Why Parties Displace Their Voters: Gentrification, Coalitional Change, and the Demise of Public Housing *

Yanxi Fang *Harvard University*

This report attempts to replicate an article of the same title, written by Winston Chou and Rafaela Dancygier and published in the *American Political Science Review* in 2021. The report offers an introduction to the article's motivation and main findings, presents various aspects of its statistical analyses, and offers additional insight into the topic. The goal is to provide a stronger understanding about the political factors contributing to the lack of affordable housing across advanced economies, using the example of the Labour Party in the UK to demonstrate the implications of political incentives for the future of urban housing.

Keywords: affordable housing, gentrification, London, Labour Party

Introduction

Gentrification and urban redevelopment have been contentious issues across the advanced, industrialized world for the past several decades. As middle- and upper-class suburban residents increasingly choose to return to newly-rehabilitated urban neighborhoods for their convenience and walkability, they displace working-class residents who had previously lived in the older, more affordable housing units. Since members of the working class provide essential urban services at relatively low wages, their displacement may limit the supply of labor, adversely affecting the long-term economic growth and prosperity of cities.

Thus far, however, many cities have actually accelerated - not limited - the pace of gentrification, while affordable housing is still a far reach for many poorer urban residents. Since housing policy is an inherently political issue, an examination of politics and electoral incentives will be necessary for understanding the lack of policy responses to the existing situation: after all, since most advanced and industrialized countries are governed by democracies, urban residents are hypothetically able to influence policy by participating in local elections.

In practice, many poor urban residents tend to vote for left-leaning parties in the hope that they will implement policies that increase the supply of affordable housing. However, the evidence shows many instances of so-called "state-led" gentrification, where local governments reduce the supply of public housing and incentivize the development of new, unaffordable housing in inner cities, have actually been implemented by those same left-leaning parties. While this may seem counterintuitive, Winston Chou and Rafaela Dancygier (2021) provide empirical evidence that left-leaning parties have gradually shifted toward a middle- and upper-class coalition, making pro-gentrification housing policy compatible with their electoral incentives.

*This is based on an article of the same title by Winston Chou and Rafaela Dancygier (2021), published in the *American Political Science Review*. The replication files for the original publication are available on the Harvard Dataverse repository (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/EAPKQX), while the files used to generate this document are available on the author's Github account (https://github.com/yanxifang/Gov-52).

In this report, I will describe my attempts to replicate the empirical work done by Chou and Dancygier (2021). Starting with a brief literature review, I will first describe the qualitative foundations of their data analysis and provide some initial insights into the policy and practical implications of their findings. I will then summarize their approach to analysis, offering a roadmap of how they start from observations about the Labour Party in London and ultimately generate regressions to determine the pro-gentrification effects that Labour has had despite its traditionally working-class base. Subsequently, I will present the results of my own analysis, and explain that apart from several key differences and an inability to replicate several tables, the replication was successful in that I was able to obtain the same set of numerical findings. Finally, I will propose and pursue an extension to the model in the form of a comparison to the United States, which should provide further clarity about the situation and open more questions for future research.

Literature Review

Since the gentrification of cities is a phenomenon that has been in place for several decades, there has been a vast amount of research on the topic, both empirically (demonstrating the impacts) and theoretically (explaining the rationales). For example, Smith (1982) focuses on the ground-level actors driving gentrification, arguing that it is a product of the stratification of work into markedly different socioeconomic classes, causing different segments of society and different types of workers to vary drastically in their preferences and financial resources. Another explanation focuses on the broader-level actors, noting that decentralization and devolution have made local governments more dependent on local property taxes, which drives them to favor and promote tax-generating new development (Hackworth and Smith 2001). Other perspectives include governments using gentrification to maintain social order in disadvantaged neighborhoods (Uitermark and Duyvendak 2007), and that different socioeconomic classes have much choice and agency in making their decisions (Lyons 1996).

However, there is considerably less literature that links pro-gentrification policies to the electoral side of urban politics. While Chou and Dancygier (2021) cite multiple pieces of work that write about the relationship between left-leaning parties and the demise of affordable urban housing, they also note that there is no work describing the connection between the electoral incentives of such left-leaning parties and the policies that are implemented. Thus, the study that I will replicate represents a new line of research for urban housing policy, making its empirical findings particularly relevant and valuable.

Paper Overview

Research and Qualitative Analysis

In their article, Chou and Dancygier (2021) research the question of left-leaning parties' political incentives with regard to housing policy. To answer this question, they focus specifically on the Labour Party, which is a major left-leaning party in the UK, and the London metropolitan area, which they describe as one of the most unaffordable cities in Western Europe. The study itself is broken down into two major components, qualitative and quantitative, with the former being used to provide background context to rationalize the authors' choice of variables in the latter. Together, these two components provide a definitive answer: left-leaning parties are increasingly looking to broaden their coalitions, choosing to prioritize the policy preferences of middle- and upper-class voters over the housing needs of its traditionally reliable working-class voters.

The qualitative component comes first in the paper, where the authors describe historical trends in the housing market. They start with a numerical comparison across the developed OECD countries, showing that the real price of houses has increased, while public investment in housing, as a percentage of GDP, has decreased. They then cite various scholarly works to note that the housing market situation in London is particularly severe, backing those assertions with data to show that units have become more expensive and homeownership has declined - even in comparison to the rest of the UK. Next, they point to a decline in the number of publicly-owned council housing units (similar to public housing units in the United States), as well as the trend that the decline has been most significant in boroughs led by the Labour Party, even though housing is locally controlled. They then describe the electoral implications of the decrease in publicly-subsidized housing: using results from the British Household Panel Survey, they show that the effects on existing Labour supporters are temporary, implying that there is relatively little cost for politicians who promote policies that result in gentrification.

Thus, the authors have established the basic assumptions of the empirical study: that housing affordability is declining in London, and that the Labour Party has been responsible for a key part of that decline. This sets the stage for the quantitative component of the paper. The authors briefly allude to a series of qualitative primary- and secondary-source interviews to understand the motivations of Labour politicians. They start by evaluating the socioeconomic demographics of the Labour Party's support base as it changes over time, finding that members of the professional class, who intuitively have higher incomes, have increased their support for Labour over time. Coupled with the interviews, the authors show that policies geared against social housing (and thus toward gentrification) are not only politically safe in terms of not losing support, but also even attractive for Labour politicians in terms of gaining new voters. They conclude the paper with a quantitative measurement of how much the council housing stock has decreased over the years, using a series of regressions with multiple related variables and controls to measure changes over a decade-long period.

Quantitative Analysis

The quantitative analysis primarily consists of two sets of OLS regressions that measure changes in the concentration of council (public) housing across London over the decade between 2001 and 2011. The authors also generate and plot a series of predictions that are based on the second set.

The first set of regressions is done using multivariate OLS, and includes three related models. Across all three models, the outcome variable is the concentration of council housing in 2011, while the explanatory variables are (a) the 2001 concentration of council housing, (b) the 2006 (midway) level of electoral support for Labour candidates, along with (c) the 2001 crime rate and (d) the 2001 median income. There are also borough fixed effects, meaning that each borough essentially has a different intercept but the same slope coefficients. Here, the two housing concentration values, along with electoral support, are measured in absolute amounts (i.e. percentages), while the crime rate and median income variables are first log-transformed and then standardized with mean 0 and standard deviation 1. The unit of analysis is the individual ward, of which there are 624 being considered in the data set. The first model contains all of these variables, while the second model adds two interactions: between Labour support and the crime rate, and between Labour support and the median income. The third model contains everything from the second model, and adds on several ward-level indicators about the quality of council housing in 2001, the demand for housing, and demographics. These three regressions are intended to test the hypothesis that higher median incomes (which correspond to more middle- and upper-class voters in a ward) and higher levels of existing Labour Party support (which corresponds to a lower risk of incumbent Labour politicians being displaced by their loyal supporters) will lead to a greater reduction in the number of council housing units in a ward relative to 2001.

The second set of regressions is also done using multivariate OLS, but includes six related models. Here, the key difference is that instead of considering all 624 wards in the greater London area, the models are based only on the 127 wards for which the Labour Party controlled the (higher-level) borough in both 2002 and 2006. The outcome variable is the same as before: the concentration of council housing in 2011. The main explanatory variables are also identical, with the exception that the 2006 level of Labour support is no longer being used - thus also eliminating the interactions from the second model above. Instead, the level of Labour support is part of the division between different individual regressions: the first model is based on all 127 wards, the second model is based on 53 wards where Labour support was relatively weak, and the third model is based on 74 wards where Labour support was relatively strong. These first three models were simple ones in the sense that they only considered the key variables listed above; there were an additional three corresponding models that were more complex and used the quality, demand, and demographic indicators that were previously used in the first set of regressions. Here, the regressions are intended to test an alternative explanation to the findings of the first set of regressions: that non-Labour boroughs (a higher level than wards) were looking to undermine Labour by reducing council housing in Labour-dominated wards. By limiting the data analysis to only the Labour-controlled boroughs, that alternative explanation can be addressed.

Finally, the authors include a panel of six different graphs that show predictions based on the second set of regressions. These graphs use the three simple models from the second set of regressions, and have either the crime rate or the median income as the explanatory variable. Since all other explanatory variables are held constant at the average value of the dataset, the graphs display the sign and magnitude of the impact of only that variable on the outcome of interest, which is the level of housing in 2011. The authors call these interaction graphs, but a simpler explanation is that they show one input variable's predicted impact on the outcome variable when holding all other variables constant.

Findings

As a result of their quantitative analysis, Chou and Dancygier find that there is a negative relationship between higher levels of crime (and of median income) in 2001 and the council housing concentration in 2011, once the existing level of council housing and the borough fixed effects are taken into account. However, once the more complex model comes into play, the crime and income coefficients are no longer statistically significant, while the interaction between those two variables and Labour strength becomes much more statistically significant. In other words, reductions in council housing in response to crime and gentrification are dependent on Labour strength, implying that Labour politicians are not particularly concerned about the issue unless they are in wards that are less politically secure for themselves.

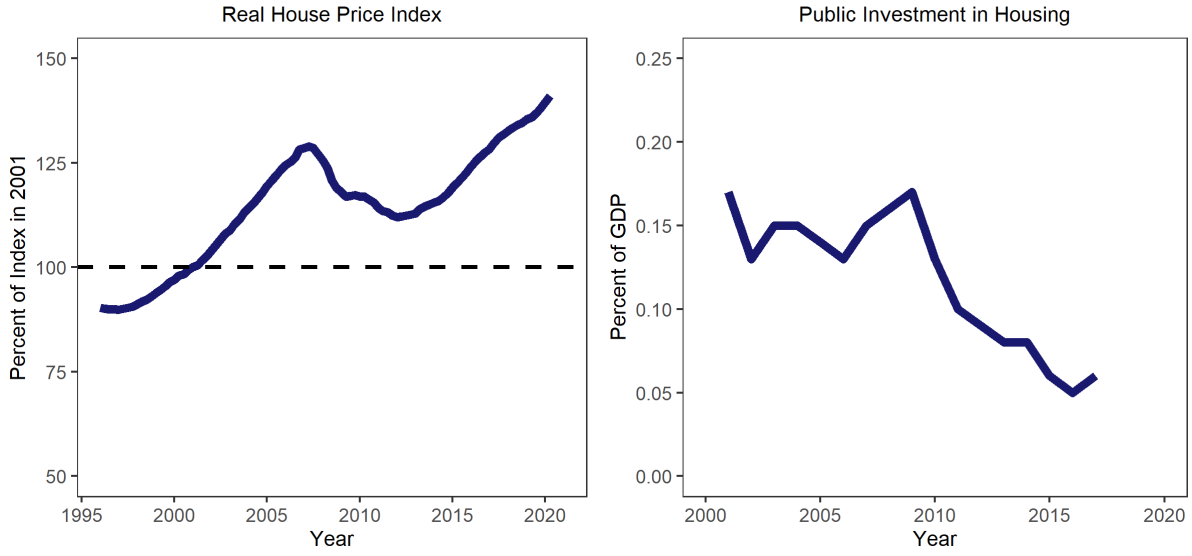
As described above, the authors attempt to address a potential alternate explanation (that non-Labour borough leadership is attempting to sabotage Labour by reducing its share of loyal supporters) by running a separate regression with only the wards that were located in Labour-controlled boroughs. This second analysis shows that the relationship between crime and median income are the strongest (i.e. largest magnitude, negative) in strong Labour wards located in strong Labour boroughs. Again, this finding is comptaible with the hypothesis that Labour politicians anticipate relatively low political costs of reducing council housing, and thus are incentivized to attract new middle- and upper-class residents and voters, which can be done through gentrification.

Replication

Qualitative Component

While it is best described as the qualitative component of the original paper, the first section consists of multiple data-heavy graphics. I started by replicating Figure 1, which depicts two side-by-side graphs showing real house prices and public investment in housing (respectively) across OECD countries, which are relatively developed and advanced. The data was readily available and was already substantially cleaned by the authors in *csv* form, and I only needed to perform a few small modifications to make the information ready to be converted into graphical form. Using 2001 as the baseline (the start date of the public investment in housing data), I generated graphs indicating that house prices increased in real terms between 2001 and the present day, while public investment in housing decreased by approximately two-thirds. Figure 1 is below; from a visual inspection, it appears to be the same as the original paper. As shown below, the index of real house prices increased while public investment in housing decreased, which suggests that broadly speaking, the average lower-income urban resident in OECD countries is having an increasingly difficult time with the affordable housing problem.

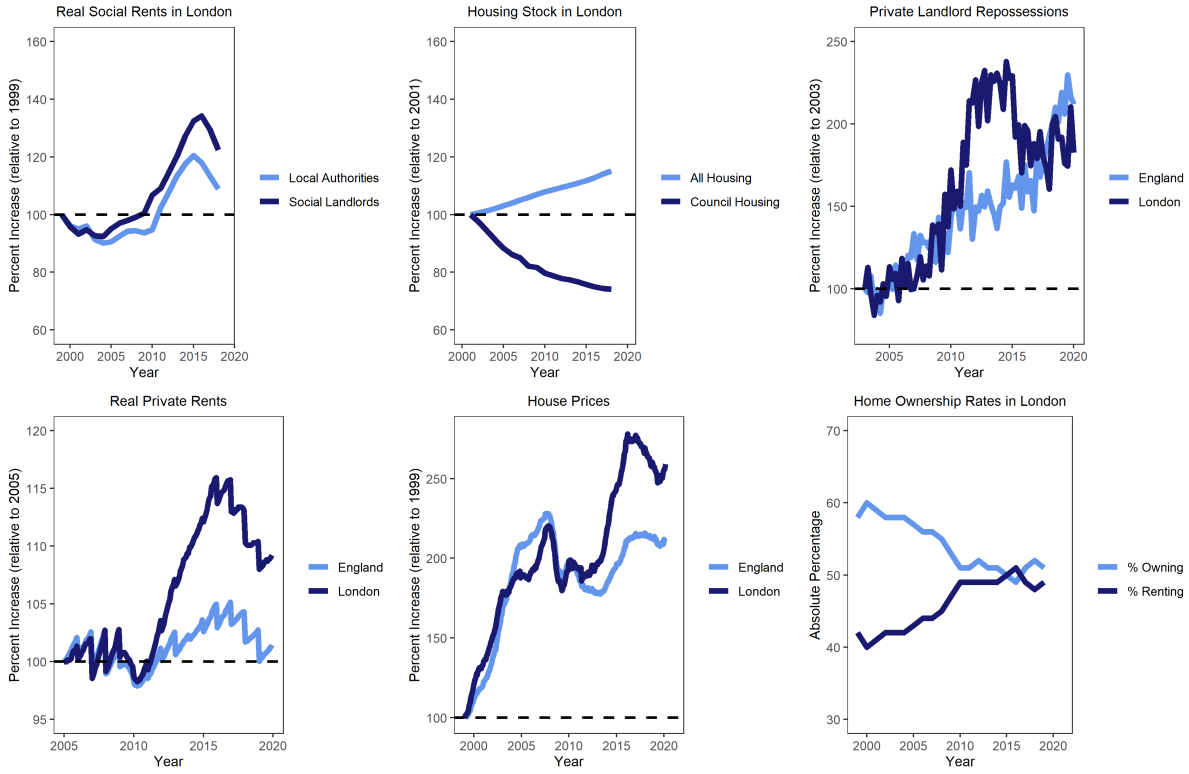
Figure 1: Real House Prices vs Public Investment in Housing in OECD Countries



Next, I replicated Figure 2, which graphically represents some of the housing trends seen in both London and the rest of England over the past two decades. Each part of the six-plot panel was based on a different dataset, so even though the data was readily available in *csv* form, I had to clean six different datasets and convert them to a format compatible with the graphing process. Five of the six data sets used the first year as the baseline, meaning that the graphs had percentage changes on the y-axis, which required additional data manipulation. Nonetheless, since these data were relatively straightforward, I was able to successfully replicate the graphs, which are shown on the next page. These graphs show a similar picture: that real rents and prices increased, market-rate housing became more available while social housing became less available, and that home ownership has decreased over time. All of these validate the authors' original motivation, which is that gentrification and housing affordability are political issues that should be better

understood.

Figure 2: Housing Trends in London and England



I then turned to Figure 3, which displays information about the stock of council homes across Labour-controlled, mixed-controlled, and Conservative-controlled boroughs in London. The underlying data was identical for both graphs, but while the one on the left uses absolute numbers, the one on the right makes the comparison using the percentage of the initial stock. Again, the dataset was not too complex to clean and manipulate, and I was able to replicate the graphs below. These show that the decline in council homes, in both absolute and relative terms, is larger in magnitude in Labour-controlled boroughs. This validates the premise that despite its historical working-class political base, the Labour Party is contributing to gentrification by removing many units of affordable council housing.

Figure 3: Decline of Publicly-Owned Council Housing in Greater London

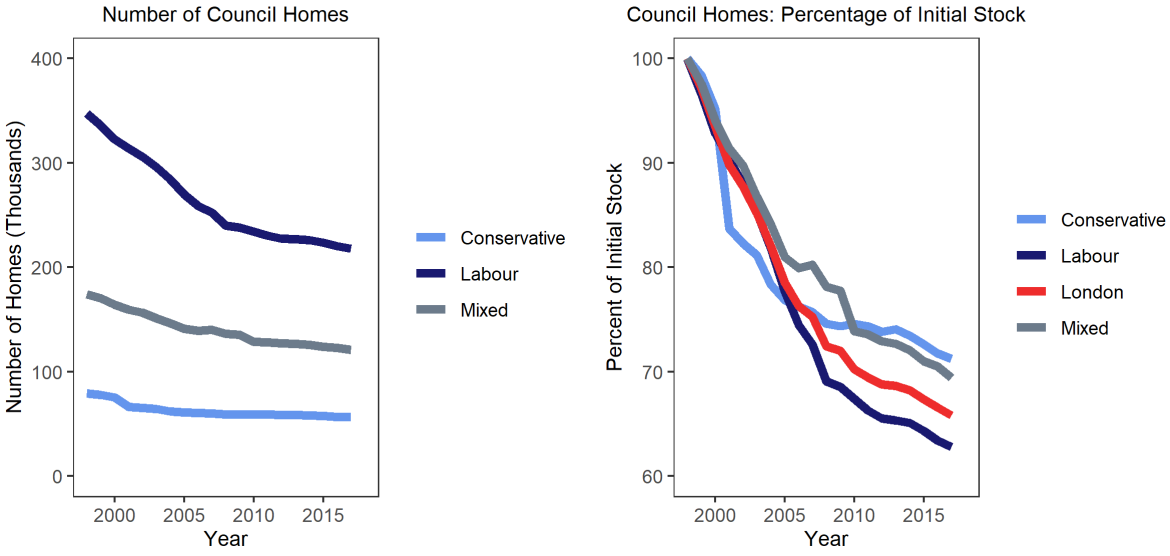
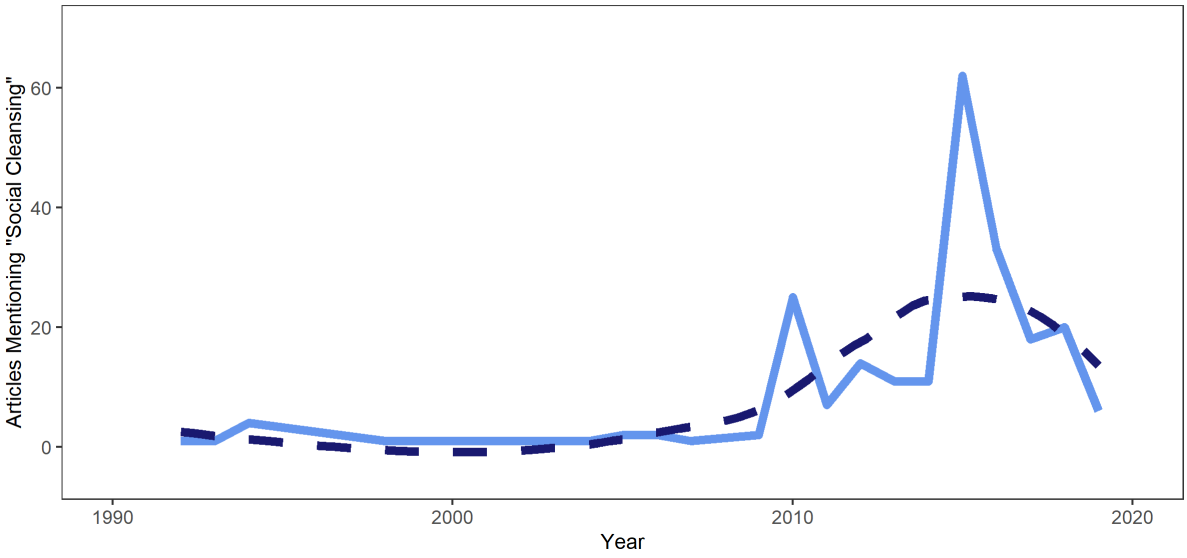


Figure 4 is intended to support a slightly more nuanced part of the authors' argument: that the decline in affordable housing in Labour-controlled wards and boroughs is caused by some desire for "social cleansing," in which poorer residents and areas with higher crime are targeted by Labour leaders for the removal of council housing in order to incentivize higher-income residents to move in (as part of Labour's gradual shift toward higher-income demographics). While the data - the number of news articles from *The Guardian* - may not indicate any causal effects whatsoever, it is a piece of qualitative evidence that supports the argument and provides further context for the quantitative parts of the paper that will be discussed later. As with the previous images, the data was relatively simple to clean, and I was able to produce a graph very similar to the original. The only difference is that the authors appear to have used a slightly different smoothing option to perform the Lowess regression than I did, since they had a significantly smoother curve.

Figure 4: Number of *Guardian* Articles Mentioning "Social Cleansing" in Relation to Housing



Following Figure 4, there were two additional figures and one table that I was unable to replicate. These three components were all based on data from the British Household Panel Survey (BPHS), which was not readily available from the Dataverse repository. Instead, I had to register with the UK government's data services branch (via https://www.ukdataservice.ac.uk/) and verify my identity as a student to gain access to the data. While this is not an issue by itself, the dependence on an external source meant that the data could be updated or changed without notice. Unfortunately, this happened to be the case.

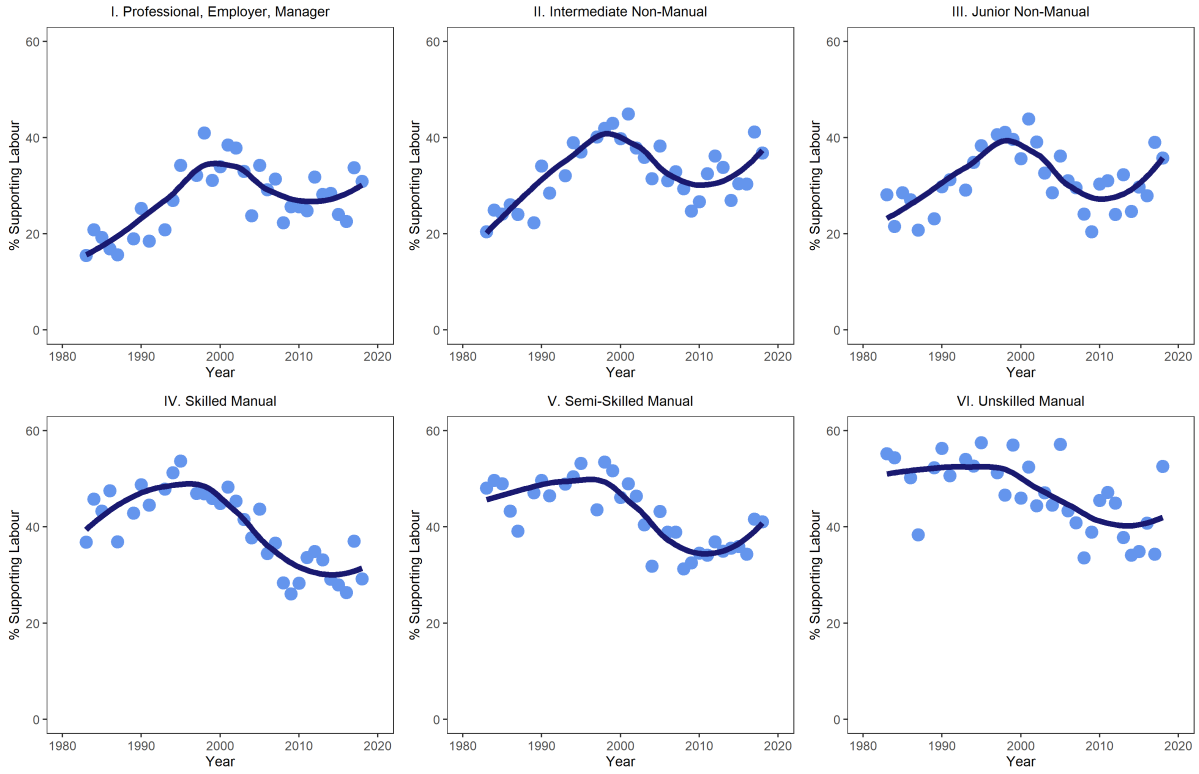
In attempting to replicate those figures, I first looked at the authors' code used to clean the data. Since the code was in excess of one thousand lines in R, and since the original data was very difficult to clean, being from an official government source, I decided to use that existing code to clean the data, although I would still create the table and graph by myself. However, after changing all of the directory references to ensure compatibility with my own setup, the code did not run successfully. I encountered many errors that started near the beginning of the script, and upon further examination, it was clear that the data itself had changed from when the authors last downloaded it. One small example is that the authors' script referred to a main folder that had a name corresponding to STATA Version 11; the data that I downloaded (which I ensured was correct because it was the same data ID number assigned by the UK government website) had a main folder with a name corresponding to STATA Version 13. Consequently, I expected differences in variable names and formats between the authors' script and the actual data available, and decided to omit those three figures.

Furthermore, the three figures only provided a measure of the political costs of urban housing unaffordability, and did not provide much insight beyond the existing qualitative analysis provided by the authors. Specifically, Table 1 showed a regression with the probability of party disaffiliation as the outcome variable, and various unfortunate events that could happen to a person (e.g. eviction, divorce, health problems, old age, unemployment); there was a statistically significant, positive relationship between eviction and disaffiliation. While the size of this relationship was small, the authors do note that even small changes can have large electoral impacts due to low voter participation. Thus, they followed up with Figures 5 and 6, where different manipulations (such as adding individual and year fixed effects) were displayed, showing that although the impact can be significant and is larger for Labour supporters than other non-Labour evictees, the effect is short-lived. The overall conclusion is that while the political costs of curtailing affordable (or council) housing may be high in the short term, the long term impacts are significantly less drastic.

After those three skipped elements, I moved on to Figure 7, which displays different socioeconomic groups' support for the Labour Party over time. For this figure (and for Table 2 and Figure 8), I used another dataset that was unavailable from Dataverse: the British Social Attitudes Survey (BSAS). Again, the data was only available for download from the UK Data Service website after registration and verification, but unlike the BPHS data, I was able to successfully run the authors' cleaning script after changing the directories to accommodate my own setup. Working off of the initial cleaned data, I performed further steps to manipulate the data and make it ready for conversion into graphs. Using the same main dataset, I subsetted it into six different socioeconomic groups based on the type of labor that individuals conduct, ranging from the highest-income category (professionals, managers, employers) to the lowest-income (unskilled manual laborers). Here, I was able to produce a panel of graphs that were identical to the original paper, based on a visual inspection. These graphs show that over time, and particularly in the past two decades, higher-income individuals have increased their support for the Labour Party, while their poorer

peers have decreased their support. This is consistent with the authors' hypothesis: that Labour is increasingly attracting voters who have not traditionally formed a part of its base, likely at the expense of alienating some of its historically loyal supporters.

Figure 7: Support for the Labour Party Over Time, by Socioeconomic Group (1983-2018)



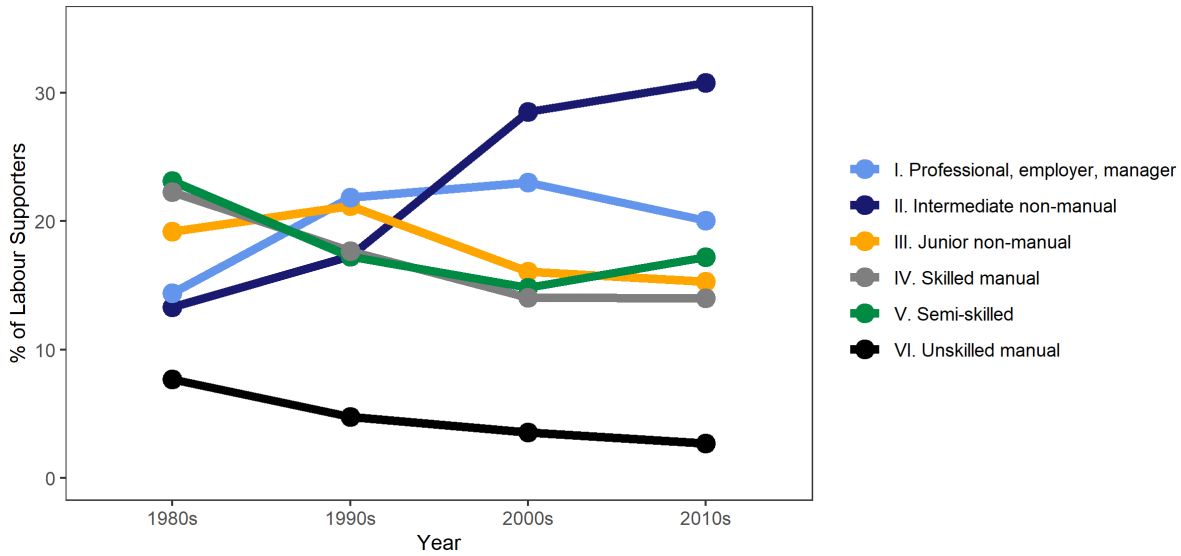
Then, I turned to Table 2 in the paper (listed as Table 1 due to R Markdown's automatic numbering), which numerically described the changing composition of socioeconomic classes in the London area and beyond. Since the dataset was still from BSAS, the data cleaning was only a minor step, and I was able to find the same numerical results as the authors did. After cleaning and compiling the data into a usable format, I used the *stargazer* package (Hlavac 2018) to generate a neat-looking table. As shown on the next page, the proportion of highly-skilled (i.e. professional/manager/employer and intermediate non-manual employees) voters has increased considerably over the past several decades in London. This supports the authors' argument: since demographics have been shifting in favor of higher-skilled people, the Labour Party will naturally attempt to secure the votes of those newcomers in order to maintain power - which would explain the anti-council housing and pro-gentrification stances that the party has taken.

Table 1: Class Composition in London and Outside London Over Time (1983-2018)

		1980s	1990s	2000s	2010s
I. Professional, employer, manager	London	20	25	22	22
I. Professional, employer, manager	Outside London	16	21	18	20
II. Intermediate non-manual	London	14	17	29	30
II. Intermediate non-manual	Outside London	13	13	22	25
III. Junior non-manual	London	24	22	19	16
III. Junior non-manual	Outside London	21	21	17	15
IV. Skilled manual	London	19	17	14	14
IV. Skilled manual	Outside London	23	20	20	18
V. Semi-skilled	London	17	14	14	16
V. Semi-skilled	Outside London	20	18	17	18
VI. Unskilled manual	London	5	4	3	2
VI. Unskilled manual	Outside London	6	7	5	4

Subsequently, I graphed Figure 8, which displays the type of electoral support that the Labour Party has received in London over time. This shows a similar trend as Table 1, since it is apparent that higher-income groups (particularly the intermediate non-manual employees) have significantly increased their support for Labour, while the lower-income groups have all reduced their support relative to the 1980s and 1990s. Consequently, the authors' argument is further reinforced: not only does the Labour Party face demographic-based incentives to appeal to voters outside its traditional working-class base (i.e. due to the significant growth in the size of the urban professional class), the Party also has proven successful in actually appealing to such voters. Again, the data cleaning process was not particularly arduous because I had adopted the authors' script to remove most of the unnecessary or irrelevant information from the original UK government dataset.

Figure 8: Labour's Coalition in London Over Time by Socioeconomic Group, 1983-2018



Quantitative Component

Having concluded the data replication for the so-called qualitative component of the paper, I move on to the actual quantitative part. Here, the authors start with a series of three related multivariate OLS regressions, with the 2011 council housing concentration (which is a percentage) used as the outcome variable. As previously described, the basic explanatory variables are the 2001 council housing concentration, the 2006 (midway point) Labour electoral strength, and the log-transformed crime rate and median income (these two variables were then standardized to have mean 0 and standard deviation 1). The first regression only had these variables, while the second added two interaction terms: between Labour strength and both the crime rate and median income. The third regression added further control measures, such as the characteristics of the council housing units in 2001, as well as housing demand and demographic indicators. All three regressions also used borough fixed effects to account for borough-wide differences that could arise between individual wards, which were the unit of analysis (and there were 624 of them). The dataset is now directly from the authors (i.e. through Dataverse), so I conducted a lengthy cleaning process myself and ran the regressions.

The results are shown in Table 2 on the next page (this is marked as Table 3 in the original paper; I was unable to manipulate R Markdown to bypass the automatic table numbering). As displayed, the 2001 concentration levels were statistically significant throughout all three regressions, and while the normalized, log-transformed crime rate and median income variables were initially statistically significant in the first regression, only their interaction terms (with Labour strength) remained statistically significant in the third regression, where many additional control variables were added. In addition, all three regressions involved a relatively large $N = 624$ observations, and has adjusted R^2 values close to 0.90, which demonstrates that while it may be simple, the multivariate OLS approach seems to make sense for the available data.

During the replication process, I was able to generate the correct (i.e. same) predicted coefficients relative to the original paper. All three regressions were simple LM functions, although I did have to standardize the log-transformed crime rate and median income by hand before performing the regressions. However, I did not get the correct standard errors at first. After realizing that the standard errors are clustered by borough, I made the necessary adjustments and got the correct results. This choice makes sense given the context: the original standard errors assumed homoskedasticity, but since local housing decisions are made at both the individual ward (i.e. unit of analysis) and the borough (i.e. higher than the unit of analysis) levels, it makes sense that the errors may depend on the specific borough that a ward is in.

Table 2: Correlates of Council Housing Reduction in Wards

	2011 Council Hsg. Concentration		
	(1)	(2)	(3)
Gentrification and Residualization Measures			
Council Housing Concentration	0.758***	0.759***	0.649***
	(0.028)	(0.029)	(0.066)
Labour Strength (2006)	0.081	0.024	0.061
	(0.162)	(0.156)	(0.179)
Log Crime Rate (SD)	−0.491***	−0.349**	−0.083
	(0.142)	(0.139)	(0.121)
Log Median Income (SD)	−0.758**	−0.558*	0.957
	(0.330)	(0.312)	(0.773)
Labour Strength x Log Crime Rate (SD)		−0.170	−0.290***
		(0.111)	(0.092)
Labour Strength x Log Median Income (SD)		−0.231	−0.492***
		(0.171)	(0.168)
Pragmatic Privatization			
% Council Homes Without Heat			−0.026
			(0.018)
% Council Homes With Own Shower			0.044*
			(0.025)
% Council Homes in Flats			−0.003
			(0.007)
% Council Homes Overcrowded			0.295
			(0.198)
Additional Covariates			
Borough Fixed Effects	✓	✓	✓
Housing Demand Indicators			✓
Demographic Indicators			✓
N	624	624	624
R^2	0.887	0.888	0.902
Adjusted R^2	0.880	0.881	0.892

*p < .1; **p < .05; ***p < .01

The second set of regressions is summarized in Table 3 (listed as Table 4 in the original paper). All six models listed are still multivariate OLS, but the key difference is that instead of considering all 624 wards in the greater London area, the models are based only on the 127 wards for which the Labour Party controlled the (higher-level) borough in both 2002 and 2006. I used the same dataset as Table 3, and only had to subset for Labour-controlled boroughs and weak-versus-strong Labour wards, both as defined by the authors.

The outcome variable is the same as before: the concentration of council housing in 2011. The main explanatory variables are also identical, with the exception that the 2006 level of Labour support is no longer being used - thus also eliminating the interactions from the second model above. Instead, the level of Labour support is part of the division between different individual regressions: all wards, weak Labour wards, and strong Labour wards. Again, as expected, the 2001 concentration levels were statistically significant throughout all six regressions; these represent a baseline against which the outcome variable can be compared. The only other interesting results are that the standardized and log-transformed crime rate and median income variables had statistically significant, negative, and large-magnitude estimated coefficients only in the "strong Labour ward" regressions. These also happened to have the highest R^2 values, and the data sample of $N = 74$ is large enough, so those two models imply that greater existing crime rates (which deter higher-income residents from moving in) or greater existing median income (which implies an existing population of higher-income residents) are predicted to lead to a greater reduction in the amount of social housing. This is consistent with the authors' main argument: that the Labour Party is prioritizing higher-income residents by implementing their policy preferences of fewer publicly-sponsored housing units in their areas.

These regressions are also intended to test an alternative explanation to the findings of the first set of regressions: that non-Labour borough leaders (who operate at a higher level of government than ward leaders) were looking to undermine Labour by reducing council housing in Labour-dominated wards, thus reducing the number of Labour votes. However, by limiting the data analysis to only the Labour-controlled boroughs, and by differentiating between wards that are weakly versus strongly controlled by Labour, that alternative explanation can be addressed. Indeed, the results show that the effect is negative and statistically significant in only the "strong Labour wards," which directly contradicts that alternative explanation and suggests that if anything, it is Labour itself that is reducing council housing.

Table 3: Correlates of Council Housing Reduction in Wards

	2011 Council Hsg. Concentration					
	(1)	(2)	(3)	(4)	(5)	(6)
Gentrification and Residualization Measures						
Council Housing Concentration	0.685*** (0.030)	0.698*** (0.075)	0.677*** (0.038)	0.700*** (0.130)	0.760*** (0.192)	0.518*** (0.134)
Log Crime Rate *Standardized*	−0.656 (0.657)	1.611*** (0.356)	−1.667*** (0.418)	−1.211** (0.527)	0.389 (1.638)	−1.008** (0.496)
Log Median Income *Standardized*	−0.729 (0.632)	0.980 (1.129)	−2.872*** (0.500)	0.753 (1.518)	3.602 (2.264)	−4.203** (1.722)
Pragmatic Privatization						
% Council Homes *Without Heat*				−0.064 (0.078)	−0.312 (0.206)	−0.016 (0.112)
% Council Homes *With Own Showers*				0.240 (0.491)	1.436* (0.759)	−0.265 (0.278)
% Council Homes *In Flats*				0.067*** (0.022)	−0.028 (0.056)	0.058** (0.026)
% Council Homes *Overcrowded*				0.122 (0.380)	−0.800 (0.602)	0.484 (0.435)
Additional Covariates						
Borough Fixed Effects	✓	✓	✓	✓	✓	✓
Hsg. Demand Indicators				✓	✓	✓
Demographic Indicators				✓	✓	✓
N	127	53	74	127	53	74
R^2	0.838	0.754	0.911	0.875	0.890	0.941
Adjusted R^2	0.825	0.703	0.898	0.843	0.780	0.908

*p < .1; **p < .05; ***p < .01

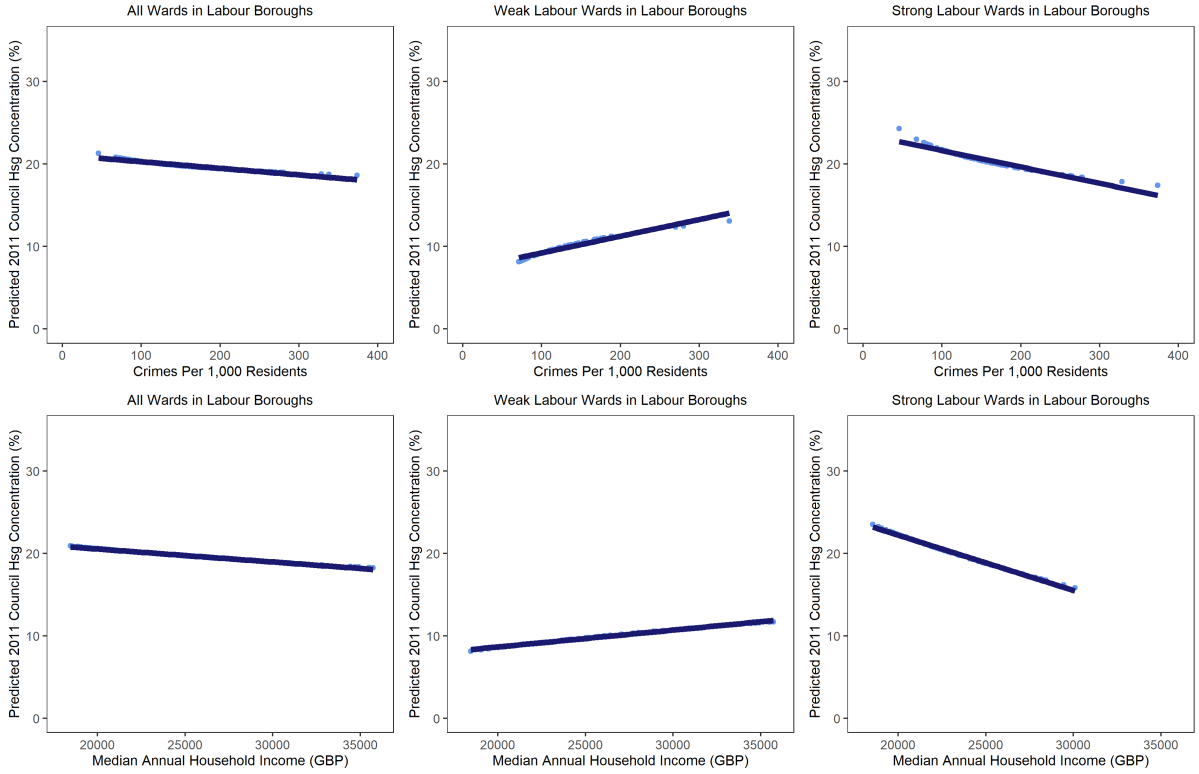
Finally, I attempted to replicate Figure 9 from the original paper. These consisted of 6 different graphs: three with the crime rate on the x-axis, and three with the median income on the x-axis, while all six had the 2011 housing concentration on the y-axis. These graphs were based on the three categories established in Table 3: all wards, weak Labour wards, and strong Labour wards, with each category receiving one plot.

The graphs themselves demonstrate predictions for the level of council housing in 2011 based on various actual values of the variable on the x-axis, when all the other variables would be held at the average value. For instance, in the upper-left plot, predictions were generated for the level of council housing in 2011 based on all the actual values of the 2001 unmodified crime rate, while median income and 2001 housing concentration was held constant at the average level of all wards. Since borough fixed effects were used, I also set the "average borough" as the borough that had the largest number of wards/observations in the dataset. In other words, the graphs display the model predictions for changing one variable when all other variables are held constant, similar to how one would interpret a standard multivariate OLS regression output.

As shown on the next page, the predicted "interaction" between both crime and median income and the outcome variable of 2011 housing was most significant for the "strong Labour wards"; the slope was most negative. Meanwhile, the "all wards" interaction was largely flat, and the "weak Labour wards" actually show a positive trend, which would indicate that Labour leaders in such wards are less willing to remove council housing, perhaps under the fear that they would be easily removed from office. Thus, Figure 9 displays similar information to the numerical estimates of Table 3.

I faced several major challenges in replicating Figure 9. First, I was unable to understand what the small bars at the bottom of the original graphs represented; an understanding might have provided more insight into the authors' approach and allowed me to replicate more successfully. I was also confused by the prediction ranges: for instance, in the "strong Labour wards," the income levels simply never reached 36,000 GBP in any of the data, but because the authors provide a line with a 95% CI reaching to that amount of median income, they must have used a different approach that incorporated fictional data as well as real data. In contrast, my approach was as follows: starting with the original data used to generate the three models, I converted the crime rate and median income variables back to their nominal values, and then generated the averages for all the variables that were not being manipulated in the graph (i.e. all the input variables except the one to be displayed on the x-axis). I used the actual, observed values of the variable in question (either crime or income) to generate predictions. Meanwhile, for borough, I know the authors used one borough for all three plots, but after evaluating the number of wards by borough, I found that a different ward - Tower Hamlets - should be used for the "weak Labour wards" instead of Newham, which the authors used for all three (and which I used for just the other two); tis would explain differences between the authors' two middle graphs and my two middle graphs. Finally, while my method did not allow me to plot the confidence intervals that the authors did, I did plot the prediction points in a light blue color; the fact that very few are visible means that the vast majority of them fall very closely to the regression line.

Figure 9: Interactions Between Labour Strength, Crime, and Median Income in Labour Boroughs



Interpretation and Summary of Replication

As extensively described above, the qualitative part of the paper provides the rationale and the background context needed to make the assumptions that the authors make in the quantitative part of the paper. Figures 1 and 2 show that public investment in housing has declined, even as prices have been increasing and as more people become renters instead of homeowners. Since poverty is still a persistent problem, the authors are able to prove a basic assumption: that housing affordability is an issue to be concerned about. Meanwhile, Figures 3 and 4 are more tailored toward the authors' argument about the Labour Party, since those two point to a more drastic decline in public/council housing and to the notion that Labour politicians could be gradually abandoning their traditionally working-class constituencies. The figures that I omitted provide further insight into the potential political costs of the Labour Party doing such a thing. Finally, Figures 7 and 8 - along with Table 2 - provide concrete evidence that not only are the right incentives present for Labour to attract higher-income voters, but also that in fact, Labour has been successfully attracting a socioeconomically-diverse coalition of voters over the past two decades.

This leads to the quantitative component of the research. The OLS regressions in Table 3 are relatively simple to interpret. In two of the three models, an increase of one percentage point (i.e. one unit) in the 2001 council housing concentration is predicted, on average, to result in a 0.76 percentage point (i.e. 0.76 units) increase in the 2011 council housing concentration; in the third and more complex model, the effect is smaller at 0.65. The log-transformed (and then standardized with mean 0, sd 1) crime-rate and median-income variables were not statistically significant (or only borderline significant) for the second and third models; they were for the first model, where an increase of one standard deviation those variables are predicted, on average, to result

in a 0.49 (and 0.76, respectively) percentage point decrease in the 2011 council housing concentration. Labour strength by itself was not a statistically significant variable in any model, but the interactions between Labour strength and both the crime-rate and median-income variables were. Since those coefficients were negative, it indicates that the decline in public housing is linked to greater existing crime and higher levels of existing middle-class residents only when the Labour Party was dominant.

These results prove the authors' original hypothesis. For instance, the crime rate and median income had statistically significant, negative coefficients by themselves in the original (simple) model, but this seems to only show the basis of gentrification: that public housing is perhaps associated with higher crime, providing political incentives to reduce the availability of such housing and thereby attract middle-class newcomers - particularly if there are higher-income residents already living in the area. However, the fact that these two variables remain statistically significant only as part of an interaction with Labour strength when more control variables are added (i.e. in the third model) shows that greater electoral support for the Labour Party is associated with greater reductions in housing. Thus, it is apparent that despite its working-class roots, Labour has become a driving factor behind gentrification.

The OLS regressions in Table 4 are also relatively simple to interpret, since they are based on the same outcome variable and a very similar set of explanatory variables. As with before, the 2001 council housing concentration remains the most statistically-significant variable across all six regressions; five of the six models had a coefficient between 0.76 and 0.80, while the remaining one had a coefficient of 0.52. These mean that an increase of one percentage point (one unit) in the 2001 housing concentration is predicted, on average, to result in a 0.76 to 0.80 (or just 0.52) percentage point increase in the 2011 concentration. Again, this makes intuitive sense because higher initial levels of housing should lead to higher levels of housing at the end of the ten-year-long period. Also, the crime-rate and median-income variables had two statistically-significant coefficients scattered across the "all wards" and "weak Labour wards" models, but they were statistically significant with large magnitudes for both of the "strong Labour wards." Since those estimated coefficients were negative, they have the same implication as before: that higher levels of support for the Labour Party - and thus more politicians in power - are associated with a greater reduction in the amount of public housing at the local level. Specifically, an increase by one standard deviation in the standardized, log-transformed crime rate is predicted to reduce the concentration of 2011 council housing by 1.67 percentage points in the simple model and by 1.01 percentage points in the more complex model. The corresponding values for the standardized, log-transformed median income are 2.87 percentage points for the simple model and 4.20 percentage points for the complex model. Here, it is also important to note that outside of the "strong Labour" wards, the crime-rate and median-income variables generally have positive coefficients (although not statistically significant), which would imply that other wards with higher crime rates and/or higher median income were able to keep more of the existing council housing stock.

To summarize, the quantitative analysis predicts that there will be less social housing in wards with higher existing levels of crime and median income, and that this relationship holds only when the Labour Party has a strong electoral base. In other words, the Labour Party is associated with more reductions in the availability of social housing, and because of the demographics analysis done in the qualitative section, it is clear that these reductions are fueled by a desire to accommodate the policy desires of higher-income residents. Applying this more broadly, left-leaning parties that are seeking to broaden their support base are predicted to actually contribute to - not deter from - the trends of gentrification and housing displacement, thus defying the com-

mon perception that they are supporters of the working class.

However, part of the issue is that lower-income voters are very loyal to left-leaning parties. This helps reduce the political risk faced by left-leaning politicians when they decide to implement policies that appeal to higher-income voters, meaning that the loyalty actually serves *against* the self-interest of the working class. Consequently, it appears that a greater willingness to switch parties might actually be beneficial: by raising the political risk of implementing pro-gentrification policies, lower-income voters may be able to persuade politicians against such actions. This has obvious implications in democracies around the world, since it suggests that among other things, traditionally non-working-class parties and political groups may be able to gain new supporters by adopting certain policies aimed at reducing the impacts of gentrification.

Limitations

Nonetheless, the study is associated with several limitations. In particular, the authors rely on social housing units as a proxy for whether politicians do or do not support the availability of affordable housing; in conjunction, they make an assumption that reductions in social housing are necessarily pro-gentrification. This limits the broader applicability of the paper because publicly-sponsored housing is only one of many forms of affordable housing. For instance, older and more dilapidated (but still within legal limits) units can be affordable for lower-income urban residents, while large-scale new development can also create affordable homes if labor, material, and land costs are low enough. Thus, reducing public housing is only one factor contributing to gentrification, and studying it cannot reasonably provide information about other political mechanisms of dealing with gentrification.

Another major limitation is that the British governing system differs from that in other countries. While I am not personally familiar with British governance, the authors seem to suggest that housing decisions are made at the local (ward and borough) levels in England, thus making locally-elected politicians relevant for consideration. However, in the United States, the federal government plays a significantly larger role in public housing decisions, particularly with regard to funding. This means that local leaders have less say about increasing or decreasing the number of available units, making the study a good start for this line of research, but not necessarily a model to directly replicate in other parts of the world.

Finally, the paper is limited by its use of a relatively small dataset. While $N = 624$ (for Table 2) and $N = 127$ (for Table 3) are considerably greater than the $N = 30$ typically required to prove statistical significance, it is unfortunate that only one time period, 2001 to 2011, is being measured. There are no historical comparisons, and while I understand the need to compare a relatively long period of time (e.g. a decade) in order to actually observe the effects of political decisions that have slow impacts, it is frustrating that there is no data for intermediate years, such as 2006, or for subsequent years, such as 2016. The limited time scale of the analysis again limits the study's applicability to other situations.

Extensions

Log-Transformed Housing Concentrations, 2001 and 2011

With those important implications in mind, I turn to my first extension of the model. I was dissatisfied by the fact that the authors used the absolute levels of the social housing concentrations in both 2001 and 2011: while these were already expressed in percentages (i.e. of the entire housing stock in a ward) rather than absolute numbers, I do not see why there is an assumption that the impact of the other explanatory variables, such as crime rate and median income, would be the same across all preexisting levels of housing. Since the 2001 concentration is used as a control variable to partly mitigate the problem from a statistical standpoint, the issue lies more with the interpretation: it predicts changes in absolute terms only. While such changes can be interpreted in the context of the mean values of the outcome variable (2011 concentration), which are indeed listed on the paper's output tables, this adds a layer of additional complexity and may be less straightforward or clear to political scientists who are less skilled with data.

Consequently, I decided to perform a logarithmic transformation on the social housing concentration variables for both 2001 and 2011. Now, the outcome variable can be interpreted as a "percentage change" in the housing concentration percentage-point value, which makes slightly more sense because by nature, it is a relative measurement of change instead of an absolute one. Using the log-transformed variables, I ran the same type of regressions as Table 2, and produced the regression results listed on the next page as Table 4.

As shown, the 2001 council housing concentration (now log-transformed) remains an important variable: a one percent increase in the percentage-point value of the 2001 council housing concentration is associated with a 0.977, 0.978, or 0.963 percent increase in the percentage-point value of the 2011 council housing concentration, depending on the specific model. This is much closer to a one-to-one ratio than Table 2, where the largest coefficient values was only 0.76.

Unsurprisingly, the increase in the magnitude of the 2001 variable meant a decrease in the magnitude of the other variables. In the simple model 1, an increase of one standard deviation in the standardized, log-transformed crime rate or median income are predicted to result in a 0.024 or 0.061 decrease in the percentage-point value of the 2011 council housing concentration. These coefficients are statistically significant (even more so than Table 2), but compared to Table 2, they are much smaller in magnitude relative to the magnitude of the 2001 concentration coefficient. Similarly, in the complex model 3, the interaction coefficients between Labour strength and the two standardized-and-log-transformed variables are now much smaller in magnitude, and one of them is no longer statistically significant.

These results are somewhat concerning. The loss of statistical significance in the third model, coupled with the small magnitude of the interaction coefficients (even in comparison to the non-interaction crime and income coefficients), mean that the authors' original points about Labour contributing to a decline in council housing are no longer proven by the data. Instead, the statistical significance of the first model points to just the crime rate and median income as the driving factors of gentrification, regardless of party. Furthermore, all three of these models fit the data in a slightly better way, with increase of approximately 0.03 in the R^2 value from 0.88 to 0.91; while this does not carry profound implications, it is yet another small bit of evidence that can be combined with my previous findings about the loss of statistical significance.

Table 4: Correlates of Council Housing Reduction in Wards, Log-Transformed

	Log of 2011 Council Hsg. Concentration		
	(1)	(2)	(3)
Gentrification and Residualization Measures			
Log of Council Housing Concentration	0.977***	0.978***	0.963***
	(0.026)	(0.026)	(0.027)
Labour Strength (2006)	0.011	0.011	0.017
	(0.011)	(0.011)	(0.011)
Log Crime Rate (SD)	−0.024***	−0.018	0.006
	(0.009)	(0.011)	(0.012)
Log Median Income (SD)	−0.061***	−0.057*	0.097*
	(0.023)	(0.030)	(0.053)
Labour Strength x Log Crime Rate (SD)		−0.008	−0.014**
		(0.006)	(0.006)
Labour Strength x Log Median Income (SD)		−0.003	−0.016
		(0.011)	(0.013)
Pragmatic Privatization			
% Council Homes Without Heat			−0.002
			(0.002)
% Council Homes With Own Shower			0.005
			(0.003)
% Council Homes in Flats			−0.0002
			(0.001)
% Council Homes Overcrowded			−0.002
			(0.006)
Additional Covariates			
Borough Fixed Effects	✓	✓	✓
Housing Demand Indicators			✓
Demographic Indicators			✓
N	624	624	624
R^2	0.917	0.917	0.922
Adjusted R^2	0.912	0.912	0.915

*p < .1; **p < .05; ***p < .01

Naturally, I turned to Table 4 next, to see if there is any evidence of the authors' original relationship still holding when there is a clearer distinction between strong and weak Labour wards. The same logarithmic transformations were applied to the 2001 and 2011 concentrations of social housing, and I ran the same models as Table 4. The results are displayed in Table 5 on the next page.

As shown, there continues to be a strong relationship between the 2001 and 2011 levels of social housing concentration, even after they were log transformed. All six of these coefficients are statistically significant, with four being slightly above 1 (meaning that a 1 percent increase in the 2001 council housing percentage-point concentration is predicted to result in more than a 1 percent increase in the 2011 percentage-point concentration), and the remaining two being at approximately 0.80 and 0.90, respectively. Again, the large magnitudes imply that the practical impact of the other variables will be much smaller.

With regard to the concerning results found in Table 4 relative to Table 2, I find similar concerning results in this table. The authors originally claimed that because the crime-rate and median-income variables had statistically significant, large magnitude estimated coefficients in both of the "strong Labour" regressions (models 3 and 6), there is then a pattern where Labour politicians are actively reducing the availability of public housing. However, with the log transformations, the magnitudes have decreased significantly (at least relative to the magnitude of the 2001 variable), and the coefficients are only statistically significant for model 3 (the simpler model); once further controls are added in model 6, those coefficients lose their statistical significance.

Thus, I am not certain if the statistical relationship that Chou and Dancygier rely upon necessarily holds true. Without the certainty about the relationship between Labour and the reduction in council housing, this would jeopardize many of the findings and implications of the paper. Specifically, in the absence of that relationship, the data presented in the original paper would only serve to confirm theories about gentrification, in the sense that the findings show a connection between both crime and median income and changes in the amount of public housing.

Table 5: Correlates of Council Housing Reduction in Wards

	Log of 2011 Council Hsg. Concentration					
	(1)	(2)	(3)	(4)	(5)	(6)
Gentrification and Residualization Measures						
Log Council Hsg. Concentration	1.057***	1.103***	1.017***	1.027***	0.793***	0.908***
	(0.019)	(0.050)	(0.050)	(0.080)	(0.132)	(0.061)
Log Crime Rate *Standardized*	−0.033	0.071*	−0.071**	−0.032	−0.065	−0.016
	(0.039)	(0.041)	(0.031)	(0.022)	(0.046)	(0.027)
Log Median Income *Standardized*	−0.020	0.019	−0.091**	0.017	0.302***	−0.259*
	(0.032)	(0.077)	(0.045)	(0.062)	(0.074)	(0.147)
Pragmatic Privatization						
% Council Homes *Without Heat*				−0.006*	−0.003	−0.008**
				(0.003)	(0.011)	(0.004)
% Council Homes *With Own Showers*				−0.006	0.111***	−0.039***
				(0.023)	(0.040)	(0.014)
% Council Homes *In Flats*				0.004**	−0.004**	0.004**
				(0.002)	(0.002)	(0.002)
% Council Homes *Overcrowded*				0.007	0.029	0.012**
				(0.010)	(0.020)	(0.005)
Additional Covariates						
Borough Fixed Effects	✓	✓	✓	✓	✓	✓
Hsg. Demand Indicators				✓	✓	✓
Demographic Indicators				✓	✓	✓
N	127	53	74	127	53	74
R^2	0.860	0.794	0.919	0.901	0.952	0.958
Adjusted R^2	0.850	0.751	0.907	0.875	0.903	0.935

*p < .1; **p < .05; ***p < .01

Proposed Extension: 2011 to 2021

In addition to performing new regressions based on log-transformed variables, I propose two additional hypothetical extensions that could enhance the findings of the original paper. The first involves extending the time scope of the study. As previously discussed, a major limitation is the decade-long nature of the data being used, which is a specific and a relatively short period of time.

Consequently, I propose collecting data about council housing concentrations at the end of this year, 2021. Since the housing concentration data was already collected for the beginning of the decade, 2011, all that remains is collecting the statistics from 2011, such as the crime rate, median income, condition of the housing, and the level of Labour support (for 2016). These data should be relatively easy to compile, since they appear to be statistics that local governments already collect and publish on a regular basis. With the new data, another set of regressions can be run, and the two sets of regressions can be compared to each other to see if there are any notable changes across the two decades. If the newer data shows similar patterns, that would reaffirm the validity of the authors' findings.

Proposed Extension: U.S. Democratic Party and Public Housing

The second hypothetical extension is extending the geographic scope beyond the London area that was covered in the study. Specifically, the United States - which is facing similar issues with gentrification in urban neighorhoods - would be a good candidate, because the policy implications, which include contentious issues surrounding race, are slightly more important.

I propose collecting data about the Democratic Party, the equivalent major left-leaning party, and its support within areas where gentrification is occurring. This approach is motivated by real-world cases: for instance, Cambridge and Somerville, Massachusetts are both heavily left-leaning cities, but the cost of housing has skyrocketed despite left-leaning policies and politicians. While public housing in the U.S. is shaped heavily by federal policy and is out of local leaders' control, a replacement for the level of housing would be the average or median real/inflation-adjusted rents paid by tenants in all housing, public and private. This would be a better proxy for affordability, since local leaders do control new-construction restrictions.

With the new data, separate but similar regressions can be run. If similar trends are found in the United States, it would again reaffirm the validity of the authors' findings, and would significantly broaden the applicability. Such a successful finding is also likely to generate much more research along this line of thinking, as scholars attempt to find other cities and metropolitan areas where such a pattern can be observed. Thus, this proposed extension also serves as a potential direction for future research.

Additional Notes

This is a listing of the tables and graphs displayed in this report:

- Figure 1: Real House Prices versus Public Investment in Housing in OECD Countries
- Figure 2: Housing Trends in London and England
- Figure 3: Decline of Publicly-Owned Council Housing in Greater London
- Figure 4: Number of *Guardian* Articles Mentioning "Social Cleansing" in Relation to Housing
- Figure 7: Support for the Labour Party Over Time, By Socioeconomic Group (1983-2018)
- Figure 8: Labour's Coalition in London Over Time, By Socioeconomic Group (1983-2018)
- Figure 9: Interactions Between Labour Strength, Crime, and Median Income in Labour Boroughs
- Table 1: Class Composition in London and Outside London Over Time (1983-2018), *listed as Table 2 in the original paper*
- Table 2: Correlates of Council Housing Reduction in Wards, *listed as Table 3 in the original paper*
- Table 3: Interactions with Labour Strength within Labour Boroughs, *listed as Table 4 in the original paper*
- Table 4: Correlates of Council Housing Reduction in Wards, *with log-transformed variables*
- Table 5: Interactions with Labour Strength within Labour Boroughs, *with log-transformed variables*

This replication was completed using the following software:

- R version 4.0.3 (2020-10-10) – "Bunny-Wunnies Freak Out"
- Platform: x86_64-w64-mingw32/x64 (64-bit)
- Running under: Windows 10 Home, version 20H2

This report was compiled using a template from Steven V. Miller (2016).

Bibliography

Chou, Winston, and Rafaela Dancygier. 2021. "Why Parties Displace Their Voters: Gentrification, Coalitional Change, and the Demise of Public Housing." *American Political Science Review* 115 (2): 429–49. https://doi.org/10.1017/S0003055421000058.

Hackworth, Jason, and Neil Smith. 2001. "The Changing State of Gentrification." *Journal of Economic and Human Geography* 92 (4): 464–77. https://doi.org/10.1111/1467-9663.00172.

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer.

Lyons, Michal. 1996. "Gentrification, Socioeconomic Change, and the Geography of Displacement." *Journal of Urban Affairs* 18 (1): 39–62. https://doi.org/10.1111/j.1467-9906.1996.tb00365.x.

Miller, Steven V. 2016. "An r Markdown Template for Academic Manuscripts." http://svmiller.com/blog/2016/02/svm-r-markdown-manuscript/.

Smith, Neil. 1982. "Gentrification and Uneven Development." *Economic Geography* 58 (2): 139–55. https://doi.org/10.2307/143793.

Uitermark, Justus, and Jan Willem Duyvendak. 2007. "Gentrification as a Governmental Strategy: Social Control and Social Cohesion in Hoogvliet, Rotterdam." *Environment and Planning A* 39 (1): 125–41. https://doi.org/10.1068/a39142.