The Relationship between Tobacco Retailer Density and Neighborhood Demographics in Ohio

Chiche Adibe1; Peter F. Craigmile, PhD2; Nathaniel Onnen, BS2; Elli Schwartz1; Megan E. Roberts, PhD3
1College of Liberal Arts and Social Sciences, DePaul University, Chicago, IL, USA
2Department of Statistics, The Ohio State University, Columbus, OH, USA
3College of Public Health, The Ohio State University, Columbus, OH, USA

Corresponding Author: Megan E. Roberts, PhD • 1841 Neil Avenue, Columbus, OH 43210 • 614-292-4647 • Roberts.1558@osu.edu

ABSTRACT

Introduction: Studies from various parts of the country suggest that tobacco-related health disparities are exacerbated by disparities in the distribution of tobacco retailers (convenience stores, tobacco shops, etc.). The purpose of the present study was to use advanced spatial modeling techniques for count data to estimate current disparities in tobacco retailer density in Ohio.

Methods: We identified and geocoded 11,392 tobacco retailers in Ohio. Next, we obtained census tract-level information on race/ethnicity, poverty, and age and obtained county-level information on whether an area was Urban, Suburban, or Rural. Finally, we used negative binomial generalized linear models, adapted for residual spatial dependence, to determine the association between per capita tobacco retailer density and demographic characteristics—summarized by adjusted rate ratios.

Results: There were more (from 1.4-1.9 times as many) retailers per capita in high-poverty vs. low-poverty tracts. Poverty also interacted with age: the association between high poverty and high retailer density was stronger for tracts with a low youth population. Density was also greater in tracts with a high (vs. low) prevalence of African Americans (1.1 times as many) and Hispanics (1.2 times as many). Finally, density was generally greater in rural (vs. suburban or urban) tracts, although the effect was modified by a three-way interaction: density was particularly high for rural tracts that also had both a high prevalence of poverty and a low youth population.

Discussion: Overall, our findings indicate that Ohio’s vulnerable populations are exposed to a greater per capita density of tobacco retailers. There is a need for state and local-level tobacco control policies that will improve equity and reduce health disparities.

Key words: Tobacco retailer density; disparities; spatial modeling

INTRODUCTION

In the United States, smoking remains the leading cause of preventable morbidity and mortality,1 with more than 480,000 deaths occurring each year as a result of cigarette smoking.2 Smoking is a particular health concern in the state of Ohio, where the smoking rate is the sixth-highest in the country: 22.5% of Ohio adults are current cigarette users, compared to a prevalence of 17.1% nationally.3

Of additional public health concern is the fact that smoking prevalences are not uniform across the population. Rather, parallel national trends,4 Ohio sees higher smoking rates among low-income populations and certain racial/ethnic minority groups.5 Furthermore, both nationally6–8 and in Ohio,9 there is a higher prevalence of adult and adolescent smokers in rural areas, compared to non-rural areas. These types of demographic differences in smoking prevalences later translate into disparities in tobacco-related disease. For example, African American individuals in the U.S., particularly African American men, have the highest rates of lung cancer incidence and mortality.10

Numerous studies have illustrated how the relationship between the density of tobacco retailers (e.g., convenience stores, grocery stores, tobacco shops) and neighborhood demographics further tobacco-related health disparities. Specifically, research with multiple cities and counties demonstrates that tobacco retailer outlets are more heavily concentrated in marginalized communities, including those with higher proportions of racial/ethnic minorities and those with lower income levels.11–13 There is also some evidence to suggest that tobacco retailer density is greater in urban communities.14 These disparities in tobacco retailer density are, in turn, associated with tobacco use.15,16 For example, higher tobacco retailer concentration within one mile of youths’ homes is associated with higher smoking frequency.17 Young adults are more than three times as likely to start using non-cigarette combustible tobacco products if they live in areas with high tobacco retailer density.18 Moreover, high tobacco retailer concentration for youth is associated with increased exposure to point-of-sale tobacco advertising and the belief that smoking looks cool.19 Additionally, a high prevalence of tobacco retailers in a community hinders tobacco cessation among adults.20

Compounding the problem of density is the problem that the quantity of tobacco advertising differs across communities. Disparities in point-of-sale marketing exist, as greater amounts of tobacco advertisements per store are found in low-income, African American neighborhoods.21,22 A high volume of tobacco advertising is a public health concern because point-of-sale marketing exposure can have a powerful impact on youth, distorting perceptions about the availability and popularity of tobacco,23 increasing curiosity about tobacco use,24 and increasing the likelihood of smoking initiation.25 For adults, tobacco marketing exposure is associated with more frequent cravings to smoke and greater difficulty quitting.20

The purpose of this study was to estimate current disparities in tobacco retailer density in Ohio. Previous research had primarily investigated disparities in tobacco retailer density at the city or county level. By researching these disparities at the state level, we would have a more thorough understanding of the distri-
bution of tobacco retailers throughout a large area that varies widely in its demographic and geographic makeup. Our first hypothesis was that tobacco retailer density would be more prevalent in low-income areas. Our second hypothesis was that there would be a higher concentration of tobacco retailers in high racial and ethnic minority neighborhoods (specifically, African American and Hispanic). Our third hypothesis was that tobacco retailer density would be higher in urban areas.

In testing these hypotheses, our methodology incorporated two of the latest approaches for investigating tobacco retailer density. First, although past literature primarily focused on the density of conventional tobacco retailers, such as convenience stores, gas stations, and grocery stores, this study additionally examined the density of alternative tobacco retailers, such as vape and hookah outlets. Second, our analyses used spatial statistical methods to account for spatial autocorrelation in the retailer counts. When estimating tobacco retailer density, there is concern over spatial dependence—that is, that nearby retailer counts cluster together. The presence of spatial dependence can violate underlying analytical assumptions of independence of observations and can produce underestimated standard errors, leading to inaccurate conclusions from confidence intervals and hypothesis tests. We therefore used a spatial modeling approach, which has been shown to sufficiently adjust for spatial autocorrelation in retailer density research.

**METHODS**

**Identification of Tobacco Retail Outlets**

**Gathering Cigarette Retailers**

A list of Ohio-based cigarette establishments was compiled from September 2017 to December 2017. Names and addresses of all establishments with active cigarette licenses were obtained from each of Ohio’s 88 county auditor offices. In instances where an address was repeated in our list (generally from a change in business ownership and thus a new cigarette license), the duplicate address was removed. In instances where an address was reported as a street intersection, locations were translated into full addresses through cross-referencing of google map street view. A subset of establishment addresses was reported without zip codes, a portion of the address necessary to properly geocode. In the case of missing zip codes, the addresses were completed through a first round of rough geocoding followed by batch reverse geocoding and exporting of the zip code. Following these steps, we had an initial list of 11,109 cigarette retailers. Retailers included venues such as gas stations, convenience stores, grocery stores, and tobacco shops.

To assess the accuracy of our list of cigarette retailers (“groundtruthing”), a random sample of 10% of the retailers were selected for phone-verification; research staff called these stores by phone to confirm they were in business and did in fact sell tobacco. Through this process, we found that over 96% of the stores on our list were indeed selling tobacco; stores verified as being out of business or not selling tobacco were removed from the list.

**Gathering Hookah and E-Cigarette Retailers**

Beyond licensed cigarette retailers, the state of Ohio does not have a formalized method for tracking other types of tobacco retailers—in particular, hookah cafés and vape shops. We therefore collected this information based on methods described by Kates et al. A database of Ohio-based hookah and e-cigarette retailers was compiled from December 2017 to April 2018 using six internet directories: Yelp, E-Cigarette-Store-Reviews.com, Hookah-Hookah, the Yellow Pages, Better Business Bureau, and Hoover directories. Search terms such as “hookah,” “hookah bar,” “hookah lounge,” “e-cigarette,” “vape,” and “vape shop” were used. Following these steps, we had an initial list of 599 vape/hookah retailers.

Duplicate retailers (n=145) were removed through comparison to the completed cigarette retailer list. All remaining hookah and e-cigarette retailers were called by phone to confirm they were still in business and did in fact sell hookah and/or e-cigarette products (73% of those contacted met this criteria). All establishments that either were no longer in business or did not sell hookah and/or e-cigarette products were removed from the list.

**Geocoding Retailers**

Our final list of tobacco retailers in Ohio comprised 11,392 locations (11,065 cigarette licenses and 327 vape/hookah). Code was written in the R software package that used the ggrep R library (https://github.com/dkahle/ggrep) to convert batches of retailer street address into latitude-longitude coordinates. In instances where the code could not locate a retailer, individual addresses were converted to latitude-longitude coordinates using http://www.latlong.net. We then wrote a R program to calculate the number of tobacco retailers in each census tract.

**Demographic Measures**

We obtained census tract-level information about race/ethnicity, poverty, age, and population size from the 2016 American Community Survey 5-year estimates. All cut-offs distinguishing “high” and “low” groups were selected a priori. Because the state, overall, is approximately 79% non-Hispanic White, we selected a somewhat low value (15%) to be sensitive to tracts where racial/ethnic minorities are concentrated. Thus, tracts were coded as having a high prevalence of African Americans [Hispanics, Asians] if 15% or more of the population was African American [Hispanic, Asian]; all other tracts were coded as having a low prevalence of African Americans [Hispanics, Asians]. Although the prevalence of Asians is low in Ohio, a different cutoff was not used because we wanted to be consistent across racial/ethnic groups in what was considered a “high” prevalence (in other words, we wanted an absolute, rather than relative, level to indicate “high”). Age is more equally distributed in Ohio, allowing a stricter criterion for classification as having a high youth population. Thus, tracts were coded as having a high prevalence of young people if 25% or more of the population was under age 18; all other tracts were coded as having a low prevalence of young people. Finally, tracts were coded as having a high prevalence of poverty if more than 15.4% of the population was below the poverty level (15.4% is the state average for Ohio); all other tracts were coded as having a low prevalence of poverty. For exploratory purposes, models were run using different cutoffs for the demographic variables; findings indicated the same patterns of effects (available from the authors upon request).

To determine whether a neighborhood was urban, rural, or suburban, we used the county-level classifications applied by the Ohio Family Health Survey (i.e., the Ohio Medicaid Assessment Survey). This system classifies all 88 counties in Ohio as either Metropolitan (urban), Suburban, Rural Non-Appalachian, or Rural Appalachian. For the purpose of this project, the two rural designations were combined.

**Statistical Analysis**

To guard against fitting statistical models to census tracts with very low populations, we removed 14 tracts with a population of less than 500 people. This left 2,937 tracts for analysis, after the removal of one further tract that was missing covariate information (poverty). These exclusions resulted in the loss of 3 tobacco retailers, leaving us with 11,359 retailers for analyses. TIGER shape files for the counties and census tracts for the State of Ohio came from the US Census Bureau (https://www.census.gov/geo-bin/geo/shpfiles/index.php). Map creation, GIS, and statistical analyses were carried out in the R software package using the maptools, MASS, mvtnorm, rgeos, sp, and SPAM R libraries.

We began our analyses with descriptive statistics and figures to explore retailer density and our demographic variables. We calculated tobacco retailer density as the number of retailers per
1,000 people in each of the 2,937 census tracts. The density variable was then log-transformed (we added a value of 0.1 to guard against taking the log of zero for the 258 tracts that were found to have no retailers). Finally, we created figures of the variability in log retailer density for both Ohio overall and—for illustrative purposes—one urban and one rural (Appalachian) subset of the state.

**Model Selection for Spatial Analyses**

We fit various Poisson and negative binomial models to formally associate tobacco retailer density with demographic variables. We considered models that included up to three level interactions between the demographic covariates (higher order interaction introduced model instability induced by small counts in cross-tabulations of covariate factor levels). The negative binomial model is a statistical model that accounts for overdispersion in count data (extra-Poisson variability arising from unexplained covariates or clustering effects). While these models account for possible independent random effects over census tracts, they do not account for possible residual spatial effects between the different spatial regions. Thus, we adapted a generalized estimated equation (GEE) approach (e.g., Gotway and Stroup) to correct the estimated standard errors for possible spatial random effects—details are given in the Supplementary Material that accompany this article. Starting with exploratory data analyses, we used analysis of deviance tables and the Akaike information criterion (AIC) to select between different statistical models. We diagnosed the fit of our models using deviance residual plots and tested for residual spatial dependence among the census tracts using Moran’s I test statistic (e.g., Waller and Gotway). For all spatial analyses, we specified a spatial neighborhood structure to relate the different tracts. We defined the so-called neighbors of each census tract to be all the tracts that shared a border with that tract. The number of neighbors for each tract varied from 1 to 27, with a median number of 6.

Comparing AIC values and residual plots, we selected a negative binomial model that includes a high prevalence of African American, Hispanics, and Asians as main effects, as well as a three-way interaction between the high prevalence of population under age 18, high prevalence of poverty, and the urban, suburban, and rural factor variables. This model had an AIC value of 13,584, which was much smaller than the corresponding AIC value for the Poisson model of 14,450. This indicated that we preferred the negative binomial model, which allows for overdispersion in the retailer counts by census tract. Further exploration of the residuals from the negative binomial model indicated significant, but weak, spatial dependence (Moran’s I statistic = 0.026, with a p-value of 0.012). To summarize the effect of the covariates upon the retailer density in the negative binomial model, we calculated adjusted rate ratios.

**RESULTS**

**Descriptive and Exploratory Results**

Across census tracts in Ohio, the median retailer density per thousand people was 0.91 and ranged from 0 to 23.99. The left panels of Figure 1 illustrate the spatial distribution of log retailer density for all of tracts in Ohio (top left panel), as well as two subsets of Ohio: Franklin county (middle left panel) and the Southeast counties of Athens, Hocking, Meigs, Noble, Perry, and Washington (bottom left panel). Whereas Franklin county is generally urban, these Southeast counties are designated Appalachian and generally rural. The maps showed substantial spatial variability in the retailer density over Ohio, with a log retailer rate ranging from -2.30 to 3.18. For example, the city of Columbus, located in Franklin county, had a retailer density that appeared to be higher in the east of the city as compared to the west; this corresponds with the distribution of racial minorities and low-income individuals in the city, as these populations are also more heavily concentrated in the east. The retailer density in south-

---

**Figure 1.** Observed and estimated log retailer rates

**Ohio**

**Observed Log Retailer Rate**

**Estimated Log Retailer Rate**

**FRANKLIN COUNTY**

**SOUTHEAST OHIO**

**[0.57, 3.18]**

**[0.19, 0.57]**

**[-0.19, 0.19]**

**[-0.75, -0.19]**

**[-2.3, -0.75]**

---

Ohio Public Health Association
Table 1 presents the demographic characteristics of Ohio’s census tracts. African Americans were the most prevalent minority in the state, with 28% of the tracts classified as having a high prevalence of African Americans. Nearly half (46%) of tracts were classified as high poverty. Approximately 27% of tracts were within rural counties, 15% were within suburban counties, and 59% were within urban counties.

Table 1. Demographic characteristics of Ohio census tracts, 2016, and corresponding median tobacco retailer density.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Prevalence (% of Ohio Census Tracts)</th>
<th>Median Retailer Density (per 1000 people)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Prevalence</td>
<td>28.0</td>
<td>1.08</td>
</tr>
<tr>
<td>Low Prevalence</td>
<td>72.0</td>
<td>0.84</td>
</tr>
<tr>
<td>Hispanic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Prevalence</td>
<td>4.5</td>
<td>1.44</td>
</tr>
<tr>
<td>Low Prevalence</td>
<td>95.5</td>
<td>0.88</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Prevalence</td>
<td>1.3</td>
<td>0.68</td>
</tr>
<tr>
<td>Low Prevalence</td>
<td>98.7</td>
<td>0.91</td>
</tr>
<tr>
<td>Population under age 18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Prevalence</td>
<td>31.3</td>
<td>0.88</td>
</tr>
<tr>
<td>Low Prevalence</td>
<td>68.7</td>
<td>0.95</td>
</tr>
<tr>
<td>Poverty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Prevalence</td>
<td>46.4</td>
<td>1.19</td>
</tr>
<tr>
<td>Low Prevalence</td>
<td>53.6</td>
<td>0.70</td>
</tr>
<tr>
<td>Neighborhood Type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>59.0</td>
<td>0.91</td>
</tr>
<tr>
<td>Suburban</td>
<td>14.6</td>
<td>0.72</td>
</tr>
<tr>
<td>Rural</td>
<td>26.5</td>
<td>1.03</td>
</tr>
</tbody>
</table>

Figure 2 relates the log retailer density to the demographic variables. For race/ethnicity, retailer density tended to be higher in tracts with a higher prevalence of African Americans and Hispanics (there was no effect for Asian; see Table 1). There tended to be more retailers in tracts with a higher prevalence of poverty and a suggestion that tracts with a higher prevalence of people under 18 years had a lower density. In terms of urban/rural characteristics, there were more retailers in rural tracts, as compared to suburban and urban tracts. Other figures not presented here suggested the possibility of high-level interactions between the retailer density and the demographic variables, which we investigated in our statistical models.

Spatial Analysis Results

The right panels of Figure 1 display the log retailer rates estimated from the negative binomial generalized linear model, over the census tracts of Ohio overall, as well as Franklin county and Southeast Ohio. As anticipated, the expected log rates in the right panels were smoother than the observed log rates shown in the left panels. These estimated rates also confirmed that there was higher retailer density in southeast Ohio, as compared to Franklin county (we estimated a median number of 1.5 retailers per thousand people in the southeast as compared to a median number of 1.2 in Franklin county). Across Ohio overall, there was a high density of retailers in the south and east. In metropolitan areas, there tended to be both areas of higher and lower retailer densities, associated with the varying demographics by tract.

*Tracts where 15% or more of the population is African American.
*Tracts where 15% or more of the population is Hispanic.
*Tracts where 15% or more of the population is Asian.
*Tracts where 25% or more of the population is under age 18.
*Tracts where more than 15.4% of the population is below the poverty level (15.4% is the state average for Ohio).
*Classification of urban, rural, and suburban is based on the county-level classifications applied by the 2008 Ohio Family Health Survey (i.e., the Ohio Medicaid Assessment Survey).
Figure 3 summarizes the adjusted rate ratios of the retailer density for different combinations of demographic variables. The left panel summarize the race/ethnicity variables. As the race/ethnic variables appeared in the model as main effects, we could interpret them independently of the other variables. Our model indicated significantly more retailers in tracts with a high (vs. low) prevalence of African Americans (11 times as many; \( z=2.37, p=0.018 \)) and Hispanics (1.2 times as many; \( z=2.23, p=0.026 \)); there was no significant effect for tracts with a high (vs. low) prevalence of Asians (\( z=-0.64, p=0.525 \)). The right panel of Figure 3 summarizes the three-way interaction in the model, focusing on the effect of poverty (higher versus lower prevalence), for different urban, suburban and rural tracts, with a lower prevalence of people under age 18 (in gray), and a higher prevalence of people under 18 (in black). Over the six combinations of levels of the variables, the rate of tobacco retailers was significantly higher in high poverty tracts versus low poverty tracts—our model estimated that the rate ratio varied between 1.4 to 1.9 as many retailers (across all six combinations of levels, all \( z \) statistics lie between 2.70 to 9.76, with \( p \)-values between <0.001 and 0.007). In populations with less people aged under 18, the effect of poverty was more pronounced, as compared to populations with more people under 18. In tracts with less people aged under 18, the effect of poverty was no different for urban and suburban tracts, but the effect of poverty may be higher for rural counties. For tracts with a higher prevalence of people under 18, there was no significant difference in the effect of poverty between the urban, suburban, and rural areas.

**DISCUSSION**

The purpose of this study was to estimate disparities in tobacco retailer density for the entire state of Ohio. Findings indicated that, in support of our first hypothesis, per capita density was greater in high-poverty (vs. low-poverty) census tracts. The effect of poverty also interacted with age: high poverty was more strongly associated with higher retailer density among tracts with a low youth population. In support of our second hypothesis, we found that per capita density was greater in census tracts with a high prevalence of African Americans, and in census tracts with a high prevalence of Hispanics. There was no significant effect for tracts with a high prevalence of Asians; this null effect may have been due to the low number tracts in Ohio with a high prevalence of Asians. Our third hypothesis was not supported, as we found that per capita density was generally greater in rural census tracts, although the effect was modified by a three-way interaction: density was particularly high for rural census tracts that also had both a high prevalence of poverty and a low youth population.

Overall, our findings replicate previous work from other parts of the country by demonstrating racial/ethnic and poverty-based disparities in tobacco retailer density. These findings likewise support the identification of African Americans, Hispanics, rural populations, and low-income individuals as vulnerable populations—i.e., populations with social characteristics that put them at risk for exposure to other risks.\(^9\) Our finding that rurality was associated with greater density was novel. As very few studies have investigated the relation between rurality and tobacco retailer density, researchers should attempt to replicate our finding in other areas to determine its generalizability. Our findings also extend previous work by presenting disparities at the state level: in particular, a state that varies widely in its demographic and geographic makeup. Our work also extends examinations beyond conventional retailers (e.g., convenience stores, grocery stores) to include vape shops and hookah cafés. Finally, our study improves on previous methodology by incorporating spatial statistical methods to account for spatial dependence in the data. Such methodology improves upon approaches that assume normal distributions and independent data points. Our analyses indicated that fitting a negative binomial model, while accounting for the...
residual spatial dependence, was able to account for the spatial dependence in the data, and we recommend that investigators consider this in the future.

Although Ohio’s licensing system provided us with a seemingly comprehensive list of cigarette retailers in Ohio, it is possible that some cigarette retailers were unlicensed or did not have an active cigarette license at the time of data collection. It is also possible that some hookah and e-cigarette retailers may not have been detected by the online searches we conducted. Another limitation to this study is that geocoding software can occasionally produce errors, which may cause the geocoded address to not correspond exactly with the actual store location. All these limitations likely introduced some random error (rather than biased, systematic differences among communities), possibly resulting in minor perturbations in our effect sizes.

PUBLIC HEALTH IMPLICATIONS

This study found disparities in how tobacco retailers are distributed in Ohio, such that retailer density was associated with a neighborhood’s racial/ethnic composition, poverty level, age distribution, and urban/rural status. Given the size and diversity of the geographical area covered in our analyses, we expect these findings to be generalizable to other parts of the country—although outcomes may differ somewhat for areas with a higher prevalence of Asian or Hispanic populations. Results from this study have concerning implications for public health, as a strong body of literature suggests that greater retailer density has a pernicious impact on local behaviors—both increasing the likelihood of youth initiation17,18 and decreasing the likelihood of adult cessation.20 Our observations of Ohio’s disparities in tobacco retailer density are thus likely contributors to the state’s disparities in tobacco-related disease.

Ultimately, our findings contribute to a strong body of literature indicating a need for tobacco control policies that will target the density and impact of tobacco retailers in communities—and thereby improve equity and reduce health disparities. Such approaches may include “content-neutral” external advertising regulations; such regulations are more likely than other advertising restrictions to withstand free-speech challenges, particularly if they are enacted for aesthetic or public safety reasons. Another promising tobacco-control approach is modifying the state’s current tobacco licensing laws to set stipulations on the number or density of retailers. Such strategies have been successfully implemented within several communities, such as San Francisco, New Orleans, and over 80 cities and towns in Massachusetts. Similar approaches should be considered at the state and local level in Ohio, and likewise evaluated for their impact on correcting disparities.

ACKNOWLEDGEMENTS

This work was supported by the National Cancer Institute under grant R21CA212308.

REFERENCES

3. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Population Health. BRFSS Prevalence & Trends Data. 2015. [https://www.cdc.gov/bfrss/bfrssprevalence/]


