Published in final edited form as:

Math Popul Stud. 2019 ; 26(4): 186–207. doi:10.1080/08898480.2018.1553399.

Neighborhood affluence protects against antenatal smoking: evidence from a spatial multiple membership model

Jennifer B. Kane, PhD,

Department of Sociology, University of California, Irvine

Ehsan Farshchi

Department of Sociology, University of California, Irvine

Abstract

A spatial multiple membership model formalizes the effect of neighborhood affluence on antenatal smoking. The data are geocoded New Jersey birth certificate records linked to United States census tract-level data from 1999 to 2007. Neighborhood affluence shows significant spatial autocorrelation and local clustering. Better model fit is observed when incorporating the spatial clustering of neighborhood affluence into multivariate analyses. Relative to the spatial multiple membership model, the multilevel model that ignores spatial clustering produced downwardly biased standard errors; the effective sample size of the key parameter of interest (neighborhood affluence) is also lower. Residents of communities located in high-high affluence clusters likely have better access to health-promoting institutions that regulate antenatal smoking behaviors.

Keywords

Health; Multi-level analysis; Markov chain; Neighborhood affluence; Prenatal smoking; USA

1. INTRODUCTION

Antenatal smoking is a modifiable cause of adverse birth outcomes such as low birth weight, preterm birth, and lower birthweight (Rosenzweig and Schultz, 1983; Kramer, 1987; Evans and Ringel, 1999; Hamilton, 2001; Cnattingius, 2004; Lien and Evans, 2005; Wehby et al., 2011). According to Lumley et al.'s (2009) analysis of randomized control trials, reductions in smoking during pregnancy are associated with a decreased risk of low birth weight and preterm birth, as well as an increase in mean birth weight. At least a portion of these effects are rooted in a physiological pathway whereby nicotine crosses the placenta, induces vasoconstriction (narrowing of the blood vessels of the umbilical cord), reducing the baby's oxygen supply and impeding fetal growth (Rogers, 2009). Identifying factors that prevent or reduce antenatal smoking could result in improved birth outcomes.

A reduction in antenatal smoking is a stated goal of the Centers for Disease Control Healthy People 2020 initiative (goal MICH 11.3, "Increase abstinence from cigarette smoking among

pregnant women”) (CDC, 2010) and is consistent with the World Health Organization’s Guidelines for Tobacco Use and Second-Hand Smoke Exposure in Pregnancy that call for increased access to psychosocial interventions for tobacco cessation among pregnant women (WHO, 2013). Reductions in antenatal smoking could also contribute to one of the targets of the United Nation’s Sustainable Development Goal 3 (“By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to at least as low as 12 per 1,000 live births and under-5 mortality to at least as low as 25 per 1,000 live births”; <http://www.un.org/sustainabledevelopment/health/>).

While formulating effective antenatal smoking cessation and prevention strategies to achieve these goals is critical, current antenatal smoking cessation efforts in the United States – the site of the present study – demonstrate limited success (for example, 5% reduction in smoking), and tend to be more successful in incentivizing lighter (versus heavier) smokers and spontaneous quitters to modify their behavior (Ershoff et al., 1999; Woodby et al., 1999; Chamberlain et al., 2017). 61% of U.S. women who smoke before becoming pregnant continue to smoke throughout the pregnancy (Curtin and Mathews, 2016). In order to meet the Healthy People 2020 goal, greater emphasis should be placed on understanding the role of social context in influencing antenatal smoking behavior (Poland et al., 2006).

We assess the role of neighborhood affluence in protecting against the risk of antenatal smoking, while simultaneously modeling the spatial patterning of neighborhood affluence. We situate our study within social disorganization theory and argue that local institutional resources that could effectively increase a woman’s chances of quitting before or upon learning of a pregnancy are more likely to be present within neighborhoods of concentrated high affluence, relative to neighborhoods of concentrated low affluence. We use all births in the State of New Jersey between 1999 and 2007 that have been geocoded and linked with U.S. census data describing the neighborhood social environment around conception (at birth year – 1). We reveal global and local spatial clustering of neighborhood affluence and demonstrate that incorporating this spatial patterning produces better model fit.

2. BACKGROUND

2.1 Prevalence and Predictors of Antenatal Smoking

The prevalence of tobacco use during pregnancy is less than 1% in Senegal, Guyana, Armenia, Ecuador, Guatemala, Honduras, and Nicaragua, while the prevalence is as high as 10-12% in the United States, 12% in Madagascar, 15% in Turkey, and 18% in Uruguay (Tong et al., 2013; WHO, 2013; Caleyachetty et al., 2014; Child Trends Databank, 2014). In the United States, upwards of 400,000 infants annually are exposed to maternal smoking in utero (Office of the Surgeon General, 2014).

Known individual-level risk factors of antenatal smoking include low levels of maternal education, low income, being unmarried, higher levels of nicotine addiction, having a spouse or a partner who also smokes, and maternal psychosocial problems; in the United States, another risk factor is White, American Indian, or Alaska Native race (Agrawal et al., 2008; Martin et al., 2008; Pickett et al., 2009; Curtin and Mathews, 2016). Neighborhood-level risk factors of antenatal smoking have not been thoroughly studied, in the United States or

around the world, despite the fact that the social environment is a known, powerful determinant of health behaviors and health outcomes (Link and Phelan, 1995; Sampson et al., 2002; Kawachi and Berkman, 2003; Sampson, 2003; Marmot and Wilkinson, 2009).

Residence in an economically deprived neighborhood may be a risk factor for antenatal smoking (Finch et al., 2001; Sellström et al., 2008; Nkansah-Amankra, 2010; Räisänen et al., 2014), although Lian et al. (2016) find no effect of neighborhood socioeconomic deprivation. Pickett et al. (2002) have shown that living in a primarily working-class neighborhood in California increases the risk of antenatal smoking, above and beyond mother's own social class status (here, a working-class neighborhood was indicated by the proportion of residents in administrative support, sales, service, operator and laborer occupations). In rural areas across the United States, higher levels of social capital are protective against antenatal smoking risk; in non-rural areas, the opposite holds (Shoff and Yang, 2013). Findings regarding United States residential segregation are mixed and depend on the racial/ethnic group under study, in terms of whether higher levels of same-ethnic density are protective or confer risk (Bell et al., 2007; Shaw et al., 2010; Yang et al., 2014).

2.2 Proposed Linkages between Neighborhood Affluence and Antenatal Smoking

Kane et al. (2017) showed that neighborhood affluence was a key protective factor of antenatal smoking, above and beyond neighborhood disadvantage and key maternal-level confounders. In medical sociology, neighborhood affluence (relative to neighborhood disadvantage) has often appeared as a superior predictor of self-reported health status and objectively-measured health status (Browning and Cagney, 2003; Browning et al., 2003; Wen et al., 2003; Cagney et al., 2005; Matthews and Yang, 2010; King et al., 2011). This effect of neighborhood affluence is also robust to the inclusion of individual-level socioeconomic status, health behaviors, and insurance coverage (Browning and Cagney, 2003; Cagney et al., 2005).

Social disorganization theory elucidates why neighborhood affluence may influence health behaviors, such as antenatal smoking. A neighborhood in the midst of decline is not only increasingly characterized by a high concentration of poverty, but is also increasingly resource-poor: local institutions providing social, economic, and health-related resources to residents flee the neighborhood, along with the more socioeconomically advantaged residents (Shaw and McKay, 1942; Kornhauser, 1978; Wilson, 1987). In contrast, neighborhoods with a high concentration of wealthy, highly educated individuals employed in professional or managerial occupations draw local institutions to their neighborhood; all residents then, regardless of their own socioeconomic resources, stand to benefit from the institution's resources and services (Robert, 1998; Browning and Cagney, 2003).

It is plausible that the rich array of local institutions often present in affluent neighborhoods regulate a woman's chances of successfully quitting smoking before or upon learning a pregnancy. Institutions within a local neighborhood can foster health by regulating opportunities for individuals to engage in health-promoting behaviors (such as regular physical activity (McNeill et al., 2006)) or regulating access to resources that people can use to procure health (Kawachi and Berkman, 2003). In the case of antenatal smoking, such resources may include access to an evidence-based intervention program—a predictor of

smoking cessation during pregnancy (Woodby et al., 1999; Melvin et al., 2000; Lee and Kahende, 2007). Additionally, participating in regular physical activity increases women's chances of not relapsing after a successful quit attempt, as physical activity regulates body weight and improves one's response to stress (Marcus et al., 1991). It is plausible that local opportunities to engage in physical activity can be influential for pregnant women who have recently quit smoking, thereby reducing the chance of relapse during pregnancy.

2.3 The Spatial Concentration of Affluence

Although many neighborhood-level characteristics are plausibly spatially concentrated (Owen et al., 2016), a strong case has been made on this point regarding affluence. The Gini index decreased in the United States by 6% from 1950 to 1975 (Massey et al., 2003), but then rose by 14% through the 1980s and 1990s; there it remained relatively constant through the early 2000s until the recession in 2007 (U.S. Census Bureau, 2017). As income inequality rose, the wealthiest individuals opted to reside in neighborhoods alongside other similarly advantaged families (Massey and Eggers, 1993; Massey, 1996; Massey et al., 2003; Dwyer, 2007; Reardon and Bischoff, 2011). In the early 2000s then, neighborhood affluence was highly spatially concentrated in the United States; this is the time period of interest in this study.

Individuals and families residing in areas of concentrated affluence stand to become increasingly privileged, as their individual-level advantages are compounded by the ecological benefits of living in safe, stable neighborhoods that contain an array of local institutions to meet the needs of all residents whereas the converse holds for those living in areas of concentrated low affluence (Massey et al., 2003). Massey (1996: 395) predicted early on that such spatial patterning would create a “deeply divided and increasingly violent social world” in the United States, wherein the day-to-day realities of individuals in different neighborhoods would become increasingly divergent.

2.4 The Incorporation of Spatial Associations into the Multilevel Model

Incorporating the spatial autocorrelation of neighborhood-level characteristics into multilevel modeling is critical, on both statistical and substantive grounds (Caicedo and Jones, 2014). On statistical grounds, not incorporating spatial autocorrelation of neighborhood-level characteristics can lead to downwardly biased standard errors and inaccurate estimations of confidence intervals (Corrado and Fingleton, 2011). Savtiz and Raudenbush (2009) compare multiple estimators and show that an empirical Bayes estimator that takes into account spatial dependence produces a more internally consistent estimate of the mean level of neighborhood collective efficacy, relative to both an ordinary least squares estimator and an empirical Bayes estimator that do not take into account spatial dependence. In the case of travel satisfaction in Beijing, Dong et al. (2016) find improved model fit within models that incorporate spatial dependence. Owen, Harris, and Jones (2016) demonstrate that taking spatial dependence into account within a multilevel model can increase the effective sample size.

Multilevel models usually rely on spatial exchangeability, according to which the same associations between a neighborhood-level risk factor and an individual-level outcome

would be observed, regardless of the physical location of the neighborhood in relation to all other neighborhoods (Jones and Subramanian, 2013). Figure 1 shows a series of neighborhood units. The four neighborhoods shaded in grey in Panel A exchange spatial locations in Panel B. Under the assumption of spatial exchangeability, the results of a multilevel model estimated using either the neighborhood structure portrayed in Panel A or Panel B are equivalent. This assumption is not satisfied however if results are sensitive to the position of a neighborhood relative to other spatial units.

Dong and Harris (2015) argue that multilevel associations comprise dependencies of observations related to the same statistical unit, which require multilevel models, and autocorrelation of neighborhood-level characteristics between geographical units. Accounting for both types of dependencies acknowledges that neighborhood processes do not operate in isolation from those within the surrounding areas, but rather that spillovers are probable (Savitz and Raudenbush, 2009). Indeed, scholars have incorporated spatial patterning into multilevel models. For example, Sampson and colleagues (1999) used this type of model to predict intergenerational closure and reciprocal exchange; Morenoff (2003) used a similar model to predict infant birth weight, as did Bartolomeo et al. (2010) who predicted hospitalization rates for lung cancer and chronic obstructive pulmonary disease.

The spatial multiple membership model incorporates spatial clustering of a neighborhood-level variable into a multilevel model. The residuals are partitioned into unstructured residuals ϵ and a spatially structured residual μ . The spatial dependence is structured as the relationship between spatial units W and the degree of dependency ρ reflected in adjacency and weight matrices:

$$Y = \beta X + \rho W\mu + \epsilon \quad (1)$$

(Browne et al., 2001; Rasbash and Browne, 2008; Jones and Subramanian, 2013). Then, within the spatial multiple membership model, features of adjacent neighborhoods are allowed to influence the referent neighborhood, thereby avoiding the assumption that each neighborhood exists in isolation from its spatial neighbors. Lower-level characteristics (in this case, individuals) are nested within higher-level features (in this case, neighborhoods) while also taking into account that adjacent geographical units (neighborhoods) influence one another. This nesting structure is portrayed in Figure 2: here, individuals are nested within their own residential neighborhood, as well as within the surrounding neighborhoods (defined here as neighborhoods within a 1 kilometer radius of their own residential neighborhood).

We examine the effect on model fit of the spatial patterning of affluence associated with antenatal smoking. If better model fit is observed here, then mothers living in neighborhoods located within an area of concentrated high affluence—that is, a highly affluent neighborhood surrounded by other highly affluent neighborhoods—likely have better access to health-promoting institutions. Mothers residing in areas of concentrated low affluence have to travel greater distances to access the same resources.

3. METHOD

3.1 Data and Sample

We use the electronic birth certificate records of 1,008,050 births from 1999 to 2007 in the State of New Jersey.¹ Ninety four percent of all records were successfully geocoded. Of the 6% of cases ($n=61,317$) that were not able to be geocoded, most cases were located outside the state of New Jersey and we therefore excluded them ($n=24,093$), had some degree of missing address information available on the birth certificate record ($n=27,833$), or were cases in which only a post office box was listed ($n=9,379$). A handful ($n=12$) had no address listed for the mother. We used the Federal Information Processing Standard code to append census (neighborhood-level) data to the birth record (individual-level) data. The sample includes all women for whom a plausible geocode could be assigned ($n = 946,733$), belonging to non-Hispanic White, non-Hispanic Black, Hispanic, or non-Hispanic Asian race/ethnicity ($n = 933,083$). We present a map of census tracts (according to 2000 census tract boundaries) in Appendix A, in which mothers giving birth during the sample period resided. We retain 99.4% of all census tracts (1,933 out of 1,944). The excluded tracts ($n=11$) had either zero population or contained populations residing within group quarters, military quarters, correctional facilities, nursing homes, or other institutions.

3.2 Measures

Maternal antenatal smoking is indicated on the birth certificate record and combines information from the medical record with maternal post-partum self-report (Reichman et al., 2008). Arendt et al. (1999) have shown that combining more than one source of information to assess antenatal substance abuse is more accurate than any one method. Here, maternal antenatal smoking is indicated by any smoking during the pregnancy (yes/no).

We use indices of neighborhood affluence and disadvantage. We lag affluence and disadvantage by one year to reflect neighborhood conditions at or around the time the mother became pregnant. Following King, Morenoff, and House (2011), neighborhood affluence is reflected by the proportion of 16+ year old civilian workers in professional/managerial occupations, the proportion of 25+ year olds with 16+ years education, and median home values. Neighborhood disadvantage is reflected by the proportion of households with income <\$15,000, the proportion of households with income <\$50,000, the proportion of families in poverty, the proportion of households receiving public assistance, the total unemployment rate, and the proportion of vacant housing units.

Tract-level information on these nine variables was drawn from United States decennial census data files (1990, 2000, 2010). Because census tract boundaries change over time, we identified 5,122 tract-based neighborhoods that were temporally unique over the 8-year study period. We found no other method of accounting for over-time boundary changes, including using normalized tract data, provide more confident results. We interpolated

¹This study focuses on 1999-2007 because income inequality, a key driver of the spatial concentration of neighborhood affluence, was high (relative to levels dating back to 1950) and remained relatively constant in the United States between 1999 and the start of the Great Recession (in December 2007) U.S. Census Bureau. (2017). Historical Income Tables: Income Inequality. Washington D.C., U.S. Census Bureau.

linearly to estimate values for intercensal years. Mayer (2002) argues that changes in economic measures at the national-level are well-approximated by an assumption of linearity, and if anything, use of linear interpolation likely produces a conservative (that is, downwardly biased) estimate of these effects.

We standardized each variable by rescaling each to have a mean of 0 and a standard deviation of 1. We performed a factor analysis; we retain the first principal component, which represents neighborhood affluence, and the second one, which represents neighborhood disadvantage. The variables comprising the affluence factor are internally consistent (Cronbach's $\alpha = 0.93$), as are the variables comprising the disadvantage factor (Cronbach's $\alpha = 0.92$).

The birth certificate record contains mother's completed years of education (range = 1-17), age (range = 12-60), race/ethnicity (non-Hispanic Black, Hispanic, Asian, and non-Hispanic White), foreign-born status (yes/no), marital status (yes/no), employment in the year prior to birth, and participation in Medicaid during the pregnancy (yes/no). This latter variable reflects the mother's household income: women are eligible to participate if their household income is below 200% of the federal poverty line. In 2003, the median year in our data, the federal poverty line for a family of 2 was \$12,120 and for a family of 4 was \$18,400 (<https://aspe.hhs.gov/2003-hhs-poverty-guidelines>). This indicator effectively compares poor and near-poor women to non-poor women.

3.3 Statistical Analyses

We present the Moran's I statistic (range between -1 and $+1$) for affluence, and a local indicator of spatial association cluster map depicting areas of high affluence surrounded by areas of high affluence (hereafter, a high-high cluster), areas of low affluence surrounded by areas of low affluence (hereafter, a low-low cluster), areas of high affluence surrounded by areas of low affluence (hereafter, a high-low cluster), and areas of low affluence surrounded by areas of high affluence (hereafter, a low-high cluster). We define neighborhoods by census tract, and the adjacency matrix has a buffer of one kilometer. A 1 km buffer (relative to that of a larger or smaller size) is suitable for assessing an individuals' exposure to local institutions present in their neighborhood, such as those that a resident would reasonably come across on a regular basis (Black et al., 2011). We found very close results using a 2 km buffer.

We first run a single-level logistic regression that ignores clustering of mothers within neighborhoods as well as the spatial clustering of neighborhood affluence. Because the nested structure of our data violates the assumption of independent observations (Raudenbush and Bryk, 2001), this regression is presumed to be misspecified. We then run a multilevel logistic regression that incorporates clustering of mothers within neighborhoods but ignores the spatial clustering of neighborhood affluence. Given that neighborhood affluence is likely spatially patterned, this regression should also be misspecified.

Therefore we use a spatial multiple membership model that incorporates clustering of mothers within neighborhoods and spatial clustering of neighborhood affluence. We follow the specification of the spatial multiple membership model formally defined by Browne et

al. (2001) (Eq. (7) on page 111) and by Jones and Subramanian (2013), who describe its implementation in MLwiN. We identify two classifications: the first denotes a woman's residential neighborhood, and the second denotes the neighborhoods surrounding her residential neighborhood. To estimate the set of multiple membership relations for the spatial multiple membership model, we created a distance-based adjacency matrix of all neighborhoods in the state of New Jersey (based on the centroid of each neighborhood) and transformed it into a Boolean matrix (less than or equal to 1 km indicated by 1; over 1 km indicated by 0) and then into an adjacency list. We distributed the weights equally between the tracts and their neighbors, summing to 1. Figure 3 presents a referent neighborhood. All neighborhoods in which the centroid (depicted with a black dot) falls within a 1 km buffer are included in the adjacency matrix. In this example, because five neighborhoods are encompassed within the buffer shown in Figure 3, each is assigned a 1/5 (or 20%) weight.

To select the best fitting model, we use the deviance information criterion, which extends the Akaike information criterion, which equals the deviance + $2p$, where p is the total number of parameters estimated in a model. We retain the model with the smallest deviance information criterion value.

We estimate the three multivariate models using Markov chain Monte Carlo. We include diffuse prior distributions for all model parameters. The chain length across each of the three models depends on the length necessary to achieve convergence. We use a burn-in of 500 iterations. Instead of p values, as usual with Bayesian techniques, we represent the marginal posterior distributions of the estimated parameters (Kruschke, 2014), the highest density intervals, which reflect the points of a distribution that are most credible (for example, 95% of the distribution), the Raftery-Lewis diagnostic (1996) and Brooks-Draper (Browne, 2015) diagnostics, which help evaluate whether the chain length is sufficient in length (Jones and Subramanian, 2013), the autocorrelation of the chain at lag k , and notably the Monte Carlo standard error, which is the ratio of the standard deviation of the chain to the effective sample size. This effective sample size reflects the ratio of the actual sample size to the amount of autocorrelation, and indicates stable estimates of the limits of the 95% highest density interval (Kruschke, 2014). The plot of the Monte Carlo standard error informs about the chain length necessary to achieve a mean estimate with a given Monte Carlo standard error.

The electronic birth certificate records contain missing data. We used complete case analysis (Allison, 2000), which produces unbiased estimates when missingness depends only on the independent variables, and not on the response variable.

4. RESULTS

Table 1 presents the mean, standard deviation, and range of all model covariates. On average, a mother completed 13.5 years of education at the time of birth and 22% of mothers participated in Medicaid during the pregnancy. 34% of women were foreign-born; this state is an important immigrant gateway in the United States (Frey, 2006). 52% of mothers were White, 23% were Hispanic, 16% Black, and 8% Asian. 69% of mothers were married at the time of birth, a proportion consistent with national trends in nonmarital childbearing during

this time period (Child Trends Databank, 2015). On average, a mother was 30 years old at the time of birth and 63% were employed in the year prior to birth.

Figure 4 indicates significant spatial autocorrelation of neighborhood affluence across census tracts (Moran's $I = 0.747$ based on a 1 km buffer adjacency matrix, pseudo $p = 0.001$ based on 999 permutations). This result is unchanged by using different weight matrices. The positive sign of the statistic indicates positive spatial autocorrelation, meaning that neighborhoods tend to be surrounded by others with a similar level of affluence. In Figure 5 indeed, most data points lie in the north-eastern quadrant, indicating high affluence neighborhoods surrounded by other high affluence neighborhoods, and the south-western quadrant, where low affluence neighborhoods are surrounded by other low affluence neighborhoods. This supports our including the spatial patterning of affluence into the multivariate analysis. The local indicator of spatial association analysis, presented in Figure 5, brings evidence of local spatial clustering of neighborhood affluence. We use the same contiguity matrix at the 5% level of significance. Local clusters of high-high affluence ($n = 29$) are indicated in black and are located primarily in the north-central, mid-western, and south-central parts of the state. Local clusters of low-low affluence ($n = 178$) are indicated in dark grey and are scattered throughout the state. We could see no clusters of high-low or low-high affluence in these data.

The aspatial single-level model, presented in Table 2, shows that neighborhood affluence (lagged by one year) is inversely associated with antenatal smoking. A one unit (standard deviation) increase in affluence is associated with a 29% reduction in the odds of antenatal smoking (odds ratio = $e^{-0.34} = 0.71$; standard error = 0.006). Neighborhood disadvantage is positively associated with antenatal smoking, such that a one unit (standard deviation) increase in disadvantage is associated with a 7% increase in the odds of antenatal smoking ($e^{0.07} = 1.07$; standard error = 0.005). The association between neighborhood affluence and antenatal smoking is greater in magnitude than that of neighborhood disadvantage and antenatal smoking. Individual-level risk factors of antenatal smoking include fewer years of completed education, residence in a poor or near poor household (indicated by participation in Medicaid during pregnancy), being native-born, non-Hispanic White, older maternal age, and not being employed in the year prior to birth. The credible interval for each model covariate does not encompass 0.

Within the aspatial multilevel model, a one unit (standard deviation) increase in neighborhood affluence is associated with a 27% decrease in the odds of antenatal smoking (odds ratio = $e^{-0.31} = 0.73$; standard error = 0.011). A one unit (standard deviation) increase in neighborhood disadvantage is associated with a 2% increase in the odds of antenatal smoking (odds ratio = $e^{0.02} = 1.02$; standard error = 0.012). The posterior means for the individual-level risk factors are in the same direction as that previously reported in the aspatial single level model. The credible interval for each model covariate does not encompass 0. The aspatial multilevel model has a higher deviance information criterion than the aspatial single level model.

In the spatial multiple membership model, a one unit (standard deviation) increase in neighborhood affluence is associated with a 24% decrease in the odds of antenatal smoking

(odds ratio = $e^{-0.28} = 0.76$; standard error = 0.014) whereas a one unit (standard deviation) increase in neighborhood disadvantage is associated with a 5% increase in the odds of antenatal smoking (odds ratio = $e^{0.05} = 1.05$; standard error = 0.013). No credible interval for model covariates encompasses 0. The between-neighborhood variance in affluence is greater than that of disadvantage; this holds for both the spatially unstructured and structured terms. The deviance information criterion values indicate that the spatial multiple membership model is the best fit of the data (214 lower in this model than in the aspatial multilevel model). On the basis that this is the best fitting model, the standard errors of the neighborhood-level characteristics (affluence and disadvantage) in the aspatial multilevel model are downwardly biased.

Figures 6 and 7 present the marginal posterior distributions of neighborhood affluence and neighborhood disadvantage, from the spatial multiple membership model. For affluence, the 95% highest density interval is $[-0.305, -0.251]$. The trace of the trajectory in the autocorrelation functions plot reached zero, and the effective sample size is 1067. (The effective sample size in the aspatial multilevel model is 285). The Raftery-Lewis quantile plot indicates a chain length of 30000-35300 (when $q = 0.025$ and 0.975); thus, the chain estimated here (of 50000) is likely to be of sufficient length. The Brooks-Draper diagnostic indicates that a chain length of 864 would be necessary to be 95% confident of estimating the posterior mean to two significant figures at the 5% level of significance; this further supports the fact that the chain is of sufficient length. The parameter plot in the upper left hand box of Figure 5 shows that the Markov chain Monte Carlo algorithm is exploring the parameter space in an even manner, and there is no evidence of trending. For neighborhood disadvantage, the 95% highest density interval is $[0.022, 0.072]$ and the effective sample size is 920. Model diagnostics indicate that the chain is of sufficient length (based on both the Raftery-Lewis quantile plot and the Brooks-Draper diagnostic) and no evidence of trending in the Markov chain Monte Carlo algorithm.

We conducted supplementary analyses (available upon request) to check the robustness of our findings. We first omitted births from the year 2007 to ascertain the effect of the recession that began nationwide in December 2007. The effect is invisible. Second, we nested women within census tracts based on census boundaries from the year 2000, on the assumption that the neighborhood social environment in 2000 did not change from 1999 to 2007. The posterior means are very close to those of Table 2, which avoids the assumption of no neighborhood change over time.

5. CONCLUSION

We have estimated the association between neighborhood affluence — a sociological construct signaling the presence of highly educated, wealthy residents, who are employed in prestigious occupations and can effectively draw local institutions to their community that both stabilize the neighborhood and meet the basic needs of all residents — and the risk of antenatal smoking, while accounting for the spatial patterning of affluence.

We have found significant spatial autocorrelation (Moran's $I = 0.747$) and local spatial clustering of neighborhood affluence (29 clusters of high-high affluence and 178 clusters of

low-low affluence). Adjacent neighborhoods tend to have similar levels of affluence. Local clusters of high-high affluence are located primarily in the north-central, mid-western, and south-central parts of the state while clusters of low-low affluence are scattered throughout the state. This spatial clustering weakens the assumption of spatial exchangeability.

Our incorporating the spatial patterning of affluence has improved the multivariate analyses. We also found an inverse association between affluence and antenatal smoking, and it is higher than that between neighborhood disadvantage and antenatal smoking. In the aspatial multilevel model, we found downwardly biased standard errors and a lower effective sample size of neighborhood affluence.

We emphasize the role of local facilities and institutions in regulating smoking behavior. Women at higher risk of smoking during pregnancy are not only socioeconomically disadvantaged, but are also living in less affluent areas, and therefore lack access to services that could encourage cessation.

Among limitations, administratively-defined neighborhood boundaries, such as tracts, only approximate the social spaces individuals encounter (Galster, 2001). Neighborhood effects may be misestimated. Antenatal smoking is not objectively verified. In vital records, not all factors used to assign women to a given neighborhood are included. The estimated effects of neighborhood affluence and disadvantage on antenatal smoking risk may be confounded by omitted variables determining both the type of neighborhood in which women live, and her antenatal health behaviors. Results may depend on the spatial unit.

Acknowledgements:

This research received support from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development: K99/R00 HD075860 (PI: Jennifer B. Kane). The authors thank Brian Frizzelle for geospatial analysis support. Opinions reflect those of the authors and not necessarily those of the granting agencies.

Appendix A.

Figure 8.

Census tracts included ($n=1,933$) and excluded ($n=11$) in analysis. The excluded tracts ($n=11$) had either zero population or contained populations residing within group quarters, military quarters, correctional facilities, nursing homes, or other institutions.

REFERENCES

Agrawal A, Pergadia ML, Waldron M, Bucholz KK, Heath AC, Madden PA, Knopik VS, and Martin NG (2008). Correlates of Cigarette Smoking During Pregnancy and its Genetic and Environmental

- Overlap with Nicotine Dependence. *Nicotine & Tobacco Research*, 10(4): 567–578. [PubMed: 18418779]
- Allison PD (2000). Multiple Imputation for Missing Data: A Cautionary Tale. *Sociological Methods & Research*, 28(3): 301–309.
- Arendt RE, Singer LT, Minnes S, and Salvator A (1999). Accuracy in Detecting Prenatal Drug Exposure. *Journal of Drug Issues*, 29(2): 203–214.
- Bartolomeo N, Trerotoli P, and Serio G (2010). Lung Cancer and COPD Rates in Apulia: A multilevel multimember model for smoothing disease mapping. *International Journal of Health Geographics*, 9(1): 15. [PubMed: 20205714]
- Bell JF, Zimmerman FJ, Mayer JD, Almgren GR, and Huebner CE (2007). Associations between residential segregation and smoking during pregnancy among urban African-American women. *Journal of Urban Health*, 84(3): 372–388. [PubMed: 17226080]
- Black JL, Carpiano RM, Fleming S, and Lauster N (2011). Exploring the distribution of food stores in British Columbia: associations with neighbourhood socio-demographic factors and urban form. *Health & Place*, 17(4): 961–970. [PubMed: 21565544]
- Browne WJ (2015). MCMC estimation in MLwiN, Version 2.32. University of Bristol, Centre for Multilevel Modelling.
- Browne WJ, Goldstein H, and Rasbash J (2001). Multiple Membership Multiple Classification (MMM) Models. *Statistical Modelling*, 1(2): 103–124.
- Browning CR and Cagney KA (2003). Moving beyond Poverty: Neighborhood structure, social processes, and health. *Journal of Health and Social Behavior*: 552–571. [PubMed: 15038149]
- Browning CR, Cagney KA, and Wen M (2003). Explaining Variation in Health Status Across Space and Time: Implications for racial and ethnic disparities in self-rated health. *Social Science & Medicine*, 57(7): 1221–1235. [PubMed: 12899906]
- Cagney KA, Browning CR, and Wen M (2005). Racial disparities in self-rated health at older ages: what difference does the neighborhood make? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 60(4): S181–S190.
- Caicedo B and Jones K (2014). Investigating neighbourhood effects on health: an approach to develop neighbourhood-level constructs using community-survey data. *Journal of Public Health*, 16(1): 88–100.
- Caleyachetty R, Tait CA, Kengne AP, Corvalan C, Uauy R, and Echouffo-Tcheugui JB (2014). Tobacco use in pregnant women: analysis of data from Demographic and Health Surveys from 54 low-income and middle-income countries. *The Lancet Global health*, 2(9): e513–e520. [PubMed: 25304418]
- CDC. (2010). Foundational Health Measures: General Health Status. Retrieved July 10, 2017, from <http://www.healthypeople.gov/2020/about/GenHealthAbout.aspx>.
- Chamberlain C, O'Mara-Eves A, Porter J, Coleman T, Perlen SM, Thomas J, and McKenzie JE (2017). Psychosocial interventions for supporting women to stop smoking in pregnancy. *Cochrane Database of Systematic Reviews*, 2: Cd001055. [PubMed: 28196405]
- Child Trends Databank. (2014). Mothers Who Smoke While Pregnant. Retrieved June 30, 2017, from <http://www.childtrends.org/?indicators=mothers-who-smoke-while-pregnant>.
- Child Trends Databank. (2015). Births to Unmarried Women. Retrieved June 30, 2017, from <https://www.childtrends.org/indicators/births-to-unmarried-women>.
- Cnattingius S (2004). The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. *Nicotine & Tobacco Research*, 6(Supplement 2): S125–S140. [PubMed: 15203816]
- Corrado L and Fingleton B (2011). Multilevel Modelling with Spatial Effects. Document Number: SIRE-DP-2011-13, Scottish Institute for Research in Economics.
- Curtin SC and Mathews TJ (2016). Smoking Prevalence and Cessation Before and During Pregnancy: Data From the Birth Certificate, 2014. *National Vital Statistics Reports*, 65(1): 1–14.
- Dong G and Harris R (2015). Spatial autoregressive models for geographically hierarchical data structures. *Geographical Analysis*, 47(2): 173–191.

- Dong G, Ma J, Harris R, and Pryce G (2016). Spatial random slope multilevel modeling using multivariate conditional autoregressive models: A case study of subjective travel satisfaction in Beijing. *Annals of the American Association of Geographers*, 106(1): 19–35.
- Dwyer RE (2007). Expanding homes and increasing inequalities: US housing development and the residential segregation of the affluent. *Social Problems*, 54(1): 23–46.
- Ershoff DH, Quinn VP, Boyd NR, Stern J, Gregory M, and Wirtschafter D (1999). The Kaiser Permanente prenatal smoking-cessation trial: When more isn't better, what is enough? *American Journal of Preventive Medicine*, 17(3): 161–168. [PubMed: 10987630]
- Evans WN and Ringel JS (1999). Can higher cigarette taxes improve birth outcomes? *Journal of Public Economics*, 72(1): 135–154.
- Finch BK, Vega WA, and Kolody B (2001). Substance use during pregnancy in the state of California, USA. *Social science & medicine*, 52(4): 571–583. [PubMed: 11206654]
- Frey WH (2006). *Immigration Goes Nationwide: Recent dispersal has made immigrants and new minorities more visible*, The Brookings Institution.
- Galster G (2001). On the nature of neighbourhood. *Urban Studies*, 38(12): 2111.
- Hamilton BH (2001). Estimating treatment effects in randomized clinical trials with non-compliance: the impact of maternal smoking on birthweight. *Health Economics*, 10(5): 399–410. [PubMed: 11466802]
- Jones K and Subramanian SV (2013). Developing multilevel models for analysing contextuality, heterogeneity and change using MLwiN 2.2.
- Kane JB, Miles G, Yourkavitch J, and King K (2017). Neighborhood context and birth outcomes: Going beyond neighborhood disadvantage, incorporating affluence. *SSM - Population Health*, 3: 699–712. [PubMed: 29349258]
- Kawachi I and Berkman LF (2003). *Neighborhoods and Health*. New York: Oxford University Press.
- King KE, Morenoff JD, and House JS (2011). Neighborhood context and social disparities in cumulative biological risk factors. *Psychosomatic Medicine*, 73(7): 572. [PubMed: 21862824]
- Kornhauser RR (1978). *Social sources of delinquency: An appraisal of analytic models*. Chicago, IL: University of Chicago Press
- Kramer MS (1987). Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the World Health Organization*, 65(5): 663. [PubMed: 3322602]
- Kruschke J (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*: Academic Press.
- Lee CW and Kahende J (2007). Factors associated with successful smoking cessation in the United States, 2000. *American Journal of Public Health*, 97(8): 1503–1509. [PubMed: 17600268]
- Lian M, Madden PA, Lynskey MT, Colditz GA, Lessov-Schlaggar CN, Schootman M, and Heath AC (2016). Geographic variation in maternal smoking during pregnancy in the Missouri Adolescent Female Twin Study (MOAFTS). *PloS One*, 11(4): e0153930. [PubMed: 27100091]
- Lien DS and Evans WN (2005). Estimating the Impact of Large Cigarette Tax Hikes The Case of Maternal Smoking and Infant Birth Weight. *Journal of Human Resources*, 40(2): 373–392.
- Link BG and Phelan J (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*(Extra Issue): 80–94. [PubMed: 7560851]
- Lumley J, Chamberlain C, Dowswell T, Oliver S, Oakley L, and Watson L (2009). Interventions for promoting smoking cessation during pregnancy. *Cochrane Database Syst Rev*, 3(3).
- Marcus BH, Albrecht AE, Niaura RS, Abrams DB, and Thompson PD (1991). Usefulness of physical exercise for maintaining smoking cessation in women. *The American Journal of Cardiology*, 68(4): 406–407. [PubMed: 1858687]
- Marmot M and Wilkinson R (2009). *Social determinants of health*: Oxford University Press.
- Martin LT, McNamara M, Milot A, Bloch M, Hair EC, and Halle T (2008). Correlates of smoking before, during, and after pregnancy. *American Journal of Health Behavior*, 32(3): 272–282. [PubMed: 18067467]
- Massey DS (1996). The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demography*, 33(4): 395–412. [PubMed: 8939412]
- Massey DS and Eggers ML (1993). The spatial concentration of affluence and poverty during the 1970s. *Urban Affairs Quarterly*, 29(2): 299–315.

- Massey DS, Fischer MJ, Dickens WT, and Levy F (2003). The geography of inequality in the United States, 1950–2000 [with comments]. *Brookings-Wharton Papers on Urban Affairs*: 1–40.
- Matthews SA and Yang TC (2010). Exploring the role of the built and social neighborhood environment in moderating stress and health. *Annals of Behavioral Medicine*, 39(2): 170–183. [PubMed: 20300905]
- Mayer SE (2002). How economic segregation affects children's educational attainment. *Social Forces*, 81(1): 153–176.
- McNeill LH, Kreuter MW, and Subramanian S (2006). Social environment and physical activity: a review of concepts and evidence. *Social Science & Medicine*, 63(4): 1011–1022. [PubMed: 16650513]
- Melvin CL, Dolan-Mullen P, Windsor RA, Whiteside HP, and Goldenberg RL (2000). Recommended cessation counselling for pregnant women who smoke: a review of the evidence. *Tobacco Control*, 9(suppl 3): iii80–iii84. [PubMed: 10982917]
- Morenoff JD (2003). Neighborhood Mechanisms and the Spatial Dynamics of Birth Weight. *American Journal of Sociology*, 108(5): 976–1017.
- Nkansah-Amankra S (2010). Neighborhood contextual factors, maternal smoking, and birth outcomes: multilevel analysis of the South Carolina PRAMS survey, 2000–2003. *Journal of Women's Health*, 19(8): 1543–1552.
- Office of the Surgeon General. (2014). *The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General*. Rockville, MD, U.S. Department of Health and Human Services: 459–521.
- Owen G, Harris R, and Jones K (2016). Under examination: Multilevel models, geography and health research. *Progress in Human Geography*, 40(3): 394–412.
- Pickett KE, Wakschlag LS, Rathouz PJ, Leventhal BL, and Abrams B (2002). The working-class context of pregnancy smoking. *Health & Place*, 8(3): 167–175. [PubMed: 12135640]
- Pickett KE, Wilkinson RG, and Wakschlag LS (2009). The psychosocial context of pregnancy smoking and quitting in the Millennium Cohort Study. *Journal of Epidemiology & Community Health*, 63(6): 474–480. [PubMed: 19258350]
- Poland B, Frohlich K, Haines RJ, Mykhalovskiy E, Rock M, and Sparks R (2006). The social context of smoking: the next frontier in tobacco control? *Tobacco Control*, 15: 59–63. [PubMed: 16436407]
- Raftery AE and Lewis SM (1996). Implementing MCMC In Markov Chain Monte Carlo in Practice, Gilks WR Spiegelhalter DJ and Richardson S (Eds). London: Chapman and Hall, 115–130.
- Räisänen S, Kramer MR, Gissler M, Saari J, Hakulinen-Viitanen T, and Heinonen S (2014). Smoking during pregnancy was up to 70% more common in the most deprived municipalities—A multilevel analysis of all singleton births during 2005–2010 in Finland. *Preventive Medicine*, 67: 6–11. [PubMed: 24983887]
- Rasbash J and Browne WJ (2008). Non-Hierarchical Multilevel Models In *Handbook of Multilevel Analysis*, Deleeuw J and Meijer E (Eds). Verlag New York: Springer, 301–334.
- Raudenbush SW and Bryk AS (2001). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.
- Reardon SF and Bischoff K (2011). Income inequality and income segregation. *American Journal of Sociology*, 116(4): 1092–1153.
- Reichman NE, Hamilton ER, Hummer RA, and Padilla YC (2008). Racial and ethnic disparities in low birthweight among urban unmarried mothers. *Maternal and Child Health Journal*, 12(2): 204–215. [PubMed: 17570044]
- Robert SA (1998). Community-level socioeconomic status effects on adult health. *Journal of Health and Social Behavior*: 18–37. [PubMed: 9575702]
- Rogers JM (2009). Tobacco and pregnancy. *Reproductive Toxicology*, 28(2): 152–160. [PubMed: 19450949]
- Rosenzweig MR and Schultz TP (1983). Estimating a household production function: heterogeneity, the demand for health inputs, and their effects on birth weight. *The Journal of Political Economy*: 723–746.

- Sampson RJ (2003). Neighborhood-level context and health: lessons from sociology In *Neighborhoods and Health*, Kawachi I and Berkman LF (Eds). New York: Oxford University Press, 132–146.
- Sampson RJ, Morenoff JD, and Earls F (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review*, 64(5): 633–660.
- Sampson RJ, Morenoff JD, and Gannon-Rowley T (2002). Assessing “neighborhood effects”: Social processes and new directions in research. *Annual Review of Sociology*, 28: 443–478.
- Savitz NV and Raudenbush SW (2009). Exploiting spatial dependence to improve measurement of neighborhood social processes. *Sociological Methodology*, 39(1): 151–183.
- Sellström E, Arnoldsson G, Bremberg S, and Hjern A (2008). The neighbourhood they live in—Does it matter to women’s smoking habits during pregnancy? *Health & Place*, 14(2): 155–166. [PubMed: 17616477]
- Shaw CR and McKay HD (1942). *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Shaw RJ, Pickett KE, and Wilkinson RG (2010). Ethnic density effects on birth outcomes and maternal smoking during pregnancy in the US linked birth and infant death data set. *American Journal of Public Health*, 100(4): 707–713. [PubMed: 20167891]
- Shoff C and Yang TC (2013). Understanding maternal smoking during pregnancy: Does residential context matter? *Social Science & Medicine*, 78(February): 50–60. [PubMed: 23246395]
- Tong VT, Dietz PM, Farr SL, D’angelo DV, and England LJ (2013). Estimates of smoking before and during pregnancy, and smoking cessation during pregnancy: comparing two population-based data sources. *Public Health Reports*, 128(3): 179–188. [PubMed: 23633733]
- U.S. Census Bureau. (2017). *Historical Income Tables: Income Inequality*. Washington D.C., U.S. Census Bureau.
- Wehby GL, Fletcher JM, Lehrer SF, Moreno LM, Murray JC, Wilcox A, and Lie RT (2011). A Genetic Instrumental Variables Analysis of the Effects of Prenatal Smoking on Birth Weight: Evidence from Two Samples. *Biodemography and Social Biology*, 57(1): 3–32. [PubMed: 21845925]
- Wen M, Browning CR, and Cagney KA (2003). Poverty, affluence, and income inequality: neighborhood economic structure and its implications for health. *Social Science & Medicine*, 57(5): 843–860. [PubMed: 12850110]
- WHO. (2013). *WHO recommendations on prevention and management of tobacco use and second-hand smoke exposure in pregnancy*. Geneva, World Health Organization.
- Wilson WJ (1987). *The Truly Disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press.
- Woodby LL, Windsor RA, Snyder SW, Kohler CL, and Diclemente CC (1999). Predictors of smoking cessation during pregnancy. *Addiction*, 94(2): 283–292. [PubMed: 10396795]
- Yang TC, Shoff C, Noah AJ, Black N, and Sparks CS (2014). Racial Segregation and Maternal Smoking during Pregnancy: A multilevel analysis using the racial segregation interaction index. *Social Science & Medicine*, 107(April): 26–36. [PubMed: 24602968]

Figure 1.
Alternative locations of four neighborhoods under the spatial exchangeability assumption

Figure 2.

Diagram of the nesting structure within a multiple membership model

Figure 3.
The multiple membership relations of a sample neighborhood using a 1 kilometer buffer

Figure 4.

Moran's I in all census tracts in the state of New Jersey, 1999-2007. The unit in the above plot is the number of standard deviations from the mean.

Figure 5.

Local indicator of spatial association cluster map for neighborhood affluence in all census tracts in the state of New Jersey, 1999-2007. Black indicates tracts with high affluence surrounded by other high affluence tracts (high-high spatial clustering; $n = 29$); dark grey indicates tracts with low affluence surrounded by other low affluence tracts (low-low spatial clustering; $n = 178$). The adjacency matrix is defined using a buffer of 1 kilometer.

Figure 6.
Markov chain Monte Carlo diagnostic plots for neighborhood affluence, generated in MLwiN based on a chain length of 50000 (effective sample size = 1067)

Figure 7.
Markov chain Monte Carlo diagnostic plots for neighborhood disadvantage, generated in MLwiN based on a chain length of 50000 (effective sample size = 920)

Table 1.

Descriptive statistics

	Mean	Standard deviation	Minimum	Maximum
Neighborhood affluence (lagged)	−0.1	1.0	−3.3	5.4
Neighborhood disadvantage (lagged)	−0.2	1.0	−2.4	7.2
Mother's years of education	13.5	2.8	0	17
Birth year	2003	2.6	1999	2007
Participation in Medicaid during pregnancy	22.4%		0	1
Foreign-born mother	34.3%		0	1
Maternal race/ethnicity				
Non-Hispanic white (reference)	52.2%		0	1
Non-Hispanic black	15.7%		0	1
Hispanic	22.9%		0	1
Asian	8.1%		0	1
Mother married	69.1%		0	1
Mother's age (years)	29.8	6.1	11.6	59.8
Mother employed in year prior to birth	63.0%		0	1

Table 2.

Results from multivariate regression models predicting antenatal smoking from neighborhood affluence and maternal covariates

	Single level		Multilevel		Spatial multiple membership	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Fixed Part						
Neighborhood affluence (lagged)	−0.339	0.006	−0.305	0.011	−0.278	0.014
Neighborhood disadvantage (lagged)	0.068	0.005	0.018	0.012	0.047	0.013
Mother's education	−0.201	0.002	−0.199	0.002	−0.198	0.002
Birth year	−0.031	0.002	−0.032	0.002	−0.035	0.002
Medicaid	0.388	0.011	0.440	0.011	0.441	0.011
Foreign-born mother	−1.781	0.017	−1.648	0.016	−1.643	0.017
Non-Hispanic black	−1.078	0.013	−0.944	0.015	−0.935	0.015
Hispanic	−1.42	0.015	−1.254	0.017	−1.241	0.017
Asian	−0.616	0.038	−0.656	0.041	−0.656	0.040
Married	−1.249	0.011	−1.192	0.012	−1.191	0.012
Mother's age	0.033	0.001	0.034	0.001	0.034	0.001
Employed	−0.147	0.009	−0.155	0.009	−0.157	0.009
Constant	63.232	3.245	63.72	3.767	70.003	3.777
Random Part						
Level: neighborhood (spatially unstructured)						
Between-neighborhood affluence			0.110	0.008	0.030	0.007
Between neighborhood disadvantage			0.087	0.007	0.002	0.001
Level: neighbors (spatially structured)						
Between neighborhood affluence					0.267	0.031
Between neighborhood disadvantage					0.233	0.022
Deviance information criterion:	422182.7		416248.2		416033.9	
Deviance information criterion change			5934.5		214.4	
Chain length:	5000		10000		50000	

Source: Electronic birth record data from the state of New Jersey, 1999-2007