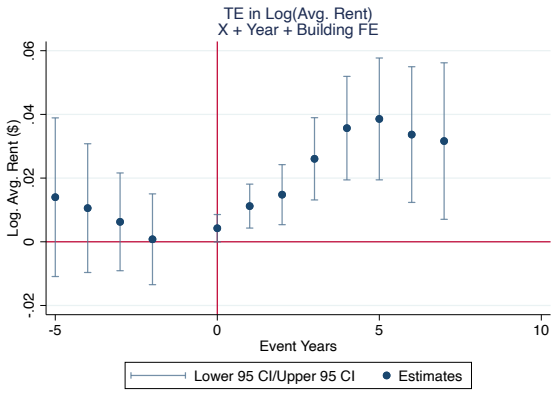
and plot $\hat{\delta}_\tau$. We use the same set of controls and control for year, metro-year, metro-price quantile-year fixed effects.

Appendix Figure B1. Dynamic Effect of Software Adoption under TWFE Model

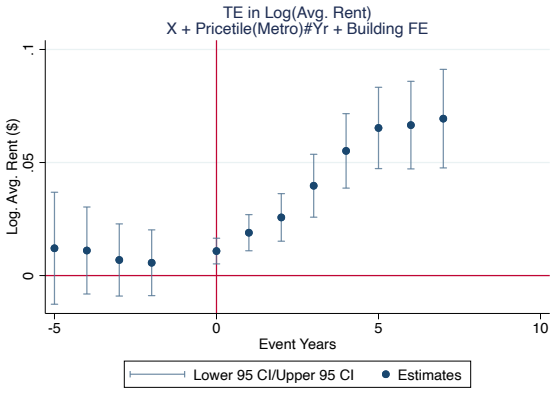
(a) No Control



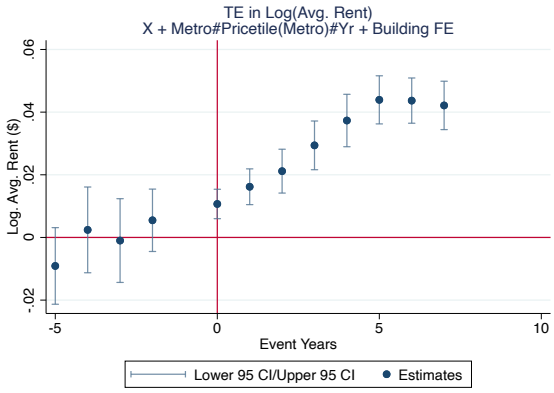
(b) Building Chars



(c) Rent Quantile + Chars



(d) Metro-Rent Quantile + Chars



As shown in Figure B1, the pretrend results and the post-period effects are very similar to Table 2.