Article

Spatially Varying Associations of Neighborhood Disadvantage with Alcohol and Tobacco Retail Outlet Rates

David C. Wheeler [1,*], **Joseph Boyle** [1], **D. Jeremy Barsell** [2], **Trevin Glasgow** [2], **F. Joseph McClernon** [3], **Jason A. Oliver** [3,4,5] and **Bernard F. Fuemmeler** [2,6]

1 Department of Biostatistics, Virginia Commonwealth University, Richmond, VA 23298, USA; boylejr@vcu.edu
2 Department of Health Behavior and Policy, Virginia Commonwealth University, Richmond, VA 23298, USA; barselldtj@vcu.edu (D.J.B.); trevin.glasgow@vcuhealth.org (T.G.); bernard.fuemmeler@vcuhealth.org (B.F.F.)
3 Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC 27705, USA; francis.mcclernon@duke.edu (F.J.M.); jason-oliver@ouhsc.edu (J.A.O.)
4 Stephenson Cancer Center, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA
5 Department of Psychiatry and Behavioral Sciences, Oklahoma State University Center for Health Sciences, Tulsa, OK 74107, USA
6 Massey Cancer Center, Virginia Commonwealth University, Richmond, VA 23298, USA
* Correspondence: dcwheeler@vcu.edu; Tel.: +1-804-828-9827

Abstract: More than 30% of cancer related deaths are related to tobacco or alcohol use. Controlling and restricting access to these cancer-causing products, especially in communities where there is a high prevalence of other cancer risk factors, has the potential to improve population health and reduce the risk of specific cancers associated with these substances in more vulnerable population subgroups. One policy-driven method of reducing access to these cancer-causing substances is to regulate where these products are sold through the placement and density of businesses selling tobacco and alcohol. Previous work has found significant positive associations between tobacco, alcohol, and tobacco and alcohol retail outlets (TRO, ARO, TARO) and a neighborhood disadvantage index (NDI) using Bayesian shared component index modeling, where NDI associations differed across outlet types and relative risks varied by population density (e.g., rural, suburban, urban). In this paper, we used a novel Bayesian index model with spatially varying effects to explore spatial nonstationarity in NDI effects for TROs, AROs, and TAROs across census tracts in North Carolina. The results revealed substantial variation in NDI effects that varied by outlet type. However, all outlet types had strong positive effects in one coastal area. The most important variables in the NDI were percent renters, Black racial segregation, and the percentage of homes built before 1940. Overall, more disadvantaged areas experienced a greater neighborhood burden of outlets selling one or both of alcohol and tobacco.

Keywords: alcohol; tobacco; retail outlets; neighborhood disadvantage; spatially varying effects

1. Introduction

More than 30% of cancer related deaths are related to tobacco or alcohol use [1]. The majority (55–66%) of cancer deaths in females that are attributable to alcohol are from breast cancer, while the majority (53–71%) of alcohol-attributed cancers in males are from the aero-digestive tract [2]. In addition, tobacco use has been linked to a variety of cancers, including cancers of the lung, urinary system, aero-digestive tract, and myeloid leukemia [3,4]. Reducing, controlling and/or restricting access to these cancer-causing products, especially in communities where there is a high prevalence of other cancer risk factors, has the potential to improve population health and reduce the risk of specific cancers associated with these substances in more vulnerable population subgroups.

One policy-driven method of reducing access to these cancer-causing substances is to monitor and regulate where and how these products are sold. Specifically, the placement

and density of businesses selling tobacco and alcohol within neighborhoods could be reviewed and better regulated to address systemic issues (over representation within certain communities) that may contribute to an increase of sales of these products. States and local communities have the legal ability to influence the sale of tobacco and alcohol through zoning policies and granting licenses. Change at the state and community levels (e.g., reduce exposure and access to the products and any marketing, target social norms that promote use of the products) can possibly lead to individual-level changes in risky health behaviors (e.g., frequent tobacco and alcohol use) [5–7]. To inform such efforts, several studies have now assessed the relationship between socio-spatial demographics and retail density [8,9], and the strength of the association between retail density and increased consumption of cancer-causing products [10]. Consistently, studies have shown that both tobacco retail outlet (TRO) and alcohol retail outlet (ARO) density relative to the population are greater in neighborhoods having a lower median income and higher percent of minority residents [10–12]. The literature generally supports an association between greater density of these retail outlets and the greater use of tobacco and alcohol products among those who live in these neighborhoods [12,13].

Within this context of examining the associations between TROs and AROs and other variables, the associations have been examined separately in most studies, rather than examining the intersection of these types of product sales. This is a limitation because stores that sell both tobacco and alcohol, tobacco and alcohol retail outlets (TAROs), could have stronger associations with neighborhood deprivation than either of TROs or AROs independently. In other words, analyzing only TROs and AROs fails to estimate the neighborhood burden of TAROs.

In previous work, Wheeler et al. aimed to overcome this limitation to better understand the extent to which neighborhood socio-demographic characteristics relate to both the overlapping and distinct spatial distributions of TROs, AROs, and TAROs. They used a novel extension of the Bayesian index model that included a shared component for the spatial pattern common to all three types of outlet relative risks, outlet-specific spatial components for relative risks, and neighborhood disadvantage index (NDI) effects that varied by outlet type [14]. The Bayesian index model is a relatively new approach for estimating neighborhood disadvantage or deprivation indices [10,14–18] that considers the association with the outcome and can adjust for covariates and residual confounding through random effects. In Wheeler et al. [14], the NDI was estimated by a weighted combination of neighborhood socioeconomic status (SES) indicators (e.g., education, income, poverty) and racial/ethnic segregation (e.g., black population segregation). The results of that study included significant positive associations between TROs, AROs, TAROs and neighborhood disadvantage in block groups in North Carolina from the Bayesian shared component index model with associations that differed across outlet types. Regarding spatial patterns of outlets, the relative risks for each outlet type varied by levels of population density (e.g., rural, suburban, urban) and the shared component of relative risk had notable spatial heterogeneity.

While the study by Wheeler et al. found varying NDI effects by outlet type and population density and spatial structure in the relative risks for AROs, TROs, and TAROs, it did not consider that the spatial structure could be partially explained by neighborhood disadvantage effects that varied over space. In fact, there are no Bayesian index models with spatially varying coefficients in the literature, as previous specifications of Bayesian index models assume a constant regression effect for the index [10,14–18] over space. Regression models with spatially varying effects have been used previously in several fields including geography [19], ecology [20], hedonic price analysis [21], and epidemiology [22–28] to model nonstationarity in regression relationships. Nonstationarity reflects differences in a statistical relationship or confounding that varies across space or time [29]. Commonly used models for spatial nonstationarity are Bayesian spatially varying coefficient models and frequentist geographically weighted regression. One example of a rationale for using spatially varying coefficient models in hedonic price modeling is that the value of an

additional bathroom means more in a highly dense market such as New York City compared with a rural market in Kansas where space is at less of a premium [19]. In a regression model of housing price, this means that the effect of the number of bathrooms would vary within a large market and over multiple markets. In an example in epidemiology, geographically weighted regression was used to identify spatially varying relationships between local urologist density and prostate cancer mortality at the county level in eastern and southern states in the US, where the strength of association varied regionally [22]. The region with the largest estimated decrease in prostate cancer mortality consisted of counties in the southern Mississippi River states of Arkansas, Louisiana, and Mississippi. In these counties, all else being equal, a one-unit increase in urologist density (one urologist per 100,000 men) was associated with a with lower prostate cancer mortality than in counties with coefficients closer to zero. These areas had higher prostate cancer mortality rates and relatively low urologist density. Therefore, these areas could be targeted for increasing the supply of urologists, as it associated with the largest predicted improvement in prostate cancer mortality.

Within the context of the present study, ignoring nonstationarity of the relationship between neighborhood disadvantage and outlet rates may be problematic because it assumes consistency of the relationship across space, and this may underestimate or inaccurately estimate the effect for neighborhood disadvantage in some areas. Therefore, in this study we propose using a new Bayesian spatially varying coefficient index model to model nonstationarity in neighborhood disadvantage effects related to TRO, ARO, and TARO rates over block groups in North Carolina. Specifically, we used a Bayesian index model with spatially varying effects to explore the hypothesis that NDI effects could vary over North Carolina to explain variation in outlet rates. We also sought to identify the geographic areas that were significantly elevated in NDI effects for each of the three outlet types to identify disparities in the neighborhood burden of tobacco and alcohol outlets.

2. Materials and Methods

2.1. Study Context

We analyzed data from the state of North Carolina for our study. North Carolina does not require a license for retailers to sell tobacco products. However, the North Carolina Alcoholic Beverage Control (ABC) regulates the general sale of alcohol in the state. Retailers seek permits to sell beer and wine, but only state-run ABC stores are able to sell liquor. When considering whether to issue beer and wine permits to retailers, the ABC Commission considers several factors, such as the number of existing permits that have been granted in a neighborhood, parking availability, potential traffic problems, the types of existing businesses in the neighborhood, and the proximity of the applying retailer to schools and churches [30].

2.2. Study Data

2.2.1. Tobacco and Alcohol Retail Outlets

We obtained data from two different sources to determine the TROs, AROs, and TAROs in North Carolina. The first data source we used was the National Establishment Time Series (NETS), which is a private, longitudinal, and geocoded record of all types of businesses in the United States. We utilized a previously developed algorithm to identify TROs [31] from this database. The algorithm filtered entries by their North American Industry Classification System codes to exclude businesses that were not TROs, deleted duplicate records, and removed records for TROs in years after state- or municipality-specific bans were enacted on the sale of tobacco products by certain business types. The second data source, for ARO data, was the North Carolina ABC Commission. The state's ABC maintains a publicly-available database [30] of businesses that have received licenses to sell alcoholic beverages, including liquor stores that are run by the ABC. We took the business entries from this database that retained active and off-premise licenses to sell alcohol, in order to avoid including different types of businesses such as bars and

restaurants. We geocoded AROs from their business address using the ArcMap software from ESRI and then standardized several variables in the data to maximize syntactical similarity between the TRO and ARO databases, such as converting all text to upper case, using a common street suffix, and excluding store identification numbers from business names. We then counted the number of TROs, AROs, and TAROs in each block group, a process which we have described in more detail previously [14]. At the end of this process, each block group had a count of businesses that were just TROs, just AROs, or TAROs. Across all the block groups in the state, we identified 7289 TROs, 4859 AROs, and 4631 TAROs. More details about the process for counting the number of outlets are also available elsewhere [14].

2.2.2. Sociodemographic Data

We constructed neighborhood disadvantage indices using five-year estimates (2014–2018) of nine social and demographic variables collected at the census block group level by the American Community Survey (ACS), an annual survey representative of the US population that includes questions related to home ownership, income, and education, among other themes. We chose the time period 2014–2018 in order to overlap with the extent of the outlet data. The social and demographic variables included: Black population segregation, Hispanic population segregation, percent with income to poverty ratio < 1, percent households receiving public assistance income, percent renter occupied housing units, percent homes constructed in 1939 or earlier, percent with less than a high school degree, percent of households in poverty, and per capita income. The Black segregation variable was defined as the percent of Black population in a block group divided by the state average percent Black population. Hispanic segregation was defined similarly. We have used similar sociodemographic variables and segregation measures [18,32–34] to estimate neighborhood disadvantage indices previously [10,16]. We began with 6127 block groups in North Carolina and removed 15 that had no population and 29 others that had missing covariates to reach an analysis set of 6083. In this set of block groups, there were a total of 7279 TROs, 4849 AROs, and 4621 TAROs.

2.3. Statistical Analysis

We fit hierarchical Bayesian regression models to explain the observed variation in the rates of the outlet types (TRO, ARO, TARO), assuming that in the ith block group, the count of the kth outlet type was distributed as $y_{ik} \sim Poisson\,(\theta_{ik}E_{ik})$. For each block group, the modeled relative risk parameter was θ_{ik} and the expected count was E_{ik}, which we defined as the overall rate r_k for the kth outlet type multiplied by the population in the block group p_i. The new Bayesian spatially varying coefficient index regression model of the log relative risk was:

$$\log(\theta_{ik}) = \alpha_k + \beta_{ik}\left(\sum_{j=1}^{C} w_j q_{ij}\right) \tag{1}$$

where α_k is the outlet-specific intercept and β_{ik} is the outlet-specific, spatially-varying effect for the neighborhood disadvantage index in the ith block group. In each block group, the neighborhood disadvantage index was defined as a weighted combination $\sum_{j=1}^{C} w_j q_j$ of the deciles $q_1, \ldots, q_c$ of the sociodemographic variables $x_1, \ldots, x_c$, with the weights $w_1, \ldots, w_C$ being estimated in the model as part of the model parameters. The weight w_j signifies the relative importance of the j_{th} variable in the neighborhood disadvantage index. The index approach for estimating neighborhood disadvantage has been used previously [16,18,34]. We used deciles to account for different scaling of and correlations between the variables to mitigate the effect of outliers, and to acknowledge the uncertainty in ACS values. We used $C = 9$ variables in the index. We defined the sociodemographic variables to reflect a hypothesized positive association of the index with relative risk (i.e., that greater disadvantage was associated with greater relative risk). We re-defined per capita income to have a positive association with rates based on univariate

Int. J. Environ. Res. Public Health **2022**, *19*, 5244

5 of 13

correlations with the outlet rates, inverting this variable with the formula $\max(x) - x_j$, where x_j is the value of the variable.

The Bayesian model specification was completed with prior distributions for the parameters. The outlet-specific intercepts followed an improper uniform distribution $\alpha_k \sim dflat()$. The NDI regression coefficients were given a multivariate intrinsic conditional autoregressive (MVCAR) prior $\beta_{ik} \sim MVCAR(\Omega)$, which has been used previously to model spatially varying effects [35,36]. The precision matrix Ω for the MVCAR received a Wishart prior, $\Omega \sim Wishart(R, K)$, where R is a KxK diagonal matrix with 0.2 on the diagonal and $K = 3$. The multivariate CAR prior allows the regression coefficients for one outlet type to be correlated over space (e.g., block groups) and for the regression coefficients to be correlated across outlet type. The inverse of Ω is the variance-covariance matrix for the regression coefficients, Σ. The index weights w_j were given a Dirichlet prior with parameters $\alpha = (\alpha_1, \ldots, \alpha_C)$. The Dirichlet prior was selected because it assures that the index weights $w_j \in (0, 1)$ and $\sum_{j=1}^{C} w_j = 1$.

We estimated all model parameters using Markov chain Monte Carlo (MCMC) sampling, burning in 50,000 iterations and retaining an additional 10,000 iterations with a thinning parameter of one from two chains. We evaluated convergence in the MCMC chains for parameters of interest using the Gelman-Rubin diagnostic, considering a parameter to have converged if its Gelman-Rubin statistic was less than 1.2 [37]. We exponentiated the regression coefficients to get relative risk estimates for the neighborhood disadvantage index. The 95% credible interval was used to determine the statistical significance of the neighborhood disadvantage index relative risks, where there was a significant positive effect if the lower bound of the interval exceeded the value of 1. We also compared the distributions of the NDI variables and the disadvantage index between the block groups with and without significantly positive NDI effects using Welch's *t*-test. To determine the amount of correlation between NDI relative risks among outlet types, we used the variance-covariance matrix components to calculate the conditional correlation between pairs of outlet types. We fit the Bayesian models using R2OpenBUGS [38] in the R computing environment [39].

3. Results

The posterior median NDI relative risks show substantial variation over space for all three outlet types (Figure 1). Some correlation in the NDI relative risks between outlet types was evident, particularly with high relative risks in the northeastern coastal area (i.e., Outer Banks) for all three outlet types. In addition, all three outlet types also have areas of elevated NDI relative risk in southwestern North Carolina. However, there are varying amounts of correlation between NDI relative risks for pairs of outlet types (Figure 2). AROs and TROs were most highly correlated (posterior median = 0.92), followed by TROs and TAROs (posterior median = 0.87), and then AROs and TAROs (posterior median = 0.69). These correlations reinforce the observed visual patterns in Figure 1. For both TROs and AROs, there are areas of elevated NDI relative risk in northeastern North Carolina and the central and southern coastal areas. In contrast, TAROs have lowered NDI relative risks in northeastern North Carolina. The block groups with significantly elevated NDI relative risks according to 95% credible intervals (Figure 3) also show that the patterns are more similar for AROs and TROs than for either of these with TAROs. Both AROs and TROs have more block groups with significantly elevated NDI effects in the southwest and northeast of North Carolina. These store types show more clustering in elevated NDI effects than do TAROs.
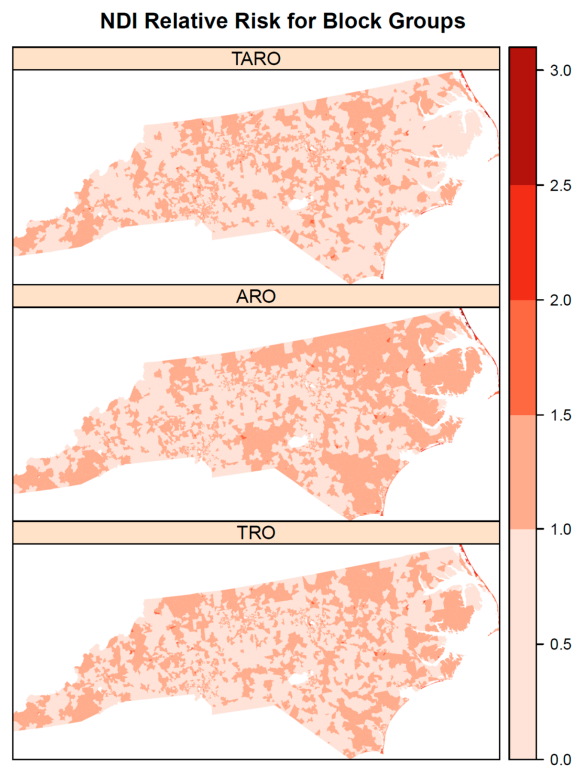
Int. J. Environ. Res. Public Health **2022**, *19*, 5244

6 of 13

Figure 1. Posterior medians for neighborhood disadvantage index relative risks by outlet type for tobacco retail outlet (TRO), alcohol retail outlet (ARO), and tobacco and alcohol retail outlet (TARO). The legend is on the relative risk scale.
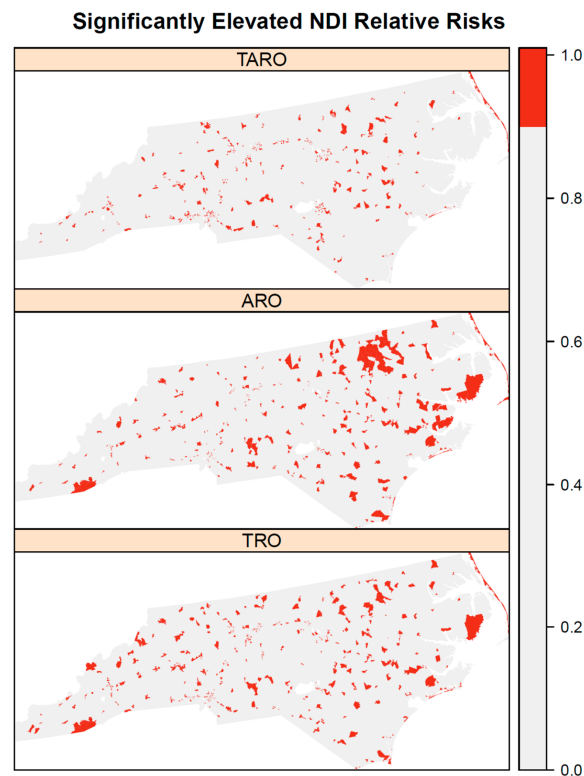

Figure 2. Areas with significantly elevated neighborhood disadvantage index effects by outlet type for tobacco retail outlet (TRO), alcohol retail outlet (ARO), and tobacco and alcohol retail outlet (TARO). The legend indicates red for 95% credible intervals that are above 1.
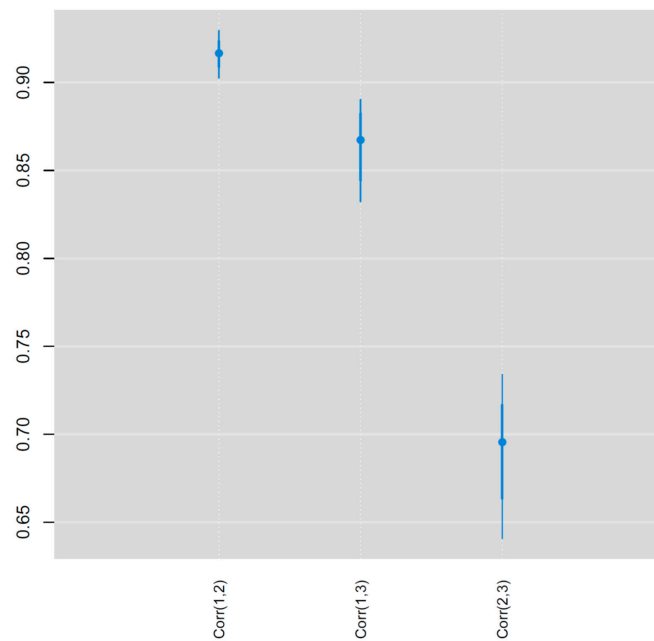
Int. J. Environ. Res. Public Health **2022**, *19*, 5244

7 of 13

Figure 3. Posterior medians (circles) and 95% credible intervals for the conditional correlation in neighborhood disadvantage index regression coefficients by outlet type, where 1 = TROs, 2 = AROs, and 3 = TAROs.

For the estimated NDI, the weights for the nine components within the NDI (Figure 4) show that percent renter (0.26), black racial segregation (0.18), and percent of homes built before 1940 (0.16) were the three most important variables in the NDI and together accounted for 0.60 of the total weight. Hispanic segregation (0.14) and percent receiving public assistance (0.12) were the other components receiving more than a priori equal weight (0.11). The estimated NDI had a greater concentration of higher values in the eastern half of the state compared with the western half (Figure 5). However, neighborhood deprivation was also high in metropolitan areas located in the center and western half of the state.
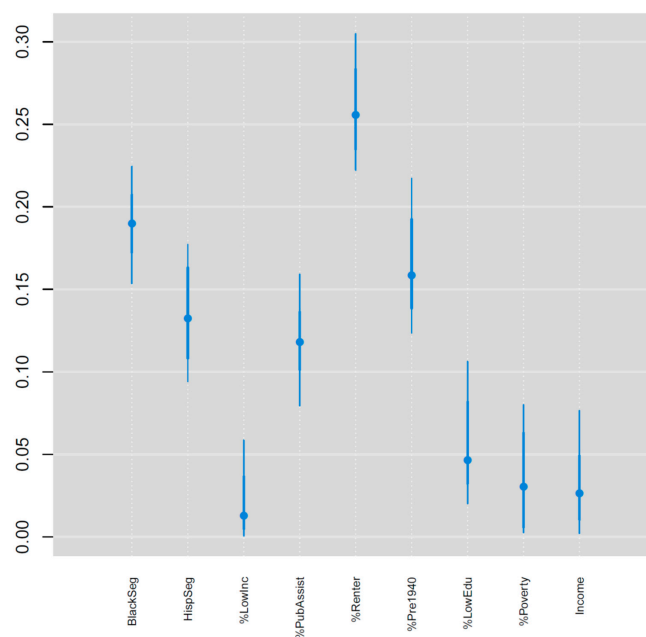

Figure 4. Posterior medians (circles) and 95% credible intervals for the neighborhood disadvantage index weights.
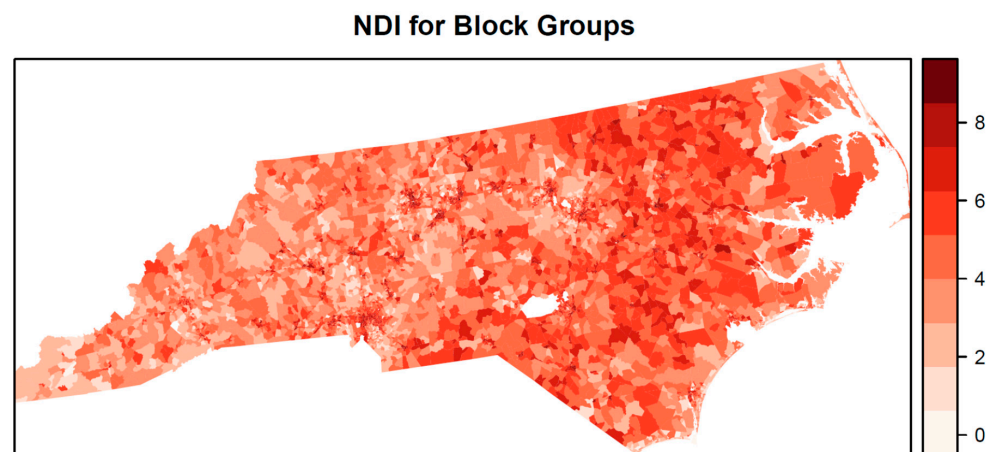
NDI for Block Groups

Figure 5. Neighborhood disadvantage index for block groups estimated from the Bayesian spatially varying coefficient index model.

To gain understanding about the areas with significantly elevated NDI relative risks, we compared the distributions of the NDI and deciles of the NDI variables between the set of block groups with significantly elevated NDI relative risks and those without. The Welch (*t*-test results showed significant differences ($p < 0.01$ for all tests) in the means of the index and all nine of the NDI variables for TROs, where the index and NDI variables were all higher in the set with significantly elevated NDI relative risks (see Supplemental Material Table S1). For AROs, test results were similar, with all tests significant ($p < 0.05$) except for Hispanic segregation, which was slightly higher in the set without elevated NDI relative risks. For TAROs, the tests were all significant ($p < 0.01$) with the NDI and NDI variable means being higher in the set with significantly elevated NDI relative risks. When looking at the distributions of the NDI and the deciles of the NDI variables, there was a shift to high values in the set with significantly elevated NDI relative risks compared with the set without significantly elevated NDI relative risks. Thus, areas with significantly elevated NDI relative risks were generally more disadvantaged areas. Therefore, while the NDI effect is not equal over space, it is generally larger in the more deprived areas. However, there were some block groups with significantly elevated NDI relative risk (≥ 1.8) and low values (<5) of the NDI (see Supplemental Material Figure S1).

4. Discussion

In this paper, we aimed to estimate spatially varying relationships between a neighborhood disadvantage index and TRO, ARO, and TARO rates using a novel Bayesian spatially varying coefficient index model to help address nonstationarity of the relationship between neighborhood disadvantage and tobacco and retail outlets. Our results showed substantial variation in the associations between the estimated neighborhood disadvantage index and the rates of each outlet type, with some notable similarities and differences in effects by outlet type. All three outlet types (AROs, TROs and TAROs) had significantly elevated NDI relative risks in the northeastern shore area of the Outer Banks. However, TROs and AROs had more significantly elevated NDI effect areas in the northeast and southeastern coast compared with TAROs, and AROs in particular had a notable cluster of significantly elevated effects in northeastern North Carolina near the state line. Overall, the NDI effects for TROs and AROs were more strongly correlated with each other than with the effects for TAROs. The block groups with significantly elevated NDI relative risks had larger values of the NDI index and NDI variables (e.g., percent renters) overall than did the block groups without significantly elevated NDI relative risks, indicating that more disadvantaged areas experienced a greater neighborhood burden of outlets selling one or both of alcohol and tobacco overall. The most important variables in the NDI were percent renters (i.e., low home ownership), Black racial segregation, and percentage of homes built before 1940 (i.e., old housing stock).

Our study considered the spatially varying associations of neighborhood deprivation with TROs, AROs, and TAROs simultaneously. Prior research has centered on only TROs [12,40] or only AROs [7,41], with few to none examining their simultaneous combination (TAROs). Research on TROs and AROs is not limited to the US. For example, in North India [42], both TROs and AROs were geocoded by physically walking in neighborhoods with a map app determining the coordinates of the outlets. While the prevalence of smoking and drinking were associated with TRO density, the at-risk health behaviors were not associated with ARO density. In Scotland, another study [8] examined both TROs and AROs and their associations with neighborhood deprivation. The result from that study that density of the retail outlets is associated with neighborhoods experiencing greater deprivation is consistent with our results. However, in that study, neighborhood deprivation was assessed only by income and by the percentage receiving government benefits and assistance. Furthermore, no previous studies have considered these associations to vary over space. By allowing the associations of NDI and the three outlet types to vary over space, we found that more disadvantaged areas overall experienced a greater neighborhood burden of outlets selling one or both of alcohol and tobacco, which highlights a disparity that is important for public health.

In this study, we employed novel analytic methods to assess the spatially varying relationships of alcohol and tobacco outlets with neighborhood sociodemographic characteristics. We extended the Bayesian index model to have spatially varying regression coefficients for multiple outcome rates. The Bayesian index regression model provides more informative output than the traditional dimension reduction techniques used in the literature [43,44], such as principal component analysis (PCA) and factor analysis. The interpretation of the components produced by PCA and factor analysis are not easy to interpret and do not account for the association with the outcome. In contrast, our Bayesian index model estimates the index weights considering associations with the outcomes and produces easily interpretable weights indicating relative importance, some of which can be effectively zero for unimportant variables. In our case, only five variables showed up as being relatively important while four had small weights. We have demonstrated previously that employing an index approach to at once estimate the NDI as well as its health effect leads to significantly improved model goodness-of-fit than constructing the index with principal components analysis [33]. Another strength of our study is its extension of the Bayesian index regression model to accommodate several related outcomes at once, as we allowed spatially varying NDI effects to be correlated across store types. We used the information in the three outlet types to estimate the neighborhood disadvantage index for all block groups together, but allowed the effect of the NDI to uniquely vary spatially for each outlet type, which in this case was useful as illustrated by the differences in patterns of effects (Figure 1) and significantly elevated effects for outlets by type (Figure 2). In our previous work, we did not allow the NDI effects to vary over space for each outlet type and therefore could not uncover our current finding that more disadvantaged areas experienced a greater neighborhood burden of outlets selling tobacco and/or alcohol.

One of the benefits of the Bayesian index model approach is that it can empirically determine the most important components in the index for explaining the outcome variable. Indeed, in our study not all variables in the index were equally important in explaining the association between neighborhood disadvantage and outlet rates (e.g., Figure 4). Select variables reflecting neighborhood socioeconomic disadvantage, such as low home ownership and older housing stock, were more important than income and poverty related variables. Racial and ethnic segregation also received more weight than the income and poverty variables. Our finding of an association with percent renters may be related to cigarette product marketing and use among youth and young adults, as youth are more likely to rent their homes [45] and be exposed to and use cigarette products [46]. Our results align to some extent with those reported from other geospatial studies conducted in New Jersey [47], Boston, Massachusetts [48], and a recent US based study [49] where race and ethnicity were important variables for TRO density. However, household income was less important in

our analysis than in the New Jersey study, where percent Hispanic population was the dominant demographic factor associated with TRO density, followed by median household income and percentage of Black population. One possible explanation for differences in variable importance across studies is that our study considered outlets selling alcohol and tobacco and alcohol together in addition to just tobacco outlets. Furthermore, the relative importance of these variables in explaining retail outlet density may reflect different local and state tobacco control policies, as well as differences in the statistical methods used. More progressive locales may better implement policies that affect retail outlet density through excise taxes and prohibitory regulations, such as restrictions on placement within certain distances of specific locations. Therefore, the continued characterization of retail outlet placement is necessary for understanding how neighborhood characteristics are associated with TRO, ARO, and TARO density over time and across geographic areas.

A discussion of the limitations of the study is also warranted. First, we operationalized the concept of neighborhood disadvantage in a block group through the use of one fixed set of sociodemographic variables at that block group. We did not consider other covariates in the block group, or variables from a larger spatial unit containing the block group, such as the census tract or county. Secondly, we emphasize that our findings are limited to the state of North Carolina. Future research should adopt the methods of our study to other geographic regions. Different relationships may hold in different contexts, such as policies regulating sales and zoning of these products, as well as of the social norms regarding alcohol and tobacco. Finally, we drew TRO and ARO business data from large registries, and the tobacco data source required an additional algorithm to classify stores as likely selling tobacco. Though we performed extensive checks to verify the accuracy of these listings and of our resulting analysis datasets, some error that is attributable to the large business databases (unregistered businesses, or those not captured by the database) is unavoidable.

5. Conclusions

In summary, our use of Bayesian spatially varying coefficient index models uncovered significant positive associations between neighborhood disadvantage and rates of tobacco, alcohol, and tobacco and alcohol outlets in North Carolina and identified clear geographic disparities in the neighborhood burden of these outlets. The factors driving the significant associations include low levels of home ownership, black racial segregation, and old housing stock. The findings in our study help to shed light on the dominant neighborhood factors that are associated with the heterogeneous distribution of tobacco and alcohol retail outlets, which are themselves environmental factors that drive smoke exposure and alcohol use. A deep understanding of both the individual- and environmental-level determinants of smoke exposure is a priority for public health, especially given the increased potential for adverse health outcomes in newborns and children [50]. Future studies should center on economic and educational factors in disadvantaged areas, as well as their connection to smoking and alcohol use, and more broadly on the structural inequalities that encourage the use of these products in these areas. The modeling approach presented in this paper should be useful in other geographic areas for uncovering disparities in neighborhood burden of tobacco and alcohol retail outlets and the factors most related to the disparities. Furthermore, they help support and justify policies and approaches that limit retail outlets in certain neighborhoods in order to reduce the apparent systemic injustice of placement of high rates of tobacco or alcohol retail outlets among minoritized communities.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijerph19095244/s1, Figure S1: Block groups with significantly elevated neighborhood disadvantage index (NDI) relative risk >= 1.8 and NDI < 5; Table S1: Two-sample *t*-test results for deciles of neighborhood disadvantage index (NDI) variables between block groups with significantly elevated NDI effects and without significantly elevated NDI effects.

References

1.　Gallaway, M.S.; Henley, S.J.; Steele, C.B.; Momin, B.; Thomas, C.C.; Jamal, A.; Trivers, K.F.; Singh, S.D.; Stewart, S.L. Surveillance for cancers associated with tobacco use—United States, 2010—2014. *MMWR Surveill. Summ.* **2018**, *67*, 1–42. [CrossRef] [PubMed]

2.　Nelson, D.E.; Jarman, D.W.; Rehm, J.; Greenfield, T.K.; Rey, G.; Kerr, W.C.; Miller, P.; Shield, K.; Ye, Y.; Naimi, T. Alcohol-Attributable Cancer Deaths and Years of Potential Life Lost in the United States. *Am. J. Public Health* **2013**, *103*, 641–648. [CrossRef]

3.　Vineis, P.; Alavanja, M.; Buffler, P.; Fontham, E.; Franceschi, S.; Gao, Y.T.; Gupta, P.C.; Hackshaw, A.; Matos, E.; Samet, J.; et al. Tobacco and Cancer: Recent Epidemiological Evidence. *J. Natl. Cancer Inst.* **2004**, *96*, 99–106. [CrossRef] [PubMed]

4.　Hecht, S.S. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat. Cancer* **2003**, *3*, 733–744. [CrossRef] [PubMed]

5.　Alaniz, M.L. Alcohol Availability and Targeted Advertising in Racial/Ethnic Minority Communities. *Alcohol Heal Res. World* **1998**, *22*, 286.

6.　Poland, B.; Frohlich, K.; Haines-Saah, R.; Mykhalovskiy, E.; Rock, M.; Sparks, R. The social context of smoking: The next frontier in tobacco control? *Tob. Control* **2006**, *15*, 59–63. [CrossRef]

7.　Berke, E.M.; Tanski, S.E.; Demidenko, E.; Alford-Teaster, J.; Shi, X.; Sargent, J.D. Alcohol retail density and demographic predictors of health disparities: A geographic analysis. *Am. J. Public Health* **2010**, *100*, 1967–1971. [CrossRef]

8.　Shortt, N.K.; Tisch, C.; Pearce, J.; Mitchell, R.; Richardson, E.A.; Hill, S.; Collin, J. A cross-sectional analysis of the relationship between tobacco and alcohol outlet density and neighbourhood deprivation. *BMC Public Health* **2015**, *15*, 1014. [CrossRef]

9.　Major, J.M.; Sargent, J.D.; Graubard, B.I.; Carlos, H.A.; Hollenbeck, A.R.; Freedman, N.D.; McGlynn, K.A. Local geographic variation in chronic liver disease and hepatocellular carcinoma: Contributions of socioeconomic deprivation, alcohol retail outlets, and lifestyle. *Ann. Epidemiol.* **2014**, *24*, 104–110. [CrossRef]

10.　Wheeler, D.C.; Do, E.K.; Hayes, R.B.; Fugate-Laus, K.; Fallavollita, W.L.; Hughes, C.; Fuemmeler, B.F. Neighborhood Disadvantage and Tobacco Retail Outlet and Vape Shop Outlet Rates. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2864. [CrossRef]

11.　Lee, J.P.; Ponicki, W.; Mair, C.; Gruenewald, P.; Ghanem, L. What explains the concentration of off-premise alcohol outlets in Black neighborhoods? *SSM-Popul. Health* **2020**, *12*, 100669. [CrossRef] [PubMed]

12.　Cantrell, J.; Pearson, J.L.; Anesetti-Rothermel, A.; Xiao, H.; Kirchner, T.R.; Vallone, D. Tobacco Retail Outlet Density and Young Adult Tobacco Initiation. *Nicotine Tob. Res.* **2016**, *18*, 130–137. [CrossRef] [PubMed]

13.　Johnson, J.L.; Canterberry, M.; Vu, U.L.; Carton, T.W. Positive association between neighborhood retail alcohol outlet access and hospital admissions for alcohol withdrawal syndrome in New Orleans. *J. Addict. Dis.* **2020**, *38*, 311–316. [CrossRef] [PubMed]

14.　Wheeler, D.C.; Boyle, J.; Barsell, D.J.; Glasgow, T.; McClernon, F.J.; Oliver, J.A.; Fuemmeler, B.F. Associations of Alcohol and Tobacco Retail Outlet Rates with Neighborhood Disadvantage. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1134. [CrossRef]

15.　Wheeler, D.C.; Raman, S.; Jones, R.M.; Schootman, M.; Nelson, E.J. Bayesian deprivation index models for explaining variation in elevated blood lead levels among children in Maryland. *Spat. Spatio-Temporal. Epidemiol.* **2019**, *30*, 100286. [CrossRef]

16.　Wheeler, D.C.; Boyle, J.; Raman, S.; Nelson, E.J. Modeling elevated blood lead level risk across the United States. *Sci. Total Environ.* **2021**, *769*, 145237. [CrossRef] [PubMed]

17.　Wheeler, D.C.; Do, E.K.; Hayes, R.B.; Hughes, C.; Fuemmeler, B.F. Evaluation of neighborhood deprivation and store characteristics in relation to tobacco retail outlet sales violations. *PLoS ONE* **2021**, *16*, e0254443. [CrossRef]

18.　Wheeler, D.C.; Boyle, J.; Nelson, E.J. Modeling annual elevated blood lead levels among children in Maryland in relation to neighborhood deprivation. *Sci. Total Environ.* **2022**, *805*, 150333. [CrossRef]

19. Fotheringham, A.S.; Brunsdon, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; John Wiley & Sons: Hoboken, NJ, USA, 2003.

20. Finley, A.O. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. *Methods Ecol. Evol.* **2011**, *2*, 143–154. [CrossRef]

21. Gelfand, A.E.; Kim, H.-J.; Sirmans, C.F.; Banerjee, S. Spatial modeling with spatially varying coefficient processes. *J. Am. Stat. Assoc.* **2003**, *98*, 387–396. [CrossRef]

22. Yao, N.; Foltz, S.M.; Odisho, A.Y.; Wheeler, D.C. Geographic Analysis of Urologist Density and Prostate Cancer Mortality in the United States. *PLoS ONE* **2015**, *10*, e0131578. [CrossRef] [PubMed]

23. Czarnota, J.; Wheeler, D.C.; Gennings, C. Evaluating geographically weighted regression models for environ-mental chemical risk analysis. *Cancer Inform.* **2015**, *14*, 117–127. [PubMed]

24. Wheeler, D.C.; Waller, L.A. Comparing spatially varying coefficient models: A case study examining violent crime rates and their relationships to alcohol outlets and illegal drug arrests. *J. Geogr. Syst.* **2009**, *11*, 1–22. [CrossRef]

25. Wheeler, D.C. Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression: The Geographically Weighted Lasso. *Environ. Plan.* **2009**, *41*, 722–742. [CrossRef]

26. Wheeler, D.C. Diagnostic Tools and a Remedial Method for Collinearity in Geographically Weighted Regression. *Environ. Plan. A* **2007**, *39*, 2464–2481. [CrossRef]

27. Wheeler, D.; Tiefelsdorf, M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *J. Geogr. Syst.* **2005**, *7*, 161–187. [CrossRef]

28. Waller, L.A.; Zhu, L.; Gotway, C.A.; Gorman, D.M.; Gruenewald, P.J. Quantifying geographic variations in as-sociations between alcohol distribution and violence: A comparison of geographically weighted regression and spatially varying coefficient models. *Stoch. Environ. Res. Risk Assess.* **2007**, *21*, 573–588. [CrossRef]

29. Rollinson, C.R.; Finley, A.O.; Alexander, M.R.; Banerjee, S.; Hamil, K.-A.D.; Koenig, L.E.; Locke, D.H.; DeMarche, M.L.; Tingley, M.W.; Wheeler, K.; et al. Working across space and time: Nonstationarity in ecological research and application. *Front. Ecol. Environ.* **2021**, *19*, 66–72. [CrossRef]

30. Retail Permits. North Carolina Alcoholic Beverage Control Commission, 2021. Available online: https://abc.nc.gov/Permit/Retail (accessed on 1 March 2021).

31. Golden, S.D.; Baggett, C.D.; Kuo, T.-M.; Kong, A.Y.; Delamater, P.L.; Tao, V.Q.; Ribisl, K.M. Trends in the Number and Type of Tobacco Product Retailers, United States, 2000–2017. *Nicotine Tob. Res.* **2021**, *24*, 77–84. [CrossRef]

32. Sparks, C.S. DEM 7263–Calculating Indices of Residential Segregation. 2019. Available online: https://rpubs.com/corey_sparks/473785 (accessed on 5 April 2022).

33. Wheeler, D.C.; Jones, R.M.; Schootman, M.; Nelson, E.J. Explaining variation in elevated blood lead levels among children in Minnesota using neighborhood socioeconomic variables. *Sci. Total Environ.* **2019**, *650*, 970–977. [CrossRef]

34. Wheeler, D.C.; Boyle, J.; Barsell, D.J.; Maguire, R.L.; Dahman, B.; Murphy, S.K.; Hoyo, C.; Zhang, J.; Oliver, J.A.; McClernon, J.; et al. Neighborhood Deprivation is Associated with Increased Risk of Prenatal Smoke Exposure. *Prev. Sci.* **2022**, 1–12. [CrossRef]

35. Srinivasan, R.; Venkatesan, P. Bayesian Multivariate Conditional Autoregressive Model for Tuberculosis and HIV in India. *Indian J. Appl. Res.* **2019**, *8*, 148–150.

36. Xie, K.; Ozbay, K.; Yang, H. A multivariate spatial approach to model crash counts by injury severity. *Accid. Anal. Prev.* **2019**, *122*, 189–198. [CrossRef] [PubMed]

37. Gelman, A.; Rubin, D.B. Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.* **1992**, *7*, 457–472. [CrossRef]

38. Sturtz, S.; Ligges, U.; Gelman, A. R2OpenBUGS: A Package for Running OpenBUGS from R. R Package Version 3.2-3.2.1. Available online: https://cran.r-project.org/web/packages/R2OpenBUGS/index.html (accessed on 16 April 2022).

39. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.

40. Loomis, B.R.; Kim, A.E.; Goetz, J.L.; Juster, H.R. Density of tobacco retailers and its association with sociodemographic character-istics of communities across New York. *Public Health* **2013**, *127*, 333–338. [CrossRef]

41. Romley, J.A.; Cohen, D.; Ringel, J.; Sturm, R. Alcohol and Environmental Justice: The Density of Liquor Stores and Bars in Urban Neighborhoods in the United States. *J. Stud. Alcohol Drugs* **2007**, *68*, 48–55. [CrossRef] [PubMed]

42. Rautela, Y.S.; Reddy, B.V.; Singh, A.K.; Gupta, A. Smoking and alcoholism among adult population and its association with outlet density in a hilly area of North India. *J. Prev. Med. Hyg.* **2019**, *60*, E361.

43. Lian, M.; Madden, P.A.; Lynskey, M.T.; Colditz, G.A.; Lessov-Schlaggar, C.N.; Schootman, M.; Heath, A.C. Geographic Variation in Maternal Smoking during Pregnancy in the Missouri Adolescent Female Twin Study (MOAFTS). *PLoS ONE* **2016**, *11*, e0153930. [CrossRef]

44. Shoff, C.; Yang, T.-C. Understanding maternal smoking during pregnancy: Does residential context matter? *Soc. Sci. Med.* **2013**, *78*, 50–60. [CrossRef]

45. Cilluffo, A.; Geiger, A.W.; Fry, R. More US Households are Renting than at Any Point in 50 Years. Pew Research Center. 2017. Available online: https://www.pewresearch.org/fact-tank/2017/07/19/more-u-s-households-are-renting-than-at-any-point-in-50-years/ (accessed on 16 April 2022).

46. Schoenborn, C.A.; Gindi, R.M. *Electronic Cigarette Use Among Adults: United States, 2014*; NCHS Data Brief No 217; National Center for Health Statistics: Hyattsville, MD, USA, 2015; pp. 1–8.

Int. J. Environ. Res. Public Health **2022**, *19*, 5244

13 of 13

47. Yu, D.; Peterson, N.A.; Sheffer, M.A.; Reid, R.J.; Schnieder, J.E. Tobacco outlet density and demographics: Analysing the relationships with a spatial regression approach. *Public Health* **2010**, *124*, 412–416. [CrossRef]

48. Duncan, D.T.; Kawachi, I.; Melly, S.J.; Blossom, J.; Sorensen, G.; Williams, D.R. Student Column Demographic Disparities in the Tobacco Retail Environment in Boston: A Citywide Spatial Analysis. *Public Health Rep.* **2014**, *129*, 209–215. [CrossRef] [PubMed]

49. Kong, A.Y.; Delamater, P.L.; Gottfredson, N.C.; Ribisl, K.M.; Baggett, C.D.; Golden, S.D. Sociodemographic inequities in tobacco retailer density: Do neighboring places matter? *Health Place* **2021**, *71*, 102653. [CrossRef] [PubMed]

50. Makadia, L.D.; Roper, P.J.; Andrews, J.O.; Tingen, M.S. Tobacco Use and Smoke Exposure in Children: New Trends, Harm, and Strategies to Improve Health Outcomes. *Curr. Allergy Asthma Rep.* **2017**, *17*, 55. [CrossRef] [PubMed]