

Responding to Gentrification?

311 Call Resolution Times and Neighborhood Change in Washington, D.C.

Danielle Corcoran *

Jeremy Joy †

2022-12-14

Abstract

In this study, we explore the relationship between gentrification and government responsiveness by investigating whether gentrification had an impact on 311 call resolution times in Washington, D.C., between 2012 and 2019. We begin by identifying census tracts that gentrified during our study period, and then use a series of difference-in-difference models to determine whether resolution times improved more in those tracts than in similar tracts where gentrification did not occur. We use matching to pair gentrified tracts with the non-gentrified tracts that most closely resemble them in order to ensure the quality of our control group. We demonstrate that, although there is some evidence that gentrification had a mild effect on average resolution times when analyzing 15 of the most common service requests together, there is no statistical significance when controlling for call type. Further, we find that there is no evidence that gentrification makes a difference when looking at the two most common service requests — bulk collection and parking enforcement — in isolation. We discuss the limitations of our study and consider its implications for equity.

*American University

†American University

1 Introduction

The tendency of 311 calls to change in both frequency and nature as neighborhoods gentrify is well-documented, but less is known about whether a city’s responsiveness to these calls changes as well. In this paper, we examine the relationship between gentrification and resolution time for non-emergency calls in Washington, D.C. Whereas most academic studies and journalistic reports on gentrification and 311 calls have explored neighbor conflict and quality-of-life complaints resulting in cultural displacement and/or police exposure (see, for example, (Reza 2022), (Beck 2020), (NYU 2015), (Vo 2018), (Misra 2018), (Harris, Rigolon, and Fernandez 2020), and (Laniyonu 2018)), our study looks at a wider variety of calls, most of which are unlikely to engender a police response, and focuses on whether gentrification has led to either a more or less equitable distribution of city resources.

Starting with the premise that there may be a link between affluence and bureaucratic responsiveness, we investigate whether census tracts that gentrified during our study period saw a greater improvement in 311 call resolution times than similar census tracts that did not. This question is particularly important in Washington, D.C., which has been characterized as one of America’s fastest-gentrifying cities in recent years — a 2015 study found that nearly 52 percent of the District’s low-income census tracts had gentrified since the turn of the millennium, the second-highest percentage in the nation (Maciag 2015a). In the sections that follow, we provide a definition of gentrification and use it to identify D.C. census tracts where gentrification has occurred. We also explore trends in 311 call data and look for variation across gentrification statuses. We then employ a series of difference-in-difference models to test whether gentrification had a statistically significant impact.

2 Defining Gentrification

Gentrification usually refers to the socioeconomic and demographic changes that occur when relatively disadvantaged neighborhoods experience an influx of more privileged residents. While there is no universal definition in the academic literature, quantitative measures of gentrification often rely upon a combination of variables, such as median household income, median home value, median rent, percent of housing stock less than 20 years old, home ownership rate, poverty rate, percent of adults over 25 with a Bachelor’s degree, race, and age (see (Bhavsar, Kumar, and Richman 2020) for metastudy). As described by Bhavsar, Kumar, and Richman (2020) and illustrated by Freeman (2005) and Maciag (2015b) many studies adopt a two-stage approach in classifying an area as gentrified or not. The first stage involves determining whether a neighborhood — usually a census tract — is eligible to gentrify, meaning that it is below a certain percentile with respect to the selected variables, when compared with the surrounding city or metro area (an area that is already affluent is not considered to have gentrification potential). The second stage involves determining whether the neighborhood has actually gentrified after a certain period of time, meaning that it has experienced relative growth that exceeds a certain percentile, when compared with the surrounding area.

Our definition uses median household income, median home value, and educational attainment as criteria for gentrification. Our cutoff for eligibility is below the 40th percentile, while the threshold for gentrification is relative growth above the 50th percentile, with economic indicators adjusted for inflation. This definition has both a theoretical and a practical basis: our criteria capture important aspects of gentrification and our eligibility cut-offs make for a rigorous test, while also resulting in a sufficient number of tracts classified as gentrified for us to proceed with the study. By our definition, there were 54 census tracts eligible for gentrification in 2012, out of a total of 179 census tracts in D.C. By 2019, 15 of these 54 had gentrified, while 39 remained ungentrified. Notably, most of the gentrification occurred in the District’s northeast quadrant, while most of eligible tracts that did not gentrify were east of the Anacostia River.

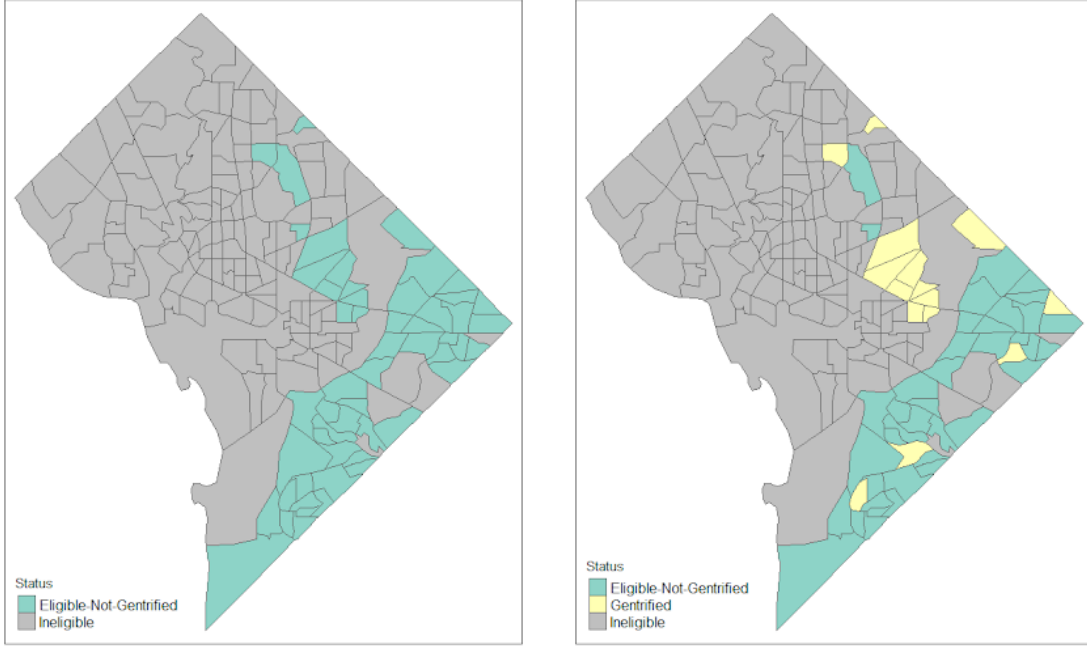


Figure 1: 2012 Census Tracts (left) and 2019 Census Tracts (right)

3 311 Call Data

We obtained data on 311 service requests in the District from Open Data DC. While data is available from 2009 through the present, we chose 2012 through 2019 as our study period. We selected 2012 because that was the year when D.C. started recording large call volumes; there were 250,000 calls or more each year from 2012 onwards, whereas the years prior to 2012 had fewer than 35,000 calls each. We selected 2019 as our end point because that was the last full year before the COVID-19 pandemic, which may have impacted both service requests and the indicators used to determine whether a neighborhood was gentrified. By using the widest time interval for which we were confident that we could make a fair comparison, we aimed to maximize the number of gentrified tracts in our study.

The 311 data contained the type of service requested, the location where the service was required, the date of the request, the date and time the request was received, and the date and time the request was resolved. We calculated resolution time by finding the difference between the date received data from the date resolved. Although there are more than 200 different service codes, some of which vary from year to year, we selected 15 of the 20 most common call types for this analysis, excluding calls that were more likely to have been placed by visitors to a tract than residents (e.g., parking meter repair) and service requests that did not require a physical response (e.g. D.C. government information). The selected call types and their combined frequencies in 2012 and 2019 are shown below in Figure 2.

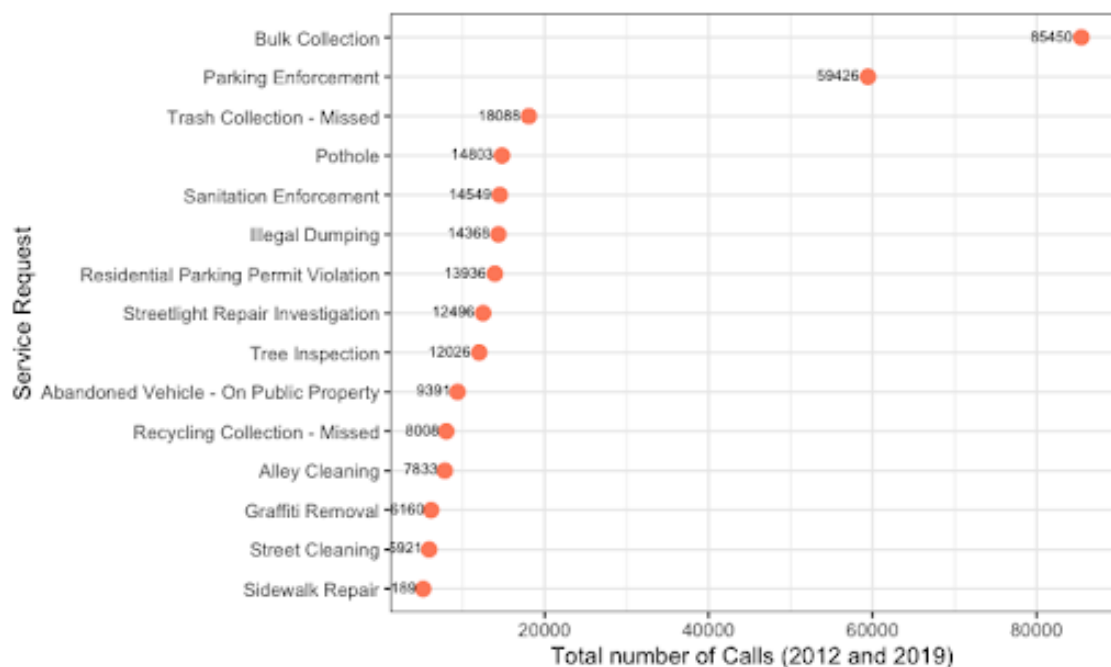


Figure 2: Selected Call Types and Frequencies (2012 and 2019)

Close inspection of resolution times across all selected call types revealed irregularities in the 2012 data — while most calls had resolution times of 32 days or less, there were more than 800 observations with resolution times of 500 days or more. Furthermore, there was a large gap, with no resolution times between 32 and 205 days (see Figure 3 (left)). Finally, many of the service requests were resolved on the same dates, suggesting that they were closed on an arbitrary timeline, rather than when the service was actually completed. Based on this discovery, we decided to filter out observations for which the resolution time was greater than 31 days in both 2012 and 2019. This decision led us to discard about 6.8% percent of our data, including all requests for sidewalk repair, and left us with about 70,000 observations for analysis. We are more confident that the remaining data are reliable, based on the distributions shown in the histograms in Figure 3 (right).

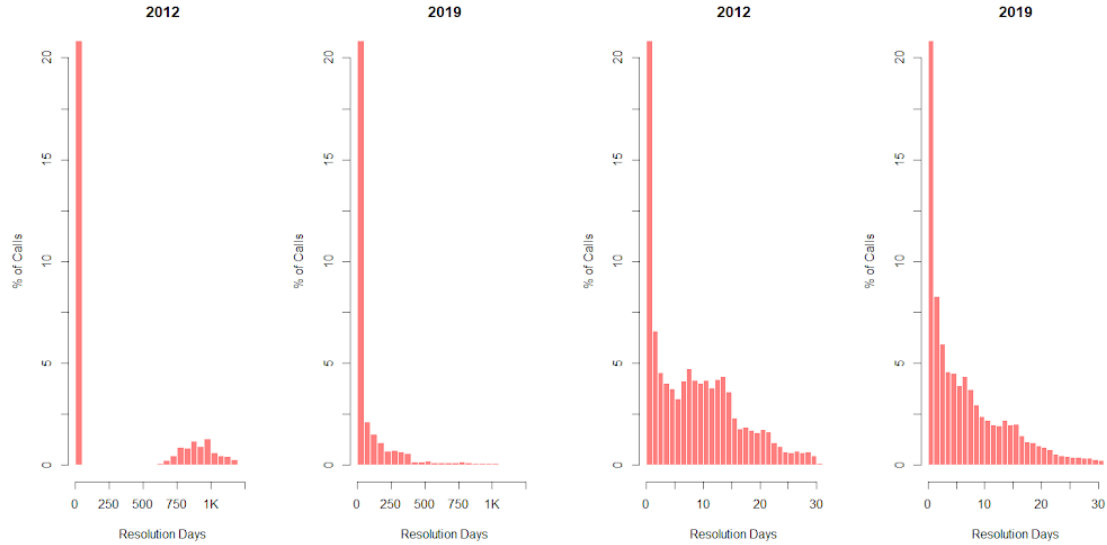


Figure 3: Before Filter (left) and After Filter (right)

4 Changes in Resolution Time by Gentrification Status

As Figure 4 illustrates, call frequency increased dramatically between 2012 and 2019, yet call resolution times improved. Whereas the call resolution times were markedly better for ineligible tracts than gentrified or eligible-not-gentrified tracts in 2012, resolution times seem to have been more equitable in 2019.

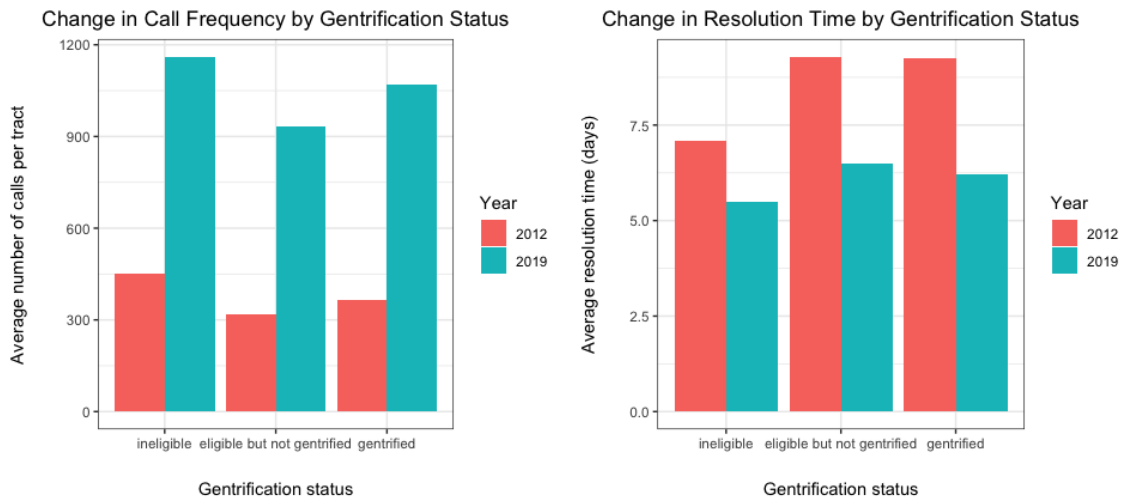


Figure 4: Change in Call Frequency (left) and Change in Resolution Time (right)

Below, Figure 5 shows that resolution times followed roughly the same trend across gentrification statuses for every year from 2012 to 2019. Ineligible tracts experienced a 1.57-day decrease in resolution time, eligible-but-not-gentrified tracts experienced a 2.80-day decrease in resolution time, and gentrified tracts showed a 3.02-day decrease in resolution time. That is to say, the resolution time for gentrified tracts improved the most. However, the ineligible tracts, which were the most affluent tracts as of 2012, started and ended with the best resolution times.

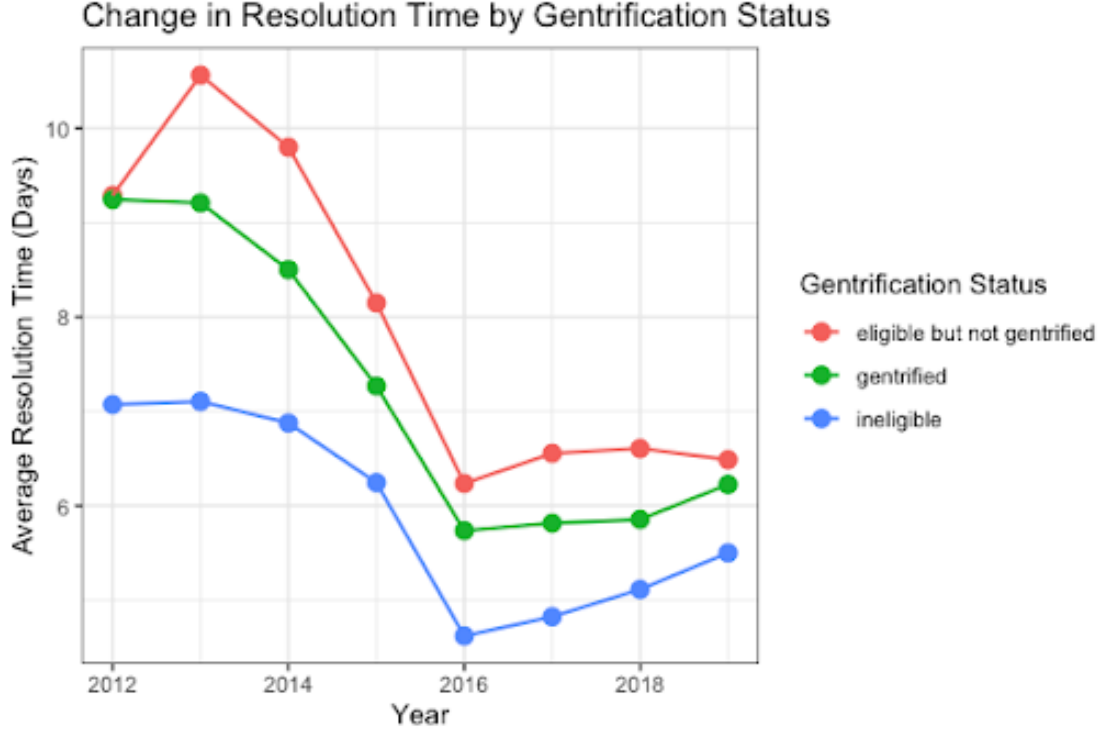


Figure 5: Change in Resolution Time by Status (2012-2019)

5 Difference-in-Difference Analysis

We used a series of difference-in-difference models to determine whether gentrification had a statistically significant effect on call resolution times. The plots below show the changes in resolution times from 2012 to 2019, for tracts that gentrified and tracts that were eligible to gentrify but did not. The counterfactual in each plot shows what would have happened in the gentrified tracts had gentrification not occurred (with the assumption that the trend in the gentrified tracts would have paralleled the trend in the eligible-but-not-gentrified tracts). The difference between the 2019 value for the gentrified tracts and the 2019 value for the counterfactual represents the effect of gentrification. (Data Science for Public Service 2019)

A difference-in-difference model requires a treatment group and a control group, as well as a pre-intervention condition and a post-intervention condition. Treating the gentrified tracts as the treatment group, the eligible-but-not-gentrified tracts as a control group, and gentrification itself as the intervention, we used regression to generate the following equation:

$$\text{Resolution Days}_i = \beta_0 + \beta_1(\text{Treatment}_i) + \beta_2(\text{Post}_i) + \beta_3(\text{Treatment:Post}_i) + \epsilon_i$$

where β_0 represents the intercept, β_1 represents the difference between the treatment and control group before gentrification, β_2 represents the difference in the control group before and after gentrification, and β_3 represents the difference-in-difference estimator (Data Science for Public Service 2019). The coefficients for all models can be found in the appendix at the end of this paper.

We first present a pair of difference-in-difference models based on 54 tracts — the 15 that were gentrified and the 39 that were eligible but not gentrified — including one model in which we controlled for call type and one in which we did not. Next, we present a similar pair of models in which 10 gentrified tracts were matched with 10 eligible-but-not-gentrified tracts based on pre-gentrification attributes, again controlling

and not controlling for call type. Finally, we present models for the two most common call types — bulk collection and parking enforcement — with matching.

5.1 Difference-in-difference models without matching

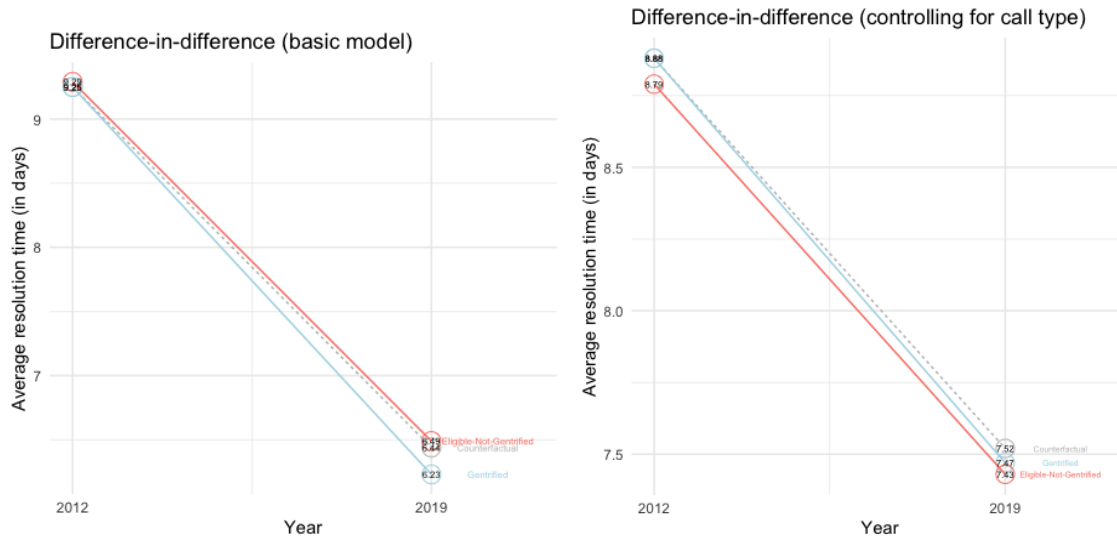


Figure 6: DiD - Basic (left) and DiD - Controlling for Call Type (right)

In the basic model (Figure 6 (left)), in which we did not control for call type, the difference-in-difference estimate is -0.22 days, suggesting that resolution times in gentrified tracts were about 5 hours and 17 minutes faster than they would have been had gentrification not occurred. However, this model has little statistical significance ($p = 0.0964$) and the confidence interval (- 0.48 days, .04 days) straddles 0.

Controlling for call type diminishes what little statistical significance there is. In Figure 6 (right), the difference-in-difference estimate is -0.05 days, or about 1 hour and 12 minutes, with a confidence interval (-0.24 days, 0.14 days) that again straddles 0. The p-value for the model that controls for call type is $p = 0.624$. Here, it is clear that gentrification may not have impacted resolution time at all, and it is likely that the mix of calls largely accounts for any difference in the rate of improvement between gentrified and eligible-but-not-gentrified tracts.

5.2 Difference-in-difference models with matching

Matching allows for clearer insights into the impact of gentrification on response times by creating a more meaningful control group. By identifying the 10 gentrified tracts and 10 eligible-but-not-gentrified tracts that were most similar to each other as of 2012, we were able to zero in on the tracts from which we could learn the most.

There are several ways to go about matching. Two of the most commonly used methods are propensity score matching and Mahalanobis distance matching. Propensity score matching, in this context, involves pairing like census tracts using a propensity score calculated by reducing pre-gentrification characteristics for each tract (MHI, median home value, and percent of adults over 25 with a Bachelor’s degree) to a single dimension. This means that matched pairs will be close in propensity scores, but will not necessarily be similar with respect to a particular characteristic. On the other hand, Mahalanobis distance matching entails matching pairs of census tracts based on a Mahalanobis distance, similar to a Euclidean distance, calculated

using the same pre-gentrification characteristics. With this method, each paired tract will be as similar as possible with respect to all of the attributes used for matching. (King and Nielsen 2019)

Our approach to matching blended the two methods above by utilizing the `MatchIt` package in R and applying the ‘genetic’ method (see (Sekhon 2011) for details on `GenMatch()`). This allowed us to balance our pre-gentrification tract attributes. Figure 7 (left) below shows the 10 matched gentrified tracts and the 10 matched eligible-but-not-gentrified tracts, while Figure 7 (right) shows the balance achieved with regard to the pre-gentrification attributes and the distance measure for the matched tracts as compared to all 54 tracts that were deemed eligible in 2012.



Figure 7: Matched Census Tracts (left) and MatchIt Plot (right)

Although matching reduced the number of tracts to 20, we still had more than 25,000 observations to use in our second pair of difference-in-difference models. The coefficients for these models can be found in the table provided in the appendix.

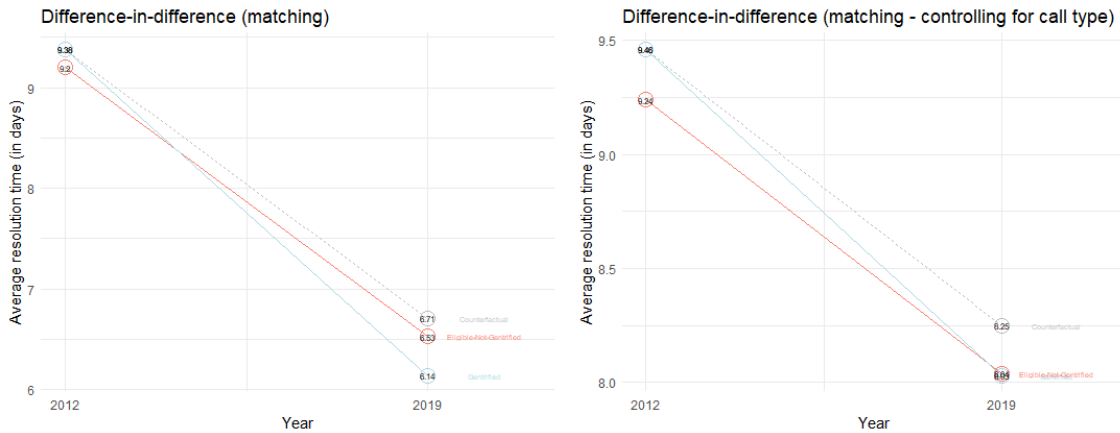


Figure 8: DiD - Basic with Matching (left) DiD - Controlling for Call Type with Matching (right)

In the basic model with matching (Figure 8 (left)), in which we did not control for call type, the difference-in-difference estimate is -0.57 days, or about 13 hours and 41 minutes faster. This model has more statistical

significance ($p = 0.0065$) than the basic model without matching (Figure 6 (left)) and has a confidence interval of $(-0.98 \text{ days}, -0.16 \text{ days})$, which does not straddle 0. This suggests that resolution times in tracts that gentrified by 2019 were a little more than a half a day faster than they would have been had gentrification not occurred, assuming that the trend in the gentrified tracts would have paralleled that of the eligible-but-not-gentrified tracts that were most similar to them in 2012.

However, when controlling for call type, the statistical significance is again diminished. In Figure 8 (right), the difference-in-difference estimate is -0.22 days, or about 5 hours and 17 minutes, with a confidence interval $(-0.52 \text{ days}, 0.08 \text{ days})$ that again straddles 0. The p -value for the model with matched tracts that controls for call type is $p = 0.149$. It again appears that gentrification may not have had any effect on resolution time, and any difference in the rate of improvement between gentrification statuses can be largely attributed to the mix of calls.

5.3 Difference-in-difference models for individual call types with matching

In addition to analyzing resolution times across all selected call types, with and without matching, we examined resolution times for each of the two most common call types — bulk collection and parking enforcement — in isolation, retaining the matched tracts from the previous pair of models.

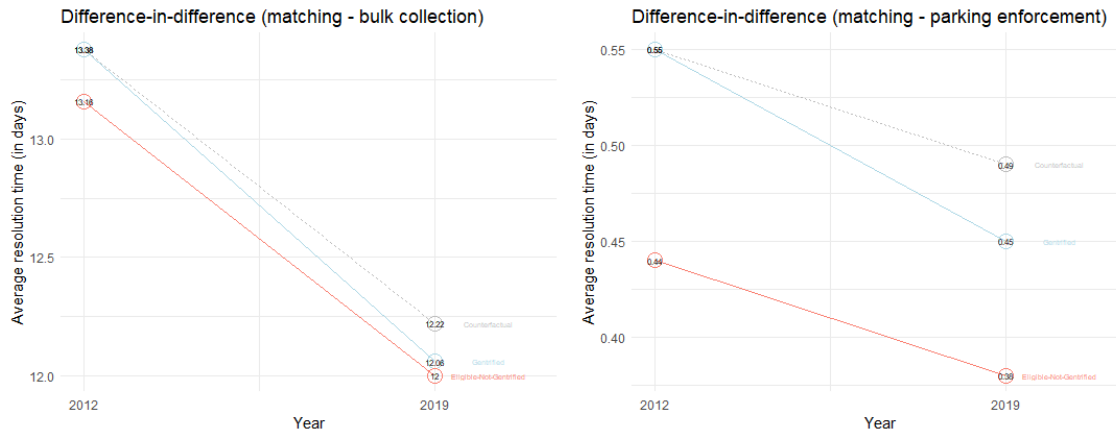


Figure 9: DiD - Bulk Collection with Matching (left) DiD - Parking Enforcement with Matching (right)

In the model for bulk collection shown in Figure 9 (left), the difference-in-difference estimate is -0.15 days, or about 3 hours and 36 minutes faster. This model has little statistical significance ($p = 0.555$) and a confidence interval of $(-0.65 \text{ days}, 0.35 \text{ days})$, which straddles 0. Likewise, the model for parking enforcement in Figure 9 (right) has low statistical significance ($p = 0.777$) and a difference-in-difference estimate of just -0.04 , or about 58 minutes faster, with a confidence interval of $(-0.33 \text{ days}, 0.25 \text{ days})$ that again includes 0. Therefore, while gentrified tracts may have seen a slightly sharper decrease in resolution times than the eligible-but-not-gentrified tracts with which they were matched, we cannot say with certainty that gentrification made a difference. Given the prevalence of bulk collection and parking enforcement in the data, it is surprising that we obtained statistically significant results with matched tracts when analyzing all of the selected call types together.

6 Discussion

Based on the foregoing analysis, we have arrived at the following conclusions:

- When comparing all gentrified and eligible-but-not-gentrified census tracts, there is weak evidence that gentrification had a mild impact on 311 call resolution times (significant at the $p = 0.1$ level). However, this effect disappears when controlling for call type, which suggests that the mix of calls was responsible for most of the variation in resolution-time improvement across gentrification statuses.
- On the other hand, when we compare 10 gentrified tracts with the 10 eligible-but-not-gentrified tracts that most closely resemble them in pre-gentrification attributes, we see evidence strong that gentrification did have an impact on call resolution times ($p = 0.0065$). Yet, when we use matched tracts and control for call type, statistical significance is again diminished ($p = 0.149$).
- When it comes to the two most common call types — bulk collection and parking enforcement — there is no evidence that gentrification has an impact on call resolution times, even when using matched tracts.

Of course, these conclusions are only as good as our data and our definition of gentrification. We would like to have established that trends in what would become the gentrified and eligible-but-not-gentrified tracts were parallel in the years leading up to 2012, but we could not do so because of data quality issues. In the absence of scholarly consensus as to what constitutes gentrification, we would also like to have been assured that our definition was meaningful, as some aspects of gentrification are difficult to capture quantitatively, and we do not know how well they correlate with the socioeconomic indicators we used. A future team of researchers might attempt to validate a common definition by conducting systematic, qualitative fieldwork to understand what census tracts that qualify as gentrified look like in real life. Then again, since gentrification is defined by relative growth, not by absolute growth or absolute wealth, perhaps newly gentrified tracts should not be expected to have much in common with previously gentrified tracts, nor with each other.

We view our results as good news for District residents. As we have shown above, resolution times for 15 of the most common 311 service requests decreased between 2012 and 2019, even while call volume skyrocketed; wait times became more equitable; and while more affluent neighborhoods continued to enjoy slightly better service, it does not appear as if certain census tracts were left far behind based on gentrification status. Gentrification may have many undesirable side effects, but exacerbating inequality with respect to everyday city service provision does not seem to be one of them, when the mix of call types is taken into account. In a city that is gentrifying as quickly as Washington, D.C., that is an encouraging thing.

7 References

- Beck, Brenden. 2020. “Policing Gentrification: Stops and Low-Level Arrests During Demographic Change and Real Estate Reinvestment.” *City & Community* 19 (1): 245–72. <https://doi.org/10.1111/cico.12473>.
- Bhavsar, Nrupen A., Manish Kumar, and Laura Richman. 2020. “Defining Gentrification for Epidemiologic Research: A Systematic Review.” Edited by Joel Msafiri Francis. *PLOS ONE* 15 (5): e0233361. <https://doi.org/10.1371/journal.pone.0233361>.
- Data Science for Public Service. 2019. “Foundations of Program Evaluation III - Regression Tools for Causal Analysis: Diff-in-Diff.” <https://ds4ps.org/PROG-EVAL-III/DiffInDiff.html>.
- Freeman, Lance. 2005. “Displacement or Succession?: Residential Mobility in Gentrifying Neighborhoods.” *Urban Affairs Review* 40 (4): 463–91. <https://doi.org/10.1177/1078087404273341>.
- Harris, Brandon, Alessandro Rigolon, and Mariela Fernandez. 2020. ““To Them, We’re Just Kids from the Hood”: Citizen-Based Policing of Youth of Color, “white Space,” and Environmental Gentrification.” *Cities* 107: 102885.
- King, Gary, and Richard Nielsen. 2019. “Why Propensity Scores Should Not Be Used for Matching.” *Political Analysis* 27 (4): 435–54. <https://doi.org/10.1017/pan.2019.11>.
- Laniyonu, Ayobami. 2018. “Coffee Shops and Street Stops: Policing Practices in Gentrifying Neighborhoods.” *Urban Affairs Review* 54 (5): 898–930. <https://doi.org/10.1177/1078087416689728>.
- Maciag, Michael. 2015a. “Gentrification in America Report,” January. <https://www.governing.com/archive/gentrification-in-cities-governing-report.html>.
- . 2015b. “Gentrification Report Methodology,” January. <https://www.governing.com/archive/gentrification-report-methodology.html>.
- Misra, Tanvi. 2018. “Yes, 311 Nuisance Calls Are Climbing in Gentrifying Neighborhoods.” *Bloomberg*, October, 3. <https://www.bloomberg.com/news/articles/2018-10-18/in-new-york-city-gentrification-brings-more-311-calls>.
- NYU. 2015. “Study Uses 311 Complaints to Track When and Where Neighborhood Conflict Emerges.pdf.” *Science Daily*, August. www.sciencedaily.com/releases/2015/08/150821082721.htm.
- Reza, Nooreem. 2022. “Problematizing Low-Level Policing’s Relationship with Racialized Gentrification.” *Va. J. Soc. Pol’y & L.* 29: 68.
- Sekhon, Jasjeet S. 2011. “Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The **Matching** Package for *R*.” *Journal of Statistical Software* 42 (7). <https://doi.org/10.18637/jss.v042.i07>.
- Vo, Lam Thuy. 2018. “They Played Dominoes Outside Their Apartment For Decades. Then The White People Moved In And Police Started Showing Up.” *Buzzfeed News*, June, 12.

8 Appendix

Table 1 - Difference-in-difference models without matching

	Resolution Days Controlling for Call Type	
	Basic	
Intercept	9.289*** (0.063)	8.785*** (0.103)
Treatment	-0.044 (0.113)	0.092 (0.084)
Post	-2.801*** (0.073)	-1.355*** (0.055)
Alley Cleaning		3.954*** (0.142)
Bulk Collection		4.071*** (0.098)
Graffiti Removal		1.431*** (0.324)
Illegal Dumping		2.831*** (0.121)
Parking Enforcement		-7.201*** (0.103)
Pothole		1.354*** (0.132)
Recycling Collection - Missed		-4.552*** (0.169)
Residential Parking Permit Violation		-7.013*** (0.124)
Sanitation Enforcement		-3.918*** (0.121)
Street Cleaning		3.429*** (0.158)
Streetlight Repair Investigation		-4.169*** (0.137)
Trash Collection - Missed		-4.782*** (0.122)
Tree Inspection		-4.787*** (0.140)
Treatment:Post	-0.218* (0.131)	-0.048 (0.098)
Observations	70,381	70,381

Note:

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

Table 2 - Difference-in-difference models with matching

	Resolution Days	
	Basic	Controlling for Call Type
Intercept	9.199*** (0.130)	9.243*** (0.184)
Treatment	0.177 (0.182)	0.217 (0.133)
Post	-2.667*** (0.150)	-1.207*** (0.112)
Alley Cleaning		3.344*** (0.245)
Bulk Collection		3.969*** (0.170)
Graffiti Removal		0.360 (0.482)
Illegal Dumping		2.534*** (0.205)
Parking Enforcement		-7.839*** (0.177)
Pothole		0.232 (0.226)
Recycling Collection - Missed		-5.219*** (0.276)
Residential Parking Permit Violation		-7.588*** (0.205)
Sanitation Enforcement		-4.351*** (0.207)
Street Cleaning		2.905*** (0.266)
Streetlight Repair Investigation		-4.749*** (0.237)
Trash Collection - Missed		-5.344*** (0.209)
Tree Inspection		-5.407*** (0.241)
Treatment:Post	-0.570*** (0.209)	-0.221 (0.154)
Observations	25,475	25,475
Note: *p<0.1; **p<0.05; ***p<0.01		

Table 3 - Difference-in-difference models for individual call types with matching

	Resolution Days	
	Bulk Collection	Parking Enforcement
Intercept	13.160*** (0.147)	0.436*** (0.097)
Treatment	0.219 (0.206)	0.114 (0.133)
Post	-1.163*** (0.183)	-0.059 (0.107)
Treatment:Post	-0.151 (0.256)	-0.041 (0.147)
Observations	8,369	4,555

Note: *p<0.1; **p<0.05; ***p<0.01