

# The Influence of Tobacco Retailer Density and Poverty on Tobacco Use in a Densely Populated Urban Environment

Shannon M. Farley, DrPH<sup>1</sup>; Andrew R. Maroko, PhD<sup>2</sup>; Shakira F. Suglia, ScD<sup>3</sup>; and Lorna E. Thorpe, PhD, MPH<sup>4</sup>

Public Health Reports  
2019, Vol. 134(2) 164-171  
© 2019, Association of Schools and  
Programs of Public Health  
All rights reserved.  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0033354918824330  
journals.sagepub.com/home/phr



## Abstract

**Objectives:** Researchers have identified associations between neighborhood-level factors (eg, income level, tobacco retailer density) and smoking behavior, but few studies have assessed these factors in urban environments. We explored the effect of tobacco retailer density, neighborhood poverty, and housing type (multiunit and public) on smoking in a large urban environment (New York City).

**Methods:** We used data on smoking prevalence and individual sociodemographic characteristics from the 2011-2013 New York City Community Health Survey, data on tobacco retailers from the 2012 New York City Department of Consumer Affairs, data on neighborhood sociodemographic characteristics and population density from the 2009-2013 American Community Survey, and data on multiunit and public housing from the 2012 New York City Primary Land Use Tax Lot Output data set. We used aggregate neighborhood-level variables and ordinary least squares regression, geographic weighted regression, and multilevel models to assess the effects of tobacco retailer density and neighborhood poverty on smoking prevalence, adjusting for sociodemographic characteristics (age, sex, race/ethnicity, and education) and neighborhood population density. We also assessed interactions between tobacco retailer density and poverty and each housing type on smoking.

**Results:** Neighborhood poverty positively and significantly modified the association between tobacco retailer density and prevalence of neighborhood smoking ( $\beta = 0.003$ ,  $P = .01$ ) when we controlled for population density, sociodemographic characteristics, and types of housing. Neighborhood poverty was positively associated with the prevalence of individual smoking ( $\beta = 0.0099$ ,  $P < .001$ ) when we adjusted for population density, sociodemographic characteristics, and type of housing.

**Conclusion:** More research is needed to determine all the environmental factors associated with smoking prevalence in a densely populated urban environment.

## Keywords

smoking, tobacco retailer density, multiunit housing, urban environment, urban health

Although smoking prevalence has declined for years in the United States, the decline has been inconsistent across geographic areas.<sup>1</sup> In 2013, the prevalence of smoking ranged from 10.3% in Utah to 27.3% in West Virginia.<sup>1</sup> Smoking prevalence in smaller geographic areas also varied. In New York State, smoking prevalence ranged from 10.2% in Rockland County to 30.6% in Cayuga County during 2013-2014, and within-county variation by neighborhoods was likely.<sup>2</sup> Neighborhood income has also been correlated with smoking; studies showed that residents of low-income neighborhoods were more likely than residents of high-income neighborhood to be smokers.<sup>3-7</sup> Hypothesized mechanisms explaining this association included varying social norms, enforcement of smoking regulation, tobacco advertising

density, and individual-level factors such as psychosocial stress.<sup>3-7</sup> However, findings were not consistent. For

<sup>1</sup> Bureau of Chronic Disease Prevention and Tobacco Control, New York City Department of Health and Mental Hygiene, Long Island City, NY, USA

<sup>2</sup> Urban GISc Laboratory, Lehman College, Bronx, NY, USA

<sup>3</sup> Rollins School of Public Health, Emory University, Atlanta, GA, USA

<sup>4</sup> Department of Population Health, New York University, New York, NY, USA

## Corresponding Author:

Shannon M. Farley, DrPH, New York City Department of Health and Mental Hygiene, Bureau of Chronic Disease Prevention and Tobacco Control, 42-09 28th St, CN-46, Long Island City, NY 11101, USA.  
Email: sfarley@health.nyc.gov

example, a study in New York City found no association between neighborhood income and smoking after adjusting for individual-level covariates.<sup>8</sup>

Other environmental factors hypothesized to influence smoking have included density of neighborhood tobacco retailers and distance to tobacco retailers. Several cross-sectional studies found positive associations between retailer density and smoking.<sup>9-12</sup> However, in several cases, results attenuated after adjusting for sociodemographic characteristics, particularly neighborhood income.<sup>9-11</sup> Distinguishing whether observed neighborhood predictors of smoking behavior were causal drivers or reflected residual confounding by population composition was a methodologic challenge.

Studies examining the potential role of neighborhood attributes on smoking have either examined correlations between neighborhood-level attributes and smoking prevalence at an ecological level or used multilevel regression models to account for clustering and adjust for neighborhood- or individual-level confounding.<sup>3-12</sup> Although using multilevel regression models avoids potential ecological fallacy interpretations when they use aggregated data, they require larger sample sizes and interpret risk only at the individual level. Both study types are potentially influenced by the range of variability among neighborhood attributes in the populations examined. The conflating of urban versus rural environments may contribute to the inconsistency of findings.

Several studies identified positive associations between living in urban environments and tobacco retailer density, but smoking prevalence was similar across urban and rural jurisdictions.<sup>3,7,13</sup> One study in New York City that examined the same data by using ecological models and multilevel regression models found no association between neighborhood income and smoking.<sup>8</sup> To our knowledge, no studies have explored the effect of tobacco retailer density on smoking in densely populated urban areas such as New York City, where residential and retail locations are likely distributed in a clustered manner. The association between tobacco retailer density and smoking might also be modified by the type of housing (multiunit or public) in neighborhoods or other factors that capture variability in residential interaction with tobacco retailers, as shown in studies that examined the relationship between density of alcohol retailers and urban violence.<sup>14-16</sup> Although the prevalence of smoking in New York City was relatively low (13.3%) in 2013-2014, a high density of housing and tobacco retailers may influence the prevalence of smoking locally rather than statewide.<sup>2,17</sup> Several studies suggested that income or race/ethnicity may modify the relationship between retailer density and smoking.<sup>9,10,18</sup>

With both a wide variety of housing types and more than 9000 tobacco retailers that expose New York City residents to tobacco purchasing opportunities, New York City is a prime setting for examining urban environmental influences on smoking. We explore the effect of tobacco retailer density, neighborhood poverty, and housing types (multiunit and public) on smoking in New York City.

## Methods

### Measures and Data Sources

**Neighborhoods.** New York City has 59 community districts (rounded mean population per district = 142 000). Community districts are an administrative neighborhood designation determined by the New York City Department of City Planning. In our study, community districts are referred to as neighborhoods, the smallest available geographic unit of analysis for examining smoking prevalence, and they have been used in previous studies.<sup>1,8,19</sup>

**Outcome measures.** The outcome measures were neighborhood smoking prevalence and individual smoking behavior among respondents to the 2011-2013 New York City Community Health Survey (CHS).<sup>20</sup> The CHS is an annual cross-sectional telephone survey on the health of adults aged  $\geq 18$  in New York City that is modeled on the Behavioral Risk Factor Surveillance System. CHS data are weighted to account for probability of selection and post-stratification weights. Post-stratification weights are created by weighting the sample to the community district populations, by age, sex, and race/ethnicity. Responses are weighted to account for the adult population distribution comprising 3 telephone usage categories (landline only, landline and cell phone, and cell phone only) by using data from the New York City Housing and Vacancy Survey.<sup>20</sup> Current smoking prevalence was based on a response of yes to 2 questions: (1) Have you ever smoked more than 100 cigarettes? and (2) Do you now smoke every day or some days? The outcome for the ecological analyses was neighborhood smoking prevalence; that is, the percentage of adult smokers per community district determined from aggregated data on self-reported smoking behavior among respondents to the 2011-2013 CHS. The outcome for the multilevel analysis was individual-level smoking behavior.

**Exposure measures.** The exposures of interest were tobacco retailer density and neighborhood poverty. The New York City Department of Consumer Affairs provided 2012 New York City tobacco retailer listings.<sup>21</sup> Of 9787 licensed tobacco retailers, we excluded 309 because of invalid addresses and 19 because they were located at airports, leaving 9459 licensed tobacco retailers with valid New York City addresses geocoded. We mapped licensed tobacco retailer locations and calculated the number of tobacco retailers per neighborhood. We converted tobacco retailer locations into a statistical surface by using a mathematical kernel function that weights retailers on the basis of proximity to each other. This process reduces "edge effects," whereby retailers on the boundary of a neighborhood will only be represented in the neighborhood into which they fall, and creates a more nuanced measure of exposure to tobacco retailers than the number of tobacco retailers per neighborhood. We used kernel density estimate (KDE) to measure tobacco retailer accessibility, turning map points into a smooth surface to provide a visual representation of retailer density. By using

KDE, we created a 50-meter raster surface showing retailer density in New York City, addressing edge effects.<sup>22</sup> We assigned the KDE bandwidth a 1-mile radius. Bandwidth defines the spatial boundaries for the geographic weighted regression (GWR) calculations. We conducted KDE sensitivity analyses by using ¼-, ½-, and 1-mile radii for tobacco retailers on neighborhood smoking prevalence. The 1-mile bandwidth explained more variance than the other radii, with a lower Akaike information criterion measure from GWR. In addition, we considered 1 mile to be a reasonable walking distance to a tobacco retailer.<sup>23</sup> We converted KDE output estimates into a density-per-square-mile scale for ease of interpretation in ecological and multilevel analyses.

We defined neighborhood poverty as the percentage of adults aged  $\geq 18$  who were living at  $< 100\%$  of the federal poverty level (FPL) per neighborhood. We used data from the 2011-2013 CHS to determine percentage of neighborhood poverty.

**Covariates.** The covariates included population density, sex, race/ethnicity, education, age, and multiunit and public housing. We used data on population density from the 2009-2013 American Community Survey (ACS) for all analyses.<sup>24</sup> We calculated population density by dividing the census unit population by the census unit area.<sup>25</sup> We divided the total ACS adult population (aged  $\geq 18$ ) by the US Census square mileage per neighborhood, creating a population density layer. We also used ACS data on the percentage of young adults aged 18-24 in each neighborhood for ecological analyses. Additional covariates included percentage of residents in each neighborhood who were female, non-Hispanic black, Hispanic, and  $\geq$  college graduate in ecological analyses. CHS individual-level variables used in multilevel regression analyses were race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic Asian, non-Hispanic other, and Hispanic) and age (18-24, 25-44, 45-64, and  $\geq 65$ ).

We focused on 2 types of housing: multiunit housing and New York City Housing Authority (NYCHA) housing. We established these categories on the basis of the New York City Department of City Planning's 2012 New York City Primary Land Use Tax Lot Output (PLUTO) data, which combines data on land use and geography at the tax lot level.<sup>26</sup> PLUTO contains information on the number of residential units that were used to determine whether a building is considered multiunit. We used the definition of multiunit from the New York City Housing and Vacancy Survey, which considers  $\geq 3$  units to be multiunit.<sup>27</sup> We determined that a building was public housing when the building owner's name in PLUTO was listed as NYCHA. We calculated the percentage of multiunit housing and the percentage of NYCHA housing in each neighborhood for all analyses. The numerator for multiunit housing was the percentage of buildings in PLUTO with  $\geq 3$  units, the numerator for NYCHA housing was the percentage of buildings in PLUTO listed as NYCHA, and the denominator for both was all residential buildings.

## Statistical Analyses

**Preliminary analyses.** We aggregated and mapped data on tobacco retailer density, neighborhood poverty, population density, and housing type (multiunit or NYCHA) at the neighborhood level by using Jenks Natural Breaks classification, which uses natural groupings within the data to identify the best breaks for grouping data.<sup>28</sup> We assessed tobacco retailers spatially by using KDE, and we aggregated the smooth surface estimates to the neighborhood level by turning the spatial estimates into statistics that could be used in regression analyses, thereby standardizing all geographic units. We used a Global Moran's *I* test for data clustering, randomly distributed or dispersed in space, to assess the spatial distribution of smoking prevalence by neighborhood.

**Ecological analyses.** We used aggregate neighborhood-level variables and ordinary least squares (OLS) regression to assess the effects of tobacco retailer density and neighborhood poverty on smoking prevalence in unadjusted and adjusted analyses. Adjusted models included sociodemographic characteristics (age, sex, race/ethnicity, and education) and neighborhood population density. A fully adjusted OLS model assessed the association of tobacco retailer density and neighborhood poverty on neighborhood smoking prevalence, adjusting for population density and neighborhood sociodemographic characteristics. We assessed the potential for effect modification in all final models. We examined tobacco retailer density and neighborhood poverty interactions across all models, and we assessed interactions between tobacco retailer density and each housing type. We assessed significant correlations between model covariates by using the variance inflation factor, a diagnostic test for multicollinearity. We removed covariates with a variance inflation factor  $\geq 7.5$  in a stepwise manner to determine final adjusted models.

We evaluated models with GWR, accounting for locally varying spatial relationships.<sup>22,29-34</sup> GWR conducts multiple local regressions, adjusting each regression point with nearby data. GWR can account for spatial non-stationarity, allowing the model to vary locally, because some relationships may vary spatially.<sup>30</sup> Spatial non-stationarity is when model coefficients are not fixed over space, and associations fluctuate based on location. Significant correlations between GWR model covariates led to diagnostic tests for multicollinearity, and we adjusted model covariates. We determined ecological model best fit by using the Akaike information criterion, the best indicator of GWR model performance.<sup>35</sup>

**Multilevel analyses.** We calculated a pseudo-intraclass correlation coefficient to estimate variability between neighborhoods.<sup>36</sup> We examined neighborhood-level influences on individual-level smoking by using multilevel regression models. Individual-level variables were smoking, age, race/ethnicity, sex, income, and education. Multilevel regression models account for non-independence of observations within

groups, the non-independence of errors, and correlations among individual-level variables nested within neighborhoods. Neighborhood-level variables aggregated into community districts were retailer density, poverty, population density, multiunit housing density, and NYCHA housing density. First, an empty model (ie, a model with no independent variables) characterized neighborhood variability in smoking. Next, we examined main exposures and adjusted for neighborhood-level factors. Then we included individual-level sociodemographic characteristics to reduce confounding. Finally, we examined possible effect modifications between tobacco retailer density and neighborhood poverty and each housing type. We did not spatially enable multi-level regression models.

We conducted the mapping, KDE, OLS regression, and GWR by using ArcGIS version 10.0.<sup>37</sup> We conducted all ecological analyses by using SAS version 9.2 and SUDAAN version 11.0.1, and we conducted all multilevel regression model analyses by using HLM version 7.<sup>38-40</sup>

## Results

### Descriptive Findings

Poverty level, population density, density of multiunit housing, and density of NYCHA housing varied across neighborhoods. The average percentage of the population living in poverty per neighborhood was 24.8% (range, 4.3%-47.5%) (Figure 1A). Population density ranged from 5000 to 97 000 persons per square mile (Figure 1B). The average percentage of multiunit housing was 17.9% (range, <1.0%-91.9%) (Figure 1C). The average percentage of NYCHA housing was 0.1% (range, 0%-2.5%) (Figure 1D).

The smoking prevalence in New York City overall was 15.4%; smoking prevalence by neighborhood ranged from 10.0% to 22.6% (Figure 1E). The number of tobacco retailers per neighborhood ranged from 76 to 365 (Figure 1F). The KDE map (Figure 1G) of licensed tobacco retailers showed high-density patterns of tobacco retailers in Manhattan, South Bronx, and parts of Brooklyn and Queens, which were similar to the patterns generated by counting the number of retailers (Figure 1E), except in Staten Island.

The Global Moran's *I* test (index = 0.16;  $P = .004$ ) indicated that neighborhood smoking prevalence was spatially clustered, and the possibility that the spatial pattern was due to random chance was <1%.

### Ecological Analyses

**OLS regression.** An adjusted interaction model between tobacco retailer density and neighborhood poverty found that for each percentage-point increase in poverty, the association between tobacco retailer density and smoking prevalence increased by 0.003 ( $P = .01$ ) (Table). The interaction model was significant overall ( $P = .02$ ) and a better fit than the adjusted model (Akaike information criterion = 303.97 vs

309.41;  $R^2 = 16\%$  vs 5%). Adding type of housing did not improve model fit. Smoking prevalence increased with increased tobacco retailer density and neighborhood poverty, but the marginal effects of tobacco retailers on smoking prevalence at 2 poverty levels (low and high) were not significant ( $P = .86$ ) (Figure 2). The fully adjusted GWR model with the tobacco retailer–poverty interaction on smoking had the same Akaike information criterion and  $R^2$  values as those of the OLS model, indicating no improvement in model fit when allowing for potential spatial non-stationarity at the neighborhood level. GWR bandwidth was large, indicating that the model operated as a global model rather than a local model.

### Multilevel Regression Analyses

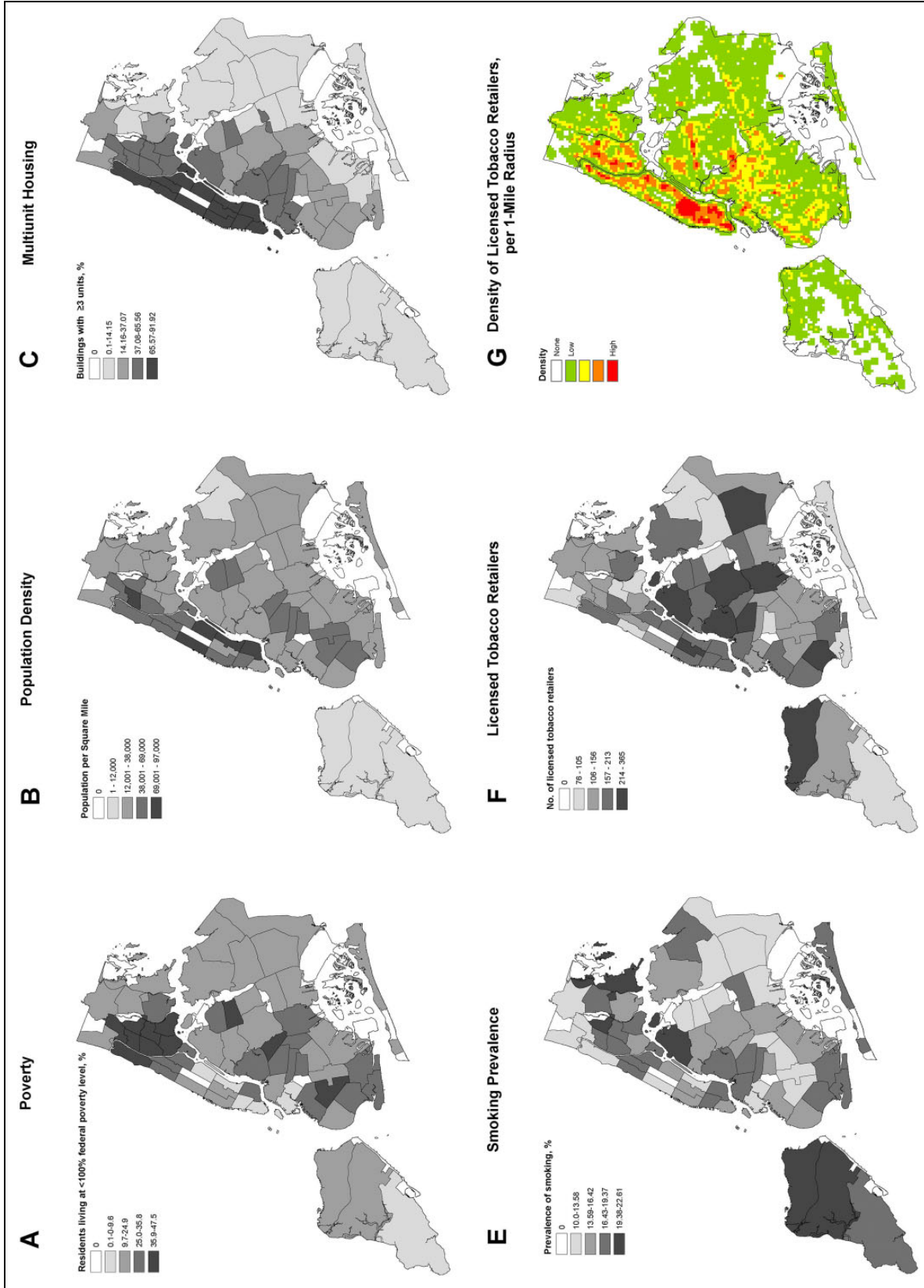
In the examination of the associations between neighborhood poverty and tobacco retailer density and individual-level smoking behavior, the pseudo-intraclass correlation estimated that neighborhood-level variance was 1%, indicating low correlation among persons residing in the same neighborhood. In the fully adjusted multilevel regression model examining the effect of tobacco retailer density and neighborhood poverty on individual smoking behavior, adjusting for population density and individual sociodemographic characteristics, the tobacco retailer–smoking association was positive and not significant ( $\beta = 0.0013$ ; SE = 0.0001;  $P = .19$ ) and the poverty–smoking association was positive and significant ( $\beta = 0.0099$ ; SE = 0.002;  $P < .001$ ). The tobacco retailer–poverty interaction was not significant ( $P = .13$ ). Additional models found that housing type did not have a strong effect on smoking.

## Discussion

We found that the interaction of retailer density and poverty positively predicted neighborhood smoking prevalence, suggesting that communities may have differing vulnerabilities to the presence and marketing of tobacco products, depending on income. Type of housing did not independently predict smoking prevalence or behavior, nor significantly influence other observed associations.

Neighborhood poverty had a marginally positive association with smoking prevalence in ecological models and was positively associated with individual smoking behavior in multilevel regression models. Some studies using multilevel regression models identified positive associations between neighborhood income and smoking,<sup>3-7</sup> other studies showed that individual income was a better predictor of smoking than neighborhood income, and yet another study found no neighborhood income association with smoking.<sup>6,7</sup> Our study was unique in finding the significant modifying effect of neighborhood poverty on the retailer–smoking association in ecological analyses. In our study, a higher level of neighborhood poverty was associated with individual smoking behavior in multilevel models.

Patterns of smoking in New York City do not always conform to well-established associations between



**Figure 1.** Distribution of income levels, population density, multiunit housing, public housing, smoking prevalence, and licensed tobacco retailers, by neighborhood, New York City, 2011-2013. (A) Prevalence of low-income residents (ie, residents living at <100% of the federal poverty level), 2011-2013 New York City Community Health Survey.<sup>20</sup> (B) Population density per square mile, 2009-2013 American Community Survey.<sup>24</sup> (C) Percentage of residents living in multiunit housing (ie, ≥ 3 units), 2012 New York City Primary Land Use Tax Lot Output.<sup>26</sup> (D) Percentage of residents living in New York City Housing Authority housing (ie, public housing), 2012 New York City Primary Land Use Tax Lot Output.<sup>26</sup> (E) Percentage of residents who smoke, 2011-2013 New York City Community Health Survey.<sup>20</sup> (F) Number of tobacco retailers, 2012 Department of Consumer Affairs.<sup>21</sup> (G) Density of licensed tobacco retailers per 1-mile radius, 2012 Department of Consumer Affairs.<sup>21</sup>

**Table.** Comparison of adjusted ordinary least squares and geographically weighted regression ecological models examining the effect of licensed tobacco retailer density and poverty on smoking prevalence across neighborhoods, New York City, 2009-2013<sup>a</sup>

Factor	Ordinary Least Squares <sup>b</sup>	Geographic Weighted Regression <sup>c</sup>				
	Coefficient (SE) P Value	Minimum	1st Quartile	Median	3rd Quartile	Maximum
Intercept	16.78 (2.88) [ $<.001$ ]	—	—	—	—	—
Density of licensed tobacco retailers <sup>d</sup>	-0.06 (0.03) [.07]	-0.057	-0.057	-0.057	-0.057	-0.057
Neighborhood poverty <sup>e</sup>	-0.10 (0.09) [.27]	-0.101	-0.101	-0.101	-0.101	-0.110
Interaction between density of licensed tobacco retailers and neighborhood poverty <sup>d,e</sup>	0.003 (0.001) [.01]	0.003	0.003	0.003	0.003	0.003

Abbreviation: —, does not apply.

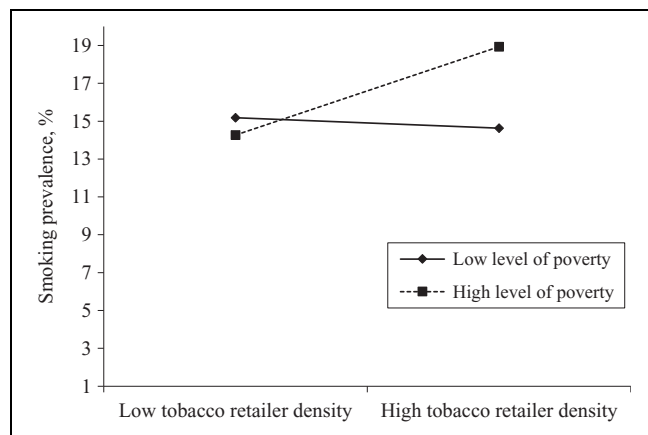
<sup>a</sup>Adjusted models included percentage of residents aged 18-24, percentage black residents, percentage Hispanic residents, and neighborhood population density. Data source: 2009-2013 American Community Survey.<sup>24</sup>

<sup>b</sup> $R^2 = 16\%$ ; Akaike information criterion = 303.97; joint  $F$  statistic = 2.59;  $P = .02$ .

<sup>c</sup> $R^2 = 16\%$ ; Akaike information criterion = 303.98; bandwidth = 297.3 miles. Bandwidth defines the spatial boundaries for the geographic weighted regression calculations.

<sup>d</sup>Data source: New York City Department of Consumer Affairs, 2012.<sup>21</sup>

<sup>e</sup>Defined as the percentage of adults aged  $\geq 18$  who were living at  $<100\%$  of the federal poverty level. "Neighborhood" is defined as a community district. Data source: New York City Community Health Survey 2011-2013.<sup>20</sup>



**Figure 2.** The marginal effects of the association between poverty level and licensed tobacco retailer density on the prevalence of smoking, New York City. Data on tobacco retailers are from the New York City Department of Consumer Affairs, 2012.<sup>21</sup> Data on poverty and smoking prevalence are from the 2011-2013 New York City Community Health Survey.<sup>20</sup> Low-level poverty was defined as the percentage of adults aged  $\geq 18$  who were living at  $<100\%$  of the federal poverty level. High-level poverty was defined as the percentage of adults aged  $\geq 18$  who were living at  $\geq 100\%$  of the federal poverty level. A low density of licensed tobacco retailers was defined as  $<53$  licensed tobacco retailers and a high density as  $>168$  licensed tobacco retailers.  $P$  value for the interaction = .86.

environmental risk factors and smoking that have been described elsewhere. There may be factors unique to densely populated environments that are not comparable with more rural environments, such as the clustered nature of retail establishments and places of residence. Research from Iowa also had unexpected results, finding associations among higher retailer density, increased smoking prevalence, and higher income levels.<sup>12</sup> A previous New York City study found no association between neighborhood median income

or individual income and individual smoking in ecological or multilevel regression analyses.<sup>8</sup> Although that study identified positive results, associations were not significant, potentially reflecting limited statistical power or varying distributions of poverty or smoking at the time. We found that neighborhood income significantly predicted individual smoking in multilevel regression models.

We expanded typical investigations of smoking predictors by accounting for locally varying spatial relationships that may affect associations between environmental factors that influence behaviors. Although GWR generally produced a better fit than OLS, the magnitude of improvement was small, suggesting that GWR is not necessary at this high level of aggregation. Future analyses using smaller neighborhood areas may demonstrate more spatial variance that might better explain environmental effects.

**Strengths and Limitations**

This study had several strengths. First, we examined relationships between neighborhood attributes and smoking prevalence at the neighborhood level and smoking behavior at the individual level. Both were important because the environment can play an important role in shaping individual health outcomes and aggregated group patterns.<sup>41</sup> Second, this study used population-based data from multiple sources and administrative data for spatial and multilevel analyses to assess multiple environmental factors simultaneously, which broadened the scope of the analysis. Third, we used KDE to represent the spatial distribution of tobacco retailers because it is not confined by administrative boundaries, which reduces edge effects. A final strength was that we built analyses on previously published work methodologically and conceptually to contribute information on a densely populated environment.



This study also had several limitations. First, we used cross-sectional data; as such, results assessed associations rather than causal relationships. Second, administrative boundaries (ie, community districts) may not accurately define neighborhoods as residents define their neighborhoods, resulting in misclassification that may introduce bias, because administrative data are based on place of domicile. Third, proximity that is spatially defined on a map may not accurately represent social or personal choices that may affect potential exposures. Using community districts, which are large geographic areas, to represent neighborhoods may not provide enough spatial variation to accurately assess spatial differences in exposures. Heterogeneity likely exists within the neighborhood areas, yet high population and housing density in New York City indicates a shared built environment and socioeconomic characteristics, making the community district a reasonable neighborhood proxy.<sup>1,8,19</sup> Analyses should be examined at smaller levels of geography, such as the ZIP code or census tract level. Fourth, because of the high density of tobacco retailers in business districts and transportation hubs, analyses of relationships with housing may have been confounded. Future analyses should include land use as a potential covariate because land use might be associated with smoking prevalence in densely populated areas. Fifth, the ecological analyses may be subject to ecological fallacy when assumptions about individuals are made from aggregated data; however, all predictors in this study were interpreted in terms of neighborhood prevalence. Sixth, this study assessed relationships between type of housing and retailers; however, exposure to retailers can happen anywhere, not just near home. Seventh, GWR does not provide a *P* value, which made it challenging to assess model significance. Eighth, our models may not have included all covariates needed to explain the associations, which may have resulted in worse model fit. Finally, the neighborhoods may not have been sufficiently numerous to show enough variation to make multilevel regression models useful because the pseudo-intracluster correlation was only 1%, indicating almost no variation within neighborhoods.

## Conclusion

New York City may have some properties unique to densely populated environments that resulted in different observations of neighborhood influences on smoking prevalence and behavior than in other settings. Our study supports the idea that further research into environmental risk factors and smaller neighborhood scales is necessary across jurisdictions to improve our understanding of which environmental factors can be addressed to further reduce smoking prevalence.

## Acknowledgments

The authors thank Drs Alfredo Morabia and Shadi Chamany for their critical review and feedback on this article.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## References

1. Nguyen K, Marshall L, Hu S, Neff L. State-specific prevalence of current cigarette smoking and smokeless tobacco use among adults aged  $\geq 18$  years—United States, 2011–2013 [published erratum appears in *MMWR Morb Mortal Wkly Rep.* 2015; 64(26):727]. *MMWR Morb Mortal Wkly Rep.* 2015;64(19): 532–536.
2. New York State Department of Health, Bureau of Tobacco Control. County level prevalence of smoking among adults. *StatShot.* 2016;9(3):1–2.
3. Ross CE. Walking, exercising, and smoking: does neighborhood matter? *Soc Sci Med.* 2000;51(2):265–274.
4. Datta GD, Subramanian SV, Colditz GA, Kawachi I, Plamer JR, Rosenberg L. Individual, neighborhood, and state-level predictors of smoking among US black women: a multilevel analysis. *Soc Sci Med.* 2006;63(4):1034–1044. doi:10.1016/j.socscimed.2006.03.010
5. Diez-Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol.* 1997;146(1):48–63.
6. Diez-Roux AV, Merkin SS, Hannan P, Jacobs DR, Kiefe CI. Area characteristics, individual-level socioeconomic indicators, and smoking in young adults: the Coronary Artery Disease Risk Development in Young Adults Study. *Am J Epidemiol.* 2003;157(4):315–326.
7. Tseng M, Yeatts K, Millikan R, Newman B. Area-level characteristics and smoking in women. *Am J Public Health.* 2001; 91(11):1847–1850.
8. Galea S, Ahern J, Tracy M, Vlahov D. Neighborhood income and income distribution and the use of cigarettes, alcohol, and marijuana. *Am J Prev Med.* 2007;32(suppl 6):S195–S202. doi: 10.1016/j.amepre.2007.04.003
9. Chuang YC, Cubbin C, Ahn D, Winkleby MA. Effects of neighbourhood socioeconomic status and convenience store concentration on individual level smoking. *J Epidemiol Community Health.* 2005;59(7):568–573. doi:10.1136/jech.2004.029041
10. Reid RJ, Peterson NA, Lowe JB, Hughey J. Tobacco outlet density and smoking prevalence: does racial concentration matter? *Drugs Educ Prev Policy.* 2005;12(3):233–238. doi:10. 1080/09687630500035485
11. Pearce J, Hiscock R, Moon G, Barnett R. The neighbourhood effects of geographical access to tobacco retailers on individual smoking behaviour. *J Epidemiol Community Health.* 2009; 63(1):69–77. doi:10.1136/jech.2007.070656
12. Peterson NA, Lowe JB, Reid RJ. Tobacco outlet density, cigarette smoking prevalence and demographics at the county level

- of analysis. *Subst Use Misuse*. 2005;40(11):1627-1635. doi:10.1080/10826080500222685
13. Rodriguez D, Carlos HA, Adachi-Mejia AM, Berke EM, Sargent JD. Predictors of tobacco outlet density nationwide: a geographic analysis. *Tob Control*. 2013;22(5):349-355. doi:10.1136/tobaccocontrol-2011-050120
  14. Ashe M, Jernigan D, Kline R, Galaz R. Land use planning and the control of alcohol, tobacco, firearms, and fast food restaurants. *Am J Public Health*. 2003;93(9):1404-1408.
  15. Pridemore WA, Grubestic TH. A spatial analysis of the moderating effects of land use on the association between alcohol outlet density and violence in urban areas. *Drug Alcohol Rev*. 2012;31(4):385-393. doi:10.1111/j.1465-3362.2011.00329.x
  16. Livingston M. A longitudinal analysis of alcohol outlet density and assault. *Alcohol Clin Exp Res*. 2008;32(6):1074-1079. doi:10.1111/j.1530-0277.2008.00669.x
  17. Laaksonen M, Rahkonen O, Karvonen S, Lahelma E. Socio-economic status and smoking: analysing inequalities with multiple indicators. *Eur J Public Health*. 2005;15(3):262-269. doi:10.1093/eurpub/cki115
  18. US Department of Health and Human Services. *The Health Consequences of Involuntary Smoking: A Report of the Surgeon General*. Atlanta, GA: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, Center for Health Promotion and Education, Office on Smoking and Health; 1986.
  19. Hyland A, Travers MJ, Cummings KM, Bauer J, Alford T, Wiczorek WF. Tobacco outlet density and demographics in Erie County, New York. *Am J Public Health*. 2003;93(7):1075-1076.
  20. New York City Department of Health and Mental Hygiene. Community Health Survey: methodology. 2019. <https://www1.nyc.gov/site/doh/data/data-sets/community-health-survey-methodology.page>. Accessed December 4, 2018.
  21. New York City Department of Consumer Affairs. Tobacco retail dealers. 2019. <https://www1.nyc.gov/site/dca/businesses/info-cigarette-retail-dealers.page>. Accessed December 4, 2018.
  22. Spencer J, Angeles G. Kernel density estimation as a technique for assessing availability of health services in Nicaragua. *Health Serv Outcomes Res Methods*. 2007;7(3-4):145-157.
  23. Maroko AR, Maantay JA, Sohler NL, Grady KL, Arno PS. The complexities of measuring access to parks and physical activity sites in New York City: a quantitative and qualitative approach. *Intl J Health Geogr*. 2009;8:34. doi:10.1186/1476-072X-8-34
  24. US Census Bureau. American Community Survey 1-year estimates. 2013. [http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS\\_13\\_1YR\\_DP05&prodType=table](http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_1YR_DP05&prodType=table). Accessed January 3, 2015.
  25. Hall SA, Kaufman JS, Ricketts TC. Defining urban and rural areas in U.S. epidemiologic studies. *J Urban Health*. 2006;83(2):162-175. doi:10.1007/s11524-005-9016-3
  26. New York City Department of City Planning. PLUTO and Map-PLUTO. 2018. <https://www1.nyc.gov/site/planning/data-maps/open-data/dwn-pluto-mappluto.page>. Accessed December 4, 2018.
  27. Lee MW. *Housing New York City, 2011*. New York: New York City Department of Housing Preservation and Development; 2013. <https://www1.nyc.gov/assets/hpd/downloads/pdf/HVS-report-2011.pdf>. Accessed November 14, 2018.
  28. ArcGIS Pro. Data classification methods. 2018. <http://pro.arcgis.com/en/pro-app/help/mapping/layer-properties/data-classification-methods.htm>. Accessed December 4, 2018.
  29. Maantay JA. Asthma and air pollution in the Bronx: methodological and data considerations in using GIS for environmental justice and health research. *Health Place*. 2007;13(1):32-56. doi:10.1016/j.healthplace.2005.09.009
  30. Maroko AR, Maantay JA, Grady K. Using geovisualization and geospatial analysis to explore respiratory disease and environmental health justice in New York City. In: Maantay JA, McClafferty S, eds. *Geospatial Analysis of Environmental Health*. New York: Springer; 2011:39-66.
  31. Chakraborty J, Maantay JA, Brender JD. Disproportionate proximity to environmental health hazards: methods, models, and measurement. *Am J Public Health*. 2011;101(suppl 1):S27-S36. doi:10.2105/AJPH.2010.300109.
  32. Spatial Structures in the Social Sciences. Spatial regression with GeoDa. S4 training modules. <https://slideplayer.com/slide/11685423>. Accessed December 4, 2018.
  33. ESRI. ArcGIS resources, ArcGIS help 10.1: interpreting GWR results. 2013. [http://resources.arcgis.com/en/help/main/10.1/index.html#/Interpreting\\_GWR\\_results/005p0000032000000](http://resources.arcgis.com/en/help/main/10.1/index.html#/Interpreting_GWR_results/005p0000032000000). Accessed May 16, 2013.
  34. Mennis J. Mapping the results of geographically weighted regression. *Cartographic J*. 2006;43(2):171-179.
  35. Fotheringham AS, Brunson C, Charlton M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. West Sussex, England: John Wiley & Sons Ltd; 2002.
  36. Browne WJ. Partitioning of variance in multilevel models. University of Nottingham. <http://seis.bris.ac.uk/~frwjb/materials/wbvp.pdf>. Accessed September 1, 2017.
  37. ESRI. *ArcMap Version 10.0 [software]*. Redlands, CA: ESRI; 2018.
  38. SAS Institute, Inc. *SAS Version 9.2 [software]*. Cary, NC: SAS Institute, Inc; 2008.
  39. RTI International. *SUDAAN Version 11.0.1 [software]*. Research Triangle Park, NC: RTI International; 2018.
  40. Scientific Software International. *HLM Version 7 [software]*. Skokie, IL: Scientific Software International; 2011.
  41. Rose G. Environmental factors and disease: the man made environment. *Br Med J*. 1987;294(6577):963-965.