

”Blockbusting” in the 21st Century?: Minority Move-ins and Neighborhood Home Value Appreciation

Emerson Schryver*

April 6, 2025

Abstract

This paper investigates the link between minority move-ins and neighborhood home value appreciation. It analyzes loan-level home purchase data to construct a dataset of move-in demographics and we find a statistically significant relationship between an increase in the share of move-ins from minorities and a suppression of home-value appreciation over the next six years. This relationship is particularly prevalent in non-republican geographies.

JEL Codes: J15, R23, R33

*Undergraduate Student, University of Toronto, emerson.schryver@mail.utoronto.ca

1 Introduction

Housing discrimination has, in one way or another, existed in the United States since independence. After reconstruction, several tactics became commonplace, including racial deed covenants, and redlining. Across the nation, there was evidence of “white flight”, or the movement of whites out of neighborhoods with minorities due to fears over home value depreciation or other factors. Real-estate agents commonly attempted to abuse these fears with a tactic known as “blockbusting” in which they would spread fear over minority-move-ins leading to a fire sale of homes in a neighborhood (Rothstein, 2017).

Much has been researched about the modern-day effects of these past tactics, whether in deed covenants leading to improved relative neighborhood quality (Sood, Ehrman-Solberg, 2024), or redlining leading to localized areas of higher poverty (Appel, Nickerson, 2016), or simply lower quality of opportunity in majority-black neighborhoods (Chetty, et al, 2014). There is also evidence of modern-day tactics still occurring in the housing market, whether in lending markets (Quillian, Lee, Honoré, 2020), or in real estates continuing to practice “racial steering”, the process of, whether knowing it or not, primarily showing people of minority groups neighborhoods that are also primarily of that minority group (Glenn, 2018).

This paper seeks to contribute literature surrounding economic effects of racism by analyzing the modern accuracy of the perceived link that 50’s era blockbusting relied on – do minority move-ins suppress home-value appreciation?

There is a long theoretical literature on discrimination, starting with a model for taste-based discrimination by Gary Becker in 1957. There now exist many other models for discrimination, most famously Ken Arrow’s 1973 ‘statistical discrimination’, but Becker’s remains the most commonly used. There are also many theoretical frameworks for home sale matching, (Badarinza, Balasubramaniam, Ramadorai, 2024), and discrimination in labor market matching (Combes et al, 2016), but there have been few models for discrimination in the housing market. The sole notable attempt was in (Combes, et al, 2018), which built a theoretical framework for racism in home sales and lease arrangements, and then empirically tested the lease framework. This paper relied on “neighbor” discrimination, which is based on perceptions of others having discriminatory tastes and their potential to decrease the value of your asset (home). This has been extended to look at a case in Moscow, where much of the racial discrimination is more overt (Avetian, 2022)

This paper contributes to this literature by applying this theoretical framework to home sales, and, in particular, analyzing the long-term effects on home values. I use loan data from Fannie Mae (FNMA) & Freddie Mac (FHLMC) to provide data on neighborhood move-ins, American Community Survey data to normalize, and Zillow ZHVI data to show the change in home prices. My study has the following methodology:

First, I select zip codes with move-ins during the 2009-2010 period (Initial Period) in metropolitan statistical areas (necessary for appreciation normalization). I take the “minority move-in share”, or the ratio of loans for new purchases made in that zip code to minorities, and then look at those same zip codes over the 2012-2013 period (Treatment Period) and analyze what “treatment” was applied (the difference in minority move-in share from the previous period), which is our main explanatory variable. Next, we analyze the association between the change in minority move-in-share and the future appreciation in home prices, relative to their MSA average, for the next 6 years (ending in 2019). This should inform us what the effect of a sudden increase in minority move-ins is on home value appreciation.

I find that an increase in the share of move-ins that are from minorities is associated with a statistically significant (though relatively minor) suppression in home value appreciation relative to the metropolitan statistical area. This effect is less present in more republican geographies, suggesting some degree of heterogeneous preferences between partisan affiliations.

2 Data

This project has loan-level data from two the main loan packagers in the United States, the Federal National Mortgage Association (Fannie Mae), and the Federal Home Loan Mortgage Corporation (Freddie Mac). This data spans from 2009 to 2023, but the analysis for this paper uses the subset from 2010-2011 and 2012-2013. This subset has over 3.5 million loans, and includes information on borrower and coborrower characteristics, home location (census tract), and loan purpose. This data is then aggregated by location (and filtered to solely 'new purchase' loans) to act as a proxy for move-ins, with the racial composition of move-ins, in a given geography.

I use data from Zillow's Zillow Home Value Index (ZHVI) to analyze the change in home values over time. This data is at the zip-code level, and is available monthly for all the years of analysis. I use quarter four averages of the ZHVI to remove some data loss (certain locations have missing values in certain months), and analyze the percent change from 2010 in each zip-code.

For normalization, and to analyze potential interactions, I use county-level election results from the 2012 presidential election from the MIT Election Data Science Lab. Many zip-codes are contained within a single county. For zip-codes that aren't, we pick the county with the lower FIPS code. Additionally, we use ZCTA (zip-code tabulation area – largely synonymous with zip-codes)-level income and demographic data (IPUMS, 2025). This provides data on the Gini index (inequality level), the per-capita income, as well as the zip-code population size, and proportion of residents who are of minority backgrounds or white backgrounds.

3 Summary Statistics

Table 1: Summary Statistics of Home Value Changes by Year (only in-MSA ZIPs)

	MMI (I)	MMI (T)	MI (I)	MI (T)
Count	19847	19847	19847	19847
Mean	8	12	100	167
Standard Deviation	20	25	142	227
Minimum	0	0	0	0
25th Percentile	0	0	6	19
Median	1	3	44	76
75th Percentile	8	14	137	230
Maximum	403	615	1460	2338
Range	403	615	1460	2338

Where MMI = Minority Move-Ins, MI = Move-Ins, (I) is the initial period, and (T) is the treatment period

Table 1 shows details on move-ins particularly for zip-codes in metropolitan statistical areas (MSAs). This subsetting is critical for robust analysis of home-value appreciation. Over a quarter of all zip-codes experienced no minority move-ins during one of the periods, meaning that the treatment is relatively concentrated.

Table 2: Table 2: Summary Statistics of Average Home Value Change (% point difference from MSA average) by Year

Statistic	count	mean	std	min	25%	50%	75%	max
2011	18246	0.003	0.039	-0.392	-0.015	0.005	0.023	0.362
2012	18248	0.003	0.062	-0.503	-0.028	0.004	0.037	0.489
2013	18248	0.001	0.083	-0.712	-0.042	0.000	0.045	0.885
2014	18248	-0.000	0.101	-0.801	-0.054	-0.002	0.049	0.973
2015	18248	-0.002	0.125	-0.907	-0.067	-0.006	0.055	1.117
2016	18248	-0.005	0.148	-1.054	-0.083	-0.012	0.060	1.669
2017	18248	-0.007	0.175	-1.147	-0.102	-0.017	0.066	1.968
2018	18248	-0.011	0.210	-1.240	-0.121	-0.024	0.072	2.496
2019	18248	-0.012	0.234	-1.251	-0.137	-0.029	0.081	3.036

Table 2 shows average home-value appreciation relative to the average for the MSA the zip-code is in. The variation is high and increases with time, which is useful for analysis.

Table 3: Summary Statistics for Income

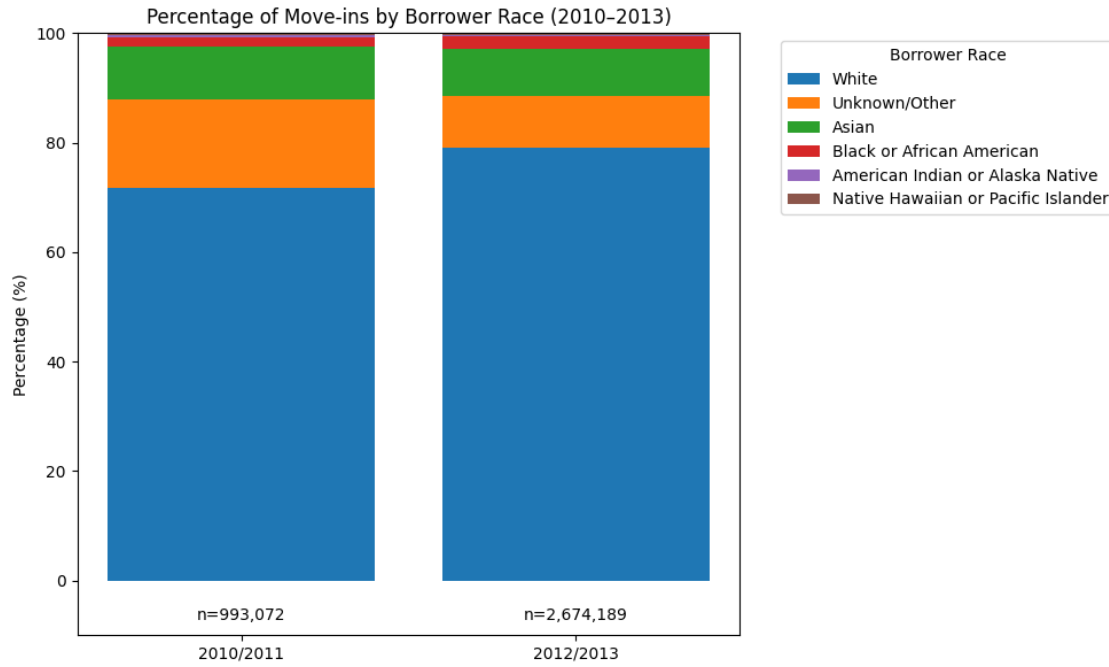
	Median Household Income	Per Capita Income	Gini Index
Count	30562	32208	32208
Mean	73170	39472	0.414700
Standard Deviation	31347	18844	0.081300
Minimum	2499	421	0.001000
25th Percentile	53500	28762	0.376500
Median	67028	35573	0.418200
75th Percentile	85316	44910	0.460100
Maximum	250001	419459	1.000000

Table 4: Summary Statistics for Demographics and Political Affiliation

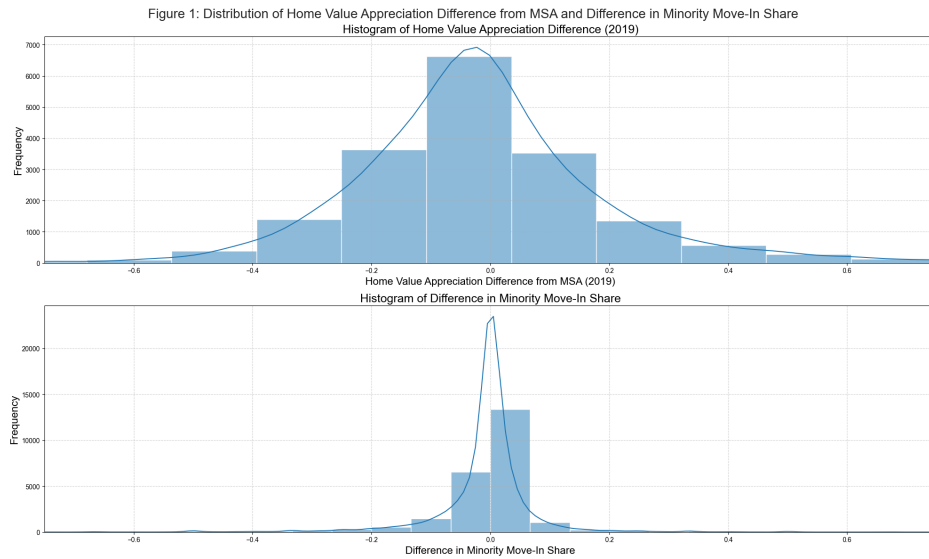
	% White	% Republican
Count	32208	29540
Mean	0.765480	0.539260
Standard Deviation	0.226515	0.152746
Minimum	0.000762	0.071934
25th Percentile	0.669562	0.432962
Median	0.859255	0.544754
75th Percentile	0.931172	0.650237
Maximum	1.000000	0.932903
Range	0.999238	0.860969

Table 3 and Table 4 provide information on control and interaction terms (income, demographics, and political affiliation).

Figure 0: Racial breakdown of move-ins



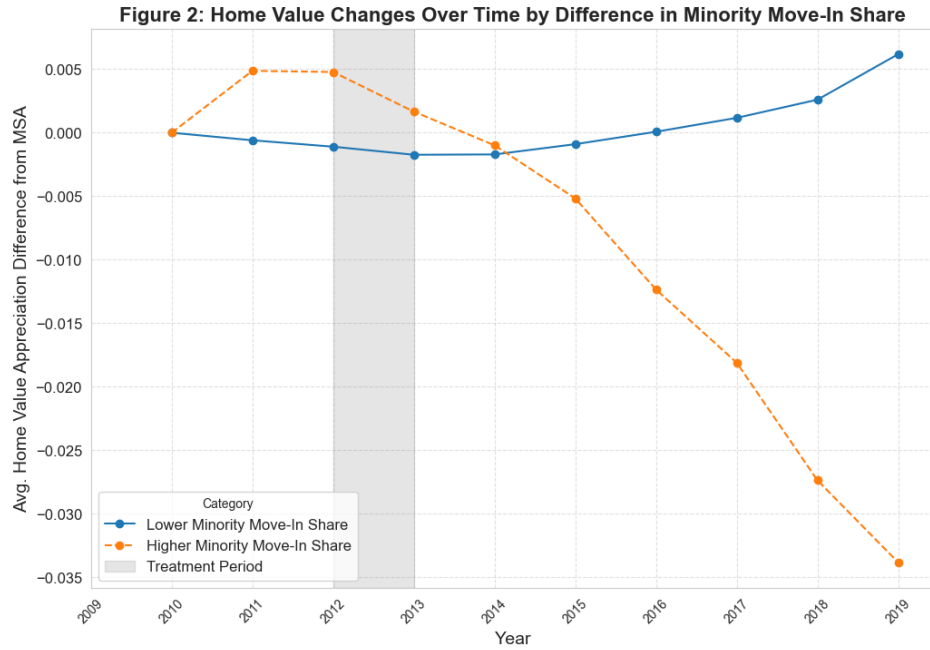
This chart shows the change in primary borrower primary racial demographics from the initial period to the treatment period over the whole sample of loans. You can see that only a slim proportion are classified as Black or African American, with the vast majority being White or Asian. This is consistent with national homeownership demographics. For the purposes of this analysis, a move in is classified as “minority” if either the borrower or coborrower are either partially African American or are hispanic.



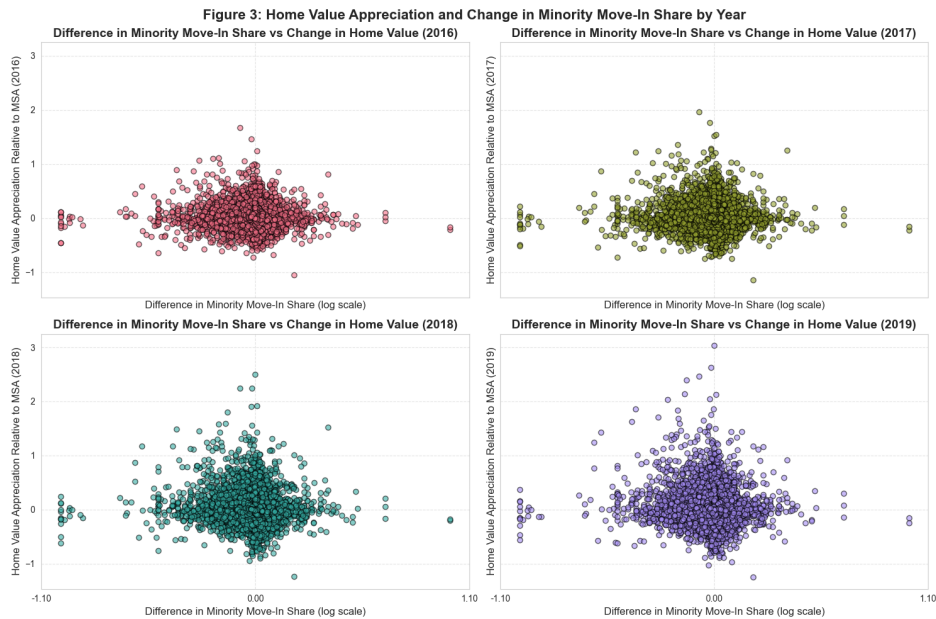
As shown in the summary tables above, the data has an incredibly large range. This plot shows how incredibly wide the range is in home value appreciation and minority move-in share (and how incredibly concentrated the difference in minority move-in share is around zero). This figure shows how large the distribution is. In the next few plots we will attempt to answer my research question and see to what extent the distribution on the left can be explained by the distribution on the right.

4 Results

4.1 Visualizations

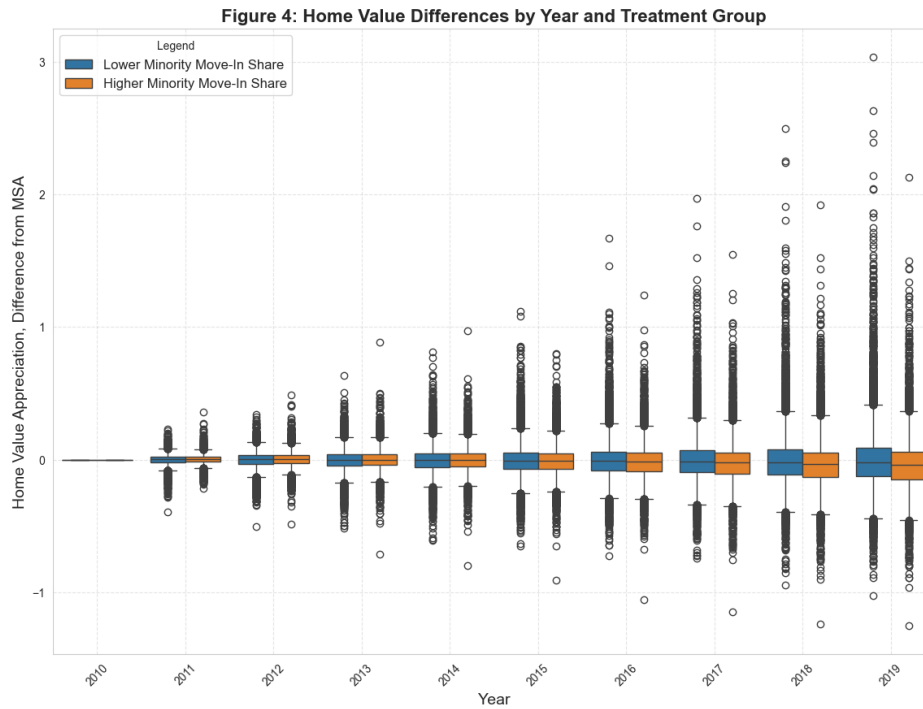


This figure shows the difference in outcome between the treatment group (higher minority move-in share) and control group (lower / same minority move-in share). It seems to indicate that the treatment experienced some home-value appreciation relative to their MSA prior to the treatment period. During the treatment period, the treatment group begins to experience a decrease in home values, which continues until the end of the graph. In contrast, the control group has every slightly increasing values for roughly the entire period. This seems to indicate that neighborhoods with a higher minority move-in share have depressed home value appreciation relative to groups without.



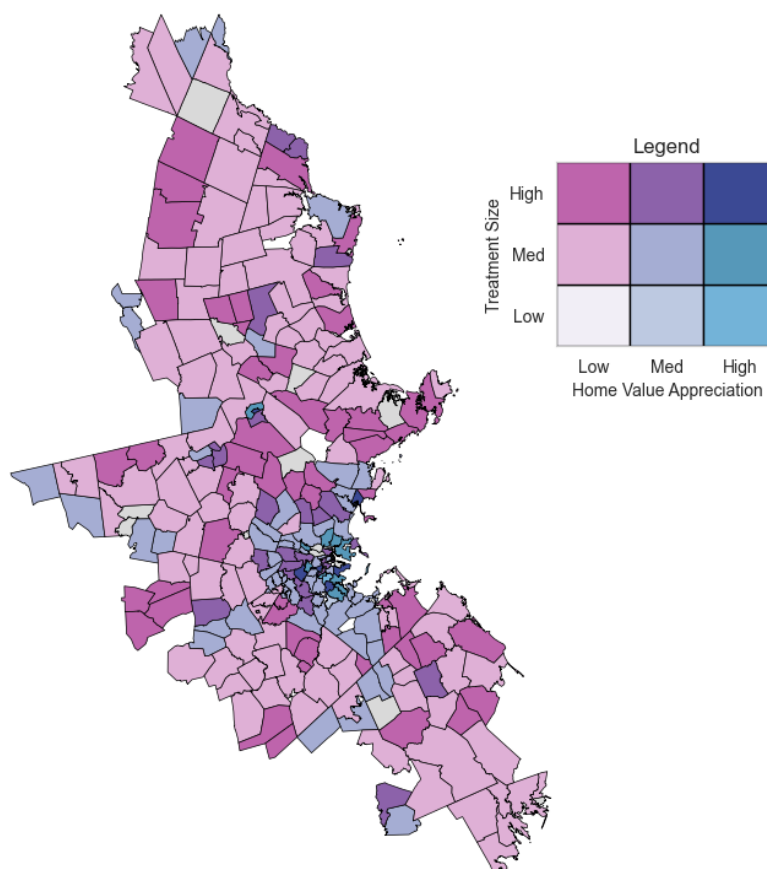
This figure shows something very important not shown in Figure 2, which is how wide the variation is. I chose a log scale to show more detail (as you can see, there are a few far outliers, with most of the data

clustered near zero). From these scatterplots, it is clear there is a high degree of variance, but it appears that areas with high treatment are less likely to receive a large increase in home-value appreciation relative to their MSA than areas with lower treatment. This is more noticable with time.



This figure shows the extent of the increase in variation across the time period, and that there is a decrease in the median for the higher minority move-in share group. There remains a high degree of variation in home values relative to the metropolitan area, which cannot be explained by the change in minority move-in share, but it does appear that the areas with an increased minority move-in share generally experienced decreased home-value appreciation.

Treatment Size & Home Value Changes (Boston Metro)



This bivariate choropleth reinforces the fact that there seems to be a relationship between the change in minority move-in share and home-value appreciation in the Boston MSA. The magenta represents areas with high minority move-ins and low home value appreciation. The teal represents areas with low minority move ins and high home value appreciation. The grey-blue represents moderate values on both. Most areas in the suburbs are either magenta, light pink, or teal, suggesting an inverse relation between change in minority move-ins and home value appreciation.

4.2 Regressions

The maps and plots suggest that there is a negative relationship between the treatment (more minority move-ins) and home-value appreciation. However, this relationship cannot be made robust without regressions. My analysis involved several regression models, controlling first for neighborhood demographics, then income, and finally political affiliation.

Table 5: Regression Specifications with Demographics

<i>Dependent variable: value ratio 2019</i>				
	Change in Average Home Value (%) (2010-2019)			
	(0)	(1)	(2)	(3)
intercept	1.450*** (0.003)	0.040*** (0.010)	0.108*** (0.010)	0.043*** (0.010)
MSA avg appreciation		0.963*** (0.007)	0.924*** (0.007)	0.961*** (0.007)
total pop			0.000*** (0.000)	
treatment size	-0.324*** (0.034)	-0.263*** (0.023)	-0.155*** (0.022)	-0.158*** (0.031)
white			-0.000*** (0.000)	
white×treatment size				-0.000*** (0.000)
Observations	15919	15919	15919	15919
R^2	0.006	0.550	0.579	0.550
Adjusted R^2	0.006	0.550	0.579	0.550

Note:

*p<0.1; **p<0.05; ***p<0.01

Model (0) shows a barebones regression with no controls – it indicates that there is a significant negative relationship between an increase in the share of minority move-ins and home value appreciation, but it’s practically useless as it has very low explanatory power. Model (1) shows a baseline model controlling only for average MSA appreciation – it indicates there is still a negative relationship between minority move-ins and home value appreciation, though it is smaller, and the model has much higher explanatory power. Model (2) indicates a similarly low positive relationship, and shows that there is little-to-no confounding occurring due to population size or whiteness of the area. Model (3) analyzes the interaction between zip-code whiteness and the change from the increase in minority move-in share and finds that there is no indication that whiteness affects the treatment effect.

Table 6: Regression Specifications with Income and Inequality

	<i>Dependent variable: value ratio 2019</i>		
	Home Value Change (%) (2010-2019)		
	(4)	(5)	(6)
Gini index		-0.013 (0.032)	-0.015 (0.032)
High Gini×MMI-share change			-0.045 (0.047)
intercept	0.126*** (0.010)	0.045*** (0.017)	0.046*** (0.017)
MSA avg appreciation	1.006*** (0.007)	0.964*** (0.007)	0.964*** (0.007)
Income (per capita)	-0.000*** (0.000)		
treatment size	-0.153*** (0.022)	-0.263*** (0.023)	-0.235*** (0.037)
Observations	15919	15919	15919
R^2	0.583	0.550	0.550
Adjusted R^2	0.583	0.550	0.550

Note:

*p<0.1; **p<0.05; ***p<0.01

Model (4) controls for income, but it does not seem that income has an economically significant effect. It decreases the coefficient on treatment size (a 1% increase in the minority move-in proportion is associated with a 0.153% decrease in home-value appreciation), similarly to controlling for whiteness (this make sense, as due to socioeconomic factors, income and whiteness are linked). (5) and (6) analyze how the effect varies over different levels of inequality in neighborhoods. The literature would suggest that those in neighborhoods with a lower Gini index would likely have preferences leading to a higher tolerance of neighborhood diversity, which would lead to a smaller change in home values after a change in MMI share, but the interaction term in (6) does not reach significance, and neither does the control term in (5).

The regression with the highest "explanatory power" involved the interaction between partisan affiliation and treatment. It has the following formula:

$$\begin{aligned}
\text{Home value appreciation} = & \beta_0 + \beta_1(\text{Increase in minority move-in share}) \\
& + \beta_2(\text{Home value appreciation of MSA}) \\
& + \beta_3(\text{Highly Republican}) \\
& + \beta_4(\text{Highly Republican} \times \text{Increase in minority move-in share}) + \varepsilon
\end{aligned}$$

Table 7: Regression Specifications with Political Affiliation

	<i>Dependent variable: value ratio 2019</i>	
	Home Value Change (%) (2010-2019)	
	(7)	(8)
Highly Republican \times Treatment Size		0.227*** (0.025)
Highly Republican		0.009** (0.004)
intercept	0.048*** (0.013)	0.047*** (0.011)
MSA avg. appreciation	0.962*** (0.007)	0.957*** (0.007)
Percent Republican	-0.013 (0.012)	
Treatment Size	-0.262*** (0.023)	-0.304*** (0.023)
Observations	15889	15889
R^2	0.549	0.552
Adjusted R^2	0.549	0.552
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01		

When evaluating regressions, I looked for significant terms. An ideal regression has clear conclusions with significant terms and a high R^2 . My preferred specification is regression (8), which estimates $\beta_1 = -0.304$ with high significance. It also estimates $\beta_4 = 0.227$ with high significance. This regression suggests that a 1% increase in treatment size is correlated with a 0.304% decrease in home-value appreciation in general. However, in particularly republican geographies, it is associated with a 0.227% increase in home-value appreciation. This regression has relatively high "explanatory power", with a R^2 of 0.552.

4.3 Machine Learning

I ran two main types of machine learning algorithms on my data: a "regression tree" and a "random forest." The regression tree is based on choosing quantities for categorical splits to minimize the mean-squared error (MSE). It has the following objective:

$$\min_{\text{splits}} \sum_{m=1}^M \sum_{i \in R_m} (y_i - \hat{y}_{R_m})^2$$

Where y_i is the actual outcome for observation i in R_m , and \hat{y}_{R_m} is the predicted outcome in the region R_m . The algorithm chooses "splits" or different ways of dividing the region of possible inputs into regions.

Figure 5 displays a regression tree with all of my X variables indicates that the income level, income inequality, and political activity are the largest predictors of home value appreciation relative to metropolitan area. Having a higher per capita income level suggests lower home value appreciation, as does being more highly republican. Higher political activity is associated with higher home value appreciation in higher-income areas, but lower home value appreciation in lower-income areas. All terms have a high level of error.

Figure 6 displays a regression tree containing my main X variables (change in minority move-ins, percent of population that voted republican, total treatment move-ins, and proportion of the zip code that is white). This tree seems to indicate that the total quantity of move-ins is of high importance, with fewer move-ins indicating higher appreciation (this is counterintuitive to economic intuition). In places with higher move-ins, the proportion of republicans is the next most important variable, with less republican areas experiencing

Figure 5: Regression Tree with All Inputs

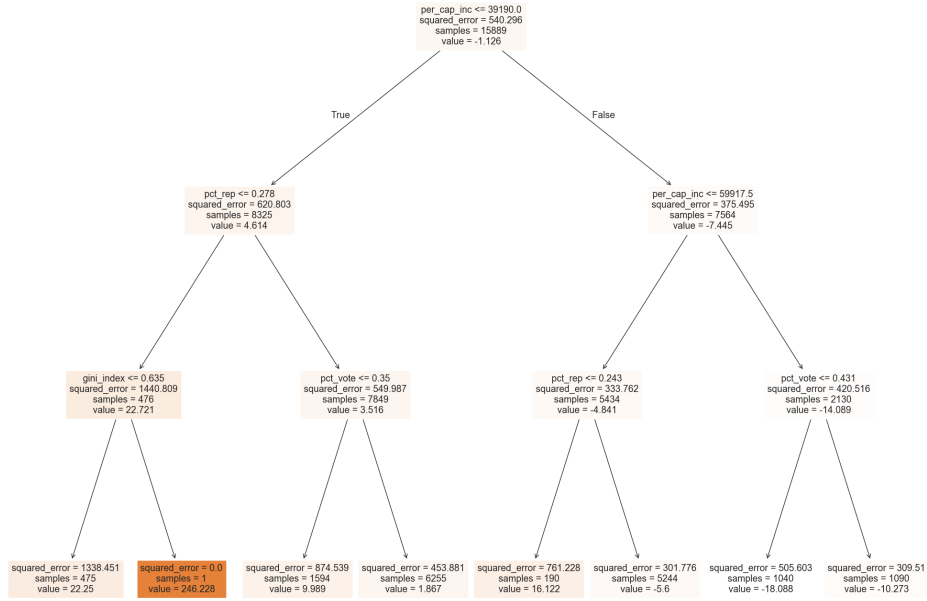
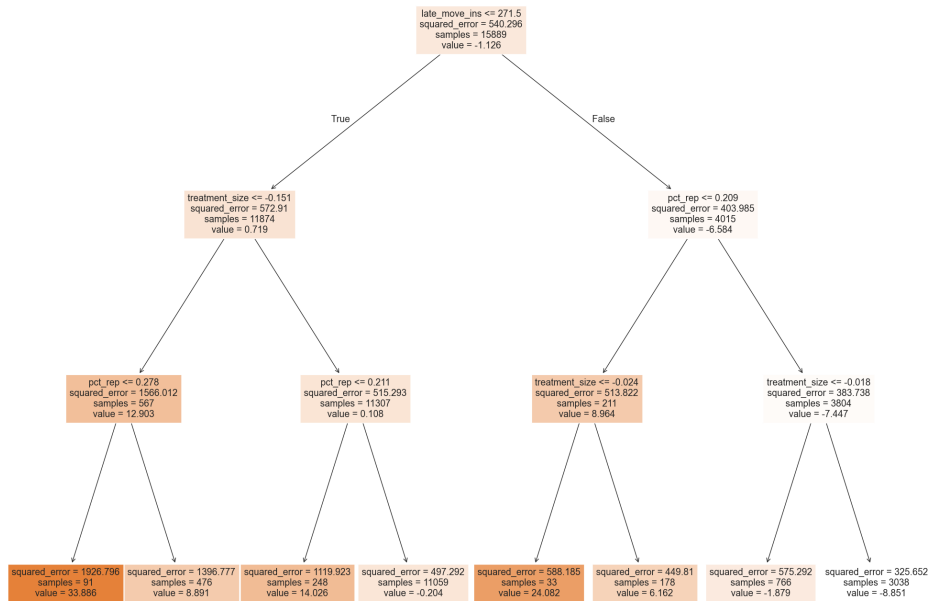


Figure 6: Regression Tree with Selected Inputs



higher home-value appreciation. Next is the treatment. ZIP Codes with a high quantity of move-ins that receive a high treatment dose experience lower home-value appreciation than those who don't. High move-in republican ZIP codes with higher treatment doses have even lower home-value appreciations. In low move-in ZIP codes, treatment dose is the next most important, followed by republican proportion. The same effects are seen. The error is higher on this model due to the removal of many control variables.

Figure 7: Importance Matrix of Variables

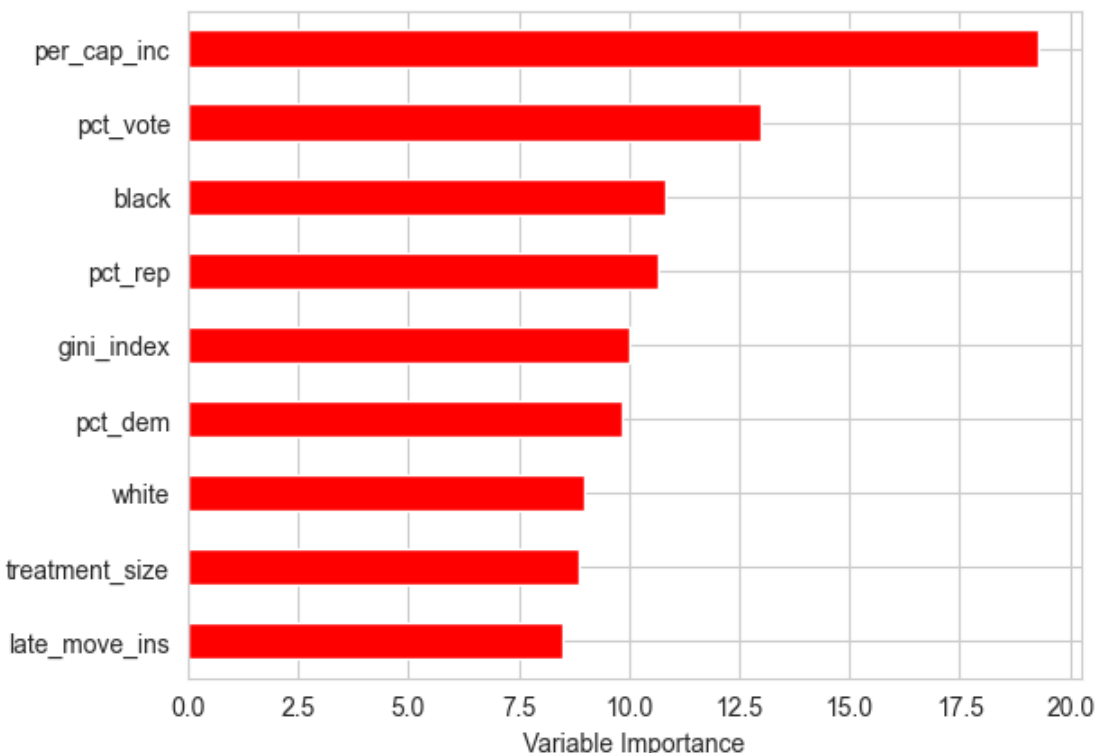


Figure 7 displays a bar chart which indicates that the most important variable is per capita income. This is followed by voting characteristics and neighborhood demographics, and then by the treatment size. It makes sense that treatment size would not be one of the most important variables, income and political affiliation could reasonably be more important for individual's preference mappings, or there could be confounding variables associated with them that changed over this period.

The results from my ML seem to run slightly contrary to our regressions, which indicated that income had little to no effect on home-value appreciation. It's possible that this is due to a slightly distinct y between our ML models and our regressions. In the regressions, we looked at the raw home value appreciation, and then added the average home-value appreciation in the MSA as a control term. In the ML models, to improve clarity, we are looking at the difference between home value appreciation and the average home value appreciation the MSA.

My regression trees (and random forest model) allow for significantly more analysis of interactions between variables than my regressions, which is what leads to slightly different results.

5 Conclusion

In this paper, I analyze the link between an increase in the share of minority move-ins (a treatment) and home values on a neighborhood level. I construct a zip-code level dataset that includes, from Fannie Mae and Freddie Mac data, the share of minority move-ins in an initial period, and the share of minority move-ins in a treatment period. The data also includes Zillow data showing the change in home values. For

much of the analysis, this change in home values is normalized to the average change in home values to the MSA the zip code is in.

My findings indicate a statistically significant negative relationship between an increase in minority move-ins and home-value appreciation. This effect is present when normalizing for MSA-level appreciation, and is consistent with webscraped listing data in Appendix A. Interestingly, this relationship is significantly less negative in highly (top 50%) Republican geographies. This could be due to different preference maps, or some other endogenous factor.

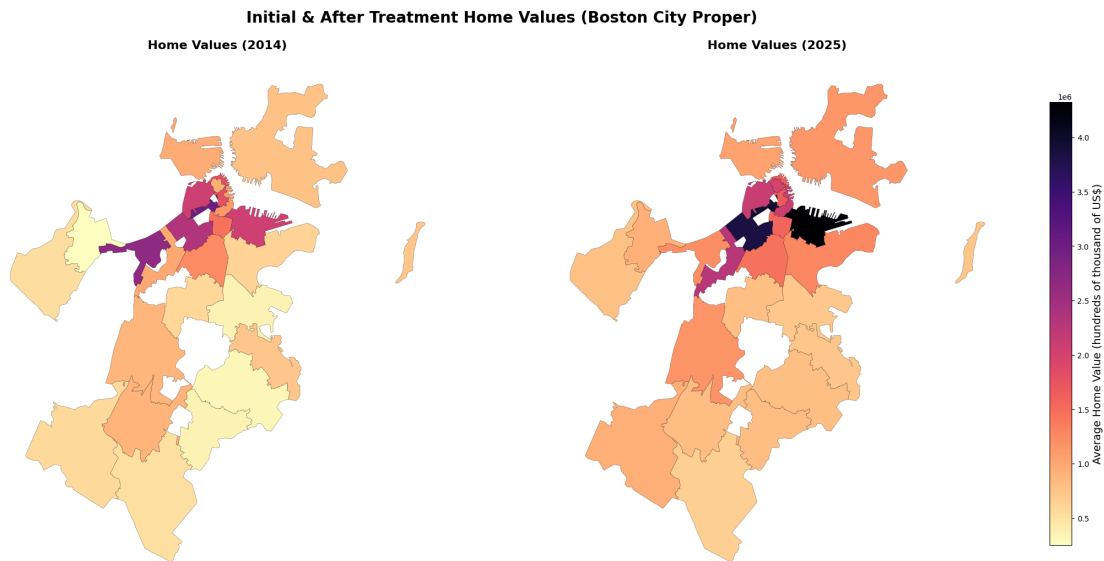
More analysis is needed, including modifying the methodology to something more similar to an event study, to allow for more years of treatment and better consideration for treatment size. Presently, there is too much endogeneity for any causal analysis. It's possible that modifying the study structure could change that. In addition, it would be insightful to use the data and regressions to estimate values for the parameters in the rational discrimination theoretical model this analysis is based upon.

6 References

- Appel, I., & Nickerson, J. (2016). *Pockets of Poverty: The Long-Term Effects of Redlining*. Retrieved from <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=2852856>
- Arrow, K. (1971). *The Theory of Discrimination*. Retrieved from <https://dataspace.princeton.edu/handle/88435/dsp014t64gn18f>
- Avetian, V. (2022). *Consider the Slavs: Overt Discrimination and Racial Disparities in Rental Housing*. Retrieved from <https://www.tse-fr.eu/sites/default/files/TSE/documents/conf/2022/echoppe/avetian.pdf>
- Badarinza, C., Balasubramaniam, V., & Ramadorai, T. (2024). *In Search of the Matching Function in the Housing Market*. SSRN Scholarly Paper No. 4594519. <https://doi.org/10.2139/ssrn.4594519>
- Becker, G. S. (1971). *The Economics of Discrimination*. University of Chicago Press.
- Combes, P.-P., Decreuse, B., Laouénan, M., & Trannoy, A. (2016). *Customer Discrimination and Employment Outcomes: Theory and Evidence From the French Labor Market*. *Journal of Labor Economics*, 34(1), 107–160. <https://doi.org/10.1086/682332>
- Combes, P.-P., Decreuse, B., Schmutz, B., & Trannoy, A. (2018). Neighbor Discrimination Theory and Evidence from the French Rental Market. *Journal of Urban Economics*, 104, 104–123. <https://doi.org/10.1016/j.jue.2018.01.002>
- Korver-Glenn, E. (2018). *Compounding Inequalities: How Racial Stereotypes and Discrimination Accumulate Across the Stages of Housing Exchange*. *American Sociological Review*, 83(4), 627–656. <https://doi.org/10.1177/0003122418781774>
- Manson, S., Schroeder, J., Van Riper, D., Knowles, K., Kugler, T., Roberts, F., & Ruggles, S. (2024). *National Historical Geographic Information System: Version 19.0* [Dataset]. Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D050.V19.0>
- Quillian, L., Lee, J. J., & Honoré, B. (2020). *Racial Discrimination in the U.S. Housing and Mortgage Lending Markets: A Quantitative Review of Trends, 1976–2016*. *Race and Social Problems*, 12(1), 13–28. <https://doi.org/10.1007/s12552-019-09276-x>
- Rothstein, R. (2017). *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liveright Publishing Corporation, a division of W. W. Norton & Company.
- Sood, A., & Ehrman-Solberg, K. (2024). *The Long Shadow of Housing Discrimination: Evidence from Racial Covenants*. Retrieved from <https://drive.google.com/file/d/1uLSaQxWiSHKMuckF2gFpATywQD2J7No5/>

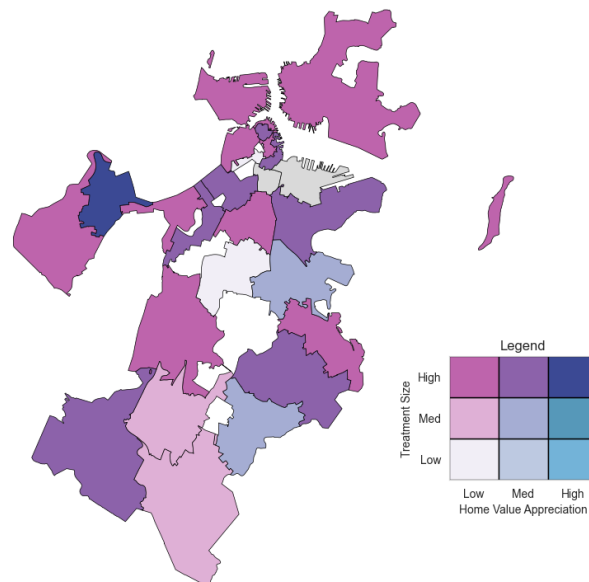
Appendix

A Webscraped Data



This map shows the set of data gathered from webscraping. Because we were scaled back to only Boston City Proper, you can see that there is significant home-value increases across the board. Unfortunately, it appears that these two times of webscraping weren't enough to remove data gaps (in fact, we have more). However, it is clear that if this script could more easily gain archival data, we would be able to fill in gaps.

Treatment Size & Home Value Changes (Boston City Proper)



This map shows less certainty around the relationship between treatment size and home value changes. This is likely due to the fact that we are using a small sample size of data, and that the "post-treatment" home values are after so much time. However, it does show that there is a relationship between treatment size and home value changes, which is consistent with the above regressions, maps, and plots.