Modeling Neighborhood Change to Mitigate Gentrification: A Case Study of Fairfax County, Virginia

Teja Pristavec, https://orcid.org/0000-0003-2021-4682, tejapris@gmail.com
Brandon Kramer, https://orcid.org/0000-0002-3233-5310
Joshua Goldstein, https://orcid.org/0000-0002-1164-1829, jrg3bs@virginia.edu
Michelle Gregory, Michelle.Gregory@fairfaxcounty.gov
Joy Tobin, joytobin11@gmail.com
Sallie Keller, https://orcid.org/0000-0001-7303-7267

Social and Decision Analytics Division
Biocomplexity Institute & Initiative
University of Virginia

June 5, 2021

Funding: This research was partially sponsored by Fairfax County, VA government under the University of Virginia award number GO12554 and by the University of Virginia Strategic Investment Fund, award number SIF160.

Citation: Pristavec T, Kramer B, Goldstein J, Gregory M, Tobin J, Keller, S. (2021). Modeling Neighborhood Change to MiXgate GentrificaXon: A Case Study of Fairfax County, Virginia, *Proceedings of the Biocomplexity Institute*, Technical Report. TR# 2021-121. University of Virginia.
https://doi.org/10.18130/bjbm-f119

UNIVERSITY of VIRGINIA
BIOCOMPLEXITY INSTITUTE

Modeling Neighborhood Change to Mitigate Gentrification: A Case Study of Fairfax County, Virginia

Abstract

Gentrification describes the transformation of working-class or vacant areas into middle-class residential or commercial zones through an influx of affluent persons and businesses displacing long-term, vulnerable populations. Local governments often lack resources to detect gentrification emergence and mitigate its negative impacts. We demonstrate how studying gentrification at granular geographies using publicly available data can provide actionable insights to stakeholders seeking to preserve neighborhood diversity and protect at-risk residents. We examine neighborhood change in Fairfax County, VA using US Census Bureau American Community Survey 5-year estimates (2008/12-2014/18). Using three multifactor dimensions to measure gentrification, we classify census tracts into those not vulnerable to gentrification at baseline; vulnerable but not gentrified, and vulnerable and gentrified over time. We employ a spatial generalized linear mixed model to examine property and population factors associated with gentrification and test the effects of a hypothetical housing policy intervention. Results suggest that 61% of Fairfax County tracts were not vulnerable and 39% were vulnerable to gentrification at baseline. Of those vulnerable, 49% did not gentrify over time. The remaining 51% experienced significant socioeconomic and investment change, gentrifying during the period. Median property values, college-educated population, and white population shares were associated with increased gentrification likelihood. Finally, we show that a 10% median property value reduction intervention would result in 26% fewer vulnerable and 50% fewer gentrified areas. We conclude with policy recommendations mitigating gentrification and highlight how public data and modeling could assist local governments with decision-making, maximizing the impact of county resources, and improving neighborhood outcomes.

Introduction

Gentrification describes the transformation of working-class or vacant areas into a middle-class residential or commercial zones through an influx of affluent persons and businesses displacing long-term, vulnerable populations (Lees et al., 2008: xv). Over the past 20 years, gentrification affected thousands of neighborhoods around the world, especially major urban centres across the US (Hyra, 2012; Lees et al., 2008). As of 2015, gentrification is estimated to impact nearly 20% of all low-income communities (i.e., census tracts in the bottom 40th percentile of median household income nationally) in the US' fifty largest cities—up from only 9% of such neighborhoods a decade earlier (Maciag, 2015). Existing scholarship focuses either on how the spread of gentrification can be modeled quantitatively to monitor displacement risk (cf. Chapple and Zuk, 2016) or to understand how gentrification impacts historically marginalized communities in places like New York (Barton, 2016; Hackworth, 2002), Los Angeles (DeVerteuil, 2011), and Washington, DC (Hyra, 2017; Hyra and Prince, 2016. We use a quantitative model to predict gentrification risk in Fairfax County, VA—the largest suburban county in the Washington, DC area. We use Fairfax County as a case study of gentrification spread in suburban areas. Moreover, we conduct a "what-if" policy-making analysis that local governments can use to consider the potential impact of housing policies before implementation.

Prior gentrification literature detailed how population dynamics and investment changes led to widespread displacement in Washington, DC. DC underwent notable demographic change over the past two decades. Since 2000, the nation's capital has been one of the few US cities to experience population growth primarily due to the in-migration of white residents—with that group expanding from ~30% to ~45% of its total population (US Census Bureau, 2018a). While DC neighborhoods experienced renovation of existing infrastructure and commerce growth that initially benefitted long-term and incoming residents, these changes eventually attracted higher-income occupants who found cheap rental and housing prices more appealing than existing upscale neighborhoods or suburbs (Hyra, 2017; Hyra and Prince, 2016; Lees et al. 2008). As a result, low-income and Black residents were disproportionately forced out of Washington (Jackson, 2015; Stancil et al., 2019). Roughly 40% of DC's census tracts gentrified during this period, leading to more than 20,000 Black residents' displacement (Richardson et al., 2019). As Walker (2018) found elsewhere in the country, many of DC's low-income and minority communities were displaced from long-time homes, moving to Southern Maryland and Northern Virginia's inner suburban rings (Stancil et al., 2019). Although US suburban areas are becoming more racially and ethnically diverse, these neighborhoods are also increasingly likely to exhibit high levels of concentrated poverty, including in the areas immediately surrounding Washington, DC (*ibid*).

To DC's west, Fairfax County, VA is one of 50 most densely populous and 100 most rapidly growing US counties (US Census Bureau, 2019b). Despite its diverse population, Fairfax County's in-demand housing stock and rising real estate prices have made it difficult for many families to afford living in the county (Xan et al., 2019). To address these concerns, Fairfax County's government advanced several initiatives promoting equitable and inclusive growth (One Fairfax, 2017). In partnership with Fairfax County, we developed a quantitative strategy for examining the impact of gentrification in the area. We build on data-driven governance techniques pioneered by academic, nonprofit, and governmental institutions seeking to understand and mitigate gentrification with predictive modeling (cf. Chapple and Zuk, 2016). Employing publicly available data sources and open source software (Keller et al., 2018), we implement a sparse spatial generalized linear model to examine the risk of gentrification in Fairfax County.

Below, we describe this collaboration and our contribution to the gentrification literature, including the implementation of a quantitative modeling approach predicting gentrification in Fairfax County VA; using a model that accounts for spatial effects; and employing a "what-if" analysis that local governments can use to test various scenarios and understand the impact they would have on gentrification risk and displacement. We first review past work on how scholars and city governments used quantitative approaches to study gentrification, including methodological considerations about how to measure the phenomenon. We also share background information for better understanding Fairfax County within the broader context of transnational capitalism and to explain the housing pressures Fairfax County is currently experiencing. Second, we detail our methodological approach to measuring gentrification. Third, we describe results, including how we classified gentrification risks at the census tract level, how the model might be used to predict future gentrification risk, and how it can inform "what-if" policy interventions. We conclude with future research suggestions and policy considerations that county governments may consider when developing similar quantitative approaches to studying gentrification.

Literature Review

Research documents gentrification as a global process in that the broader structures of capitalism shape the composition and consumption patterns of populations in cities across the world (Lees et al., 2016; Liu et al., 2021). As Lees et al. (2008) argue, "local clusters of transnational corporate services and headquarters not only generate demand for local gentrified residential space, but also serve to weave this local demand into transnational circuits of labor or migration amongst itinerant professionals" (80). The ongoing capital investment into central business districts, alongside the integration and deregulation of financial markets, continued to put higher demand on low-income housing markets across US metropolitan areas over the past two decades (Fainstein, 2000; Sassen, 2000). As a result, historical factors like deindustrialization, public housing demolition, subprime mortgage lending, and local and federal policies designed to promote growth in developing areas (Hackworth, 2002; Hyra, 2012; Rothstein, 2017; Taylor, 2019), led to more US neighborhoods becoming gentrified over time (Maciag, 2015; Richardson et al., 2019).

Most scholarship focused on how gentrification affects inner city housing dynamics, providing a number of production and consumption-based theories to explain why investors and/or incoming residents are attracted to these neighborhoods (Hyra, 2017; Lees et al., 2008). Recently, research identified novel ways in which gentrification geographies have shifted toward suburban areas. As urban renewal spread across city centres (Hyra, 2012; Taylor, 2019), many Black and low-income residents moved from inner city to inner suburbs (Walker, 2018). Oftentimes, these groups are segregated into undesirable neighborhoods via exclusionary housing policies or displaced through predatory market factors (Hyra, 2012;

Rothstein, 2017; Taylor, 2019). Despite racial and ethnic minorities reporting that they experience more equitable conditions in the suburbs now than in the early 2000s (Pfeiffer, 2012), these suburban areas are more likely to have higher levels of concentrated poverty and forced displacement (Stancil et al., 2019).

We focus on predicting gentrification in Fairfax County, VA and ultimately developing a policymaking tool that government officials can use to model counterfactual scenarios that help mitigate the impacts of this process. Part of the suburban ring surrounding Washington, DC, Fairfax County has a population of 1.1 million people, making it one of US' 50 most populous counties. Close to DC, Fairfax County is both a hub for workers associated with government agencies and contractors, and a burgeoning tech and healthcare sector hotspot. Notably, Amazon recently built its second headquarters in nearby Arlington County, opening a pipeline for thousands of new tech workers to move to the area (Arcieri, 2018). Moreover, the Dulles Corridor has roughly 600 data centers serving over 3,000 tech companies, making Fairfax County a geographic bridge between the nation's political capital to its east, and to its west what some dubbed the "Silicon Valley of the East" (St. Germain, 2019). Given this confluence of factors, it is unsurprising that Fairfax County is one the most expensive US counties to live in, with residents making a median household income of ~$115,000 and median home values among the country' highest. Despite this, Fairfax County's population continues to grow larger, older, and wealthier while also becoming more diverse as more Asian and Hispanic residents migrate to the region (Xan et al., 2019).

To better understand and mitigate the issues identified above, we worked with Fairfax County to develop a quantitative approach to examine gentrification (Keller et al., 2018). We built on past work studying neighborhood change in New York, Los Angeles, Seattle, Portland, Harris County, TX, and the Bay Area (Choudhary et al., 2018; Chapple and Zuk, 2016). While the literature is rich in examples using quantitative modeling to examine gentrification, there is considerable variation in the data and variables included in models, leading to extensive debate about what gentrification means and how to appropriately measure its spread across different geographical contexts (Barton, 2016; Chapple and Zuk, 2016; Reades et al., 2019). Most approaches include three core components: baseline vulnerability for the populations or regions of interest, change in socio-demographic variables, and changes in neighborhood investment (Choudhary et al., 2018). Yet, as Reades et al. (2019) point out, this work is largely historical and rarely aims to address ongoing policymaking by predicting gentrification's emergence in developing neighborhoods. Building on this foundational work, we aim to fill the gap, using a gentrification typology capturing neighborhood change and developing a predictive approach that can inform local policymakers' decision-making processes.

Data and Methods

Data

To construct our gentrification typology and predictive model, we use the US Census Bureau's American Community Survey (ACS) data on Fairfax County population and housing units. ACS is an ongoing, nationally representative household survey. It annually releases household, family, and individual demographic and socioeconomic characteristics estimates, as well as select property characteristics (US Census Bureau, 2018a). Using the R statistical programming environment tidycensus package (Walker et al., 2018), we retrieved 5-year estimates of census tract-level ACS data based on 2010 Decennial Census geographies (US Census Bureau, 2018a). Census tracts are statistical subdivisions smaller than counties and larger than block groups, ranging in size from 1,200 to 8,000 individuals (US Census Bureau, 2019c). Compared to single-year estimates, 5-year estimates provide larger sample sizes, greater estimate reliability, and better geographic granularity. They provide data for all geographic areas regardless of population size, and are suitable for analyzing small areas for which single-year estimates are frequently unavailable. We use 2008/12 estimates for our baseline period and 2014/18 estimates as our period end-point.

Dependent Variable: Gentrification Status

We draw on Choudary et al. (2018) to construct three criteria defining our gentrification status outcome variable: tract vulnerability in the base year, tract socio-demographic change, and tract investment change over time. We describe the criteria used to assign gentrification status below. Table 1 lists ACS source tables for variables associated with each criterion.

[Table 1 about here]

First, we define a census tract as *vulnerable to gentrification in the base year* using median household income, percent individuals over age 25 without a Bachelor's degree, percent non-white population, and percent renter households. We retrieve relevant information from ACS tables shown in Table 1, and calculate median household income, percent population without a Bachelor's degree, percent non-white population, and percent of renter households at both the tract and county level. Finally, we construct a dichotomous tract gentrification vulnerability indicator, with tracts categorized as vulnerable if they exhibit three or more of the following, compared to the county median value: lower median household income, higher percent individuals over age 25 without a Bachelor's degree, higher percent non-white population, and higher percent renter households.

Second, we define a census tract as having experienced considerable *socio-demographic change over time* using information on change in percent population with a Bachelor's degree or higher, change in median household income, and change in percent non-Hispanic white population. We retrieve tract-level ACS table information and calculate percent population aged over 25 with at least a Bachelor's degree, Annual Average Consumer Price Index Research Series data inflation-adjusted median household income, and percent non-Hispanic white population using both ACS 2008/12 and 2014/18 estimates. After calculating each variable at both time points, we subtract the base year values from end year values to obtain percent change over time on each variable. We repeat the procedure at county level to obtain county-level change. Finally, we construct an indicator for dichotomous tract socio-demographic change over time with a tract categorized as having experienced such change if it satisfies at least one of the following conditions: the tract change over time in percent population over age 25 with at least a Bachelor's degree is greater than the county change; or the tract change in median household income and the tract change in percent non-Hispanic population are both greater than the county change over the same time period.

Third, we define a census tract as having experienced *investment change over time* using information on change in median gross rent and change in median property values. We retrieve tract-level information on monthly median gross rent in dollars for renter-occupied housing units and median property value in dollars for owner-occupied housing units. We again create each variable at the base and end years using the ACS 2008/12 and 2014/18 estimates respectively before adjusting all values to 2018 constant dollars using Annual Average Consumer Price Index Research Series data. We repeat the procedure at the county level to calculate county-level change. Lastly, we construct a dichotomous tract investment change over time indicator, with tracts categorized as having experienced such change if either their change in monthly median gross rent or their change in median property values over time were greater than the county change during the same time period.

We finally use the constructed tract baseline vulnerability, socioeconomic change, and investment change variables to define our classification and model outcome variable, *tract gentrification status*. Gentrification status is a three-category variable with tracts coded as not vulnerable to gentrification if they do not satisfy the baseline vulnerability criterion. For tracts that do satisfy the baseline vulnerability criterion, we consider the remaining two criteria in determining gentrification status. Tracts that satisfy the baseline vulnerability criterion but not the socioeconomic change and investment change criteria are coded as vulnerable, but not gentrified. Tracts that satisfy the baseline vulnerability, as well as both the socioeconomic change and investment change criteria are coded as vulnerable and gentrified.

Property- and Population-Level Independent Variables

Our models control for seven property-level factors and ten population-level factors relevant to neighborhood change and predicting tract gentrification status. We calculate all variables using ACS data. Our property-level factors capture housing change and include differences over time in percent multi-unit residential properties; percent vacant housing; percent single family properties; percent in renter-occupied units; percent change in median property value; percent change in median gross rent; and percent change in housing density. Our population-level factors capture sociodemographic change and include differences over time in percent population living poverty; percent rent burdened; percent living in different house than in previous year; percent taking public transit to work; percent unemployed; percent with Bachelor's degree or higher; percent non-Hispanic white; percent non-family households; percent change in median household income; and population growth. Supplementary File 1 provides source ACS table numbers and detailed operationalization descriptions for each property- and population-level variable included.

Missing Data

Data was available for all variables at the county level. At tract level, of 258 tracts in Fairfax County, information was missing for a total of 22 tracts. Among the 22, information was missing on median property value for 7 tracts in 2008/12 estimates and for 15 in 2014/18; on median gross rent for 6 tracts in 2008/12 and for 15 in 2014/18; and on median household income for 3 tracts in 2008/12 and in 2014/18. Data for these tracts were suppressed due to margins of error being larger than medians themselves, making estimates statistically unreliable (US Census Bureau, 2016). Where available, we used estimates from preceding years to represent the base or end year, filling in information for 13 tracts. The remaining 9 tracts had valid missingness; they are areas occupied by military installations, airports, or large green spaces. We code gentrification status for these tracts as unavailable, and exclude them from our model. No data was missing on model covariates.

Methods

Despite past work identifying spatial proximity as a key factor in the likelihood of adjacent neighborhoods' gentrification (Guerrieri et al., 2013), not all quantitative approaches have appropriately accounted for spatial autocorrelation (Reades et al., 2019). Spatial autocorrelation refers to the presence of systematic variation in a mapped variable, leading to positive or negative residual clustering based on spatial proximity. To test for spatial autocorrelation and to select the most appropriate geographic modeling approach, we use the permutation and the joint count tests with a spatial weights matrix. Both tests reject the null hypothesis of no spatial clustering ($p<0.001$), indicating that tract gentrification status is similar in adjacent tracts more often than would be expected with spatial randomness.

To account for spatial autocorrelation, we employed a sparse spatial generalized linear mixed model (SSGLMM) with a logistic link function in predicting tract gentrification using the ngspatial package in the R 3.6.3 open source statistical programming environment (Hughes, 2014). This model relies on a Bayesian framework with Markov chain Monte Carlo

(MCMC) simulation for inference. Our neighborhood matrix is defined by whether two tracts share a border, and the model accounts for spatial confounding using a restricted spatial regression method. Restricted spatial regression removes confounding from spatial effects, constraining spatial random effects to be orthogonal to fixed effects (Hodges and Reich, 2010). The model uses Metropolis-Hastings random walks with normal proposals to update regression parameters (Hughes and Haran, 2014). We run models with standard deviation for spatial random effects set at 0.01, and the prior standard deviation for regression coefficients at 1000. We specify a minimum sample size of 5,000 to permit accurate Monte Carlo standard error estimation, a Monte Carlo standard error tolerance of 0.01, and a maximum of one million iterations (Flegal et al., 2008). We base interpretations on combined samples from three parallel Markov chains.

Given the difficulty of establishing MCMC model convergence, we conduct sensitivity analyses by running analogous multivariate spatial binomial likelihood models (MSBLMM) with a logistic link function predicting tract gentrification using the CARBayes package (Lee, 2013). As was the case with the SSGLMM, the MSBLMM uses a MCMC simulation in a Bayesian setting with Metropolis adjusted Langevin algorithm updates for regression parameters and spatial effects (Roberts and Rosenthal, 1998). We use 210,000 samples, discard 10,000 as the burn-in period, and thin the remaining samples by 100 to reduce temporal autocorrelation. MSBLMM results did not differ in coefficient direction or statistical significance pattern from the SSGLMM.

We determine convergence for both the SSGLMM and MSBLMM using visual inspection of traceplots for the three chains, the potential scale reduction factor (PSRF)

(Gelman et al., 2003), and the Geweke diagnostic (Geweke, 1992). For each model type, traceplots of the posterior distribution for all chains by variable suggested good Markov chain mixing, with similar regression parameter means, no apparent trend, and good sample space exploration. Multivariate PSRF less than 1.1 suggests model convergence, and values did not exceed this threshold. Covariate PSRF values were at or close to 1, also suggesting convergence and equal between- and within-chain variance. Geweke diagnostic z-scores (obtained for the MSBLMM model) under +/-1.96 further suggest convergence; the diagnostic, testing for the equality of the means of the first 10% and last 50% of a Markov chain, did not exceed this threshold on any covariate except on change in tract housing density ($z = 2.1$), change in population living in a different house than a year ago ($z = -2.1$), and change in multiunit housing ($z = -2.1$). Each of these covariates only exceeded the score in one of the three chains, and we retain them in the model on conceptual grounds.

We compare model fit between SSGLMM chains and MSBLMM chains using the deviance information criterion (DIC), a fit criterion with model complexity penalization (Spiegelhalter et al., 2002). Smaller DIC values suggest a better fitting model. We obtained near identical values in the two models (values of 226, 225.9, and 215.1 for each of the three SSGLM chains; values of 226.4, 226.3, and 226.3 for the three MSBLMM chains). We present results from the SSGLMM given its advantage in coefficient interpretability over the MSBLMM.

Results

Classifying Tract-Level Gentrification Status

Overall, results suggest that gentrification is a serious risk for low-income residents throughout much of Fairfax County, VA. Of the 249 census tracts in our model, 151 (61%) were not vulnerable and 98 (39%) were vulnerable to gentrification at baseline. For vulnerable tracts, we considered whether they experienced significant socio-demographic change over time and whether they experienced significant investment change during the observation period to assign gentrification status. Of the 98 tracts vulnerable to gentrification, 48 (49%) did not gentrify during the observation period. The remaining 50 tracts (51%) experienced both significant socioeconomic and investment change over time, and were classified as having gentrified during the first half of the decade. Figure 1 shows gentrification status for Fairfax County census tracts.

[Figure 1 about here]

Figure 1: Fairfax County census tract gentrification status.

Table 2 shows descriptive statistics for all criteria considered in constructing the typology at both the baseline and end years, as well as comparisons with county values. Both vulnerable tracts that gentrified and vulnerable tracts that did not gentrify over time exhibit lower median household income than tracts not vulnerable to gentrification at baseline. Among vulnerable tracts, those that gentrified had a higher percent population without a Bachelor's degree and percent non-white population, but experienced larger decreases in these population shares over time compared to vulnerable tracts that did not gentrify. Percent non-Hispanic white population decreased over time in vulnerable non-gentrified tracts, and increased in vulnerable tracts that did gentrify. Similarly, while median gross rent increased somewhat and median property values remained approximately the same in vulnerable non-gentrified tracts, vulnerable tracts that gentrified experienced large increases in both domains. Between 64% and 86% of gentrified tracts experienced changes in percent population with a Bachelor's degree, non-Hispanic white population, median household income, median gross rent, and median property value that were greater than the county change in comparison with between 17% and 48% of vulnerable but not gentrified tracts.

[Table 2 about here]

Predicting Tract-Level Gentrification Status

Table 3 shows model predictor descriptive statistics by tract gentrification status. Table 4 and Table 5 present model results, showing posterior medians (PM) and 95% credible intervals (CI) for each covariate predicting tract gentrification pooled across chains. Both models performed well at classifying tracts with 0.88 area under curve (AUC) of the receiver operating characteristic (ROC) probability curve (Sullivan Pepe, 2000). AUC over 0.80 indicates good class separability. For further detail, Supplementary Table 1 provides individual chain results.

[Table 3 about here]

Predicting vulnerable and gentrified status. Table 4 shows that property values, college-educated population, and white population shares are statistically significantly associated with the likelihood of tract gentrification over time. Tracts that experienced a larger percent change in median property value, exhibited a larger difference in percent population with a Bachelor's degree, or exhibited a larger difference in percent non-Hispanic white population between baseline and end of observation period were more likely to gentrify compared to other tracts. More specifically, a vulnerable tract is 13% (95%CI = 0.07-0.21) more likely to gentrify for every 1% increase in median property value over time; 18% (95%CI = 0.09-0.30) more likely to gentrify for every 1% larger difference in percent population with a Bachelor's degree; and 11% (95%CI = 0.04-0.18) more likely to gentrify for every 1% larger difference in percent non-Hispanic white population.

[Table 4 about here]

Predicting vulnerable but not gentrified status. Table 5 shows that property values, median rent, and college-education population share are statistically significantly associated with the likelihood of a vulnerable tract not gentrifying over time, in a way consistent with results of the model predicting gentrified status. Vulnerable tracts that experienced larger percent change in median property value, larger percent change in median rent, and larger differences in percent population with a Bachelor's degree between baseline and observation period end were less likely to not gentrify compared to other tracts. Specifically, we expect to see a 7% decrease (95%CI = -0.12 - -0.03) in the likelihood of a tract being vulnerable but not gentrifying for every 1% increase in median property value; a 5% decrease (95%CI = -0.08 - -0.02) in the likelihood of a tract being vulnerable but not gentrifying for every 1% increase in median gross rent; and a 9% decrease (95%CI = -0.17 - -0.02) in the likelihood of a tract being vulnerable but not gentrifying for every 1% larger difference in percent population with a Bachelor's degree.

[Table 5 about here]

Examining effects of a hypothetical housing intervention. One of Fairfax County's principal interests in advancing this modeling approach was to garner more insights about "what-if" or counterfactual policy-making scenarios. To address this objective, we used model estimates to predict tract gentrification status after a hypothetical intervention that would reduce the change in median property value over the observed time period by 10%, holding all other factors constant. Table 6 displays the tract classification determined by our gentrification typology, model prediction, and model prediction post-intervention. Figure 2 displays these results. Gentrification typology classification is included for reference; we compare post-intervention and model-predicted gentrification status, as they rely on the same model parameters. Tracts with gentrification predicted probability ≥0.50 were coded as gentrified. Results suggest that with a hypothetical affordable housing intervention, we would observe a 50% reduction in the number of tracts that gentrified over time (a reduction from 34 to 17 of the total tracts), a 19% drop in the number of tracts that would have been vulnerable but not gentrified over time (10 fewer tracts), and a 14% increase in the number of tracts that would *not* have been vulnerable to gentrification (27 more tracts).

[Table 6 about here]

[Figure 2 about here]

Figure 2: Typology classified, model-predicted, and post-intervention tract gentrification status.

Discussion and Limitations

In this paper, we advance an open science quantitative model to predict the risk of census tracts gentrifying across Fairfax County, VA and, in turn, implement a “what-if” policy-making tool to compare the effects of the different housing proposals designed to mitigate gentrification. Drawing on the results of a sparse spatial generalized linear model to examine property- and population-level factors associated with neighborhood change and test the effects of a hypothetical policy intervention, results suggest that nearly 40% of Fairfax County census tracts were vulnerable to gentrification at baseline. Of the vulnerable tracts, 51% gentrified while experiencing significant socio-economic and investment change over time. We found that median property values as well as the college-educated and white population shares were associated with significantly increased likelihood of tract gentrification. In the final step, we modeled the effect a 10% reduction in median property value would have on gentrification risk, finding that the number of vulnerable areas would decline by 26%, and the number of gentrified areas would be cut by half.

This paper contributes to the gentrification literature in several important ways. First, we show the ways that suburban areas are impacted by gentrification. As we detailed above, many urban centres across the US have already become gentrified, leading to low-income and racial and ethnic minority residents being pushed out to suburban rings surrounding major cities like Washington, DC (Hyra, 2016; Richardson et al., 2019; Stancil et al., 2019). Results show that gentrification processes have clearly proliferated throughout Fairfax County over the past half decade. While these trends and approach likely generalize to other contexts, there are also a particular set of social and economic forces affecting how gentrification could shape housing conditions in Fairfax County over the coming years.

To this point, quantitative modeling must account for spatial effects to understand not only how suburban areas are impacted by cites, but also how spatial fixes within larger areas create spillover effects into adjacent communities. While our model does well in accounting for spatial autocorrelation, we believe that future work must also consider how more specific spatial effects may have shaped housing dynamics in the area. Given Fairfax County's proximity to Washington, DC to its east and Dulles Corridor to its west, Fairfax County is to see spillover effects where wealthier residents move to areas close to already gentrified neighborhoods (Guerrieri et al., 2013), driving up housing prices along the main travel routes that connect to employment clusters in the region (Kahn, 2007). For example, Figure 2 suggests that high-risk gentrification tracts cluster along two main transportation lines. The first is in the northern part of the county along Highway 66, as well as the Orange and Silver Metro lines, which connect the densely populated (sub)urban centres of Washington DC, Arlington, Falls Church, Tyson's Corner, the Reston/Herndon area, and Dulles International Airport to the data storage centres in Loudoun County. The second is along the southern border of the county connecting Crystal City, where Amazon's new headquarters is located, as well as Fort Belvoir, Fairfax County's largest employer. While we do not directly test the neighborhood spillover effects discussed in Guerrieri et al.'s (2013) endogenous gentrification theory, future work will need to account for Fairfax County's (and other major suburban areas') proximity to political and financial capital in more systematic ways than we able to here.

While the growth of gentrification documented in this paper is concerning, one shortcoming is not being able to account for the probable acceleration of gentrification resulting from the COVID-19 pandemic; this is a data availability limitation that can be addressed as ACS estimates covering the pandemic timeframe are released. Hackworth's (2002) analysis of post-recession New York housing conditions suggests that corporate developers as well as federal- and state-based policies accelerated certain types of neighborhood change. Given the parallels between how the Great Recession and the COVID-19 pandemic have impacted low- and middle-income unemployment trends as well as affordable housing (Pattath and Landen, 2021), coupled with the growing corporate influence of the tech industry in Fairfax County, it is likely that models underestimate the potential severity of gentrification and displacement in the coming years.

Our key contribution was examining "what if" policy-making scenarios. Results suggests that a 10% reduction in median property value change over time would reduce the number of vulnerable areas by about a quarter, and reduce the number of gentrified areas by half. Obviously, most residents would oppose artificially reducing property values, but there are several policies that have been successfully implemented across US cities that are designed to lower costs for low-income residents. Lubell (2016) generally places these policy approaches into six categories: preservation, protection, inclusion, revenue generation, incentives, and property acquisition. Concrete examples of such interventions include rent control in San Francisco (Diamond et al., 2019), real estate tax relief programs for long-term and low-income residents in Philadelphia (City of Philadelphia, 2021), eviction moratoriums like those set in place during the COVID-19 pandemic (Benfer et al., 2021; Pattath and Landen, 2021), and various housing voucher programs. In fact, Seattle and King County's recent Moving to Opportunity experiments provide mentorship to residents that have been allocated housing vouchers in order to promote residential mobility into "high opportunity" areas (Bergman et al., 2019). This would theoretically lower racial and/or economic segregation while exposing low-income residents to more social, political, and economic capital. Given that housing vouchers and inclusionary zoning do not necessarily protect against segregation (Chaskin 2013; Glaeser 2002), local governments may need to become more proactive about shaping how housing protections are implemented to ensure their residents gain access to equal opportunities.

While the housing intervention that we modeled may offer short-term benefits to current residents, the broader mission of Fairfax County is to support equitable and inclusive growth for all its residents (One Fairfax, 2017). The approach offered here is a useful tool to classify, predict, and model "what-if" scenarios that help advance evidence-based policies, especially in locations that are spatially-dependent on large metropolitan areas (like Washington, D.C.) that draw in residents from around the world. That said, the use of data science models must also continue to be informed by local residents, especially those that are at highest risk of displacement. In this regard, Hyra et al. (2019) remind us that community-based participatory action will also be integral to advancing equitable policies by strengthening the social and political capital of long-term residents. Like Reades et al. (2020), we advocate that quantitative approaches may help forecast which areas are most at risk of gentrification, but also recognize that proactive collaboration with diverse groups of local community members is needed to counteract the negative effects of public investment into (sub)urban renewal (Hyra et al., 2019). In this sense, qualitative and quantitative scholars could share complementary roles in shaping policy initiatives on gentrification and displacement moving forward.

References

Arcieri K (2018) Virginia's Amazon HQ2 win wasn't just based on traditional incentives. Washington Business Journal. Available at: https://www.bizjournals.com/washington/news/2018/11/13/virginias-win-of-amazon-hq2-wasnt-just-based-on.html (accessed 2/10/2021).

Barton M (2016) An exploration of the importance of the strategy used to identify gentrification. *Urban Studies* 53(1):92-111.

Benfer EA, Vlahov D, Long MY, Walker-Wells E, Pottenger J L, Gonsalves G and Keene DE (2021) Eviction, health inequity, and the spread of COVID-19: Housing policy as a primary pandemic mitigation strategy. *Journal of Urban Health* 98(1):1-12.

Bergman P, Chetty R, DeLuca S, Hendren N, Katz LF and Palmer C (2019) Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice *National Bureau of Economic Research* No. w26164.

Chaskin RJ (2013) Integration and exclusion: Urban poverty, public housing reform, and the dynamics of neighborhood restructuring. *The Annals of the American Academy of Political and Social Science* 647(1):237-267.

Chapple K and Zuk M (2016) Forewarned: The use of neighborhood early warning systems for gentrification and displacement. *Cityscape* 18(3):109-130.

De Verteuil G (2011) Evidence of gentrification-induced displacement among social services in London and Los Angeles. *Urban studies* 48(8):1563-1580.

Diamond R, McQuade T, and Qian F (2019) The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco. *American Economic Review* 109(9): 3365-3394.

Fainstein S (2000) *The City Builders: Property, Politics, and Planning in London and New York*. Oxford, UK: Blackwell Press.

Flegal J, Haran M, and Jones G (2008) Markov chain Monte Carlo: Can we trust the third significant figure? *Statistical Science* 23:250-260.

Gelman A, Carlin J, Stern H and Rubin D (2003) *Bayesian Data Analysis, 2nd edition*. London, UK: Chapman and Hall/CRC.

Guerrieri V, Hartley D, and Hurst E (2013) Endogenous gentrification and housing price dynamics. *Journal of Public Economics* 100:45-60.

Glaeser EL (2002) Does rent control reduce segregation? *Swedish Economic Policy Review* 10:179-202.

Hackworth J (2002) Postrecession gentrification in New York City. *Urban Affairs Review* 37(6):815-843.

Hodges J and Reich B (2010) Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician* 64: 325-334.

Hughes J (2014) Ngspatial: A package for fitting the centered autologistic and sparse spatial generalized linear mixed models for areal data. *The R Journal* 2:81-95.

Hughes J and Haran M (2013) Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. *Journal of the Royal Statistical Society, Series B* 75:139-159.

Hyra DS (2012) Conceptualizing the new urban renewal: Comparing the past to the present. *Urban Affairs Review* 48(4):498-527.

Hyra DS (2017) *Race, class, and politics in the Cappuccino City*. Chicago, IL: University of Chicago Press.

Hyra D, Moulden D, Weted C, and Fullilove M (2019) A method for making the just city: Housing, gentrification, and health. *Housing Policy Debate* 29(3):421-431.

Hyra DS and Prince S (eds.) (2015) *Capital Dilemma: Growth and Inequality in Washington, DC*. New York, NY: Routledge.

Jackson J (2015) The consequences of gentrification for racial change in Washington, DC. *Housing Policy Debate* 25(2):353-373.

Kahn ME (2007) Gentrification trends in new transit-oriented communities: Evidence from 14 cities that expanded and built rail transit systems. *Real Estate Economics* 35(2): 155-182.

Keller S, Nusser S, Shipp S, and Woteki CE (2018) Helping communities use data to make better decisions. *Issues in Science and Technology* 34(3):83-89.

Lee D (2013) CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors. *Journal of Statistical Software* 55:1-24.

Lees L, Slater T, and Wyly EK (eds.) (2008) *The Gentrification Reader*. London, UK: Routledge.

Leroux B, Lei X, and Breslow N (2000) Estimation of disease rates in small areas: A new mixed model for spatial dependence." In S*tatistical Models in Epidemiology, the Environment and Clinical Trials*, eds. M. Halloran and D. Berry, pp.179-191. New York, NY: Springer-Verlag.

Liu C, Deng Y, Song W, Gong J, and Zeng J (2021) Rethinking the geography of gentrification: From a scale perspective. *Geoforum* 118:23-29.

Maciag M (2015) Gentrification in America Report. Available at: https://www.governing.com/archive/gentrification-in-cities-governing-report.html (accessed 2/10/2021).

Pattath P and Landen M (2021) Eviction Moratoriums and COVID-19. Virginia Department of Health. Available at: https://www.vdh.virginia.gov/coronavirus/2021/01/20/ eviction-moratoriums-and-covid-19/ (accessed 3/31/2021).

Pfeiffer D (2016) Racial equity in the post-civil rights suburbs? Evidence from US regions 2000–2012. *Urban Studies* 53(4):799-817.

Philadelphia, City of. (2021) Longtime Owner Occupants Program (LOOP). Available at: https://www.phila.gov/services/payments-assistance-taxes/income-based-assistance-programs/longtime-owner-occupants-program/ (accessed 3/31/2021).

Fairfax County (2020) One Fairfax Policy. Available at: https://www.fairfaxcounty.gov/topics/one-fairfax (accessed 2/9/2021).

Reades J, De Souza J, and Hubbard P (2019) Understanding urban gentrification through machine learning. *Urban Studies* 56(5):922-942.

Richardson J, Mitchell B, Francho J (2019) Shifting Neighborhoods: Gentrification and cultural displacement in American cities. Available at: https://ncrc.org/gentrification/ (accessed 2/9/2021).

Rofe MW (2003) 'I want to be global': Theorising the gentrifying class as an emergent élite global community. *Urban Studies* 40(12):2511-2526.

Rothstein R (2017) *The Color of Law: A Forgotten History of How Our Government Segregated America*. New York, NY: Liveright Publishing.

Sassen S (2000) *Cities in a World Economy*. Thousand Oaks, CA: Pine Forge Press.

Spiegelhalter DJ, Best NG, Carlin BR, and van der Linde A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64:583-639. doi:10.1111/1467-9868.00353

St. German J (2019) Why is Ashburn the Data Center Capital of the World? Available at: https://www.datacenters.com/news/why-is-ashburn-the-data-center-capital-of-the-world. (accessed 2/9/2021).

Stancil W (2019). American Neighborhood Change in the 21st Century. Available at: https://www.law.umn.edu/sites/law.umn.edu/files/metro-files/american_neighborhood_change_in_the_21st_century_-_full_report_-_4-1-2019.pdf (accessed 2/9/2021).

Sullivan Pepe M (2000) Receiver operating characteristic methodology. *Journal of the American Statistical Association* 95(449):308-311. doi:10.1080/01621459.2000.10473930

Taylor KY (2019) *Race for Profit: How Banks and the Real Estate Industry Undermined Black Homeownership*. Chapel Hill, NC: UNC Press Books.

US Census Bureau (2019b) County Population Totals: 2010-2019. Available at: https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html#par_textimage_739801612 (accessed 3/30/2021).

US Census Bureau (2019c) Glossary: Census Tract. Available at: https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13 (accessed 5/12/2020).

US Census Bureau (2018a) What All Data Users Need to Know: Understanding and Using ACS Single-Year and Multiyear Estimates. Washington, DC: US Government Printing Office. Available at: https://census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018_ch03.pdf (accessed 5/28/2020).

US Census Bureau (2018b) Understanding and Using American Community Survey Data: What All Data Users Need to Know. Washington, DC: US Government Printing Office. Available at: https://www.census.gov/content/dam/Census/library/publications/2018/acs/acs_general_handbook_2018.pdf (accessed 5/28/2020).

Walker K (2018) Locating neighbourhood diversity in the American metropolis. *Urban Studies* 55(1):116-132.

Walker K, Eberwein K, and Herman M (2018) tidycensus: Load US census boundary and attribute data as tidyverse and sf-ready data frames. R package version 0.9(6).

Xan, X, Khaja F, and Lamsal M (2019) Demographic Reports 2019: County of Fairfax, Virginia. Available at:

https://www.fairfaxcounty.gov/demographics/sites/demographics/files/assets/demographicreports/fullrpt.pdf (accessed 2/9/2021).

Table 1. Tract Gentrification Criteria.

Criterion and Variables	*ACS Source Table(s)*
Criterion 1: Vulnerability in the base year.	
A census tract is considered vulnerable to gentrification if it exhibits 3 out of 4 characteristics compared to the county median in the base year. • Lower inflation-adjusted median household income • Higher percentage of individuals 25 years+ without a Bachelor's degree • Higher percentage of non-white population • Higher percentage of renter households	B19013 (Median Household Income) B15003 (Educational Attainment for the Population 25 Years and Over) B02001 (Race) B25003 (Housing Tenure)
Criterion 2: Sociodemographic change over time	
A census tract is considered changing in sociodemographics if at least 1 tract change was greater than the county's change from the base year to the end of a given period. • Change in percent population 25 and over with at least a Bachelor's degree OR • Change in inflation-adjusted median household income AND change in percent non Hispanic white population	B15003 (Educational Attainment for the Population 25 Years and Over) B19013 (Median Household Income) B03002 (Hispanic or Latino Origin by Race)
Criterion 3: Investment change over time	
A census tract's investment has changed over time if the tract's change in at least 1 is greater than the county's change. • Change in inflation-adjusted median monthly gross rent OR • Change in inflation-adjusted median home value	B25064 (Median Gross Rent) B25077 (Median Property Value in Dollars)

Note: ACS = American Community Survey 2008/12 and 2014/18 5-year estimates.

Table 2. Census Tract-Level Descriptive Statistics for Variables Related to Gentrification Typology by Gentrification Status.

		Not Vulnerable N = 151				Vulnerable but not Gentrified N = 48				Vulnerable and Gentrified N = 50			
Variable	*Year*	*mean*	*sd*	*min*	*max*	*mean*	*sd*	*min*	*max*	*mean*	*sd*	*min*	*max*
Median household income ($)	2008/12	155,897.28	40,630.43	70,524.97	273,927.02	93,513.40	19,275.31	55,283.73	119,759.32	90,375.99	19,344.26	35,475.60	122,444.89
	2014/18	156,626.97	39,031.91	68,889.00	250,001.00	91,232.19	19,617.23	41,859.00	143,688.00	98,447.96	20,441.77	42,382.00	140,250.00
% population age 25+ without BA	2008/12	33.37	10.86	13.02	61.24	49.64	14.04	13.09	76.95	55.38	12.98	20.33	86.85
	2014/18	30.76	10.37	9.58	56.17	49.89	14.16	18.72	83.97	48.62	14.63	14.65	81.47
% population non-white	2008/12	26.33	9.71	3.45	61.15	48.6	10.67	18.33	74.98	48.86	10.72	29.49	71.67
	2014/18	29.31	11.35	7.76	66.51	49.67	9.04	32.17	71.91	45.85	9.31	27.88	68.49
% rental properties	2008/12	16.75	14.81	0.51	82.84	43.12	16.96	9.11	81.92	46.17	20.98	5.79	90.44
	2014/18	17.03	15.4	0.74	79.28	45.24	18.78	6.82	79.62	47.27	22.48	6.29	87.06
% population non-Hispanic white	2008/12	66.66	11.71	36.64	93.75	40.32	12.09	12.11	71.01	38.45	11.78	15.27	64.17
	2014/18	63.6	12.66	31.38	88.99	36.08	12.52	6.81	53.58	38.92	12.04	12.1	64.83
Median gross rent ($)	2008/12	2,024.73	319.49	211.47	2,192.50	1,771.07	262.48	1,191.03	2,192.50	1,662.09	284.01	667.28	2,192.50
	2014/18	2,388.09	614.71	241.00	3,501.00	1,823.65	282.89	1,098.00	2,548.00	1,838.62	346.57	672.00	2,704.00
Median property value ($)	2008/12	615,281.20	186,914.51	279,514.01	1,095,704.80	393,138.49	86,783.47	212,456.95	643,835.50	405,962.61	74,647.01	243,794.07	564,616.12
	2014/18	633,144.37	209,075.52	251,900.00	1,444,200.00	392,895.83	99,195.14	165,000.00	689,300.00	458,172.00	79,232.02	274,200.00	665,700.00
% tracts where change in % population with BA degree > county median change	2008/12 - 2014/18	47				17				86			
% tracts where % change in median household income > county median change	2008/12 - 2014/18	43				40				70			
% tracts where mean change in % population non-Hispanic white > county median change	2008/12 - 2014/18	48				48				74			
% tracts where % change in median gross rent > county median change	2008/12 - 2014/18	65				42				64			
% tracts where % change in median property value > county change	2008/12 - 2014/18	37				38				84			

Notes: BA = Bachelor’s degree; % = Percent; $ = Inflation-adjusted United States dollars. > = Is greater than.
Data source: American Community Survey 5 year estimates, 2008/12 and 2014/18.

Table 3. Census Tract-Level Descriptive Statistics for Model Predictors by Tract Gentrification Status.

	Not Vulnerable N = 151				Vulnerable, did not Gentrify N = 48				Vulnerable, Gentrified N = 50			
	mean	*sd*	*min*	*max*	*mean*	*sd*	*min*	*max*	*mean*	*sd*	*min*	*max*
Property characteristics												
Difference in % multi-unit residential properties	0.28	3.92	-10.82	16.39	0.90	6.42	-9.58	18.57	-0.52	6.18	-14.46	14.29
Difference in % vacant housing	0.04	3.30	-12.48	12.31	-0.52	4.89	-11.75	8.51	-1.05	5.06	-13.37	9.56
Difference in % single family properties	-0.40	4.02	-16.11	11.99	-1.17	6.20	-18.15	11.60	0.38	5.79	-15.99	13.54
% change in median property value	2.66	8.18	-23.51	31.81	0.02	12.81	-51.66	24.52	14.51	20.63	-16.02	127.53
Difference in % in renter-occupied units	0.28	5.64	-17.60	13.21	2.13	7.75	-12.97	23.10	1.09	6.95	-13.94	15.77
% change in median gross rent	18.97	32.18	-44.58	281.94	3.18	8.30	-15.40	22.66	11.88	22.60	-16.87	149.37
% change in housing density	-0.28	4.35	-10.01	24.26	3.31	11.91	-8.98	62.88	2.32	14.07	-9.14	82.05
Population characteristics												
Difference in % rent burdened residents	-1.21	25.20	-100.00	100.00	0.28	14.45	-42.64	26.74	-0.47	14.77	-46.92	31.68
Difference in % living in different house than 1 year ago	0.14	5.23	-22.25	13.22	0.24	6.13	-14.64	21.37	1.29	7.79	-16.54	18.50
Difference in % taking public transit	0.77	3.81	-15.72	16.84	-0.72	5.23	-18.34	11.44	0.55	5.18	-13.89	13.24
Difference in % unemployed	-0.53	2.70	-7.56	10.01	-0.93	3.40	-10.01	5.17	-1.58	3.42	-11.26	5.58
Difference in % in poverty	0.19	3.37	-19.72	9.32	1.06	6.56	-15.78	14.07	-0.29	5.01	-17.19	14.85
% change in median household income	1.50	12.52	-23.12	54.34	-1.69	14.43	-40.64	42.21	10.66	18.47	-26.20	82.57
Difference in % non-family households	0.12	5.35	-20.70	13.01	-1.73	7.40	-16.44	16.08	-2.88	6.36	-15.27	12.91
Difference in % population with Bachelor’s degree	2.61	5.87	-14.11	16.36	-0.25	5.16	-11.46	12.50	6.76	6.61	-13.21	24.57
Difference in % non-Hispanic white population	-3.06	6.99	-23.78	17.96	-4.25	8.93	-28.61	11.81	0.47	7.12	-17.71	17.41
Population growth	3.13	8.97	-16.88	34.06	10.81	12.60	-17.29	43.61	9.16	15.48	-13.22	69.44

Notes: Difference = Difference between 2014/18 and 2008/12 estimate. % = Percent.
Data source: American Community Survey 5-year estimates, 2008/12 and 2014/18.

Table 4. Logistic Link Sparse Spatial Generalized Linear Model Results Predicting Tract Vulnerable and Gentrified Status.

	Posterior Mean	2.5% Credible Interval	97.5% Credible Interval
(Intercept)	-3.63	-5.42	-2.61
Property characteristics			
Difference in % multi-unit residential properties	0.14	-0.10	0.40
Difference in % vacant housing	-0.04	-0.20	0.10
Difference in % single family properties	0.19	-0.07	0.46
% change in median property value	**0.13**	**0.08**	**0.21**
Difference in % in renter-occupied units	0.01	-0.07	0.09
% change in median gross rent	-0.01	-0.03	0.01
% change in housing density	-0.01	-0.10	0.08
Population characteristics			
Difference in % rent burdened residents	0.00	-0.02	0.02
Difference in % living in different house than prior year	0.04	-0.04	0.12
Difference in % taking public transit	0.00	-0.11	0.12
Difference in % unemployed	-0.12	-0.28	0.03
Difference in % in poverty	0.00	-0.11	0.12
% change in median household income	0.01	-0.03	0.05
Difference in % non-family households	-0.08	-0.17	0.01
Difference in % population with Bachelor's degree	**0.18**	**0.09**	**0.30**
Difference in % non-Hispanic white population	**0.11**	**0.04**	**0.18**
Population growth	0.03	-0.04	0.10

Note: Statistically significant predictors bolded to facilitate interpretation % = Percent.

Table 5. Logistic Link Sparse Spatial Generalized Linear Model Results Predicting Tract Vulnerable but Not Gentrified Status.

	Posterior Mean	2.5% Credible Interval	97.5% Credible Interval
(Intercept)	-1.48	-2.14	-0.89
Property characteristics	0.10	-0.09	0.29
Difference in % multi-unit residential properties	0.03	-0.09	0.15
Difference in % vacant housing	0.15	-0.05	0.36
Difference in % single family properties	-0.07	-0.12	-0.03
% change in median property value	**0.02**	**-0.04**	**0.09**
Difference in % in renter-occupied units	-0.05	-0.08	-0.02
% change in median gross rent	**0.00**	**-0.08**	**0.07**
% change in housing density	0.00	-0.02	0.02
Population characteristics	-0.02	-0.09	0.05
Difference in % rent burdened residents	-0.05	-0.14	0.04
Difference in % living in different house than prior year	-0.10	-0.24	0.04
Difference in % taking public transit	0.02	-0.07	0.12
Difference in % unemployed	-0.02	-0.06	0.01
Difference in % in poverty	-0.05	-0.12	0.02
% change in median household income	-0.09	-0.17	-0.02
Difference in % non-family households	-0.03	-0.09	0.02
Difference in % population with Bachelor's degree	**0.05**	**0.00**	**0.11**
Difference in % non-Hispanic white population	-1.48	-2.14	-0.89
Population growth	0.10	-0.09	0.29

Note: Statistically significant predictors bolded to facilitate interpretation % = Percent.

Table 6. Number of Tracts by Status Determined with Classification, Model Prediction, and Post-Intervention Model Prediction.

Tract Type/Model	*Classification*	*Model prediction*	*Post-intervention model prediction*
Not vulnerable to gentrification	151	162	189
Vulnerable but not gentrified	48	53	43
Vulnerable and gentrified	50	34	17
Total	*249*	*249*	*249*

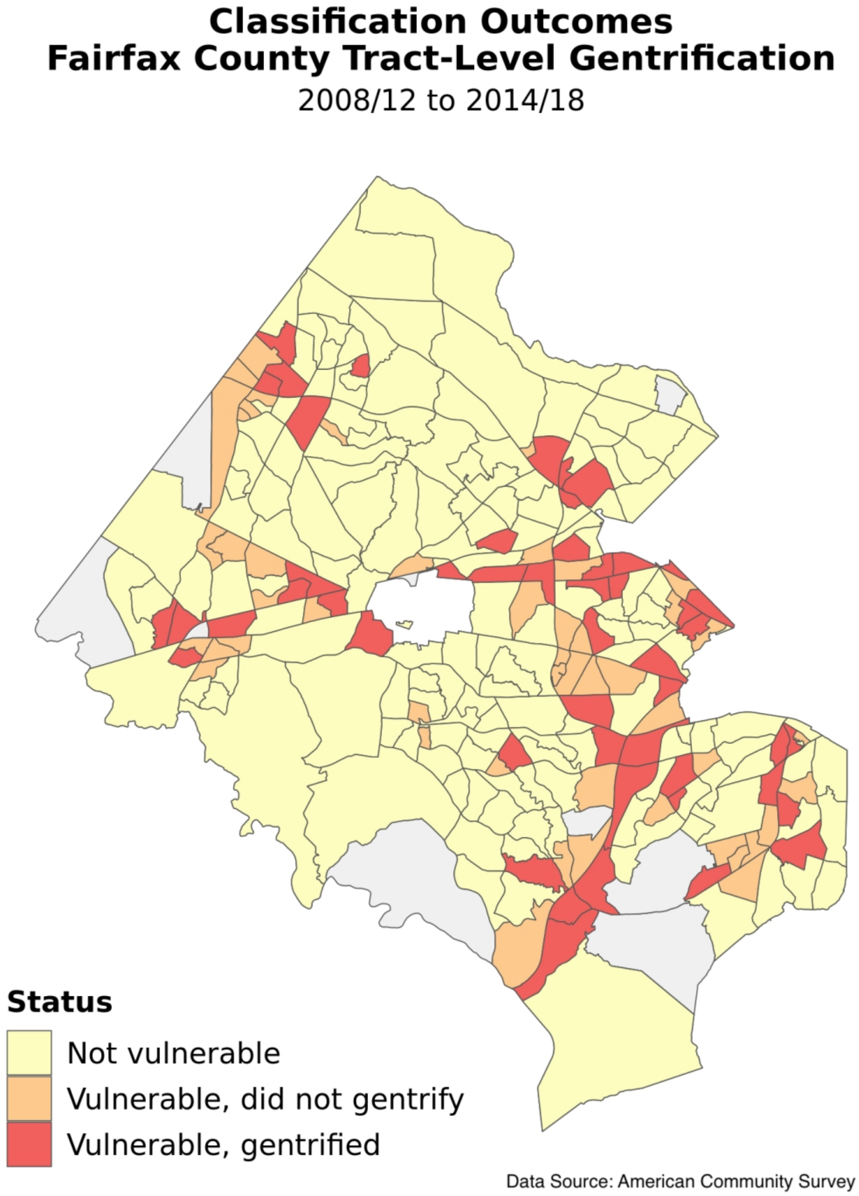

Figure 1. Fairfax County Census Tract Gentrification Status

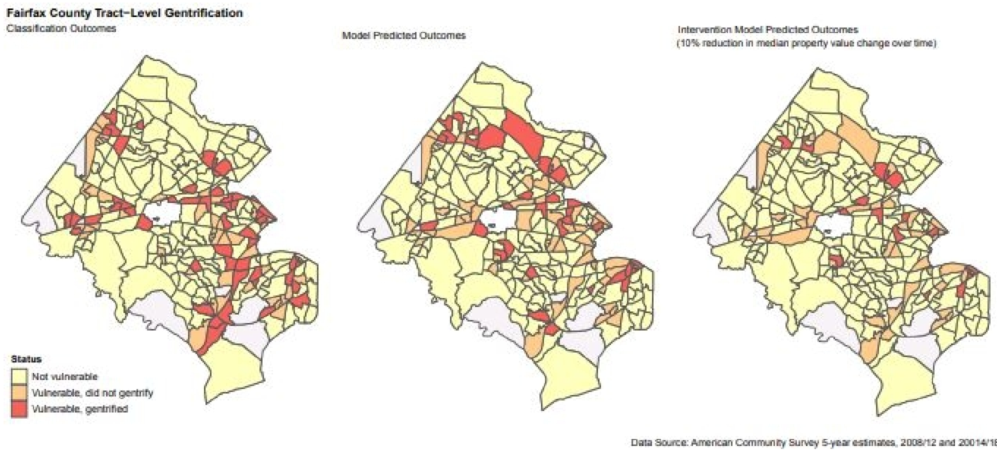

Figure 2: Typology classified, model-predicted, and post-intervention tract gentrification status

180x80mm (96 x 96 DPI)

SUPPLEMENTARY MATERIAL (FOR ONLINE PUBLICATION ONLY)

Property- and Population-Level Independent Variables

Property Characteristics

Difference in percent multi-unit residential properties over time. We sum the number of housing units in a tract with five or more units from table B25024 (Units in Structure), divide by the total number of housing units, and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent residential properties between baseline and observation period end-year.

Difference in percent vacant housing over time. We divide the number of vacant units by the total number of housing units from table B25002 (Occupancy Status) and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent vacant housing between baseline and observation period end-year.

Difference in percent single family properties over time. We divide the number of single family detached units by the total number of housing units from table B25024 (Units in Structure) and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent single family housing between baseline and observation period end-year.

Percent change in median property value over time. We retrieve median property value in dollars for owner-occupied housing units from table B25077 (Median Value in Dollars). We create the variable using both 2008/12 and 2014/18 estimates, adjust all values to 2018 constant dollars using Annual Average Consumer Price Index Research Series data, and calculate percent change between the two time periods.

Difference in percent in renter-occupied units over time. We divide the number of renter-occupied housing units by the total number of housing units from table B25003 (Tenure) and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent renters between baseline and observation period end-year.

Percent change in median gross rent over time. We retrieve tract-level information on monthly median gross rent in dollars for renter-occupied housing units from ACS table B25064 (Median Gross Rent). We create the variable using both 2008/12 and 2014/18 estimates, adjust all values to 2018 constant dollars using Annual Average Consumer Price Index Research Series data, and calculate percent change between the two time periods.

Percent change in housing density over time. We retrieve the total number of housing units in a tract from table B25024 (Units in Structure) from 2008/12 and 2014/18 estimates, and calculate the percent change over time.

Population Characteristics

Difference in percent in poverty over time. We divide the number of individuals below 100 of the poverty level by the total number of individuals from table B06012 (Place of Birth by Poverty Status in the Past 12 Months in the United States) and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent in poverty between baseline and observation period end-year.

Difference in percent rent burdened over time. We sum the number of renter-occupied housing units paying 35% or more of household income for rent in the past 12 months, divide by the total number of renter-occupied housing units from table B25070 (Gross Rent as a Percentage of Household Income in the Past 12 Months), and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent rent burdened population between baseline and observation period end-year.

Difference in percent living in different house than 1 year ago over time. We sum the number of males and females who moved within county, across counties, across states, or from abroad, divide by the total population aged 1 year or over from table B07003 (Geographical Mobility in the Past Year by Sex for Current Residence in the United States), and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent living in different house than in the previous year between baseline and observation period end-year.

Difference in percent taking public transit over time. We sum the number of workers taking public transportation (excluding taxicab) to work, divide by the total number of workers age 16 years and over from table B08101 (Means of Transportation to Work by Age), and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent taking public transit between baseline and observation period end-year.

Difference in percent unemployed over time. We divide the number of unemployed individuals by the number of individuals aged 16 years and over who are in the labor force from table B23025 (Employment Status for the Population 16 Years and Over), and multiply by 100. We calculate the variable using both 2008/12 and 2014/18 estimates, and subtract the calculated values to obtain the difference in tract percent unemployed between baseline and observation period end-year.

Difference in percent population with Bachelor's degree over time. We sum tract-level educational attainment categories from Bachelor's degree through Doctorate degree from table B15003 (Educational Attainment for the Population 25 Years and Over), divide by the total number of individuals, and multiply by 100 to calculate percent population aged over 25 with at least a Bachelor's degree. We create the variable using both 2008/12 and 2014/18 estimates and subtract the calculated values to obtain the difference in tract percent population with a Bachelor's degree between baseline and observation period end-year.

Difference in percent non-Hispanic white population over time. We divide the number of non-Hispanic or Latino white individuals from table B03002 (Hispanic or Latino Origin by Race), divide by the number of all individuals, and multiply by 100 to calculate percent non-Hispanic white population in a tract. We calculate the variable using 2008/12 and 2014/18 estimates and subtract base year values from end year values to obtain percent change over time.

Difference in percent non-family households over time. We divide the number of non-family households by the total number of households in a tract from table B11001 (Household Type Including Living Alone) and multiply by 100. We calculate the variable using 2008/12 and 2014/18 estimates and subtract base year values from end year values to obtain the difference in percent non-family households over time.

Percent change in median household income over time. We retrieve tract-level median household income information from ACS table B19013 (Median Household Income). We create the variable using both 2008/12 and 2014/18 estimates, adjust all values to 2018 constant dollars using Annual Average Consumer Price Index Research Series data, and calculate percent change between the two time periods.

Population growth over time. We retrieve the total number of individuals residing in a tract from table B01003 (Total Population) from 2008/12 and 2014/18 estimates, and calculate the percent change over time.