Women and social class: a methodological study comparing individual, household, and census measures as predictors of black/white differences in reproductive history.

Nancy Krieger

Abstract

Study objective—The aim was to compare four different measures of women's social class (interview and census based) as predictors of well known social gradients in reproductive outcomes. The intent was to address two obstacles that confront research in the United States regarding social gradients in women's health: how women's social class should be measured, and the absence of socioeconomic data in most health records.

Study design—The study was a retrospective cohort analysis, using a community based random sample.

Setting—Alameda County, California, 1987.

Subjects—51 black and 50 white women, ages 20 to 80 years, identified by random digit dialling, were interviewed by telephone.

Measurements and main results—Census data were linked to individual records via the respondents' addresses. Using number of full term pregnancies as an example, multiple linear regression analyses demonstrated that individual class was not significantly associated with this outcome, whereas household class was: women from non-working-class households had 0·8 fewer such pregnancies than women from working class households (95% confidence interval [CI] = −1·4, −0·1). The block group measure functioned most like the household class measure (beta = −0·7, 95% CI = −1·4, 0·1), while the census tract measure was non-significant (beta = −0·4, 95% CI = −1·2, 0·4). Similar results were obtained for the outcomes: age at first full term pregnancy, percent of early terminated pregnancies, and yearly income.

Conclusions—These results suggest block group data may offer a uniform source of social class information that can be appended to individual health records, and that this strategy is not invalidated by concerns regarding ecological fallacy.

Investigations of social gradients in disease and health among women in the United States are hampered by two problems: lack of consensus regarding appropriate measures of women's social class[1–15] and the absence of social class data both in US vital statistics and in most disease registries and medical records.[16 17] One solution employed by many US epidemiologists has been to approximate individual level socioeconomic data with census derived social indicators, typically from the census tract level.[1–3 18 19] Despite concerns about using ecological data, however, little research has been conducted within the US to assess how this methodology may be marred by problems associated with ecological fallacy, ie, the erroneous inference of causal relations at the individual level on the basis of grouped data.[20–22]

To address these issues, the present methodological pilot study examined the association between reproductive history and social class indicators of both black and white women at the individual, household, census tract, and census block group levels. Reproductive outcomes were selected because of: (1) their known association with both race and socioeconomic status, and (2) their relevance to women's lives.[23–26] The specific questions the study sought to answer were: (1) which individual level social class measure best predicts a woman's reproductive history: her individual of her household class? (2) Which type of census data best approximates the individual level class measure: data from the census tract, or data from a potentially more homogeneous subdivision of the census tract, known as the census "block group"? (3) Are census data primarily a proxy for, or do they also supplement, individual level data?

Methods

SAMPLE

The target population consisted of black and white women, ages 20 to 80, who lived in Alameda County, California in 1987. Random digit dialling[27] was used by the staff of the Bay Area Resource for Cancer Control to identify and conduct telephone interviews with 51 black and 50 white women meeting these criteria. Respondents living in households containing more than one eligible woman were selected by the Kish procedure.[28] Among eligible respondents, black and white women were equally likely to refuse to be interviewed. The overall effective completion rate was 72°₀.

INSTRUMENT

The 20 minute interview included questions regarding each respondent's age, reproductive history, health status, educational level, marital or partner status, usual work conditions plus occupation and employment status (and that of her partner or other head of household, if relevant), class self identification, family size and age structure, 1986 family income and wealth, housing conditions, and address. Because reproductive patterns may be influenced by social class origin as well as current class position,[26]

University of California at Berkeley, Epidemiology Program, School of Public Health
N Krieger

Correspondence address: Dr N Krieger, Kaiser Foundation Research Institute, Division of Research, 3451 Piedmont Avenue, Oakland, CA 94611–5463, USA

additional questions addressed the usual work conditions and educational level of each respondent's mother and father (or other head of household, if appropriate), plus family structure and household conditions, when the respondent was 13 years old (the age of menarche for many US adolescents).

OPERATIONALISATION OF THE INDIVIDUAL AND HOUSEHOLD CLASS MEASURES

Individual and household class measures were based on four questions regarding "usual type of employment" (see appendix). "Usual" rather than "current" or "last" type of employment was chosen because it often provides a more representative sense of a person's lifetime work history.[29] Answers were assigned to eight mutually exclusive categories, which were then grouped to create three class positions: "working class" (non-professional supervisory and non-supervisory employees); "not working class" (professional supervisory and non-supervisory employees, self employed professionals and non-professionals, those who own a business and employ others); and "other" (not in the paid workforce). Underlying the formation of these categories was a relational approach to social class, which holds that "class" is determined by people's necessary relations to others and to property through the work process, and as such is conceptually distinct from "occupation" and "status".[1-7] For example, an employee presupposes an employer; one cannot exist without the other. In contrast, the category "data analyst" provides no information about class position, because someone in this occupational group could be employed on someone else's grant, self employed, or own a consulting firm.

The schema by which the household class measure was constructed is presented in table I. In this algorithm, widows were treated like women currently living with a partner or other head of household, since widows' usual household position is typically that which they shared with their deceased partner.[30] The same method was used to determine each respondent's past household class, based on the individual class position of her mother, father, or other head of household. Socioeconomic status scores, such as the Duncan Socioeconomic Index, were not employed, since these measures (which are based on US white male employment patterns) predict US women and men to have the same average

occupational status, despite the well known concentration of women in lower wage occupations.[6 12 31-32]

CATEGORISATION OF THE INDIVIDUAL LEVEL POVERTY AND EDUCATIONAL MEASURES

The federally defined poverty line, rather than family income, was used to assess the respondents' economic level, because the poverty line takes into account the number and age of the people supported by a given family income.[1 33-35] In 1986, the poverty line for one adult equalled $5572, and was $11 113 for a family of two adults and two children.[35] Educational level was categorised according to the last year of completed education.

CONSTRUCTION OF CENSUS BASED VARIABLES

Sociodemographic data from the 1980 census were matched to each respondent's record via her address at the time of the interview. These neighbourhood data were obtained at two different census levels: the census tract and the census block group. Census tracts typically contain 4000 persons, and usually consist of four block groups, each with an average of 1000 residents.[36] The 97 respondents (48 white, 49 black) with sufficient information to assign census tract and block group numbers lived in 77 census tracts, subdivided into 287 block groups.

Although the census does not utilise a relational approach to class, it does contain occupational data. A census based class measure therefore could be constructed by identifying which occupations predominantly include people whose relation to others through work, as described above, characterises them as "working class".[5-6 37] These "working class" occupations were selected by: (1) dichotomising the respondents' individual class position ("working class" v "not working class"), and (2) ascertaining which combination of their known individual occupations least misclassified the respondents' individual class position. Predominantly working class tracts and block groups were defined as regions where 66% or more of the employed population belonged to these "working class" occupations.

Additional census based measures characterised federally defined "poverty areas",[38] where 20% or more of the population lived below the poverty line, and "undereducated areas", where at least 25% of persons $\geqslant 25$ years old had not completed high school.

Table I Determination of household level class,[a] based on individual level class[b] of respondent and, if relevant, of her partner, other head of household, or deceased spouse.

| Respondent's individual level class | Individual level class of the respondent's partner, other head of household, or deceased spouse | | | |
	None[c]	Not in paid labour force	Working class	Not working class
Not in paid labour force	Other	Other	Working class	Not working class
Working class	Working class	Working class	Working class	Not working class
Not working class	Not working class	Not working class	Not working class	Not working class

[a] Each cell of the table represents the household level class, as determined by the respondent's position and that of the other relevant person, if any.
[b] Working class = non-professional employee (supervisory/non-supervisory)
Not working class = professional employee (supervisory/non supervisory), self employed, owner who employs others
Other = not in the paid labour force
[c] Respondent lives alone

STATISTICAL METHODS

Comparisons of the black and white respondents used the χ^2 test for categorical variables, and the Student's t test for continuous variables. Multiple linear regression was used to assess the relative predictive value of individual, household, census tract, and census block group class measures. Dependent variables included "number of full term pregnancies", "percent of early terminated pregnancies", "age at first full term pregnancy", and "yearly income". Using the same dependent variables, contextual regression analyses[39 40] were employed to assess the presence, if any, of individual and group level effects for the working class variable.

Table II Selected sociodemographic characteristics of the respondents, present (1987) and past (when respondent was 13 years old), by race.

Characteristic	White			Black			p value[b]
	Mean	Percent	(n)[a]	Mean	Percent	(n)[a]	
Age (years)	38·6		(48)	40·5		(51)	0·482
Educational level			(50)			(51)	
<high school		0·0			9·8		
≥high school, <4 yrs college		58·0			66·7		0·107
≥4 yrs college		42·0			23·5		
Head of household							
Present:			(50)			(50)	
Respondent		48·0			66·0		
Partner		26·0			18·0		0·081
Respondent + partner		26·0			12·0		
Other		0·0			4·0		
Past:			(50)			(50)	
Mother		10·0			22·0		
Father		86·0			70·0		0·154
Other		4·0			8·0		
Number in family household							
Present	2·5		(50)	3·0		(51)	0·091
Past	4·7		(50)	6·3		(50)	0·000
Annual family income (median)	$32 500		(42)	$29 000		(46)	0·059
Poverty level[c]			(49)			(49)	
≤100%		2·0			12·2		
101–199%		8·2			18·4		0·039
≥200%		89·8			69·4		

[a] The total number of respondents varies due to either missing values or non-applicability of the category
[b] Two tail p value for differences in distribution of black and white respondents
[c] The n for poverty level is greater than for family income because some respondents gave categorical, rather than exact, answers

Results

SAMPLE CHARACTERISTICS: INDIVIDUAL, HOUSEHOLD, TRACT AND BLOCK GROUP

Tables II and III present selected sociodemographic and social class characteristics of the respondents, while table IV provides data pertaining to their reproductive histories. Nearly 96% of the respondents, as compared to 40% of their mothers, had worked in the paid labour force (p < 0·001). Among both black and white women, the individual class measure categorised a higher proportion as "working class" than did the household class measure (82% v 65% for the black respondents, and 56% v 43% for the white respondents). Differences in the individual v household class measure stemmed from intrahousehold heterogeneity: of the 67 respondents individually identified as "working class", 25% lived in "not working class" households. At both the individual and household levels, black women were more likely to be "working class" than white women (p < 0·05).

Table III Individual class position of respondents and partners, percentage distribution by race

Class position	Respondents		Partners[a]	
	White (n=50)	Black (n=50)	White (n=35)	Black (n=30)
Working class				
Employee:				
Non-supervisory, non-professional	32·0	48·0	22·9	43·3
Supervisory, non-professional	24·0	34·0	20·0	13·3
Subtotal	56·0	82·0	42·9	56·6
Not working class				
Employee:				
Non-supervisory, professional	14·0	4·0	5·7	10·0
Supervisory professional	18·0	10·0	22·9	20·0
Self-employed:				
Non-professional	0·0	0·0	5·7	3·3
Professional	4·0	0·0	8·6	3·3
Owns a business and employs others	2·0	2·0	14·3	0·0
Subtotal	38·0	16·0	57·2	36·6
Other				
Not in paid workforce	6·0	2·0	0·0	6·7
Total[b]	100·0	100·0	100·1	99·9

[a] Only 35 white and 30 black respondents were living with a partner at the time of the interview. The "race" heading refers to the respondent; the racial identity of the respondents' partners was not ascertained.
[b] The sums do not all equal 100·0 due to rounding error

With regard to occupational distribution, nearly 30% of the white respondents, as compared to 15% of the black respondents, typically were employed in executive, managerial, or professional occupations (p < 0·05). In contrast, 25% of the black women, but only 4% of the white women, usually worked in service occupations (p < 0·05). Technical, sales, or clerical positions comprised most of the remaining occupations among the black and white respondents (53% and 61%, respectively); the rest (7%) worked chiefly as machine operators or were in the armed forces.

The dichotomised grouping of occupational categories that least misclassified the respondents' individual "working class" designation is shown in table V. Using this grouping, table VI presents data regarding the working class, poverty, and educational composition of the black and white respondents' census tracts and block groups. The range of values at the block group level typically exceeded that at the tract level. On average, black respondents were twice as likely as white respondents to live in predominantly working class, impoverished, and undereducated neighbourhoods (p < 0·01).

COMPARISON OF INDIVIDUAL, HOUSEHOLD, TRACT, AND BLOCK GROUP MEASURES AS PREDICTORS OF REPRODUCTIVE HISTORY

Table VII presents multiple linear regression results regarding three different reproductive outcomes, and provides a means of comparing individual, household, tract, and block group measures of social class. Three patterns are evident. First, household class position more often served as a significant predictor and explained more variance than did the individual class measure. Second, in comparison to the census tract measures, coefficients of the block group variables more accurately approximated those of the corresponding individual level variables in terms of both absolute value and statistical significance. Third, block group models consistently explained more variance than models using tract data. These patterns held true for models in which "yearly income" was the

Table IV Respondents' reproductive history, by race

Characteristic	White			Black			
	Mean	Percent	(n)[a]	Mean	Percent	(n)[a]	p value[b]
Ever pregnant		68·0	(50)		84·3	(51)	0·828
Age at first pregnancy	23·4		(34)	20·2		(41)	0·003
Age at first full term pregnancy	24·3		(28)	20·7		(38)	0·002
Number of total pregnancies	2·8		(34)	3·2		(43)	0·364
Number of full term pregnancies	1·6		(34)	2·3		(43)	0·040
Percent of total pregnancies that were terminated early[c]	39·0		(34)	31·0		(43)	0·349

[a] The total number of respondents varies due to either missing values or non-applicability of the category
[b] Two tail p value for differences in distribution of black and white respondents
[c] Includes spontaneous and induced abortions

Table V Specificity, sensitivity, and predictive value of respondent's occupational "class" measure as compared to respondent's individual class position

Individual class position[b]	Working class position by occupational "class"[a]	
	Working class	Non-working class
Working class	59	7
Non-working class	2	25

Sensitivity = 59/(59 + 7) = 89·4%
Specificity = 25/(2 + 25) = 92·6%
Predictive value (+) = 59/(59 + 2) = 96·7%
Predictive value (−) = 25/(7 + 25) = 78·1%

[a] Census based occupations in the occupational "class" measure:
Working class: Administrative support occupations including clerical; Sales occupations; Private household service occupations; Service occupations, except protective and private household; Precision production, craft, and repair occupations; Machine operators, assemblers, and inspectors; Transportation and material moving occupations; Handlers, equipment cleaners, helpers, and labourers.
Non-working class: Executive, administrative, and managerial occupations; Professional specialty occupations; Technicians and related support occupations; Protective service occupations; Farming, forestry, and fishing occupations
[b] Working class: non-professional employee (supervisory/non-supervisory)
Non-working class: professional employee (supervisory/non-supervisory), self employed, owner who employs others

Table VI Working class, poverty, and educational composition of respondents' census tracts (CT) and census block groups (BG), range of values and percent of respondents residing in type of census region, by race

Measure of composition	Census level	Respondents	
		White	Black
Percentage of population (minimum and maximum values):			
% Working class	CT	26–87	49–92
	BG	25–87	32–100
% Below poverty	CT	1–39	2–43
	BG	0–52	1–59
% Less than high school education	CT	1–41	0–70
	BG	2–58	0–90
Percent of respondents residing in census areas defined as: [a,b]			
Working class	CT	55	81
	BG	54	80
Impoverished	CT	17	50
	BG	20	47
Undereducated	CT	23	56
	BG	23	51

[a] Working class = ⩾66% working class
Impoverished = ⩾20% below the poverty line
Undereducated = ⩾25% not completed high school
[b] All two tailed p values for black/white differences were ⩽0·01

dependent variable (not shown); as expected, a significant positive association existed between higher income and both non-working class and non-poor status. Models containing the respondents' educational level are not presented, since education was a non-significant predictor of reproductive outcomes in both univariate and multivariate analyses. The association of increased educational level with these outcomes, however, paralleled that of non-working class and non-poor status.

With regard to the number of full term pregnancies, the triad of age, poverty and

household—but not individual—class were all significant predictors in models containing individual level data (N1 N2). For example, adjusting for all other variables, women above 200% of the poverty line had 1·3 fewer such pregnancies than those below the poverty line (95% CI = −2·5, −0·1), and women from non-working class households had 0·8 fewer children than those from working class households (95% CI = −1·4, −0·1) (N2). If these poverty and class variables were replaced by census tract measures (N3), only "age" remained significant. In contrast, the coefficient of the class variable in the analogous block group model (N4) was significant and also closely resembled the household class measure: women from block groups with less than a 66% working class composition averaged 0·7 fewer full term pregnancies than women from block groups in which more than 66% of the population was working class (95% CI = −1·4, −0·1). The initially significant crude excess of 0·7 full term pregnancies among black as compared to white respondents was reduced and rendered statistically non-significant in all four models by adjusting for age, class, and poverty.

In models with "percent of early terminated pregnancies" serving as the dependent variable (P1-P2), only age and household class—not individual class—were significant predictors at the individual level, with women from non-working class households having 20% more pregnancies terminate early than those from working class households (P2). The tract and block group class measures both yielded estimates comparable to the household class variable (P3, P4), with the block group model explaining more variance than the tract model. Race was non-significant in all four models.

In contrast, the younger age of black respondents at first full term pregnancy remained statistically significant in all four models (A1-A4), even after adjusting for age, class, and poverty status. A positive but non-significant association between increasing age at first full term pregnancy and increasing elevation above the poverty line was evident in both individual-level models (A1, A2). As compared to the census tract version (A3), the block group model (A4) had a significant overall F value, explained more variance, and yielded estimates for the effect of age that more closely resembled those of the indivdual level models.

CONTEXTUAL ANALYSIS: HOUSEHOLD AND BLOCK GROUP CLASS AS JOINT PREDICTORS OF REPRODUCTIVE HISTORY
Finally, table VIII offers evidence that data from the block group level might not only serve as a proxy for, but actually supplement, individual

Table VII Linear regression models with poverty and working class variables at individual, census tract, and census block group levels, for "number of full term pregnancies", "percent of early terminated pregnancies", and "age at first full term pregnancy"[a]

	Coefficient of variable in model (95% confidence interval)													
Model[b] a		RACE	AGE1	AGE2	POV1	POV2	RCLASS	HCLASS	CPOV20	CWC66	BPOV20	BWC66	Sig F	R^2
N1	2·5	0·3 (−0·4,1·0)	1·1 (0·4,1·8)	0·0 (−1·4,1·4)	−0·6 (−2·0,0·8)	−1·2 (−2·4,−0·1)	−0·2 (−0·7,0·4)						0·003	0·25
N2	2·9	0·3 (−0·4,0·9)	1·1 (0·5,1·8)	0·1 (−1·3,1·4)	−0·8 (−2·2,0·6)	−1·3 (−2·5,−0·1)		−0·8 (−1·4,−0·1)					0·000	0·32
N3	1·7	0·4 (−0·3,1·1)	1·0 (0·2,1·6)	−1·2 (−2·9,0·4)					−0·3 (−1·0,0·4)	−0·4 (−1·2,0·4)			0·010	0·21
N4	1·9	0·2 (−0·5,0·9)	1·0 (0·3,1·7)	−0·3 (−1·5,1·0)							−0·4 (−1·1,0·3)	−0·7 (−1·4,−0·1)	0·005	0·23
P1	0·2	0·0 (−0·2,0·2)	−0·2 (−0·3,0·0)	−0·2 (−0·6,0·1)	0·1 (−0·3,0·4)	0·2 (−0·1,0·5)	0·3 (−0·1,0·2)						0·276	0·11
P2	0·2	0·0 (−0·1,0·2)	−0·2 (−0·3,0·1)	−0·2 (−0·6,0·1)	0·1 (−0·3,0·5)	0·2 (−0·1,0·5)		0·2 (0·1,0·4)					0·040	0·18
P3	0·4	0·0 (−0·2,0·2)	−0·2 (−0·3,0·0)	−0·0 (−0·4,0·4)					0·0 (−0·2,0·2)	0·2 (0·0,0·4)			0·332	0·08
P4	0·3	0·0 (−0·1,0·2)	−0·2 (−0·3,0·0)	−0·1 (−0·4,0·3)							0·1 (−0·1,0·2)	0·2 (0·0,0·4)	0·120	0·12
A1	22·1	−2·9 (−5·5,−0·4)	−1·1 (−3·6,1·3)	2·3 (−2·6,7·1)	1·4 (−3·3,6·1)	2·8 (−1·2,6·8)	−0·4 (−2·6,1·7)						0·037	0·20
A2	21·6	−2·7 (−5·2,−0·2)	−1·4 (−3·9,1·1)	2·2 (−2·7,7·1)	1·7 (−3·3,6·8)	3·1 (−1·3,7·5)		0·1 (−2·4,2·6)					0·047	0·20
A3	24·7	−3·1 (−5·6,−0·5)	−1·2 (−3·5,1·2)	−1·7 (−7·9,4·5)					−1·1 (−3·7,1·6)	1·4 (−1·4,4·1)			0·115	0·15
A4	24·0	−2·9 (−5·4,−0·3)	−1·1 (−3·6,1·3)	2·0 (−2·9,6·9)							−0·3 (−3·0,2·3)	2·1 (−0·6,4·7)	0·032	0·19

[a] Variables used in the regression models:
a:　　　　constant
RACE:　　0 = White, 1 = Black
AGE1:　　0 = 20–39, 1 = 40–59
AGE2:　　0 = 20–39, 1 = 60–80
POV1:　　0 = ≤ Poverty level, 1 = 101–199% Poverty level
POV2:　　0 = ≤ Poverty level, 1 = 200 + % Poverty level
RCLASS: 0 = Working class (individual), 1 = Non-working class
HCLASS: 0 = Working class (household), 1 = Non-working class
CPOV20: 0 = 20 + % below poverty (census tract), 1 = <20% below poverty
CWC66:　0 = 66 + % working class, (census tract), 1 = <66% working class
BPOV20: 0 = 20 + % below poverty (block group), 1 = <20% below poverty
BWC66:　0 = 66 + % working class (block group), 1 = <66% working class
[b] Model definition: Dependent variable: N = "number of full term pregnancies"; P = "percent of early terminated pregnancies"; A = "age at first full term pregnancy";
Data level: 1 = individual level, individual class, 2 = individual level, household class, 3 = census tract level, 4 = block group level.

level data. In these contextual regression analyses, both the respondent's household class and block group class measure were simultaneously entered into the equations by means of dummy variables. With regard to the number of full term pregnancies (N5), women from non-working class households in non-working class block groups had significantly fewer children (adjusting for age, race, and poverty) than women from working class households in working class block groups. Women from non-working class households in working class block groups and women from working class households in non-working class block groups occupied an intermediate position. The reverse pattern held for "percent of early terminated pregnancies" (P5). No significant contextual effects appeared when "age at first full term pregnancy" served as the dependent variable (A5).

Discussion

The results of this study suggest that it is feasible and meaningful to employ a relational approach to characterise a women's social class standing, and that a household class measure based on this method serves as a better predictor of reproductive outcomes and economic level than does individual class standing. The findings further indicate that it is valid to approximate this household class measure with census block group data, whereas comparable indicators from the census tract level perform less well. Beyond this, they provide preliminary evidence that a person's social context, as described by block group data, may provide additional information about social gradients not captured by individual level data.

That categorising women's social class in terms of individual versus household position leads to different results is consistent with other public

Table VIII Contextual linear regression models, with joint household/census block group class variable, for "number of full term pregnancies", "percent of early terminated pregnancies", and "Age at first full term pregnancy"[a]

	Coefficient of variable in model (95% confidence interval)									
Model[b] a		RACE	AGE1	AGE2	BPOV20	BWC1	BWC2	BWC3	Sig F	R^2
N5	2·4	0·2 (−0·5,0·9)	1·0 (0·3,1·7)	0·3 (−1·2,1·7)	−0·2 (−1·0,0·6)	−0·8 (−1·8,0·3)	−0·8 (−1·7,0·1)	−1·3 (−2·2,−0·4)	0·003	0·28
P5	0·2	0·0 (−0·1,0·2)	−0·1 (−0·3,0·0)	−0·3 (−0·6,0·1)	0·0 (−0·2,0·2)	0·2 (0·0,0·5)	0·2 (0·1,0·4)	0·4 (0·1,0·6)	0·027	0·22
A5	24·5	−2·9 (−5·6,−0·3)	−1·0 (−3·6,1·6)	1·6 (−3·6,6·8)	−0·6 (−3·5,2·4)	2·7 (−1·0,6·5)	0·6 (−2·7,4·0)	1·8 (−1·9,5·4)	0·101	0·19

[a] Variables used in the regression models:
a:　　　　constant
RACE:　　0 = White, 1 = Black
AGE1:　　0 = 20–39, 1 = 40–59
AGE2:　　0 = 20–39, 1 = 60–80
BPOV20: 0 = 20 + % below poverty, 1 = <20% below poverty (block group)
BWC1:　　0 = working class household, 66 + % working class block group
　　　　　1 = non-working class household, 66 + % working class block group
BWC2:　　0 = working class household, 66 + % working class block group
　　　　　1 = working class household, <66% working class block group
BWC3:　　0 = working class household, 66 + % working class block group
　　　　　1 = non-working class household, <66% working class block group
[b] Model definition: Dependent variable: N5 = "Number of full term pregnancies"; P5 = "Percent of early terminated pregnancies"; A5 = Age at first full term pregnancy."

health and sociologic research regarding women's overall mortality,[10] [11] parity,[13] and income plus educational level.[8] [12–15] A recent British investigation based on national survey data, for example, observed that the traditional method of assigning wives their husband's rank was misleading with regard to predicting not only income and voting behaviour, but also household size.[13] Specifically, in those families where the woman held a non-manual job and her husband was a manual worker, the number of children was related primarily to the woman's, not the husband's, occupational status. This latter finding further suggests that it may be important to categorise not just women, but also men, by their household as well as individual class position.

Evidence that the sociodemographic characteristics of people's immediate neighbourhoods can not only approximate but also modify the risk associated with their individual level socioeconomic position has likewise been found in public health investigations regarding overall mortality[19] and also birthweight, height, and respiratory illness among children.[8] A recent prospective study comparing overall mortality in a poverty versus non-poverty area of Alameda County, for example, detected a 1·7 times greater risk of mortality (95% CI = 1·2, 2·4) among residents of the poverty area during the nine year follow up period even after controlling for a large number of demographic, physiological, and psychosocial characteristics.[19] Contextual effects have also been documented for such phenomena as voting behaviour[41] [42] and students' test scores.[43]

To date, no US studies have examined directly whether problems associated with ecologic fallacy can be reduced by using census data from the block group, as opposed to tract, level. This investigation's findings in favour of the block group level are plausible. Since at least the 1950s, researchers have noted the greater heterogeneity among tract as compared to block group inhabitants,[44] [45] with block group data able to identify pockets of poverty or affluence lost by aggregating up to the census tract level.[18] One graphic example concerns a well known phenomenon involving hills: people who live at the top tend to be more affluent than those living at the bottom. If a hill were included in a census tract, block group data would identify this gradient, whereas tract data would not.

Most discussions concerning census data and ecological fallacy, however, have focused on the "classic" situation, in which *both* the independent and dependent variables are grouped data, and underlying factors associated with the grouping process confound the results.[20–22] [39] [40] In contrast, the present study design used individual level dependent variables in conjunction with individual and census based independent variables. If the "classic" type of ecological fallacy were at issue, then the census tract regression models should have yielded more, not less, "significant" coefficients (albeit spuriously so), and also should have "explained" more, not less, variance, than the corresponding block group models. That this did not occur, and that neither the tract nor block group versions produced statistically significant "class" effects when such effects were absent in the individual level models, further suggests that this investigation's census based strategy was not rendered invalid by the usual concerns regarding ecological fallacy.

Nevertheless, several caveats do apply to the results of this study. With regard to census data, possible secular changes in neighbourhood composition, along with the respondents' migration patterns, might have resulted in 1980 census data mischaracterising the respondents' 1987 residential areas. Such misclassification, however, most likely would have biased the census based comparisons toward the null, and so would not account for this study's findings. Another concern is that the individual and household class measures presented here have not yet been validated in other public health studies. Even so, the occupations they characterised as containing chiefly "working class" respondents were virtually identical to those identified for black and white women in prior sociological research utilising a property/control model of social class.[5] [6]

Moreover, unlike traditional occupational status scales which are based on perceived prestige,[8] [31] [32] the method of evaluating class position used in this study primarily relied on non-subjective criteria (eg, employee *v* self-employed *v* owner; supervisee *v* supervisor). These demarcations also directly reflect job control and other social dimensions now thought to contribute to class gradients in disease.[1–3] [46] [47] To the extent that this approach incorporated subjective factors (selecting the "usual" type of employment, choosing between "professional" and "non-professional"), status considerations would most likely have deflated, not inflated, membership in the "working class" category, producing a conservative bias. The reported reproductive outcomes also were unlikely to have been affected substantially by recall bias, since evidence indicates that women accurately remember their age at first full term pregnancy and the number of full term and early terminated pregnancies.[48–50]

Further concerns pertain to the study's small sample size, exclusive focus on reproductive outcomes, and limited number of variables. While the sample size restricts the degree to which the findings can be generalised, it is important to note that the respondents constituted a random, population based sample, and that significant household class and poverty effects were found despite the small sample size. If other health outcomes had been considered, education and individual class might have proved to be significant risk factors.

The small number of independent variables used in the regression equations also does not invalidate the results. This is because the purpose of the study was to determine whether it is possible to stratify by the selected class and poverty measures, rather than elucidate the causal pathways through which class and poverty exert their effects. Additionally, the finding that the class and poverty variables independently contributed toward predicting reproductive events, whereas educational level did not, further bolsters the view that single concept, theoretically

derived measures may be more useful and appropriate for public health research than empirically constructed composite variables.[1-3]

The limited number of variables, however, may temper interpretation of the contextual regression analyses. As noted by some critics, if the individual level model is not fully specified, the variance explained by group level variables potentially could be better accounted for by additional individual level data.[40 51] That the contextual estimates nonetheless produced plausible trends in the expected direction indicates that this technique may offer public health researchers an important means by which to avoid what has been termed the "individualistic fallacy",[20] that is, the assumption that individual level data are sufficient to explain social phenomena.[20 52] If a given disease or health event is influenced by factors associated with a community's overall standard of living (eg, employment, sanitation, crowding, crime, literacy), whether apart from or in conjunction with individual occupational or environmental exposures, then neighbourhood characteristics could conceivably modify the risk associated with a person's individual level risk factors.

In sum, the results of this study suggest that the obstacle posed by the absence of social class data in most individual health records could perhaps be addressed through judicious use of census block group data. These data constitute a uniform source of sociodemographic variables available to investigators throughout the United States. As such, they readily can be incorporated into both case-control and cohort studies, and provide a common measure of neighbourhood class composition equally applicable to women and men of all ages, from infancy through their senior years.[9] Beyond this, block group data could also be used to construct population based incidence or prevalence rates stratified by social class. This is because the denominators for these rates are typically census derived, and so can be characterised by the same census based social class measures as the relevant numerator data.[53] If future research substantiates the validity of this approach, then block group data may offer a powerful and consistent means to describe and analyse social gradients in disease among different populations. In doing so, this technique could potentially provide new avenues for aetiological inquiry and renewed impetus to eliminate socially determined disparities in health.

Appendix
In the interview, the "usual type of employment" was defined as "the kind of job you've most typically held", which could be "the type of job you've held the longest, or the type of job you think best represents the kind of work you've done to earn your income". Respondents first were asked if they "ever worked outside of your home and either earned some income or contributed to your family's income by working as an unpaid employee in a family business". Those responding "yes" were then asked whether they typically were: (1) "an employee of someone else;" (2) "self employed in your own business or professional practice;" or (3) "owner of your own business and employer of workers other than yourself". Those selecting "an employee of someone else" were asked whether they typically were: (1) "a person who supervises or employs other workers" v "a non-supervisory worker", and (2) "a professional employee" v "some other type of worker". "Self employed" respondents were asked whether they typically were self employed: (1) in "a professional practice", (2) "as an artisan or craft worker", or (3) in "some other type of job".

1 Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev* 1988; **10:** 87–121.
2 Susser M, Watson W, Hopper K. *Sociology in medicine.* 3rd ed. New York: Oxford University Press, 1985.
3 Morgenstern H. Socioeconomic factors: concepts, measurements, and health effects. In: Ostfeld AM, Eaker ED, eds. *Measuring psychosocial variables in epidemiologic studies of cardiovascular disease.* NIH Publication No 85–2270. Bethesda, MD: National Institutes of Health, 1985: 3–35.
4 Haug M. Measurement in social stratification. *Annu Rev Sociol* 1977; **3:** 51–77.
5 Wright EO, Perrone L. Marxist class categories and income inequality. *Am Sociol Rev* 1977; **42:** 32–55.
6 Wright EO, Costello C, Hacken D, Sprague J. The American class structure. *Am Sociol Rev* 1982; **47:** 709–26.
7 Parkin F. *Marxism and class theory: a bourgeois critique.* New York: Columbia University Press, 1979.
8 Morgan J, Clinn S. ACORN group, social class, and child health. *J Epidemiol Community Health* 1983; **37:** 196–203.
9 Alexander FE, O'Brien F, Hepburn W, Miller M. Association between mortality among women and socioeconomic factors in general practices in Edinburgh: an application of small area statistics. *BMJ* 1987; **295:** 754–6.
10 Morgan M. Measuring social inequality: occupational classifications and their alternatives. *Community Med* 1983; **5:** 116–24.
11 Arber S. Social class, nonemployment, and chronic illness: continuing the inequalities in health debate. *BMJ* 1987; **294:** 1069–73.
12 Haug M. Social class measurement and women's occupational roles. *Social Forces* 1973; **52:** 87–98.
13 Britten N, Heath A. Women, men and social class. In: Gamarnikow E, Morgan DHJ, Purvis J, Taylorson DE, eds. *Gender, class and work.* London: Heineman, 1983: 46–60.
14 Goldethorpe J. Women and class analysis: in defense of the conventional view. *Sociology* 1983; **17:** 465–88.
15 Stanworth M. Women and class analysis: a reply to John Goldethorpe. *Sociology* 1984; **18:** 159–70.
16 Westermeyer J. Problems with surveillance methods for alcoholism: differences in coding systems among Federal, State, and private agencies. *Am J Public Health* 1988; **78:** 130–3.
17 US Department of Health and Human Services. *Health status of minorities and low income groups.* DHHS Publication No (HRSA) HRS-P-DV 85–1. Washington, DC: US Government Printing Office, 1985.
18 US Bureau of the Census. *Census use study: health information system-II,* Report No 12. Washington, DC: US Government Printing Office, 1976.
19 Haan M, Kaplan G, Camacho T. Poverty and health: prospective evidence from the Alameda County study. *Am J Epidemiol* 1987; **125:** 989–98.
20 Alker HR. A typology of ecological fallacies. In: Doggan M, Rokkan S, Eds. *Social ecology.* Cambridge, MA: MIT Press, 1969: 69–86.
21 Morgenstern H. Uses of ecologic analyses in epidemiologic research. *Am J Public Health* 1982; **72:** 1336–44.
22 Selvin S. Two issues concerning the analysis of grouped data. *Eur J Epidemiol* 1987; **3:** 284–7.
23 US Bureau of the Census. *The social and economic status of the Black population in the United States: an historical view, 1790–1978.* Current Population Reports, Special Studies Series P-23, No 80. Washington, DC: US Government Printing Office, 1980.
24 National Survey of Family Growth. *Socioeconomic differentials and trends in the timing of births. Vital and health statistics.* Series 23, No 6, DHHS Publication No (PHS) 81–1982. Hyattsville, MD: National Center for Health Statistics, 1981.
25 National Center for Health Statistics. *Induced terminations of pregnancy: reporting states, 1980.* Monthly Vital Statistics Report, 32(8), Suppl DHHS Publication No (PHS) 84–1120. Hyattsville, MD: National Center for Health Statistics, 1983.

26 Hayes CD, *Risking the future: adolescent sexuality, pregnancy, and childbearing*. Vol I. Washington, DC: National Academy Press, 1987.

27 Hartage P, Brinton LA, Rosenthal JF, Caholt JI, Hoover RN, Waksberg J. Random digit dialing in selecting a population-based control group. *Am J Epidemiol* 1984; 120: 825–33.

28 Kish L. A procedure for objective respondent selection within the household. *J Am Stat Assoc* 1949; 44: 380–3.

29 Illis WR, Swanson H, Satariano ER, Schwartz AG. Summary measures of occupational history: a comparison of last occupation and industry with usual occupation and industry. *Am J Public Health* 1987; 77: 1532–4.

30 Carter H, Glick PC. *Marriage and divorce: a social and economic study*. Cambridge, MA: Harvard University Press, 1970.

31 Mutchler JE, Poston DL. Do females necessarily have the same occupational status scores as males?: a conceptual and empirical examination of the Duncan Socioeconomic Status Index and Nam-Powers Occupational Status Scores. *Soc Sci Res* 1983; 12: 353–62.

32 Stevens G, Hoisington E. Occupational prestige in the 1980 US labor force. *Soc Sci Res* 1987; 16: 74–105.

33 Congressional Budget Office. *Trends in family income: 1970–1986*. Washington, DC: US Government Printing Office, 1988.

34 Beeghley L. Illusion and reality in the measurement of poverty. *Social Problems* 1984; 31: 322–3.

35 US Bureau of the Census. *Money income and poverty status of families and persons in the United States 1986: advanced data from the March 1987 Current Population Survey*. Consumer Income Series P-60. Washington, DC: US Government Printing Office, 1987.

36 Kaplan CP, Van Valey TL. *Census '80: continuing the fact finding tradition*. Washington, DC: US Government Printing Office, 1980.

37 Bassett M, Krieger N. Social class and Black-White differences in breast cancer survival. *Am J Public Health* 1986; 76: 1400–03.

38 US Bureau of the Census. *Poverty areas in large cities*. PC80–2–8D. 1980 Census of the Population, Vol 2, Subject Reports. Washington, DC: US Government Printing Office, 1985.

39 Boyd LH, Iversen GR. *Contextual analysis: concepts and statistical techniques*. Belmont, CA; Wadsworth Publishing Co, 1979.

40 Blalock HM. Contextual-effects models: theoretic and methodological issues. *Annu Rev Sociol* 1984; 10: 353–72.

41 Przeworski A. Contextual models of political behaviour. *Political Methodol* 1974; 1: 27–61.

42 Segal DR, Meyer MW. The social context of political partisanship. In. Doggan M, Rokkan S, eds. *Social ecology*. Cambridge, MA: MIT Press, 1969: 217–32.

43 Langbein LI. Schools or students: aggregation problems in the study of achievement. *Evaluation Studies Annu Rev* 1977; 2: 270–98.

44 Foley DL. Census tracts and urban research. *J Am Stat Assoc* 1953; 48: 733–42.

45 Myers JK. Note on the homogeneity of census tracts: a methodological problem in urban ecological research. *Social Forces* 1954; 32: 364–6.

46 Syme SL. Social epidemiology and the work environment. *Int J Health Services* 1988; 18: 635–45.

47 Karasek RA, Theorell T, Schwartz JE, Schanll PL, Pieper CF, Michela JL. Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Survey (HANES). *Am J Public Health* 1988; 78: 910–18.

48 Harlow SD, Linet MS. Agreement between questionnaire data and medical records: the evidence for accuracy of recall. *Am J Epidemiol* 1989; 129: 233–48.

49 Tilley BC, Barnes AB, Bergstralh E, *et al*. A comparison of pregnancy history recall and medical records. *Am J Epidemiol* 1985; 121: 269–81.

50 Wilcox AJ, Horney LF. Accuracy of spontaneous abortion recall. *Am J Epidemiol* 1984; 120: 727–33.

51 Hauser RM. Context and consex: a cautionary tale. *Am J Sociol* 1970; 75: 645–64.

52 Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985; 14: 32–8.

53 Krieger N. Social class and the Black/White crossover in the age-specific incidence of breast cancer: a study linking census-derived data to population-based registry records. *Am J Epidemiol* 1990; 131: 804–14.