American Journal of Epidemiology

Vol. 171, No. 6
DOI: 10.1093/aje/kwp435
Advance Access publication:
February 5, 2010

Practice of Epidemiology

Effects of Socioeconomic and Racial Residential Segregation on Preterm Birth: A Cautionary Tale of Structural Confounding

Lynne C. Messer*, J. Michael Oakes, and Susan Mason

* Correspondence to Dr. Lynne C. Messer, Center for Health Policy, Duke Global Health Institute, Duke University, CB 90392, 2812 Erwin Road, Suite 403, Durham, NC 27705 (e-mail: lynne.messer@duke.edu).

Initially submitted February 2, 2009; accepted for publication October 1, 2009.

Confounding associated with social stratification or other selection processes has been called structural confounding. In the presence of structural confounding, certain covariate strata will contain only subjects who could never be exposed, a violation of the positivity or experimental treatment effect assumption. Thus, structural confounding can prohibit the exchangeability necessary for meaningful causal contrasts across levels of exposure. The authors explored the presence and magnitude of structural confounding by estimating the independent effects of neighborhood deprivation and neighborhood racial composition (segregation) on rates of preterm birth in Wake and Durham counties, North Carolina (1999–2001). Tabular analyses and random-intercept fixed-slope multilevel logistic models portrayed different structural realities in these counties. The multilevel modeling results suggested some nonsignificant effect of residence in tracts with high levels of socioeconomic deprivation or racial residential segregation on adjusted odds of preterm birth for white and black women living in these counties, and the confidence limit ratios indicated fairly consistent levels of precision around the estimates. The results of the tabular analysis, however, suggested that many of these regression modeling findings were off-support and based on no actual data. The implications for statistical and public health inference, in the presence of no data, are considered.

confounding factors (epidemiology); multilevel analysis; premature birth; residence characteristics; social class; social environment

Abbreviations: CI, confidence interval; NDI, neighborhood deprivation index.

Editor's note*: An invited commentary on this article and a related article appears on page* ***674****, and these authors' response appears on page* ***680****.*

Confounding associated with social stratification or other selection processes has been called structural confounding (1, 2). The term ''structural'' implies that the confounding results from the social sorting mechanism at work and thus cannot be overcome by the collection of more data. In neighborhood effects research, for example, investigators typically adjust for individual-level covariates (e.g., income, education, race) in order to make subjects comparable in all respects except for the neighborhood exposure of interest, but the more completely a given set of individual-level covariates predicts residence in one neighborhood (i.e., the better one controls for selection into neighborhoods), the more difficult it is to identify persons in any *other* neighborhood with the same set of characteristics (3). In other words, the covariate distributions of the exposed and unexposed populations do not overlap (1). This means that certain covariate strata will contain only subjects who could never be exposed, a violation of the positivity (4) or experimental treatment effect (5) assumption. Thus, structural confounding prohibits the exchangeability necessary for meaningful causal contrasts across exposures (2, 6, 7).

Statistical inferences based on observed data may be called ''on-support,'' whereas inferences resulting from cells for which no data exist, or model-dependent inferences, have been called ''off-support'' (8). This paper addresses structural confounding and how far off-support a typical modeling exercise may be. Given both the importance of and recent attention to neighborhood effects, we do so by estimating the independent effects of neighborhood deprivation and neighborhood racial composition (segregation) on preterm birth.

MATERIALS AND METHODS

Data sources

Three consecutive years of birth records (1999–2001) provided information on birth outcomes and individual-level characteristics for women delivering live singletons in Wake and Durham counties, North Carolina. Maternal addresses were geocoded with latitude and longitude values by Geographic Data Technology, Inc. (Lebanon, New Hampshire) and assigned to year 2000 US Census tracts. Of the 98.6% of birth files with complete addresses sent to Geographic Data Technology for geocoding, 93.2% achieved an exact match using the company's methods.

Tract-level data on deprivation and racial composition were obtained from the 2000 US Census (http://www.census.gov/). Although respondents could report more than 1 race on the 2000 Census, only about 3% did (9), and for simplicity, only those reporting 1 race were used for this analysis. Census data were merged with birth-record data by means of census-tract Federal Information Processing Standards codes.

Neighborhood definition

This research was designed to mirror the prevailing approach to neighborhood effects research. Therefore, we used census tracts to represent the neighborhood environment, both because they are commonly employed in neighborhood research (10–13) and because they are considered reasonable approximations of the immediate physical environment (14).

Outcome definition

Preterm birth was defined as a singleton livebirth occurring at a clinically estimated gestational age of $\geq$20 weeks and $<$37 weeks with a birth weight less than 3,888 g (15). Because of the limited numbers of Hispanic births during the study period (approximately 10% in each county) and for comparability with previous research, only births to non-Hispanic black or non-Hispanic white women were used in this analysis.

Exposure definitions

Racial residential segregation and neighborhood deprivation were the 2 exposures of interest. For simplicity and comparison with prior work, we represented racial residential segregation using the percentage of a census tract's population reporting their race as black in the 2000 Census ("tract-level percent black"). We estimated area-level deprivation using the neighborhood deprivation index (NDI). The NDI, described in detail elsewhere (16), is calculated using 8 census variables from 5 sociodemographic domains (poverty, education, employment, housing, and occupation), empirically summarized using principal components analysis. To avoid imposing any particular functional form on the relation between the exposures (segregation, neighborhood deprivation) and the outcome (preterm birth), and in an attempt to ensure adequate numbers in each exposure stratum, we used county-specific quartiles of census tract-level percent black and NDI for these analyses. Therefore, while the specific numeric value of any given quartile will differ between Wake and Durham counties, the rank ordering of the quartiles is consistent and correspondent.

Data analyses

Pearson correlations and county-specific race-stratified tabulations were our main analytic tools, but we also fitted a multilevel random-effects logistic regression model, with a fixed slope and random tract-level intercepts. While this is a 2-stage model, the results from the combined model are presented here.

$$\ln\left(\frac{Py_{ij}}{1 - Py_{ij}}\right) = \gamma_{00} + \gamma_{01}Z_j + \gamma_{02}Z_j + \beta_1 X_{ij} + \mu_{0j}.$$

The race-stratified model predicted a dichotomous preterm birth outcome for woman i residing in tract j, as a function of γ_{00} (the grand mean, derived from the distribution of random intercepts), $\gamma_{01}Z_j$ (representing quartiles of tract-level neighborhood deprivation), and $\gamma_{02}Z_j$ (representing quartiles of tract-level percent black), adjusted for 1 individual-level covariate $\beta_1 X_{ij}$ (tertiles of maternal education) and μ_{0j} (the tract-specific random deviations). Linear combinations of the estimated coefficients were also computed. All analyses were performed using Stata 9.2 (Stata Corporation, College Station, Texas).

RESULTS

Approximately 5% of birth records were excluded because of missing values and improper codes, resulting in 31,715 singleton livebirths to non-Hispanic women residing in Wake and Durham counties between 1999 and 2001. Black mothers in the study area were younger, less educated, and less likely to be married than were white women (Table 1). Both black women and white women residing in Durham County were younger, less educated, and less likely to be married than their counterparts in Wake County. Rates of preterm birth were related to maternal characteristics, as expected, with black women having 1.5–2 times' the rate of preterm birth as white women; preterm birth rates were generally higher among women residing in Durham County. As reported elsewhere (13), neighborhood-level percentage of black residents was positively associated with odds of preterm birth among both black and white women. Previous work in this population has also demonstrated that white women tend to live in the least deprived neighborhoods while black women tend to live in the most deprived neighborhoods in both Wake and Durham counties (10).

The mean percent black for Durham was 43, as compared with 22% for Wake County (Table 2). One-quarter of

Table 1. Distribution of Maternal Characteristics by County and Race, Durham and Wake Counties, North Carolina, 1999–2001

	Durham County						Wake County					
	White Women (*n* = 5,365)			Black Women (*n* = 4,275)			White Women (*n* = 16,517)			Black Women (*n* = 5,558)		
	No.	%	% With PTB	No.	%	% With PTB	No.	%	% With PTB	No.	%	% With PTB
Individual-Level Characteristics												
Maternal age, years												
<20	407	7.6	14.0	690	16.1	16.7	458	2.8	8.7	726	13.1	10.9
20–24	1,002	18.7	6.0	1,308	30.6	14.8	1,700	10.3	7.2	1,583	28.5	11.7
25–29	1,535	28.6	8.0	1,089	25.5	14.9	4,505	27.3	7.4	1,443	26.0	11.4
30–34	1,582	29.5	6.9	782	18.3	12.5	6,288	38.1	6.0	1,125	20.2	13.7
≥35	839	15.6	8.1	406	9.5	17.2	3,566	21.6	6.7	681	12.3	15.6
Education, years												
>12	3,170	59.5	6.5	1,972	46.3	12.1	13,647	82.8	6.3	2,835	51.2	10.4
12	818	15.4	9.9	1,232	28.9	15.4	2,159	13.1	8.6	1,738	31.4	13.5
<12	1,337	25.1	9.3	1,054	24.8	19.5	685	4.2	8.3	970	17.5	15.9
Marital status												
Married	5,119	96.2	8.0	3,886	91.3	16.0	15,072	94.3	6.8	5,046	90.9	12.9
Single	205	3.9	18.5	369	8.7	35.2	908	5.7	13.8	508	9.2	28.0
Area-Level Characteristics												
Quartile of percent black[a]												
%BL1 (low)	1,636	30.5	7.0	304	7.1	12.8	4,820	29.2	6.1	259	4.7	7.7
%BL2	1,866	34.8	7.6	841	19.7	11.3	5,311	32.2	6.8	894	16.1	11.6
%BL3	1,479	27.6	8.2	1,652	38.6	15.0	4,682	28.4	6.8	1,540	27.7	11.2
%BL4 (high)	383	7.1	10.4	1,478	34.6	17.4	1,704	10.3	8.2	2,865	51.6	13.7
Quartile of NDI[b]												
NDI1 (low)	2,328	43.4	7.1	575	13.5	11.8	5,711	34.6	6.1	402	7.2	4.4
NDI2	1,315	24.5	7.5	997	23.3	12.2	6,207	37.6	6.6	1,201	21.6	12.2
NDI3	963	18.0	9.0	1,187	27.8	16.0	3,150	19.1	7.1	1,406	25.3	11.6
NDI4 (high)	758	14.1	8.6	1,516	35.5	17.0	1,449	8.8	8.9	2,549	45.9	13.7

Abbreviations: %BL, percent black; NDI, neighborhood deprivation index; PTB, preterm birth.
[a] Percentage of black residents in a census tract. The first quartile represents the lowest percentage.
[b] NDI in a census tract. The first quartile represents the smallest deprivation value.

Durham County tracts were at least 69% black; the highest quartile for percent black in Wake County began at 28% black. For each county, the NDI was standardized to a mean of 0, but its range was much smaller in Durham County (−1.5 to 2.7) than in Wake County (−1.3 to 5.0), indicating that the deprived Wake County neighborhoods were especially disadvantaged. The correlation between socioeconomic deprivation and racial residential segregation was high ($r = 0.75$).

Of the 53 Durham County and 105 Wake County census tracts (Table 3), either 1 tract or no tracts fell into the highest deprivation and lowest percent black quartiles (upper right cells of the table) or into the lowest deprivation and highest percent black quartiles (lower left cells of the table). In short, there were almost no poor white tracts or wealthy black tracts in Durham or Wake county. In further tabular representations, these tracts were eliminated from consideration; their ‘‘elimination’’ is represented by dashes. The distribution of tracts across the remaining deprivation and segregation cells varied, with the bulk of the tract-level evidence occurring along the positive diagonal, from low deprivation–low percent black (i.e., the upper left cells representing wealthy white tracts) to high deprivation–high percent black (the lower right cells representing poor black tracts).

Table 4 shows the mean numbers of women per cell in the remaining percent black–deprivation categories. We also provide (in parentheses) the number of census tracts representing each combination to show the number of contexts across which the women were distributed in each county. In general, most white women lived in low-deprivation and relatively low-percent-black census tracts (NDI1–%BL1 and NDI1–%BL2 cells), while most black women lived in higher-deprivation and higher-percent-black tracts (NDI3–%BL4 and NDI4–%BL3 cells) in both Durham and Wake counties.

While one important strength of regression modeling in general, and multilevel modeling in this case, is the ability to borrow strength from adjacent cells to help inform and smooth over areas of sparse data, some authors suggest that

Table 2. Census Tract-Level Percentage of Non-Hispanic Black Residents and Neighborhood Deprivation Index Score, Durham and Wake Counties, North Carolina, 1999–2001

Characteristic	Durham County (n = 53)		Wake County (n = 105)	
	No. of Census Tracts	% or NDI Score	No. of Census Tracts	% or NDI Score
Percentage of black residents				
Mean %BL (continuous variable)		43.1 (18.8–65.0)[a]		21.9 (6.8–28.2)
Quartile				
%BL1 (low)	14	4.5–18.8	27	0.7–6.8
%BL2	13	19.2–37.0	27	7.1–15.2
%BL3	13	40.0–65.0	27	15.3–28.2
%BL4 (high)	13	68.6–97.8	27	28.2–92.7
NDI score				
Mean NDI score (continuous variable)		0.0 (−0.8 to 0.6)		0.0 (−0.7 to 0.4)
Quartile				
NDI1 (low)	14	−1.6 to 0.8	27	−1.3 to −0.7
NDI2	13	−0.7 to −0.3	27	−0.7 to −0.2
NDI3	13	−0.1 to 0.6	27	−0.2 to 0.4
NDI4 (high)	13	0.8–2.7	27	0.4–5.0

Abbreviations: %BL, percent black; NDI, neighborhood deprivation index.
[a] Numbers in parentheses, interquartile range.

each context should contain a minimum number of observations per context for estimate stability (17). The cells with italic text in Table 4 represent those with sparse numbers of observations, defined here as fewer than 30 women per census tract context, on average. In further tabular representations, these cells are excluded from consideration (as indicated by dashes).

Table 3. Distribution of Racial Segregation (Number of Census Tracts per Cell[a]) According to Level of Neighborhood Deprivation in Wake and Durham Counties, North Carolina, 1999–2001[b]

County and Quartile of Percent Black	Quartile of NDI			
	NDI1 (Low)	NDI2	NDI3	NDI4 (High)
Durham County (n = 53 tracts)				
%BL1 (low)	10	2	*1*	*1*
%BL2	4	6	3	*0*
%BL3	*0*	5	4	4
%BL4 (high)	*0*	*0*	5	8
Wake County (n = 105 tracts)				
%BL1 (low)	23	4	*0*	*0*
%BL2	3	12	10	*1*
%BL3	*1*	8	12	5
%BL4 (high)	*0*	2	4	20

Abbreviations: %BL, percent black; NDI, neighborhood deprivation index.
[a] Cells are defined as the intersection between quartile of NDI and quartile of percent black.
[b] Cells with italicized numbers represent those with too few contexts (≤1 tract per cell) for meaningful comparisons.

Standard multilevel modeling practice assumes that individual-level characteristics, such as low levels of education, both influence a person's selection into poor and/or black neighborhoods and also cause poor health outcomes. On the basis of this assumption, individual-level covariates are considered confounders of the relation of primary interest—in this case, the relation between area-level segregation and/or deprivation and preterm birth—and are controlled for in the statistical model. Consistent with this practice, we adjusted for education here. Table 5 presents the numbers of preterm births (our health outcome) to women in each category of our 3-level education covariate (more than high school, high school, and less than high school), for each combination of segregation and deprivation in Wake and Durham census tracts. The number of women within each educational level across which the preterm birth observations are distributed is provided in parentheses. The number of census tracts across which these women are distributed, per education stratum, is available from prior tables.

Generally, in Durham County, most of the preterm births to white women occurred in the upper left section of the diagonal, representing the less-deprived and less-black census tracts, because these were the neighborhoods in which most white women resided. Similarly, most of the preterm births to black women were located in the bottom right quadrant of the diagonal, corresponding to maternal residence in the more-deprived and more-black census tracts. A similar pattern is noted for Wake County, but Wake County is noteworthy for the more stark distribution of

Table 4. Mean Census Tract-Level Numbers and Ranges of Women in Each Combination of Racial Segregation and Economic Deprivation, by County and Race, Durham and Wake County, North Carolina, 1999–2001[a]

County and Quartile of Percent Black	White Women								Black Women							
	NDI1 (Low)		NDI2		NDI3		NDI4 (High)		NDI1 (Low)		NDI2		NDI3		NDI4 (High)	
	Mean No. of Women (Tracts[b])	Range	Mean No. of Women (Tracts)	Range	Mean No. of Women (Tracts)	Range	Mean No. of Women (Tracts)	Range	Mean No. of Women (Tracts)	Range	Mean No. of Women (Tracts)	Range	Mean No. of Women (Tracts)	Range	Mean No. of Women (Tracts)	Range
Durham County																
%BL1 (low)	188 (10)	89–250	45 (2)	3–48	—	—	—	—	53 (10)	7–87	*3 (2)*	*2–3*	—	—	—	—
%BL2	290 (4)	87–428	105 (6)	64–154	173 (3)	140–214	—	—	100 (4)	41–149	60 (6)	32–96	88 (3)	50–101	—	—
%BL3	—	—	160 (5)	76–221	95 (4)	56–131	157 (4)	15–188	—	—	164 (5)	68–222	138 (4)	37–202	177 (4)	32–225
%BL4 (high)	—	—	—	—	30 (5)	19–40	61 (8)	6–79	—	—	—	—	158 (5)	74–206	143 (8)	70–217
Wake County																
%BL1 (low)	275 (23)	60–594	146 (4)	44–175	—	—	—	—	*22 (23)*	*1–55*	*12 (4)*	*7–18*	—	—	—	—
%BL2	732 (3)	103–883	418 (12)	49–722	141 (10)	22–275	—	—	46 (3)	20–63	60 (12)	14–96	33 (10)	4–62	—	—
%BL3	—	—	468 (8)	54–790	220 (12)	26–386	126 (5)	62–178	—	—	66 (8)	13–104	86 (12)	1–132	73 (5)	47–94
%BL4 (high)	—	—	30 (2)	12–18	126 (4)	63–151	71 (20)	5–114	—	—	124 (2)	101–141	121 (4)	86–148	138 (20)	4–180

Abbreviations: %BL, percent black; NDI, neighborhood deprivation index.
[a] Cells with italicized text represent those with sparse observations (<30 women per census tract on average); cells with dashes are those eliminated from presentation based on prior exclusions.
[b] Number of census tracts representing that particular combination of NDI and percent black.

preterm births by level of education, which suggests that few low-educated white women lived in the least-deprived and least-black tracts in Wake County.

The literature suggests that 5 outcome observations per cell—in our case, at least 5 preterm births—represents a sufficient quantity of data for stable statistical inference (18). The italic numbers represent the tracts we would exclude as having too few outcome observations per level of education to produce stable estimates. Note that too few preterm births in *any* education stratum within a cell resulted in exclusion of the entire cell, because these cells violated the positivity assumption.

Despite the dearth of data revealed in the tabular analyses, the race- and county-stratified models (Table 6) provided results for the independent effects of tract-level percent black and tract-level deprivation and suggested that there was little effect of tract-level deprivation on odds of preterm birth for either white or black women residing in Durham County, after accounting for tract percent black and maternal education. The one exception to the otherwise modest findings was for white women in the fourth quartile of percent black, where the odds ratio was 1.6 (95% CI: 0.9, 3.0).

Among Wake County white women, however, some suggestion of increased odds of preterm birth appeared in association with the fourth quartile of tract-level deprivation (odds ratio = 1.4, 95% CI: 1.0, 1.9). Increased odds of similar magnitude were observed for black women in all but the first quartile of tract deprivation, but these estimates did not exclude null findings. The results for the association of quartiles of tract percent black with preterm birth, adjusting for tract-level deprivation and maternal education, were largely null for women of both races residing in Wake County, with estimates ranging from 1.0 (95% CI: 0.8, 1.3) to 1.6 (95% CI: 0.9, 2.7).

Given the empty cells highlighted by the tabular exploration, our obtaining seemingly reasonable results from the multilevel models testifies to the models' potent capacity to smooth over areas of data sparseness. We next sought to explore whether obtaining estimates for the specific cell combinations of racial segregation and neighborhood deprivation, particularly those combinations for which we knew we had few or no actual observations, was possible and, if it was possible to obtain the estimates, whether the data scarcity was represented in the 95% confidence intervals. Table 7 presents the linear combinations estimable from the prior multilevel model for preterm birth, including both quartiles of NDI and quartiles of percent black, adjusted for maternal education category. Italic numbers indicate the cells (geographic spaces) in which few or no women resided or in which there was no overlap in the covariate (education) distribution of exposed and unexposed women. If we compare the measures of precision using the confidence limit ratio by dividing the upper 95% confidence limit by the lower 95% confidence limit, we find little difference in precision between estimates based on sparse data and those based on abundant observations. For instance, comparing the Wake County confidence limit ratios of 1.6 and 3.6 for the NDI1–%BL4 cell for white and black women, respectively, a cell that we know to contain zero census tracts, with the confidence limit ratios of 1.6 and 2.8 for NDI4–%BL4 in Wake

Table 5. Numbers of Preterm Births and Numbers of Women in Each Combination[a] of Racial Segregation and Economic Deprivation, by County, Race, and Level of Maternal Education, Durham and Wake Counties, North Carolina, 1999–2001[b]

County, Percent Black, and Maternal Educational Level	White Women								Black Women							
	NDI1		NDI2		NDI3		NDI4		NDI1		NDI2		NDI3		NDI4	
	No. of PTBs	No. of Women	No. of PTBs	No. of Women	No. of PTBs	No. of Women	No. of PTBs	No. of Women	No. of PTBs	No. of Women	No. of PTBs	No. of Women	No. of PTBs	No. of Women	No. of PTBs	No. of Women
Durham County																
%BL1 (low)																
>HS	75	1,287	*3*	*36*	—	—	—	—	*23*	*217*	—	—	—	—	—	—
=HS	19	164	*0*	*12*	—	—	—	—	*10*	*47*	—	—	—	—	—	—
<HS	10	85	*0*	*3*	—	—	—	—	*2*	*18*	—	—	—	—	—	—
%BL2																
>HS	*52*	*699*	20	373	11	165	—	—	*16*	*196*	18	161	*18*	147	—	—
=HS	*6*	*49*	5	84	12	89	—	—	*13*	*65*	8	84	*3*	68	—	—
<HS	*2*	*34*	12	121	20	241	—	—	*4*	*28*	12	59	*3*	30	—	—
%BL3																
>HS	—	—	26	396	7	91	*4*	*72*	—	—	55	486	21	152	25	206
=HS	—	—	19	149	8	74	*5*	*113*	—	—	18	146	29	142	34	188
<HS	—	—	13	137	16	171	*22*	*262*	—	—	11	52	21	99	30	172
%BL4 (high)																
>HS	—	—	—	—	*3*	*16*	*3*	*16*	—	—	—	—	34	222	26	176
=HS	—	—	—	—	*0*	*20*	*6*	*53*	—	—	—	—	28	163	47	324
<HS	—	—	—	—	*4*	*45*	*24*	*226*	—	—	—	—	27	144	94	445
Wake County																
%BL1 (low)																
>HS	*233*	*4,003*	*24*	*431*	—	—	—	—	—	—	—	—	—	—	—	—
=HS	*22*	*266*	*10*	*52*	—	—	—	—	—	—	—	—	—	—	—	—
<HS	*1*	*41*	*2*	*16*	—	—	—	—	—	—	—	—	—	—	—	—
%BL2																
>HS	*65*	*984*	165	2,601	42	658	—	—	*6*	*77*	26	281	15	118	—	—
=HS	*10*	*93*	37	458	9	123	—	—	*4*	*24*	28	180	7	59	—	—
<HS	*0*	*18*	10	152	9	63	—	—	*0*	*10*	7	68	7	32	—	—
%BL3																
>HS	—	—	*108*	*1,678*	75	1,352	*31*	*374*	—	—	24	230	36	406	19	173
=HS	—	—	*22*	*262*	26	361	*20*	*120*	—	—	17	105	27	222	16	116
<HS	—	—	*2*	*59*	15	125	*3*	*50*	—	—	11	51	11	100	5	59
%BL4 (high)																
>HS	—	—	*28*	*401*	*37*	*335*	35	457	—	—	16	163	38	310	97	789
=HS	—	—	*2*	*69*	*5*	*100*	19	196	—	—	6	54	16	120	106	787
<HS	—	—	*0*	*21*	*3*	*29*	10	94	—	—	7	24	5	36	98	573

Abbreviations: %BL, percent black; HS, high school; NDI, neighborhood deprivation index; PTB, preterm birth.

[a] Cells are defined as the intersection between quartile of NDI, quartile of percent black, and maternal education category (less than high school, high school, more than high school).

[b] Cells with italicized text represent those with sparse observations (<5 outcomes per education category); cells with dashes are those eliminated from presentation based on prior exclusions.

County (white and black women, respectively), a cell that we know contains abundant observations, leaves one with the impression that the estimates are of comparable precision.

DISCUSSION

The multilevel modeling results suggested a modest but not statistically significant effect of residence in tracts with high levels of socioeconomic deprivation on adjusted odds of preterm birth for both white and black women living in Wake County, but no association for women living in Durham County. Increased odds of preterm birth were also noted for the effect of residence in tracts with the highest quartile of percent black for white women living in Durham County, adjusted for maternal education. The linear combinations of

Table 6. Adjusted[a] Odds Ratios for Preterm Birth According to Census Tract-Level Deprivation and Percent Black From Multilevel Logistic Regression Models, by County and Race, Durham and Wake Counties, North Carolina, 1999–2001

Quartile	Durham County				Wake County			
	White Women		Black Women		White Women		Black Women	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Tract deprivation								
NDI1	1.0	Referent	1.0	Referent	1.0	Referent	1.0	Referent
NDI2	0.8	0.6, 1.2	0.9	0.6, 1.5	1.1	0.8, 1.3	1.6	0.9, 2.7
NDI3	0.9	0.6, 1.4	1.1	0.7, 1.7	1.1	0.8, 1.4	1.4	0.8, 2.5
NDI4	0.7	0.4, 1.2	1.0	0.6, 1.7	1.4	1.0, 1.9	1.5	0.8, 2.6
Tract percent black								
%BL1	1.0	Referent	1.0	Referent	1.0	Referent	1.0	Referent
%BL2	1.1	0.8, 1.6	0.8	0.5, 1.3	1.0	0.8, 1.3	1.2	0.6, 2.2
%BL3	1.2	0.8, 1.9	1.1	0.6, 1.9	1.0	0.8, 1.3	1.1	0.6, 2.1
%BL4	1.6	0.9, 3.0	1.1	0.6, 2.0	1.1	0.8, 1.3	1.3	0.7, 2.5

Abbreviations: %BL, percent black; CI, confidence interval; NDI, neighborhood deprivation index; OR, odds ratio.

[a] Models included quartiles of both tract deprivation and tract percent black; odds ratios were adjusted for maternal education category (less than high school, high school, more than high school).

tract-level deprivation and racial residential segregation indicated that specific arrangements of racial segregation and neighborhood deprivation are associated with increased risk. In particular, residence in tracts with the highest quartiles of neighborhood deprivation, almost irrespective of tract percent black, was associated with increased odds of preterm birth for both black women and white women.

The results of the tabular analysis, however, suggested that many of these regression modeling findings were off-support and based on no actual data. Of Durham County's 53 census tracts, zero existed for those combinations representing low deprivation–high percent black and 1 represented high deprivation–low percent black. The same pattern was observed for Wake County's 105 census tracts, except that there were zero poor white tracts. After elimination of cells for which there were too few women delivering singletons and too few preterm births per stratum of maternal educational level, out of 16 possible combinations of racial and deprivation quartiles in each county, 5 Durham County cells contained minimum numbers of white women and 6 contained sufficient numbers of black women; the Wake County tabular analysis resulted in 4 and 8 cells for white and black women, respectively. In short, we were left with relatively few cells with which we could examine the unique contribution of racial segregation and neighborhood deprivation, adjusted for maternal education, for statistical or public health inference.

Finding so few cells representing racial and economic heterogeneity was unexpected. Previous work using these county-level data have suggested that Durham and Wake counties are neither particularly affluent nor deprived, but rather centrally located on the affluence–deprivation continuum (16). Further, while it is well-established that white and black women live in relatively different socioeconomic environments, very few census tracts in either Wake County or Durham County are hypersegregated. Only 10 of Durham County's 53 census tracts and 5 of Wake's County's 105 census tracts have greater than 80% non-Hispanic black residents; similarly, only 9 of Durham's and 38 of Wake's tracts have more than 80% non-Hispanic white residents. These findings suggested that Wake and Durham counties would be appropriate sociodemographic locations for disentangling the effects of racial segregation from the effects of neighborhood deprivation and would be comparatively free of the structural confounding that is likely in more hypersegregated areas. The tabular exploration, however, highlighted the extent to which structural confounding limits the data for such inference even in these less-segregated areas. In light of these results, we suggest that stratified data analysis should almost certainly precede multilevel modeling and that the limitations posed by areas of thin data should be explicitly noted and discussed.

One possible way to overcome the data sparseness highlighted in this study is to use a different quantile, such as tertiles, for the exposure categorization in stratified analyses. Cutting the data into larger and more heterogeneous units, however, would probably trade structural confounding for residual confounding, since changing the categorization does not change the fact that black and white populations experience separate and nonoverlapping economic environments. Although quartiles may not represent the most substantively meaningful unit of differentiation, they were chosen for this exercise because they represent what has been done in other multilevel modeling exercises. Similarly, census tracts are unlikely to exactly correspond to the salient neighborhood environment for most women, and much has been written about the limitations of administrative units for approximating neighborhoods (19); however, we were

Table 7. Adjusted[a] Odds Ratios and Confidence Limit Ratios for Preterm Birth According to Linear Combinations of Quartiles of Census Tract-Level Deprivation and Census Tract-Level Percent Black From Multilevel Logistic Regression Models, by County and Race, Durham and Wake Counties, North Carolina, 1999–2001[b]

County and Quartile of Percent Black	White Women				Black Women			
	NDI1	NDI2	NDI3	NDI4	NDI1	NDI2	NDI3	NDI4
Durham County								
%BL1								
OR	1.0	*0.8*	*0.9*	*0.7*	*1.0*	*0.9*	*1.1*	*1.0*
95% CI	1.0, 1.0	*0.6, 1.2*	*0.6, 1.4*	*0.4, 1.2*	*1.0, 1.0*	*0.6, 1.5*	*0.7, 1.7*	*0.6, 1.7*
CLR	1.0	*2.0*	*2.3*	*3.0*	*1.0*	*2.5*	*2.4*	*2.8*
%BL2								
OR	*1.1*	0.9	1.0	*0.8*	*0.8*	0.8	*0.9*	*0.9*
95% CI	*0.8, 1.6*	0.7, 1.3	0.7, 1.5	*0.5, 1.3*	*0.5, 1.3*	0.5, 1.2	*0.6, 1.5*	*0.5, 1.4*
CLR	*2.0*	1.9	2.1	*2.6*	*2.6*	2.4	*2.5*	*2.8*
%BL3								
OR	*1.2*	1.0	1.1	*0.9*	*1.1*	1.0	1.2	1.1
95% CI	*0.8, 1.9*	0.7, 1.5	0.8, 1.6	*0.6, 1.3*	*0.6, 1.9*	0.7, 1.5	0.8, 1.8	0.7, 1.7
CLR	*2.4*	2.1	2.0	*2.2*	*3.2*	2.1	2.3	2.4
%BL4								
OR	*1.6*	*1.4*	*1.5*	*1.1*	*1.1*	*1.0*	1.2	1.2
95% CI	*0.9, 3.0*	*0.8, 2.3*	*0.9, 2.5*	*0.7, 1.8*	*0.6, 2.0*	*0.6, 1.7*	0.8, 1.9	0.8, 1.7
CLR	*3.3*	*2.9*	*2.8*	*2.6*	*3.3*	*2.8*	2.4	2.1
Wake County								
%BL1								
OR	*1.0*	*1.1*	*1.1*	*1.4*	*1.0*	*1.6*	*1.4*	*1.5*
95% CI	*1.0, 1.0*	*0.8, 1.3*	*0.8, 1.4*	*1.0, 1.9*	*1.0, 1.0*	*0.9, 2.7*	*0.8, 2.5*	*0.8, 2.6*
CLR	*1.0*	*1.6*	*1.8*	*1.9*	*1.0*	*3.0*	*3.1*	*3.3*
%BL2								
OR	*1.0*	1.1	1.1	*1.4*	*1.2*	1.8	1.7	*1.7*
95% CI	*0.8, 1.3*	0.9, 1.3	0.9, 1.4	*1.1, 1.9*	*0.6, 2.2*	1.0, 3.2	0.9, 3.0	*1.0, 3.2*
CLR	*1.6*	1.4	1.6	*1.7*	*3.7*	3.2	3.3	*3.2*
%BL3								
OR	*1.0*	*1.0*	1.1	*1.3*	*1.1*	1.7	1.6	1.6
95% CI	*0.8, 1.3*	*0.8, 1.3*	0.9, 1.3	1.0, 1.7	*0.6, 2.1*	1.0, 3.1	0.9, 2.7	0.9, 2.9
CLR	*1.6*	*1.6*	1.4	*1.7*	*3.5*	3.1	3.0	3.2
%BL4								
OR	*1.1*	*1.2*	*1.2*	1.5	*1.3*	2.0	1.9	1.9
95% CI	*0.8, 1.3*	*0.9, 1.5*	*0.9, 1.6*	1.2, 1.9	*0.7, 2.5*	1.1, 3.6	1.1, 3.3	1.2, 3.3
CLR	*1.6*	*1.7*	*1.8*	1.6	*3.6*	3.3	3.0	2.8

Abbreviations: %BL, percent black; CI, confidence interval; CLR, confidence limit ratio; NDI, neighborhood deprivation index; OR, odds ratio.

[a] Odds ratios were adjusted for maternal education category (less than high school, high school, more than high school).

[b] Italicized text indicates the geographic spaces in which few or no women resided or in which there was no overlap in the covariate (education) distribution of exposed and unexposed women.

trying to replicate how neighborhood effects are typically modeled in a multilevel context, and we chose the census tract because it is one of the most commonly employed units of administrative geography.

Likewise, much of the multilevel model-based research on neighborhood effects adjusts for individual-level covariates, treating them as confounders of the neighborhood-health relation (11, 20–23). As documented by other researchers (24, 25), individual-level covariates may function instead as intermediates between contextual-level exposures and disease, in which case they should not be included in statistical models (26). We chose to include individual-level covariates in our model because it has become standard practice, not because we advocate a particular causal structure. Our findings regarding data scarcity remain relevant to investigators who choose not to control for individual-level covariates; as Table 4 demonstrates, estimation of the independent effects of racial segregation and neighborhood deprivation is challenged by a lack of poor white and rich black tracts, even without adjustment for individual-level education.

Placing these findings in the context of other research is challenging, because much of the existing neighborhood effects literature focuses on multilevel models showing the estimated effects of racial residential segregation or deprivation following adjustment for individual-level covariates, without clarifying the underlying population structure or causal counterfactuals under study (12, 27). Exceptions exist, however. In work exploring the effect of neighborhood-level concentrated poverty on verbal ability, Sampson et al. (28) were forced to restrict their analysis to black children, because no white children lived in the most disadvantaged neighborhoods. Other recent work (29–32) has explicitly considered the exchangeability assumption and the presence of structural confounding.

If the findings of this analysis are comparable to the underlying social and demographic structure of other areas, structural confounding poses a significant threat to epidemiologists' ability to make sound inferences from multilevel regression model results, particularly as relates to health disparities and public health policy. Regression models are useful tools that use regions for which ample data exist to inform and smooth over data cells for which data are limited or sparse. Unfortunately, the lines drawn through empty space may be completely incorrect (33). The utility and appropriateness of these models, especially when we understand how social stratification processes influence many of the individual- and area-level features of public health interest, are unclear.

''Solutions'' to structural confounding are not immediately apparent. Statistical techniques like bootstrapping can be used to generate pseudo-populations, thereby enabling investigators to estimate what the association between an exposure of interest and a particular health outcome might look like if real humans resided within these social/geographic spaces, but such techniques do not change the reality that some potential locations are not currently populated. Data collection targeted to sparse levels of exposures and covariates might be useful, but simply collecting more data from within the same social structure is unlikely to yield significant benefit. The key design issue for investigators attempting to study racial disparities in particular health outcomes is to identify exchangeable (homogeneous) subjects across heterogeneous environments. Neighborhood selection, and the differential neighborhood sorting patterns that result in structural confounding, is more than a ''nuisance'' or a special case of individual-level decision-making or confounding. Rather, it is part of a social allocation process in the United States, with its particular history of race relations, resulting in acute sorting along racial-economic lines that uniformly disadvantages blacks (34).

Extrapolation beyond existing data is an important statistical function but is a questionable practice for public health policy, since the accuracy of such extrapolation cannot be tested (35). Researchers attempting to explore contextual effects on health should make explicit the underlying data distribution across the range of exposures and covariates from which inference will be made, thereby facilitating increased understanding of where the estimates are grounded and where they are off-support. All models rest on assumptions, but causal research, particularly that used to inform public health policy, will benefit from more explicit identification and assessment of them.

ACKNOWLEDGMENTS

Author affiliations: Health Inequalities Program, Center for Health Policy, Duke Global Health Institute, Duke University, Durham, North Carolina (Lynne C. Messer); Epidemiology and Community Health Division, School of Public Health, University of Minnesota, Minneapolis, Minnesota (J. Michael Oakes); and Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina (Susan Mason).

This work was supported by contracts awarded by the Maternal and Child Health Bureau of the Health Resources and Services Administration.

The authors acknowledge the essential role played by their public health partner in the North Carolina State Center for Health Statistics, Dr. Paul Buescher, and their Maternal and Child Health Bureau colleagues, Drs. Mary Kay Kenney and Michael Kogan. Colleagues who commented on early drafts of the manuscript, including Drs. Rich MacLehose, Jay S. Kaufman, and Danelle Lobdell, are also appreciated.

Conflict of interest: none declared.

REFERENCES

1. Oakes J, Johnson P. Propensity score matching for social epidemiology. In: Oakes J, Kaufman J, eds. *Methods in Social Epidemiology*. San Francisco, CA: Jossey-Bass, Inc; 2006: 370–392.
2. Oakes JM. Commentary: advancing neighbourhood-effects research—selection, inferential support, and structural confounding. *Int J Epidemiol*. 2006;35(3):643–647.
3. Oakes JM. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med*. 2004;58(10):1929–1952.
4. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008; 168(6):656–664.
5. Mortimer KM, Neugebauer R, van der Laan M, et al. An application of model-fitting procedures for marginal structural models. *Am J Epidemiol*. 2005;162(4):382–388.
6. Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol*. 2002;31(2):422–428.
7. Draper D, Hodges J, Mallows C, et al. Exchangeability and data analysis (with discussion). *J R Stat Soc Ser A*. 1993; 156(1):9–38.

8. Manski C. Identification problems in the social sciences. In: Marsden P, ed. *Sociological Methodology.* San Francisco, CA: Jossey-Bass, Inc; 1993:1–56.
9. Bureau of the Census, US Department of Commerce. *Overview of Race and Hispanic Origin 2000.* (Census 2000 brief). Washington, DC: US Census Bureau; 2001. (Publication no. C2KBR/01-1). (http://www.census.gov/prod/2001pubs/cenbr01-1.pdf). (Accessed May 2007).
10. O'Campo P, Burke JG, Culhane J, et al. Neighborhood deprivation and preterm birth among non-Hispanic black and white women in eight geographic areas in the United States. *Am J Epidemiol*. 2008;167(2):155–163.
11. Kaufman JS, Dole N, Savitz DA, et al. Modeling community-level effects on preterm birth. *Ann Epidemiol*. 2003;13(5): 377–384.
12. Messer LC, Vinikoor LC, Laraia BA, et al. Socioeconomic domains and associations with preterm birth. *Soc Sci Med.* 2008;67(8):1247–1257.
13. Mason SM, Messer LC, Laraia BA, et al. Segregation and preterm birth: the effects of neighborhood racial composition in North Carolina. *Health Place*. 2009;15(1):1–9.
14. Krieger N, Chen J, Waterman P, et al. Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: the Public Health Disparities Geocoding Project (US). *J Epidemiol Community Health*. 2003;57(3):186–199.
15. Alexander GR, Himes JH, Kaufman RB, et al. A United States national reference for fetal growth. *Obstet Gynecol*. 1996; 87(2):163–168.
16. Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *J Urban Health*. 2006;83(6):1041–1062.
17. Kreft I, De Leeuw J. *Introducing Multilevel Modeling*. Thousand Oaks, CA: Sage Publications; 1998.
18. Hinkle D, Wiersma W, Jurs S. *Applied Statistics for the Behavioral Sciences*. 4th ed. Boston, MA: Houghton Mifflin Company; 1998.
19. Laraia BA, Messer L, Kaufman JS, et al. Direct observation of neighbourhood attributes in an urban area of the US South: characterizing the social context of pregnancy [electronic article]. *Int J Health Geogr.* 2006;5:11.
20. Ahern J, Pickett KE, Selvin S, et al. Preterm birth among African American and white women: a multilevel analysis of socioeconomic characteristics and cigarette smoking. *J Epidemiol Community Health*. 2003;57(8):606–611.
21. Buka SL, Brennan RT, Rich-Edwards JW, et al. Neighborhood support and the birth weight of urban infants. *Am J Epidemiol.* 2003;157(1):1–8.
22. O'Campo P, Xue X, Wang MC, et al. Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. *Am J Public Health*. 1997;87(7):1113–1118.
23. Pearl M, Braveman P, Abrams B. The relationship of neighborhood socioeconomic characteristics to birthweight among 5 ethnic groups in California. *Am J Public Health*. 2001; 91(11):1808–1814.
24. Geronimus AT. Damned if you do: culture, identity, privilege, and teenage childbearing in the United States. *Soc Sci Med*. 2003;57(5):881–893.
25. Geronimus AT. Teenage childbearing as cultural prism. *Br Med Bull.* 2004;69:155–166.
26. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology.* 1992;3(2): 143–155.
27. Messer LC, Kaufman JS, Dole N, et al. The neighborhood crime and deprivation environment and preterm birth. *Ann Epidemiol*. 2006;16(6):455–462.
28. Sampson RJ, Sharkey P, Raudenbush SW. Durable effects of concentrated disadvantage on verbal ability among African-American children. *Proc Natl Acad Sci U S A*. 2008;105(3): 845–852.
29. Berk R. *Regression Analysis: A Constructive Critique*. Thousand Oaks, CA: Sage Publications; 2004.
30. Hearst MO, Oakes JM, Johnson PJ. The effect of racial residential segregation on black infant mortality. *Am J Epidemiol.* 2008;168(11):1247–1254.
31. Johnson PJ, Oakes JM, Anderton DL. Neighborhood poverty and American Indian infant death: are the effects identifiable? *Ann Epidemiol*. 2008;18(7):552–559.
32. Oakes JM, Forsyth A, Schmitz KH. The effects of neighborhood density and street connectivity on walking behavior: the Twin Cities Walking Study [electronic article]. *Epidemiol Perspect Innov.* 2007;4:16.
33. Vandenbroucke JP. Should we abandon statistical modeling altogether? *Am J Epidemiol*. 1987;126(1):10–13.
34. Sampson RJ. Moving to inequality: neighborhood effects and experiments meet social structure. *Am J Sociol*. 2008;114(1): 189–231.
35. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health*. 2006; 60(7):578–586.