



## Structural Covariates of Homicide Rates: Are There Any Invariances Across Time and Social Space?

Kenneth C. Land; Patricia L. McCall; Lawrence E. Cohen

*The American Journal of Sociology*, Vol. 95, No. 4. (Jan., 1990), pp. 922-963.

Stable URL:

<http://links.jstor.org/sici?sici=0002-9602%28199001%2995%3A4%3C922%3ASCOHRA%3E2.0.CO%3B2-W>

*The American Journal of Sociology* is currently published by The University of Chicago Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# Structural Covariates of Homicide Rates: Are There Any Invariances across Time and Social Space?<sup>1</sup>

Kenneth C. Land  
*Duke University*

Patricia L. McCall  
*North Carolina State University*

Lawrence E. Cohen  
*University of California, Davis*

This study demonstrates that the empirical literature on the structural covariates of homicide rates contains inconsistent findings across different time periods and different geographical units. This apparent variance of findings may be due to statistical or methodological artifacts of particular studies, such as different time periods covered, units of analysis, samples, model specifications, and problems of statistical analysis and inference. A baseline regression model using 11 structural covariates is estimated for cities, metropolitan areas, and states in 1960, 1970, and 1980. The empirical estimates of this model exhibit instability because of high levels of collinearity among several regressors. Principal components analysis is applied to simplify the dimensionality of the structural covariate space. Reestimation of the regression model then indicates that the apparent inconsistencies across time and social space are greatly reduced. The theoretical significance of the findings for substantive theories of violent crime are discussed.

The past two decades have seen an upsurge in the number of empirical studies that address the question of why some cities, metropolitan areas,

<sup>1</sup> The research reported here was supported, in part, by National Science Foundation grant SES-8800595. We are grateful to five anonymous *AJS* referees for stimulating and useful comments on an earlier version of this paper. Requests for reprints should be sent to Kenneth C. Land, Department of Sociology and Center for Demographic Studies, Duke University, Durham, North Carolina 27706.

or states have higher homicide rates than others.<sup>2</sup> Sociologists and criminologists have attempted to explain interunit variation in these rates through application of a variety of statistical techniques using many different social, demographic, ecological, and economic covariates.<sup>3</sup> Initially, much of this research employed data from the 1960 census for the 48 contiguous states in order to assess the validity of the "southern subculture-of-violence" hypothesis. More recently, however, empirical studies using data from the 1970 and 1980 censuses have focused on cities and metropolitan aggregate units as well as states and have expanded their theoretical interests to include a wider range of possible covariates such as poverty, economic inequality, racial composition, and family disorganization.

Given this accumulated body of research literature, it now is appropriate to ask whether or not the empirical results have produced a consistent or invariant set of findings. That is, Do we now have a set of consistent findings about the structural covariates of homicide rates that allow for reliable generalizations, or do the accumulated results differ from one time period to the next, or one level of analysis to another?

In this article, we demonstrate that, instead of the accumulated literature depicting an invariant set of findings across different time periods and levels of analysis, it indicates a general pattern of inconsistent results. For example, a structural covariate found in one study to have, say, a statistically significant positive effect on homicide rates for a particular unit of analysis (i.e., city, metropolitan area, or state) may be found to have statistically null or even negative effects at other time periods or levels of analysis. The reader familiar with this literature will immediately recall the many debates and publications that have resulted from disagreements over these reported inconsistencies.

How can these empirical inconsistencies be explained? Are there no covariate invariances for homicide rates? That is, Are the effects of structural covariates on homicide rates so context specific that they appear to be variable over different periods of time and at various levels of analysis? Note that a finding of no covariate consistency over social space and

<sup>2</sup> More generally, this question has been posed for various crime rates, with violent crime rates and homicide of particular interest. Because of space limitations, and because homicide is considered the most accurately reported crime rate for interunit comparisons, our study will focus only on homicide rates.

<sup>3</sup> Literature citations are deferred to the next section of the paper. We note that a number of studies have been directed at other levels of analysis, e.g., census tracts (Schuerman and Kobrin 1986), suburbs (Stahura and Sloan 1988), and counties (Breault 1987). But the bulk of research has focused on cities, metropolitan areas, and states. For this reason, and for purposes of comparability, we limit our attention here to these levels.

time would cast considerable doubt on the possibility of developing a generalizable theory to explain community, metropolitan, or ecological structural effects on homicide or, more generally, any other type of crime rate. At best, theoretical efforts to explain crime could become entangled in a potentially unending specification of context-specific relationships. Sociologists and criminologists might then be well advised to pursue the development of general explanations of crime-rate variation at the level of individual differences in criminality, as recently advocated by Wilson and Herrnstein (1985).<sup>4</sup>

Before reaching this radical conclusion, however, we explore the possibility that the apparent variance in findings of the significance of the covariates of homicide (and perhaps other types of crimes as well) over social space and time is a statistical or methodological artifact. We begin in the next section of the article by examining the theoretical perspectives, time periods, units of analysis, samples, model specifications, and statistical procedures of a number of previous studies. We conclude that numerous differences in research procedures used in extant studies, coupled with data-analysis and statistical inference problems, may indeed account for the inconsistencies in findings.

We then define a theoretically defensible "baseline model" for which it is possible to obtain empirical data across all three levels of analysis (cities, metropolitan areas, and states) for the three most recent decennial census years (1960, 1970, and 1980). Results from initial estimation of this model reveal the statistical problems we expected from examining extant studies. To deal with these problems, we respecify and reestimate our baseline model. This produces a number of empirical findings that do indeed generalize across all levels of analysis and time periods as well as some results that appear to challenge accepted wisdom regarding structural covariates of homicide rates. We then discuss the implications for subsequent research and for theories of the structural covariates of homicide rates.

## PREVIOUS RESEARCH

### Theoretical Perspectives

Much of the recent interest in the explanation of differential homicide rates was generated by two papers published in different disciplinary journals by Hackney (1969) and Gastil (1971). Both papers attempted to explain why homicide rates in southern states were significantly higher year after year than those of states in other regions. Inspired by the theoretical views of Wolfgang and Ferracuti (1967), these papers ap-

<sup>4</sup> For a critique of this position, see Cohen (1987) and Cohen and Machalek (1988).

peared to confirm the hypothesis that, within the southern states, there exists a value system different from that in other regions of the country. This value system is said to reinforce violent behavior and is termed a "subculture of violence."<sup>5</sup> The subculture-of-violence thesis also has been applied to the explanation of high black crime rates in urban areas (Wolf-gang and Ferracuti 1967; Curtis 1975; Silberman 1978). Specifically, it is posited that proportionally large populations of blacks are characterized by cultural values that promote violence independently of structural factors such as poverty.<sup>6</sup>

Typically, the southern subculture-of-violence thesis is evaluated against what we term a social control model of the causes of crime rates, which has its roots in the classical Chicago school studies of urbanization and delinquency by Wirth (1938) and Shaw and McKay (1942). Wirth (1938) chronicled how the large population size, high density, and great racial and ethnic heterogeneity of urban areas produce anonymous, superficial, and transitory human relations. This, in turn, increases friction among urban dwellers, which necessitates increasing formal regulation and control and leads to a complicated stratification of social relations that breaks down shared cultural understandings. Shaw and McKay (1942) posited that it is the social disorganization created by the urban characteristics cited by Wirth, as well as by two additional aggregate features of some urban neighborhoods—economic hardship (poverty) and population mobility—that leads to the weakening of social control and the development of delinquent subcultures and increased delinquency rates. As explicated and reformulated by Kornhauser (1978) and Sampson (1987*a*), the central hypothesis of this neoclassical community-control theory is that these and related community-level characteristics (e.g., family disintegration) directly or indirectly affect informal social control networks, community attachment, anonymity, and the capacity for surveillance and guardianship. A weakening of these dimensions of social organization is posited to lead to increased rates of deviance and crimes such as homicide.<sup>7</sup>

<sup>5</sup> Additional discussions of the subculture-of-violence thesis can be found in Ball-Rokeach (1973), Dixon and Lizotte (1987), Doerner (1975), Erlanger (1976), and Reed (1971, 1972).

<sup>6</sup> Recent research by Sampson (1987*b*) casts serious doubt on the applicability of the subculture-of-violence thesis to the explanation of urban black violence.

<sup>7</sup> Although the list of structural covariates given here is justified in terms of the control, strain, and criminal opportunity theoretic arguments typical of recent studies of structural covariates of homicide, it should be noted that application of differential association theory (Sutherland 1947) would lead to a similar list of structural covariates to explain interunit variability in crime rates. But the two theories differ in terms of the individual-level mechanisms by which indicators of social disorganization (control theory) or differential social organization (differential association theory) are translated

A similar theoretical model explaining crime rates also has been the baseline for a sequence of studies examining the importance of economic inequality/relative-deprivation/strain-theoretic factors. Notable among these studies, Blau and Blau (1982) revitalized Merton's (1938) strain theory of deviance by arguing that a particular type of socioeconomic inequality—that associated with ascribed positions—is a key cause of metropolitan violence and more important than urban poverty levels per se in explaining crime. This article has stimulated a great deal of research and debate. For example, a series of articles by Messner (1982, 1983*a*, 1983*b*) and Bailey (1984) debate the relative merits of poverty versus relative deprivation/economic inequality in the explanation of homicide rates for cities and metropolitan areas.<sup>8</sup>

The theoretical orientations of recent research on covariates of homicide rates are also often informed by one or both perspectives on the effects of age. The first perspective is based on the long-standing observation of a greater propensity for teenagers and young adults to commit more crimes than individuals at other ages and is most recently articulated by Hirschi and Gottfredson (1983). Much debate has been aired over whether or not this relationship is a basic fact beyond the pale of contemporary criminological theory (see, e.g., Baldwin 1985; Greenberg 1985; Hirschi and Gottfredson 1985*a*, 1985*b*). Regardless of the outcome of this debate, however, the existence and invariance of the age and crime-propensity relationship is well established. The consequence is that most studies of interunit variation in homicide rates posit a positive relationship between the concentration of population in the teenage and young adult ages and the homicide rate.<sup>9</sup>

The second perspective is articulated by Cohen and Land (1987), who

---

into criminal behavior. As explicated, e.g., by Matsueda and Heimer (1987), the key mechanisms in the former are various dimensions of social bonding, whereas the latter emphasizes one critical dimension—an excess in the legal code of definitions favorable to crime over those unfavorable. Kornhauser (1978, p. 26) adds an additional criterion that must be satisfied in order for a model to be classified as a cultural deviance/differential association theory and not merely as a variant of a social disorganization model (see n. 33).

<sup>8</sup> For theoretical discussions of the relative-deprivation conceptualization of poverty, see Miller and Roby (1970) and Townsend (1974).

<sup>9</sup> There is substantial variability in the particular age-structure index used by researchers to operationalize this proposition. We employ the percentage of the population in the age range 15–29 because this was found by Cohen and Land (1987) to be the combination of five-year age groupings with the highest propensity for homicide in recent decades in the United States. Other frequently used age ranges are 15–24 and 20–34. We have experimented empirically with several variations on the definition of this percentage and find that they are highly correlated for the units of analysis of concern in the present paper. Thus, the particular age range used in a given study is not likely to be of great importance.

note that teenagers and young adults not only commit crimes more frequently than those in other age groups but also are more likely to be victims. This presumably is due to the participation of young persons in lifestyles at greater risk of victimization. Thus, at the level of an aggregate unit, the age-structure variable can have an effect both through the supply of victims of crimes and the relative frequency of offenders. This alternative interpretation of the age-structure variable is due to the criminal opportunity/routine-activities/life-style theory of crime and victimization rates developed especially by Cohen and Felson (1979), Cohen, Felson, and Land (1980), Cohen, Kluegel, and Land (1981), Hindelang, Gottfredson, and Garofalo (1978), and Cook (1986). A basic proposition derived from this theory is that an increase in the concentration of work and leisure activities outside households is associated with an increase in crime rates.

This theory often leads to interpretations of the effects of covariates on crime rates that are at variance with predictions from traditional criminological theories that focus solely on the supply of motivated offenders. For instance, control and strain theories would lead to the expectation that a higher unemployment rate would be associated with a higher crime rate owing to an increase in motivated offenders. By contrast, criminal opportunity theory would point to the effects a higher unemployment rate has on routine-activity patterns and the concentration of activities in or near the household and at less risk of criminal victimization, thus implying a negative effect on crime rates.

### Contradictory Empirical Findings

Based on various combinations of arguments from the foregoing theoretical perspectives, numerous empirical findings have accumulated over the past two decades from studies that estimate the effects of structural covariates on interunit variation in homicide rates. Do these findings exhibit the consistency one would expect on the basis of the above theories?

To address this question, we have assembled table 1, which summarizes the empirical results obtained in some 21 studies employing various combinations of 11 covariates that are hypothesized to have significant effects on homicide rates. For each of these covariates, comparable definitions and data are available at various time points and levels of analysis (cities, metropolitan areas or SMSAs, as designated by the Bureau of the Census, and states). The covariates are population size, population density, percentage of the population that is black, percentage of the population ages 15–29, percentage of the population of males ages 15 and over that is divorced, percentage of children 18 years old or younger not living

TABLE 1

## STRUCTURAL COVARIATES OF HOMICIDE CRIME RATES—RESULTS FROM PREVIOUS STUDIES OF CITIES, SMSAs, AND STATES

PREVIOUS STUDIES	STRUCTURAL COVARIATES										
	Population Size	Population Density	Percentage Black	Percentage Ages 15-29	Percentage Divorced	Percentage Kids Not with Both Parents	Median Family Income	Percentage Families in Poverty	Gini Index of Income Inequality	Percentage Unemployed	South
City level using 1960 data:											
Bailey 1984:											
Model 1 ( <i>N</i> = 73) . . . .	0 <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	+	0	*	0
Model 2 ( <i>N</i> = 73) . . . .	0 <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	+	*	*	0
Model 3 ( <i>N</i> = 138) . . . .	0 <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	*	0	*	+
Model 4 ( <i>N</i> = 138) . . . .	0 <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	+ (low inc)	*	*	+
Gastil 1971 ( <i>N</i> = 130) . .	0	*	+ (white)	— (% not 20-34)	*	*	+ (adj.) (sic?)	*	*	*	— (sic) <sup>b</sup>
City level using 1970 data:											
Bailey 1984:											
Model 1 ( <i>N</i> = 153) . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	—	*	*	*	+	0	*	0
Model 2 ( <i>N</i> = 153) . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	—	*	*	*	+	*	*	0
Model 3 ( <i>N</i> = 153) . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	—	*	*	*	*	0	*	0
Model 4 ( <i>N</i> = 153) . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	—	*	*	*	0 (low inc.)	*	*	0
Jackson 1984:											
Model 1 ( <i>N</i> = 408) . . . .	+	0	+	0 (15-24)	*	*	*	+	0 <sup>c</sup>	0	0
Model 2 ( <i>N</i> = 265) . . . .	+	0	+	— (15-24)	*	*	*	+	0 <sup>c</sup>	0	0
Model 3 ( <i>N</i> = 143) . . . .	—	+	+	0 (15-24)	*	*	*	0	0 <sup>c</sup>	0	0
Loftin and Parker 1985											
( <i>N</i> = 49) . . . . .	+ <sup>a</sup>	0	+ (non-white)	0 <sup>a</sup> (18-24)	*	*	*	+ (infant mortality)	*	*	0
Messner 1983b:											
Model 1 (Nonsouth) ( <i>N</i> = 256) . . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	*	0	*	*
Model 2 (South) ( <i>N</i> = 91) . . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	*	0	*	*
Model 3 (Nonsouth) ( <i>N</i> = 256) . . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	+	*	*	*



Model 4 (South) (N = 91) . . . . .	+ <sup>a</sup>	0 <sup>a</sup>	+	0	*	*	*	0	*	*	*	*	*	*
Sampson 1985 (N = 54) . . . . .	+	*	+	*	*	*	*	+	*	*	0 <sup>c</sup>	-	*	*
City level using 1980 data:														
Sampson 1986 (N = 171)	+	*	+	*	*	*	*	*	+	(di- vorce rate)	+	*	*	*
Williams and Fiewelling 1988 (N = 168) . . . . .	*	+ <sup>d</sup>	+ <sup>d</sup>	*	*	*	*	+ <sup>d</sup>	*	*	*	*	*	0
SMSA level using 1960 data:														
Gastil 1971 (N = ?) . . . . .	0	*	-	(non- white)	0 (20-34)	*	*	0 (adj.)	*	*	*	*	*	+ <sup>b</sup>
SMSA level using 1970 data:														
Blau and Blau 1982 (N = 125) . . . . .	0 <sup>d</sup>	*	+	*	*	+	*	0	*	*	+	*	*	*
Blau and Golden 1986 (N = 125) . . . . .	+ <sup>d</sup>	*	+ <sup>d</sup>	*	*	+ <sup>d</sup>	*	+ <sup>d</sup>	*	*	*	*	*	+ <sup>d,e</sup>
Crutchfield, Geerken, and Gove 1982 (N = 65) . . . . .	0	*	+	-	(m, 10-18)	*	*	0	*	*	*	-	*	*
DeFronzo 1983:														
Model 1 (N = 39) . . . . .	+ <sup>a</sup>	*	+	0 (m, 15-24)	*	*	*	0 (adj.)	*	*	*	0	0	0
Model 2 (N = 39) . . . . .	+ <sup>a</sup>	*	+	0 (m, 15-24)	*	*	*	0 (adj.)	*	*	0	0	0	0
Messner 1982:														
Model 1 (N = 204) . . . . .	+ <sup>a</sup>	- <sup>a</sup>	+	0	*	*	*	*	*	*	0	*	+	+
Model 2 (N = 204) . . . . .	+ <sup>a</sup>	- <sup>a</sup>	+	0	*	*	*	*	*	*	0	*	+	+
Model 3 (N = 204) . . . . .	+ <sup>a</sup>	- <sup>a</sup>	+	0	*	*	*	*	*	*	*	*	+	+
Messner 1983a:														
Model 1 (N = 204) . . . . .	+ <sup>a</sup>	- <sup>a</sup>	+	(& SQ = - <sup>y</sup> )	0 (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	*	0	*	+	+
Model 2 (N = 204) . . . . .	+ <sup>a</sup>	- <sup>a</sup>	+	(& SQ = - <sup>y</sup> )	0 (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	*	0	*	+	+ <sup>b</sup>
Rosenfeld 1986:														
Model 1 (N = 125) . . . . .	+	*	*	*	*	*	*	*	*	*	+	0	+	+
Model 2 (N = 125) . . . . .	0	*	+	*	*	*	*	*	*	*	+	0	*	*
Simpson 1985:														
Model 1 (N = 125) . . . . .	0 <sup>d</sup>	*	+	(logit)	0 (wrm, logit)	+	*	*	*	*	+	*	*	0
Model 2 . . . . .	0 <sup>d</sup>	*	+	(logit)	0 (wrm, logit)	0 (logit)	*	*	*	*	+	*	*	+ <sup>b</sup>

TABLE 1 (Continued)

PREVIOUS STUDIES	STRUCTURAL COVARIATES										
	Population Size	Population Density	Percentage Black	Percentage Ages 15-29	Percentage Divorced	Percentage Kids Not with Both Parents	Median Family Income	Percentage Families in Poverty	Gini Index of Income Inequality	Percentage Unemployed	South
Williams 1984											
Model 1 (N = 125)	0 <sup>d</sup>	*	+ (& SQ = -) <sup>f</sup>	*	+	*	*	+	0	*	*
Model 2 (N = 125)	+ <sup>d</sup>	0 <sup>d</sup>	+ (& SQ = -) <sup>f</sup>	*	*	*	*	+	0	*	0
State level using 1960 data											
Gastil 1971 (N = 48)	0	*	+	+ (20-34)	*	*	-(adj)	*	*	*	+ <sup>b</sup>
Loflin and Hill 1974											
Model 1 (N = 48)	*	*	0 (nonwhite)	+ (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	+	*	0
Model 2 (N = 48)	*	*	0 (nonwhite)	+ (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	+	*	0 <sup>b</sup>
State level using 1970 data											
Huff-Corzine et al. 1986											
Model 1 (N = 48)	*	*	+ (non-white)	0 (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	0	*	+ <sup>b</sup>
Model 2 (N = 48)	*	*	+ (non-white)	0 (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	0	*	+ <sup>a,e</sup>
Parker and Smith 1979 (N = 48)											
Model 1 (N = 48)	*	*	0 (nonwhite)	0 (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	*	*	0
Model 2 (N = 48)	*	*	0 (nonwhite)	0 (20-34)	*	+ <sup>g</sup>	*	+ <sup>g</sup>	0	*	0 (non-south)

Notes — + denotes a statistically significant positive effect of the covariate indicated, - denotes a significant negative effect, 0 denotes a nonsignificant or null effect, \* indicates that the covariate was not included in the model. The dependent variable is in its metric form for these studies unless specified in App A. For notes on specific studies used in this table, see App A

- <sup>a</sup> The natural logarithm transformation was applied to the structural covariate
- <sup>b</sup> Gastil's (1971) Southernness Index was used
- <sup>c</sup> An alternative measure of relative deprivation was used
- <sup>d</sup> The log base 10 transformation was applied to the structural covariate
- <sup>e</sup> Percentage southern born was used
- <sup>f</sup> The squared value of the variable was also used
- <sup>g</sup> Loflin and Hill's Structural-Poverty Index was used

with both parents, median family income, percentage of families living below the official poverty line, the Gini index of family income inequality, the unemployment rate (percentage), and a variable indicating those cities, metropolitan areas, or states that are located in the South. Entries in the table denote whether each variable was included in each study (\* denotes not included) and, if so, whether its effect was statistically null (denoted "0") or significantly positive (+) or negative (-). To conserve manuscript space, we included details on the samples, units of analysis, and model specifications of each study as Appendix A (a comparative analysis of these research design specifications is given in the next subsection of the article). While we make no claim that the list of studies cited in table 1 exhausts the empirical literature on this topic, it does include most of the recent major statistical studies of U.S. data that have taken interunit variation in homicide rates as a problem to be explained.

It can be concluded from the empirical findings cataloged in table 1 that virtually no structural covariate exhibits estimated effects that are statistically significant and of invariant algebraic sign across all studies. The one exception to this statement is the percentage of children under age 18 not living with both parents, which appears as only one component in the structural poverty index used in a few studies. Otherwise, the empirical findings are quite variable. For instance, not only do population size, population density, and population heterogeneity as indexed by the percentage black (or white or nonwhite in some cases)—three components of the classical urbanization/control theory model—fail consistently to exhibit a significantly positive relationship to homicide rates, but they actually are found to have significantly negative effects in some cases.<sup>10</sup> Another component of the community-control theory model—family disintegration as indexed by the percentage divorced variable—fares a bit better in that it is either significantly positive or null. Again, however, this structural covariate appears in only a small number of studies. The age-structure index, taken as tapping "a factor with . . . 'inherent pathogenic qualities' " (Hirschi and Gottfredson 1985*b*, pp. 1332–33), also exhibits a disturbing pattern of positive, null, or negative effects across studies. The pattern of results for the unemployment rate is more often null than statistically significant. When it is significant, however, the negative algebraic sign is consistent with that deduced in criminal opportunity/routine-activities theory as opposed to strain or relative-deprivation theory.

<sup>10</sup> In the case of the percentage black or nonwhite variable, however, the algebraic signs almost always are significantly positive, as predicted by control theory.

How well do the critical variables of the major theoretical arguments cited in the previous subsection fare? Results for the various indices of the southern subculture-of-violence thesis—usually an indicator (dummy) variable for those cities, metropolitan areas, and states located in the South but also Gastil's southernness index and the percentage of the unit's population that is southern born—appear to be consistent with either a significantly positive or a null effect.<sup>11</sup> The same can be said for the Gini index of income inequality and, except for the significantly negative findings in Messner (1982), for the percentage of families living below the poverty line. When it is significant, median family income displays a negative algebraic sign, indicating that it moves in a direction opposite to that of the poverty and inequality indicators. But median family income also is insignificant in some studies.

Not much can be salvaged by limiting the foregoing comparisons to specific time periods (1960 or 1970; there are only two studies using 1980 data) or levels of analysis. Reversals of algebraic sign or levels of significance also appear frequently in these more restricted domains.

In brief, findings from empirical studies of the effects of structural covariates on the homicide rates of cities, metropolitan areas, and states summarized in table 1 exhibit little invariance. Some variables—unit population size, density, racial heterogeneity, and percentage young—exhibit relationships to homicide rates that can be positive, negative, or null. Most of the remaining variables in the table appear to be capable of either having a statistically significant relationship to homicide in the theoretically expected direction or not.

### Issues of Research Design and Statistical Inference

What inferences can be drawn from the information about structural covariates of homicide rates depicted in table 1? Must sociological and

<sup>11</sup> These three indexes of the South are highly intercorrelated. For instance, Loftin and Hill (1974) reported that Gastil's (1971) southernness index, defined for the 48 contiguous states, is correlated .68 with Hackney's (1969) dummy variable for states located in the South. Independently, we accumulated data on the percentage of the population of a unit that is southern born for all those cities, SMSAs, and states for which such data are available from the 1960, 1970, and 1980 decennial censuses and computed correlations with the South indicator/dummy variable. The correlations ranged from .885 to .912 for cities, from .925 to .939 for SMSAs, and from .981 to .983 for states. Apparently, there is not much to be gained from using any of these alternative indexes relative to the others. The indicator variable for the South is employed here not only because of these high correlations but also because of its use in most previous studies in this research tradition.

criminological theories of the structural covariates of homicide rates become more context specific, thereby foreclosing the possibility of developing a general theory that is applicable across time and social space? Or is it possible that many of the inconsistencies displayed in table 1 are the consequence of differences in research design, data analysis, or problems of statistical inference? Consider the following design characteristics and problems of inference in these studies.

*Time periods.*—All but two of the studies in table 1 use data from 1960 and 1970. The necessity for using data from a decennial census year is obvious, since only in these years are comparable data on structural covariates available for numerous subnational units of analysis. However, to date only two studies include 1980, despite the fact that data from the 1980 census have been available for several years.<sup>12</sup> Theoretically invariant relationships should hold across time, and the greater the number of time periods in which this is found to be the case, the more confidence one has in conclusions regarding invariance.

*Units of analysis.*—Studies of regional differences in homicide and, in particular, of the southern subculture-of-violence thesis in table 1 have tended to focus on states as the unit of analysis, partly because this is a theoretically reasonable choice for this topic and partly because data on structural covariates often are readily available and require less data file-building time at the state level. Other studies, however, disagree about which level of aggregation is most appropriate for theoretical concerns. Messner (1982), for example, argued that SMSAs are more appropriate aggregations than cities for some theoretical concerns because they are more likely to coincide with the boundaries of genuine communities. Bailey (1984) disagreed with Messner (1982), stating that aggregating to the SMSA level has the effect of ignoring substantial within-unit variability in homicide rates and sociodemographic characteristics and, therefore, cities constitute a more theoretically appropriate unit of analysis. City and SMSA boundaries, like those of states, generally are *arbitrary* with respect to social theory. A general theory of structural covariates of homicide should be capable of accommodating *all* these levels of analysis.<sup>13</sup>

<sup>12</sup> We are aware of one other study, Farley (1987), that uses 1980 census data on structural covariates to analyze SMSA crime rates. However, the structure of Farley's analyses—which are concerned with the effects of suburbanization on central-city crime rates—is quite different from those reviewed in table 1.

<sup>13</sup> As noted in n. 3, a number of studies of structural covariates of homicide rates (or violent crime rates including homicide) have examined other levels of analysis. A general theory of how structural covariates affect homicide rates also should be applicable at these levels, but attention to this is beyond the scope of the present paper.

*Samples.*—The studies in table 1 that employ state-level data usually include all 48 contiguous states, but studies based on city- and SMSA-level data tend not to include all available cases. In fact, the sample sizes tend to be relatively small, with few city-level studies based on more than 200 cases and few SMSA-level studies based on more than 125 cases. Most of the city- and SMSA-level studies also use data only on relatively large units, with minimum populations of 100,000 (for cities) and 250,000 (for SMSAs). Of course, population size tends to determine the number of cases, as there are only a limited number of large cities and SMSAs in any decennial census year. Only four studies (Jackson 1984; Messner 1982, 1983*a*, 1983*b*) are based on all cities or SMSAs for which data are available (in 1970). We find no rationale in the theories guiding research on the structural covariates of homicide rates for restricting sample sizes or constraining them only to large cities or SMSAs.

*Model specifications.*—As table 1 reveals, studies of structural covariates of homicide rates have exhibited many variations in model specification. Several analysts have noted nonlinearities in the bivariate relationship of the homicide rate to one or more structural covariates. The typical response to this is to apply a nonlinear transformation (logarithmic, polynomial, or logit) to one or more of the variables or to the entire regression function. Yet another major variation in model specification pertains to whether or not a study used the structural poverty index developed by Loftin and Hill (1974) or the southernness index of Gastil (1971) or the Blau and Blau (1982) SES inequality in race index. Different model specifications may account for some of the variance in empirical results noted in table 1.

*Problems of statistical inference.*—Each of the studies in table 1 employs multiple-regression analysis. Some of the studies use the more recent methodological innovations in regression analysis such as diagnostic plots and residual analysis (see, e.g., Cook and Weisberg 1982; Neter, Wasserman, and Kutner 1983), but most do not. As a consequence, the estimated regression coefficients are taken at face value.

Pitfalls in the application of multiple-regression analysis to area-based data on delinquency (and, by extension, to crime) were identified by Gordon (1967, 1968) over two decades ago. The most insidious of these pertain to the influence of multicollinearity on the substantive inferences that can be drawn from partialing in regression analysis. Collinearity among regressors is associated with (*a*) large changes in the estimated regression coefficients when a variable is added or deleted, or when an observation is altered or deleted; (*b*) wide confidence intervals, nonsignificant test statistics, and algebraic signs opposite to those expected from theoretical considerations or previous experience for important inde-

pendent variables; and (c) a corresponding instability of the regression-coefficient estimates from sample to sample. The presence of multicollinearity has been noted in several previous studies of homicide rates (Huff-Corzine, Corzine, and Moore 1986; Loftin and Hill 1974; Messner 1982; Smith and Parker 1980), although its possible influence on the pattern of inconsistent findings across studies (depicted in table 1) has to our knowledge not been discussed.

## A BASELINE MODEL

### Variables, Time Periods, Units, and Samples

In order to examine the extent to which the foregoing issues could account for the divergent empirical findings noted in table 1, we now describe the estimation of what we will refer to as a "baseline" regression model using the 11 explanatory variables identified in table 1 as regressors. As noted in the previous section, these variables have appeared in critical positions in the various theoretical arguments that have motivated recent studies of the structural covariates of homicide rates. Although not a comprehensive list, in the sense that previous studies often have used additional structural covariates or one or more variations on these, they will serve to facilitate cross-level and cross-time period comparisons.

In contrast to previous studies, which have tended to focus on one or at most two of the levels of analysis identified above, our baseline model is estimated at all three conventional levels: cities, SMSAs, and states. Our model also is estimated for 1980 as well as the 1960 and 1970 time periods, which have been the focus of most previous studies. Decennial census data are available for the application of standard definitions of all 11 variables for all three levels of analysis for all three time periods. The operational definitions we apply are identical to those used in many of the previous studies reviewed in table 1 (see App. B for details on definitions and data sources).

Whereas previous studies using city or SMSA data have been based on relatively small samples, ours is more inclusive. Our data file includes *all* cities and SMSAs identified by the Bureau of the Census in 1960, 1970, and 1980 for which decennial census data are available on all 11 explanatory variables in our baseline model and for which the corresponding Federal Bureau of Investigation's *Uniform Crime Report* contains a homicide rate. Thus, our sample comprises 528 cities in 1960, 729 in 1970, and 904 in 1980, while the corresponding numbers for SMSAs are 182, 187, and 259. In brief, the sample *Ns* on which our city and SMSA analyses are based are substantially larger than those of most previous

studies.<sup>14</sup> For estimation at the state level, we use data on *all* 50 states for all three time periods.<sup>15</sup>

### Nonlinearities and Data Transformations

There has been considerable discussion over the presence of nonlinearities in the relationships between structural covariates and homicide rates, especially insofar as they were present but undetected in the Blau and Blau (1982) study of 1970 SMSA data (see, e.g., Messner 1983*a*; Simpson 1985; Williams 1984; Blau and Golden 1986). Sometimes this has resulted in logarithmic, logit, or polynomial transformations of specific explanatory variables or regressors (structural covariates) or the dependent variable or regressand (homicide rate), and sometimes all variables have been logarithmically transformed. Often the statistical case for the transformation of specific variables is not stated, and sometimes the stated rationale is inappropriate.<sup>16</sup>

The contemporary view in statistics is that transformations of the explanatory variables or the dependent variable to “linearize” the relationships are appropriately applied before carrying out a regression analysis (see, e.g., Neter et al. 1983, p. 134). Examinations of univariate and bivariate data plots often are useful in suggesting an appropriate transformation.

<sup>14</sup> The total number of SMSAs in our sample for 1970 is 187, which is smaller than the 204 SMSAs in Messner’s (1982, 1983*a*) analyses because, in order to preserve continuity in the definition of SMSA boundaries across time periods, we applied the definitions used by the Bureau of the Census in 1980 consistently to 1970 and 1960 as well. This results in the combination of data on some SMSAs identified in 1960 and 1970 that were no longer distinct in 1980 (e.g., the Dallas and Fort Worth SMSAs are combined into a Dallas-Fort Worth SMSA). We also used the 1980 specifications of the New England Consolidated Metropolitan Areas as given in the *Uniform Crime Reports*—which generally are the same as those of the Bureau of the Census. In the construction of our cities files, we could not include five cases/jurisdictions for which homicide rates are reported in the *Uniform Crime Reports* because of unclarities about the corresponding census units.

<sup>15</sup> It can be argued that Alaska and Hawaii are sufficiently distinct from the 48 contiguous states that they should be excluded from the state-level analyses, as previous state-level studies have done. We find, however, that these two states actually are not statistical outliers on any of the variables in our baseline model and do not exclude them.

<sup>16</sup> For example, in justifying the application of logarithmic transformations to certain variables in his regressions, Williams (1984, pp. 286–87) states that logarithmic models consistently yielded “more efficient” parameter estimates than metric or polynomial alternatives. Although the sense in which Williams uses the term “efficient” in this context is not clear, the conventional application of the term “efficiency” in statistics (e.g., the efficiency of ordinary least squares relative to other linear estimators in the general linear-regression model) has nothing to do with data transformations.



In the absence of theoretical or a priori considerations to guide the selection of a transformation, some bivariate curvilinear data patterns can be linearized by transforming either the dependent or the explanatory variables, or both. Hence, the question immediately arises about which variables to transform. Current statistical guidelines (Neter et al. 1983, p. 138) suggest that transformation of the dependent variable is appropriate to reduce either skew or unequal variability at different levels of the independent variables (heteroscedasticity); otherwise, transformations of explanatory variables involved in curvilinear patterns may be more appropriate.

In the present case, we began, therefore, with an examination of the sample marginal frequency distributions of our dependent variable (the homicide rate) at each of three time periods for each level of analysis. For our largest units of analysis—states and SMSAs—these distributions are bell shaped and display little skew. Hence, unless evidence is found for heteroscedasticity in specific relationships, there is no apparent rationale for transforming the homicide rate distribution in the state and SMSA analyses. The frequency distributions of the city-level homicide data also are bell shaped and nonskewed, except for the presence of a substantial density at the zero homicide-rate level.<sup>17</sup> Since no simple transformations exist for such frequency distributions, and since the relatively larger sizes of our city samples will allow the Central Limit Theorem to come into play and induce approximately normal sampling distributions of regression coefficients for city-level models (which is what is really important for statistical inference in regression models [see, e.g., Hanushek and Jackson 1977, pp. 60–65]), we similarly find no compelling reason to transform the homicide rate at the city level. Nonetheless, in order to ensure that this decision does not affect substantive inferences, we examine the residuals of our regression analyses for possible heteroscedasticity problems created by the zero homicide-rate cases and report below any substantive differences created by corresponding data adjustments.

In the absence of substantial skew in the dependent variables of our analyses, we focus our attention on transformations of the independent variables to induce linearity in relationships. On examination of bivariate scatter plots of each of these, with the homicide rates for each level of

<sup>17</sup> This lumping at the zero homicide-rate level results from our cities samples' not being restricted to arbitrarily large cities. Because homicide, even at its relatively high U.S. levels in recent decades, still is a relatively rare event, a number of smaller cities may have zero homicide rates in any given year. Similarly, the absence of substantial skew in the frequency distributions of homicide rates in our city and SMSA samples, in contrast to what other studies have found (e.g., Williams 1984; Williams and Flewelling 1988), appears to be due to our samples' including smaller as well as larger SMSAs.

analysis and time period, we found nonlinearities in relationships only for unit population size, population density, and median family income. Relationships between homicide rate and these variables were suitably linearized by applying logarithmic transformations to the latter. Note that the finding of logarithmic relationships of the homicide rate with population size and density is consistent, for example, with what Messner (1982) and Bailey (1984) found with SMSA- and city-level data.

On the other hand, the absence of a nonlinear relationship of homicide rates with the percentage of the population that is black is inconsistent with Messner (1983*a*), Williams (1984), Simpson (1985), and Blau and Golden (1986) for SMSAs but consistent with Bailey (1984), Huff-Corzine et al. (1986), and others for analyses of city- and state-level data. Similarly, the absence of a nonlinear relationship with the poverty rate is inconsistent with Williams's (1984) findings for 1970 SMSA data. Since our samples generally are more inclusive than these earlier studies and extend across three decennial time periods, we infer that these apparent inconsistencies with earlier studies are due to the happenstance of the particular time periods, samples, and levels of analysis available to other researchers. Finally, we find no data-based rationale for logarithmic transformations of all dependent and independent variables.<sup>18</sup> Indeed, our analyses would suggest that the application of logarithmic transformations to relationships that already are linear would induce artifactual nonlinearities.

### Model Estimation

On the basis of the foregoing preliminary data analysis, regression models were estimated with the homicide rate as dependent variable and the 11 structural covariates identified in table 1 as explanatory variables for each of three time periods and levels of analysis. All variables are in metric form except the three identified in the preceding subsection (population size, population density, and median family income); natural logarithmic transformations were applied to transform them. Table 2 reports the resulting estimated metric regression coefficients, standardized regression coefficients, and *t*-ratios.<sup>19</sup> Coefficients of determination, adjusted for

<sup>18</sup> When a study is restricted to a sample only of large cities or SMSAs (e.g., Blau and Golden 1986; Williams and Flewelling 1988), however, the sample distribution of homicide rates is more likely to be skewed. Under such circumstances, there may be a statistical rationale for logarithmic transformation of the homicide rate as well as one or more regressors.

<sup>19</sup> Note that specific directional hypotheses have been derived earlier in the text for all explanatory variables except the unemployment rate, for which theory predicts contradictory directional relationships that may cancel out. Hence, one-tailed critical

degrees of freedom (adjusted  $R^2$ ) and sample sizes ( $N$ s) also are shown at the bottom of the table.

Do these empirical results—estimated for standardized definitions of variables applied to the complete samples of each level of analysis—exhibit any more consistency and invariance than those summarized in table 1? The answer, at least to some extent, is yes. Two structural covariates—the percentage of the population that is black and the percentage of the male population that is divorced—have regression coefficients that show algebraic signs that are consistently significant (.05 level) statistically in the theoretically expected direction (positive) for two levels of analysis (cities and SMSAs) across all three time periods. Each of these also has positive but not always statistically significant estimated coefficients at the state level in all three time periods. Two additional variables—unit population size and the indicator (dummy) variable for location in the South (coded “1” if in the South, “0” otherwise)—are always statistically significant and in the theoretically expected direction (positive) for one level of analysis (cities) across all three time periods and have positive but not always significant coefficients across all periods for the two other levels. While these results are somewhat more consistent across time and social space than those from the studies in table 1, they are not highly invariant. Furthermore, the fact remains that estimated relationships for six of the structural covariates in table 2 exhibit little consistency (in terms of expected algebraic signs and statistical significance) across time or level of analysis.

Is this an accurate assessment of the relationships in the data, or are there still some research-design and inference problems in the regressions underlying table 2? The failure of the percentage of families living in poverty to reach the tolerance level in the computer program used to estimate the regressions for states in 1960 and 1980 suggests the latter. This implies that less than 1% of the sample variance of this structural covariate is not accounted for by the other explanatory variables in our baseline model. This, of course, is another way of saying that the percentage poverty variable is highly collinear with the other structural covariates.

While space limitations preclude exhibiting the sample correlation matrices for all nine combinations of time periods and levels of analysis contained in table 2, suffice it to say that inspection of these suggests that collinearity among the structural covariates is a substantial problem that cannot be ignored at *all* levels of analysis and time periods. In each correlation matrix, there are several structural covariates with intercorre-

---

values are applied to evaluate the statistical significance of the  $t$ -ratios in table 2 for all variables except the unemployment rate.

TABLE 2

METRIC REGRESSION COEFFICIENTS, STANDARDIZED REGRESSION COEFFICIENTS (in Brackets), AND *t*-RATIOS (in Parentheses) OF BASELINE STRUCTURAL COVARIATES OF HOMICIDE RATES FOR U.S. CITIES, SMSAs, AND STATES: 1960, 1970, AND 1980

STRUCTURAL COVARIATE	1960			1970			1980		
	City	SMSA	State	City	SMSA	State	City	SMSA	State
Constant . . . . .	-14.636 (.613)	32.723 (.586)	-68.101 (3.620)**	-52.403 (1.998)*	25.956 (.465)	-184.428 (2.234)*	-74.318 (2.092)*	-21.116 (.390)	-39.160 (1.908)*
Population size <sup>a</sup> . . . . .	.506 [.092] (2.616)**	.318 [.077] (.993)	.349 [.104] (1.449)	1.641 [.176] (6.877)**	1.185 [.219] (3.186)**	.770 [.214] (2.072)*	2.451 [.199] (8.412)**	1.303 [.231] (3.906)**	1.453 [.331] (3.569)**
Population den- sity <sup>a</sup> . . . . .	-.225 [-.032] (.913)	-.736 [-.191] (2.449) <sup>+</sup>	-.880 [-.376] (3.015) <sup>+</sup>	.024 [.002] (.080)	-.869 [-.157] (2.209) <sup>+</sup>	-1.038 [-.407] (2.345) <sup>+</sup>	.413 [.031] (1.208)	-.240 [-.040] (.587)	-.748 [.244] (1.398)
Percentage black . . . . .	.201 [.517] (8.691)**	.159 [.386] (3.535)**	.132 [.383] (3.106)**	.302 [.515] (11.091)**	.197 [.339] (3.783)**	.145 [.350] (2.056)*	.289 [.466] (10.452)**	.276 [.481] (4.823)**	.034 [.069] (.346)
Percentage ages 15-29 . . . . .	-.031 [-.030] (.916)	-.070 [-.047] (.889)	.181 [.084] (1.041)	.021 [.018] (.752)	-.091 [-.666] (1.367)	.024 [.009] (.094)	-.130 [-.094] (4.186) <sup>+</sup>	-.105 [-.078] (1.766) <sup>+</sup>	.185 [.058] (.553)
Percentage di- vorced . . . . .	.684 [.154] (4.177)**	.805 [.162] (2.784)**	.207 [.052] (.520)	.965 [.159] (5.137)**	.701 [.108] (1.949)*	.250 [.059] (.460)	.827 [.158] (5.606)**	1.166 [.261] (4.481)**	.992 [.294] (2.211)*
Percentage kids not with both parents . . . . .	-.058 [-.073] (.890)	.113 [.113] (.823)	.358 [.400] (2.306)*	.109 [.107] (1.570)	.462 [.374] (3.440)**	.370 [.388] (1.942)*	-.051 [-.054] (.839)	-.127 [-.117] (1.059)	.617 [.613] (2.745)**

Median family income <sup>a</sup> . . . . .	1.451 [.062] (.541)	-4.452 [-.165] (-.677)	6.396 [.362] (2.976)**	3.469 [.094] (1.244)	-3.681 [-.105] (.595)	18.224 [.790] (1.878)*	4.218 [.101] (1.194)	-1.622 [-.040] (.292)	-200 [-.008] (.072)
Percentage families below poverty level . . . . .	.109 [.193] (1.281)	-.252 [-.455] (1.272)	. . . . <sup>b</sup>	.166 [.136] (1.428)	-.324 [-.319] (1.593)	.357 [.548] (1.083)	.414 [.306] (2.864)**	-.073 [-.056] (.328)	. . . . <sup>b</sup>
Gini index . . . . .	-3.909 [-.046] (.803)	36.283 [.289] (1.556)	18.355 [.192] (1.370)	-5.857 [-.039] (.981)	2.277 [.013] (.134)	13.187 [.083] (.394)	.893 [.005] (.137)	78.004 [.354] (4.264)**	23.903 [.105] (.871)
Unemployment rate . . . . .	-.202 [-.762] (1.975)*	-.140 [-.052] (.965)	-.098 [-.041] (.729)	-.095 [-.020] (.674)	-.483 [-.128] (2.216)*	-.313 [-.103] (.963)	.130 [.038] (1.189)	-.031 [-.012] (.215)	-.300 [-.104] (1.489)
South . . . . .	2.156 [.204] (3.951)**	3.067 [.360] (4.390)**	1.977 [.259] (2.689)**	1.606 [.092] (2.721)**	2.792 [.257] (3.143)**	1.200 [.148] (1.069)	1.202 [.055] (1.802)*	1.082 [.095] (1.239)	.471 [.050] (.375)
Adjusted R <sup>2</sup> . . . . .	.507	.630	.896	.628	.656	.772	.584	.595	.792
N . . . . .	528	182	50	729	187	50	904	259	50

<sup>a</sup> Variables to which a natural logarithmic transformation have been applied; all other variables are in metric form.

<sup>b</sup> Variables exceeding tolerance limits and therefore not entered into the regression model.

\*  $P \leq .05$  (one-tailed).

\*\*  $P \leq .01$  (one-tailed).

+  $P \leq .05$  (two-tailed).

lations of greater magnitude than their respective correlations with the homicide rate, thus violating Klein's rule of thumb for judging when collinearity is harmful (see Klein 1962, p. 101; Maddala 1977, p. 185).<sup>20</sup> This, of course, creates circumstances under which all the statistical estimation problems associated with multicollinearity cited earlier are likely to be present to some extent. In addition, violations of Klein's rule may lead to what Gordon (1968) called the *partialing fallacy*, namely, the allocation of all explained variance to that particular regressor (among a highly intercorrelated set of regressors) that possesses the (possibly very slightly) higher correlation with the dependent variable—even though there may be no theoretical or substantive basis for doing so. As noted earlier in this article, several analysts have observed and commented on high multicollinearity among the structural covariates of homicide rates; few have done anything in response. The application of ridge-regression techniques to reduce sampling variances by introducing small amounts of bias into the ordinary-least-squares estimators by Huff-Corzine et al. (1986) is an exception. However, the interpretation of ridge-regression estimates of regression coefficients is unclear (Johnston 1984, p. 252). Furthermore, in the presence of a large number of explanatory variables, as in our baseline model, it may be more parsimonious to examine the variable space spanned by the regressors for redundancies and possible simplification.<sup>21</sup> This is the approach we pursue.

## A REVISED MODEL

### Reduction of the Structural Covariate Space

Analyses of the dimensionality (by principal components analysis) of the structural covariate spaces spanned by the nine sample correlation matrices (excluding the indicator variable for location in the South) in our data files revealed two clusters of variables that consistently hang together over all levels of analysis and time periods.<sup>22</sup> The first is what we will call

<sup>20</sup> There is little agreement on the meaningfulness of the statistical tests for multicollinearity that have appeared in the literature (see, e.g., Maddala 1977, p. 186).

<sup>21</sup> There are long-standing and relatively independent traditions of research in the factorial classification of cities (e.g., Berry 1972; Janson 1980) and in the factorial study of urban correlates of crime rates (e.g., Schuessler and Slatin 1964; Wilks 1967; Flango and Sherbenou 1976) based on components/clustering analysis. However, studies of the structural covariates of homicide rates seem to have developed relatively independently of these literatures, perhaps because studies of homicide have focused on SMSA and state levels of analysis as well as cities.

<sup>22</sup> It should be emphasized that we applied principal components analysis to the structural covariates of our regressor space to ascertain whether or not this would give some insight into how many independent sources of variation there really are in this space. Maddala (1977, pp. 193–94) states that if the number of components is small relative

a *population structure component*, consisting of the unit population size and population density variables (both in natural log form). This, of course, is another way of saying that units with greater population sizes also tend to have greater population density. While this proposition may appear obvious, it is by no means clear a priori that this relationship must be sufficiently strong to define a dimensional component consistently across all levels of analysis and time periods.

The second dimensional component includes three income variables (median family income, the percentage of families living below the official poverty line, and the Gini index of family income inequality) and two other social indicators (the percentage of the unit population that is black and the percentage of children age 18 or under not living with both parents). The last social indicator appears in Loftin and Hill's (1974) structural poverty index, which also contains a low-income index similar to the poverty rate. But the present dimensional component also contains the median family income variable, which can indicate either the poverty or affluence of a city, SMSA, or state. However, all the variables in the present component pertain to the composition of the human resources within a given population. Therefore, using terminology similar to that of Williams and Flewelling (1988), we will refer to this as a *resource-deprivation/affluence component*.

Note that this component contains both the usual structural covariate taken as indicative of absolute deprivation (percentage of families living below the official poverty line) and the usual structural covariate taken as

---

to the number of regressors and if the components can be given a substantive interpretation, then this can be very useful in solving the collinearity problem. Note that the application of principal components analysis in this way and the construction of corresponding component scores should be sharply distinguished from the specification of a latent-variable/factor-analytic/covariance-structure model (on the distinction, see Kim and Mueller 1978). In the analysis described in the text, we are not asserting the existence of latent variables that are characteristic of cities, SMSAs, and states and that account for the high intercorrelations among certain structural covariates, as extant theory simply is not sufficiently strong to support such an assertion. Rather, we are stating only that the empirical correlations of our structural covariates are sufficiently similar across time and levels of analysis so that component scores that simplify the geometric space spanned by these covariates can be consistently defined and so that these component scores can be substantively related to recent developments in social theory. Operationally, our components analysis proceeded in the usual manner (Kim and Mueller 1978). Only two components accounting for substantial variance in the regressor space and having substantial component loadings for two or more regressors were identifiable. As a rule for defining a component score, we included only those variables that possessed component loadings of .5 or greater across levels of analysis and time periods. This relatively conservative criterion resulted in a parsimonious number of meaningful component scores (described in the text). Their success in simplifying the regressor space is corroborated by the small correlations remaining among the regressors of our revised model (see n. 28).

indicative of relative deprivation (the Gini index of family income inequality). In other words, while these covariates may be conceptually and operationally distinct, component analyses of the variable spaces they span suggest that the variables are not very distinct empirically; cities, SMSAs, and states with high (low) levels of absolute deprivation also tend to have high (low) levels of relative deprivation. This collinearity and corresponding lack of empirical identifiability may account for some of the inconsistencies regarding the influences of absolute and relative deprivation on homicide rates reported previously in the literature and summarized earlier.<sup>23</sup>

Similar comments about the lack of empirical identifiability of substantial independent variance apply also to the variables “percentage black population” and “percentage of children not living with both parents.” In brief, these indicators of racial heterogeneity and family living arrangements are substantially collinear with the economic variables in the resource-deprivation/affluence component. The invariance of this fact across time periods and levels of analysis merits comment and theoretical interpretation. The racial composition of a city, metropolitan area, or state may reflect several economic and social forces, and the percentage of children not living with both parents is most directly interpretable in terms of capacity for supervision and guardianship.

Wilson’s (1987) idea of *concentration effects* may help explain why these social indicators—while conceptually distinct from each other and from community economic structure—are invariably associated with the economic deprivation indicators. Wilson (1987, pp. 46–62) argues that the social transformation of the inner city in recent years (through segregation, selective out-migration by the middle class, and racial discrimination) has resulted in a disproportionate and criminogenic concentration of the disadvantaged. For example, opposition from organized community groups to public housing projects and the decisions of governments to neglect rehabilitation of existing single-family housing have led to the

<sup>23</sup> Our structural covariates do not include the SES inequality in race index constructed by Blau and Blau (1982) for the 125 largest SMSAs in 1970 and used in several studies in table 1. This is because we have limited our analyses to those covariates for which data are available across all three levels of analysis and three time periods studied herein. On the basis of a comparison of the correlation pattern of the variables in our resource-deprivation/affluence index for our larger SMSA sample in 1970 with the correlation pattern of the same variables reported in Blau and Blau (1982), however, it is plausible to hypothesize that their SES inequality in race variable would cluster into a component along with the other variables in our index. This implies that the SES inequality in race index, while conceptually distinct, is not highly distinct empirically from other deprivation/affluence components. Thus, attempts to estimate its distinct effects on homicide rates empirically may suffer from the same sensitivity to research design and collinearity found for our baseline model in table 2.



massive segregation of housing projects. These projects have become ghettos for the disadvantaged—especially poverty-stricken black households headed by females. The consequences for racial differences in living conditions are striking; according to Wilson's (1987) calculations, in 1980 less than 10% of urban poor whites lived in poverty areas, but 40% of poor blacks did. In addition to sheer economic deprivation, the corresponding decreasing vertical integration of middle- and lower-class black families may thus contribute—through a process Wilson (1987, p. 61) calls *social isolation*—to the disintegration of community social control institutions and the supervision capacity of adults.

Wilson's (1987) perspective is consistent with our empirical results for the resource component. Indeed, our results go beyond Wilson's (1987) by suggesting that the clustering of economic and social indicators documented in his analyses appears not only in 1980 and not only for neighborhoods of large cities but also for the two previous decennial periods and at the levels of cities as a whole, metropolitan areas, and states. That is, those cities, metropolitan areas, and states that have low median family incomes, large absolute poverty levels, and great relative economic inequality in 1960, 1970, and 1980 also tend to have large concentrations of blacks and children living in broken families. In addition, however, evidence corroborating Wilson's arguments that concentration effects grew more severe from 1970 to 1980 in large cities is found in our city-level data: the numerical values of the component loadings of percentage poverty, percentage black, and percentage of children under age 18 not living with both parents are larger in 1980 than in 1970.

None of the three other explanatory variables in our baseline regression model (percentage of males divorced, percentage of the population ages 15–29, the unemployment rate) exhibit a consistent pattern of clustering among themselves or with the variables in the foregoing components across levels of analysis or time periods.<sup>24</sup> In other words, each of these not only is conceptually distinct but also exhibits a pattern of empirical independence from the other structural covariates.<sup>25</sup>

<sup>24</sup> Additional principal components analyses at the SMSA level show that the percentage-divorced-males variable consistently clusters with the natural-log-transformed absolute value of the net migration rate for the preceding 10 years. But the clustering of these two variables into a component is not consistent across time at the state level, and data on net migration are not available at the city level. Nonetheless, this result can assist in giving a substantive sociological interpretation to the percentage-divorced variable, the interpretation of which has not been consistent in previous studies. In brief, with use of terminology similar to Williams and Flewelling's (1988), the percentage-divorced variable can be taken as indicative of *social disintegration* (i.e., the breakdown of family and social network relationships).

<sup>25</sup> Using 1970 data on 59 social, demographic, economic, and ecological variables defined on a sample of 840 U.S. cities, Flango and Sherbenou (1976) extracted six

### Model Estimation

On the basis of the foregoing component analyses, component scores, or indexes, were calculated in the conventional manner as weighted sums of the component variables, with weights equal to their component loadings (Kim and Mueller 1978, p. 72).<sup>26</sup> Table 3 reports the results of regressing the homicide rate on component scores for the population structure and resource-deprivation/affluence components together with the percentage divorced, percentage ages 15–29, unemployment rate, and South indicator variables for all nine combinations of levels of analysis and time periods studied herein. These ordinary-least-squares estimates were computed after we eliminated a few cases of extreme residual outliers in the city and SMSA analyses identified in an initial regression (there were no outliers in the state-level analyses).<sup>27</sup>

Adjusted  $R^2$  statistics at the bottom of the table indicate that this simple six-variable regression model can account for 50%–60% of the variance of the homicide rate at the city level, 55%–60% at the SMSA level, and 65%–85% at the state level. With the use of conventional structural covariates, these percentages of explained variance seem about typical of what can be expected in national samples without increased

---

principal components that are similar to those of Berry (1972) and that they related to various crime rates (robbery, aggravated assault, burglary, and auto theft, but not homicide). One of their components (which they label urbanization) is similar to our population-structure component, whereas our resource-deprivation/affluence component appears to combine elements of two of their components (which they call urban affluence and culture of poverty). The composition of components is, of course, always conditional, in part, on the number and substance of the variables entered into an analysis. Accordingly, it is possible that the inclusion of data on additional structural covariates would bring our components more into line with those of Flango and Sherbenou (1976). However, the variables dealt with here are limited to those that have appeared in one or more of the studies summarized in table 1 and that are available at all three levels of analysis and three time periods studied.

<sup>26</sup> The component scores are used in metric regression analyses with other nonstandardized variables. Therefore, we did not scale the component loadings by their corresponding eigenvalues.

<sup>27</sup> Because each of the deleted cities and SMSAs corresponds to a unique local situation not covered by the regression model, and because our samples are sufficiently large so that the loss in degrees of freedom is not significant, modern statistical practice suggests that they can be discarded (Neter et al. 1983, p. 409). The cities deleted are Gulfport, Mississippi, and Greenville, South Carolina, in 1960; Anniston, Alabama; Atlanta, Georgia; Augusta, Georgia; East St. Louis, Illinois; and Greenville, South Carolina, in 1970; and Compton, California; Inglewood, California; Miami, Florida; Gary, Indiana; Highland Park, Michigan; and St. Louis, Missouri, in 1980. The SMSAs deleted are Charlotte, North Carolina, in 1970 (there were no SMSA outliers in 1960) and Miami, Florida, and Houston, Texas, in 1980. Each of these units merits intensive study of the unique factors responsible for its extremely high homicide rate (e.g., Lundsgaarde 1977).

collinearity among regressors and a corresponding increase in sample variability of the regression coefficients, as illustrated in table 2 above.

Note that the larger coefficients of determination at the state level, coupled with a failure of the regression coefficients for the same variables to increase systematically when compared with the city and SMSA levels, suggest that the state-level aggregation tends to reduce the relative importance of neglected variables in the error or disturbance terms of the regressions (Blalock 1964, pp. 97–114; Theil 1971, p. 181). But this gain is counterbalanced by the smaller sample sizes and correspondingly larger standard errors at the state level.

Substantively, the results reported in table 3 are much more readily interpretable and consistent than those summarized or reviewed earlier.<sup>28</sup> Specifically, three structural indexes/covariates—population structure, resource deprivation/affluence, and percentage of the male population divorced—now exhibit statistically significant relationships to the homicide rate in the theoretically expected positive direction across all time periods and levels of analysis.<sup>29</sup>

The relationships of the homicide rate to the other three variables in the model of table 3 are somewhat more complicated and less invariant. The percentage of the population ages 15–29 consistently has a significant relationship in the expected positive direction at the state level but a null effect at the SMSA level and works in the opposite direction at the city level in the most recent time period. For 1960 and 1970, the regression coefficients for this structural variable increase roughly in proportion to the increase in the coefficient of determination. This is consistent with a hypothesis that the positive state-level relationship is spuriously due to the coincidence of proximity groupings (i.e., by states) with groupings on the dependent variable (Blalock 1964, pp. 97–114). But this does not appear to be the case for 1980. Moreover, the presence of a significant negative relationship of homicide to the percentage ages 15–29 at the city level in 1980 is inconsistent with extant theoretical arguments about the relationship of age structure to homicide rates.

By comparison, the unemployment rate shows a statistically significant negative effect, as postulated in criminal opportunity theory, at the city and SMSA levels across most time periods but has a null effect at the state

<sup>28</sup> The consistency and interpretability of the estimates in table 3 are very much ascribable to the success of our components analysis in reducing levels of collinearity among the regressor variables. For cities and SMSAs, most correlations among the regressors are less than .2 in absolute value, whereas for states the corresponding level is .4.

<sup>29</sup> The statistically insignificant effect of the population-structure index at the SMSA level in 1960 becomes significant under the square-root transformation to reduce heteroscedasticity in the homicide residuals described later.

TABLE 3

METRIC REGRESSION COEFFICIENTS, STANDARDIZED REGRESSION COEFFICIENTS (in Brackets), AND *t*-RATIOS (in Parentheses)  
OF STRUCTURAL INDEXES/COVARIATES OF HOMICIDE RATES FOR U.S. CITIES, SMSAs, AND STATES: 1960, 1970, AND 1980

STRUCTURAL INDEX/ COVARIATE	1960			1970			1980		
	City	SMSA	State	City	SMSA	State	City	SMSA	State
Constant . . . . .	-6.070 (2.586)**	-1.784 (.537)	-19.872 (4.334)**	-22.877 (7.801)**	-14.415 (3.831)**	-23.098 (3.325)**	-26.738 (7.413)**	-13.404 (3.423)**	-35.808 (3.368)**
Population struc- ture . . . . .	.566 [.107] (3.280)**	.124 [.045] (.869)	.291 [.154] (1.912)*	1.843 [.200] (7.577)**	1.028 [.247] (4.899)**	.385 [.194] (2.009)*	3.085 [.192] (7.747)**	.685 [.183] (3.842)**	.767 [.267] (2.660)**
Resource depri- vation/ affluence . .	.116 [.555] (10.924)**	.120 [.446] (5.530)**	.107 [.598] (5.346)**	.195 [.583] (17.337)**	.273 [.571] (7.045)**	.206 [.773] (5.160)**	.178 [.528] (15.613)**	.242 [.545] (8.423)**	.283 [.802] (6.124)**

Percentage divorced . . . . .	.425 [.101] (3.094)**	1.184 [.239] (4.878)**	1.444 [.364] (4.645)**	.925 [.176] (6.298)**	1.747 [.278] (5.582)**	1.331 [.316] (3.166)**	.558 [.121] (5.027)**	1.113 [.263] (6.168)**	1.970 [.584] (7.012)**
Percentage ages 15-29 . . . . .	-.057 [-.059] (1.872)	-.020 [-.013] (.243)	.689 [.321] (4.212)**	.003 [.003] (.115)	.016 [.012] (.234)	.625 [.239] (2.704)**	-.129 [-.107] (4.699)+	-.040 [-.032] (.703)	.549 [.171] (1.883)*
Unemployment rate . . . . .	-.270 [-.108] (3.074)**	-.324 [-.121] (2.398)**	-.027 [-.011] (.161)	-.368 [-.088] (3.100)**	-.838 [-.227] (4.090)**	-.043 [-.014] (.135)	.057 [.185] (6.654)	-.214 [-.086] (1.735)*	-.257 [-.089] (1.118)
South . . . . .	2.007 [.201] (4.241)**	3.067 [.359] (4.405)**	1.842 [.242] (2.193)*	1.942 [.128] (3.983)**	1.510 [.143] (1.695)*	.002 [.000] (.002)	3.149 [.165] (5.778)**	1.512 [.141] (2.060)*	-.786 [-.083] (.664)
Adjusted R <sup>2</sup> . . . . .	.500	.591	.832	.594	.587	.680	.547	.546	.722
N . . . . .	526	182	50	722	186	50	896	257	50

\*  $P \leq .05$  (one-tailed).  
 \*\*  $P \leq .01$  (one-tailed).  
 +  $P \leq .05$  (two-tailed).

level in all time periods. Finally, the South regional indicator variable is statistically significant consistently across all time periods for cities and SMSAs but is significant for states only in 1960.

### Heteroscedasticity Analyses

We noted earlier that the sample frequency distributions of the homicide rate for our city-level samples contain substantial density at the zero level. The effects of this were revealed in a slight heteroscedasticity of the residuals in the estimated regression models of table 3. Since the variance of the residuals appears to be proportional to the expected value of the homicide rate, the regression models for cities were reestimated after a square-root transformation of the homicide rate (Neter et al. 1983, p. 138). As might be expected given the relatively large sample sizes of the city samples and the associated effects of the Central Limit Theorem on the sampling distributions of the regression coefficients (Hanushek and Jackson 1977, pp. 60–68), this produced no changes in statistical significance levels and substantive inferences.

Because of similar evidence of a low level of heteroscedasticity in residuals at the SMSA level (due to a tendency of SMSAs to have somewhat larger homicide rates than expected by the regression models), we also introduced a square-root transformation of the homicide rate and reestimated the SMSA models. The main effect of this was to reduce to statistical insignificance some effects that were marginally significant in table 3. Specifically, after transformation, the effect of the unemployment rate falls to insignificance in 1980, the effect of South becomes insignificant in 1970, and the unemployment rate falls to just below the .05 level while population structure becomes strongly significant in 1960. For the most part, these changes do not affect the major implications of the results reported in table 3, as will be elaborated in the discussion section below.

### Replication on 1950 Data

One limitation of the model described above is that it was formulated and estimated on the basis of all available data on the city, metropolitan area, and state levels in the three most recent decennial census periods. A question can be posed, therefore, about the extent to which the results reported above are due to overfitting of available data. If this is the case, it is possible that the relationships we have estimated are not invariant across time periods.

Data from the 1990 decennial census that would allow this possibility to be examined will not be available for several years. A limited repli-

cation analysis, nonetheless, can be performed with 1950 data. Specifically, data on all the explanatory variables we have been dealing with except the percentage of children under age 18 not living with both parents are available for 434 cities of 25,000 population and over and all 48 states in 1950.<sup>30</sup> A principal components analysis revealed that the four other structural covariates in our resource-deprivation/affluence index again form a dimensional component in 1950 at both the city and state levels. The clustering of unit population size and density in a population-structure component also is apparent but is weaker at the state than the city level.

Table 4 displays the results of regressing homicide rates on the corresponding structural indexes/covariates. It can be seen that the percentages of variance in homicide explained by the models are of the same order of magnitude as for 1960, 1970, and 1980. The statistically significant explanatory variables also are similar. At the city level, population structure, resource deprivation/affluence, the unemployment rate, and location in the South again have significant effects, but the percentage divorced males falls to statistical insignificance. At the state level, resource deprivation/affluence, percentage divorced, and percentage of the population ages 15–29 again are statistically significant, but population structure (see remark in preceding paragraph) and location in the South are not.

## DISCUSSION AND CONCLUSION

The empirical results described above have a number of implications for comparative research on the structural covariates of homicide rates. Perhaps the most important of these is that the theoretical presumption of relatively invariant relationships across time periods and social space appears to be corroborated by use of large samples, standard definitions, and the reduction of collinearity among the structural covariates. Much of the inconsistency apparent in earlier studies must be ascribed to collinearity. Future studies on this topic should not proceed without recognizing and accommodating this complication of inference.

Second, among our six structural indexes/covariates, three effects stand out in strength and invariance. By far, the strongest and most invariant effect is due to the resource-deprivation/affluence index; consistently across the four decennial census periods, cities, metropolitan areas, or states that are more deprived have higher homicide rates, and those that are more affluent have lower rates. Our population-structure index and

<sup>30</sup> In 1950, Alaska and Hawaii were not yet states, and homicide rates for SMSAs are not available for that decennial census year.

TABLE 4

METRIC REGRESSION COEFFICIENTS, STANDARDIZED REGRESSION COEFFICIENTS  
(in Brackets), AND *t*-RATIOS (in Parentheses) OF STRUCTURAL  
INDEXES/COVARIATES OF HOMICIDE RATES FOR U.S. CITIES AND STATES: 1950

STRUCTURAL INDEX/COVARIATE	1950	
	City	State
Constant . . . . .	- 11.202 (3.596)**	- 44.431 (2.724)*
Population structure . . . . .	.792 [.125] (3.635)**	.375 [.151] (1.602)
Resource deprivation/affluence . . . . .	.153 [.514] (8.752)**	.134 [.583] (3.829)**
Percentage divorced . . . . .	.178 [.046] (1.175)	1.906 [.410] (2.743)**
Percentage ages 15-29 . . . . .	.022 [.017] (.434)	1.227 [.339] (2.435)**
Unemployment rate . . . . .	- .303 [-.111] (2.982)**	- .402 [-.129] (1.304)
South . . . . .	3.132 [.514] (4.465)**	2.322 [.232] (1.472)
Adjusted <i>R</i> <sup>2</sup> . . . . .	.519	.690
<i>N</i> . . . . .	434	48

\* *P* ≤ .05 (one-tailed).  
\*\* *P* ≤ .01 (one-tailed).

the percentage-divorced covariate (for which there is some evidence supporting interpretation as a social disintegration indicator; see n. 22) also show strong relationships to homicide at virtually all levels of analysis and time periods. Cities, metropolitan areas, or states with larger population size/density or a greater percentage of males who are divorced tend to have higher homicide rates; those that are smaller/less dense or have a smaller percentage of divorced males tend to have lower homicide rates. The effects of these three structural indexes/covariates also are largely independent of each other, as their intercorrelations typically are .3 or less in absolute value.<sup>31</sup>

<sup>31</sup> Space considerations do not permit a full comparative analysis of the differential effects of structural covariates across various index crimes. It can be briefly noted, however, that for robbery (which, next to homicide, is the most completely reported



Somewhat less invariance can be attributed to the relationship of the homicide rate to the unemployment rate and the percentage of the population ages 15–29. The effects of the former are negative and limited to the city and metropolitan levels of analysis.

By contrast, a positive relationship to the age-structure indicator appears only at the state level of analysis. We noted earlier that the changing pattern of correlation and regression coefficients across levels of analysis for this structural covariate suggests that the state-level positive age-structure effect cannot generally be attributed to aggregation bias. In addition, the significant negative relationship of homicide to the percentage 15–29 at the city level in 1980 needs to be accounted for. These facts suggest that there is one or more other factors that must be present in order for the age-structure variable to have the theoretically expected positive relationship to homicide at the lower levels of disaggregation, but we defer a more complete discussion of these issues to another research report (Land et al. 1989). In the absence of controls for such factors, it appears that the state level of analysis is the best level at which to estimate the theoretically expected positive relationship.

Third, there are a number of theoretical implications of our empirical results. With respect to the southern subculture-of-violence thesis, we find the expected positive effect of location in the South to be strong and invariant only at the city level of analysis. For metropolitan areas, this effect is highly significant in 1960 but weakens in statistical significance in 1970 and 1980. For states, the South effect is present only in 1960—at which point it is not highly significant. Interestingly, it was the 1960 period that was analyzed in the widely cited state-level study of Gastil (1971). Had any other recent decennial period been studied, a South location effect would not be significant in the presence of controls for other variables such as those studied here. Should these contradictory effects be taken as indicative of a southern subculture of violence? Recent research suggests that this is debatable (see, e.g., Dixon and Lizotte 1987).

Concerning the theoretical question of the effects of relative versus absolute deprivation on homicide rates that has enlivened much recent research, our empirical results imply that these effects are not individually identifiable with correlational data on cities, metropolitan areas, and states. We consistently find that those units that are high (low) on abso-

---

violent crime) the effects of the resource-deprivation/affluence index declines both in size and level of statistical significance compared with those of the population-structure component and the percentage-divorced variable (which, as observed in n. 24, behaves as an indicator of social disintegration). A full discussion of these cross-crime comparisons is reported in McCall, Land, and Cohen (1988).

lute deprivation also tend to be so high (low) on relative deprivation that attempts to estimate separate effects will be extremely sensitive to slight differences from one sample to the next. In consequence, attempts to estimate distinct regression effects for absolute and relative deprivation will result in an accumulation of null or contradictory findings, as we discussed earlier.

Similar remarks apply to the other elements of our resource-deprivation/affluence index. Among these elements, the single element that is most consistently (although not invariably, as documented earlier in tables 1 and 2) associated with homicide rates is the percentage of the population of a city, metropolitan area, or state that is black. This might lead an analyst to insist that this covariate be introduced into a regression separately from a reduced resource-deprivation/affluence index consisting of the four other elements.<sup>32</sup> Doing so, however, would reintroduce extreme collinearity between the percentage-black variable and the reduced-deprivation/affluence index to the covariance matrix of the regressors. In fact, this collinearity would be so extreme as to violate Klein's rule (as we noted earlier). The numerical consequence often would be to allocate explained variance solely to the percentage-black variable or to the reduced-deprivation/affluence index in accordance with whichever of the two has the numerically larger (even marginally so) correlation with the homicide rates in a particular time period for a particular level of analysis. But this is tantamount to committing what Gordon (1968) labeled the "partialling fallacy" because there is no theoretical reason for so apportioning the redundant explanatory power of the two covariates.

In brief, attempts to estimate separate effects for individual elements of resource deprivation/affluence may be tenuous at best and misleading at worst. Similar comments may apply to attempts to disentangle the effects of racial inequality from those of poverty, median income level, and so on (see n. 23). We do not advocate a moratorium on such efforts. However, it seems clear that fruitful new ideas are needed as much as additional empirical estimates. As noted in our discussion of the principal components derived from the correlation matrices of regressors in our analyses, Wilson's (1987) ideas of concentration effects and social isolation may be such new ideas. Applied at the levels of cities, metropolitan areas, and states, these concepts make meaningful the persistent covariation of the five components of our resource-deprivation/affluence index and its relationship to the homicide rate.

<sup>32</sup> Indeed, because of the similarity of the composition of the structural poverty index to this reduced form of our resource-deprivation/affluence index, and because they entered the percentage of the population of states that is nonwhite separately into their state-level regressions, it can be noted that this is close to what Loftin and Hill (1974) did in their analyses.

Do the concentration effects and social isolation concepts also help to adjudicate among the competing control and differential association theoretic interpretations (see n. 7) of the mechanisms by which structural circumstances affect delinquency and crime? According to Wilson (1987, p. 137), the social isolation associated with concentration effects does not imply that cultural traits (such as an excess of definitions favorable to delinquency/crime) are irrelevant; rather, culture is a response to social structural constraints and opportunities. In Kornhauser's (1978, p. 26) taxonomy, this proposition places Wilson's theory within the social disorganization/control/strain theoretic tradition (as opposed to the cultural deviance/differential association theoretic tradition) of criminogenic theorizing.<sup>33</sup> Additional research is necessary, however, to examine the generalizability of this theoretical conceptualization and the empirical findings reported here to other violent and property crimes and to clarify the effects of the age structure and South regional indicator variables. Questions about the extent to which the invariances established here for cross-sectional covariate relationships to the total homicide rate apply to disaggregated rates (as in Williams and Flewelling 1988; and Sampson 1987*b*) and to over-time changes in homicide rates also merit attention.

## APPENDIX A

### Notes on Specific Studies Cited in Table 1

Bailey 1984: For 1970 study, sample consists of cities with populations  $\geq$  100,000; 1960 study sample consists of the same cities (where data were available) as in the 1970 sample. The natural logarithm transformation was applied to the dependent variable. Variables not reported in above table: (1) % low income ( $\leq$  \$1,000), which had a positive effect in 1960 and a null effect in 1970.

<sup>33</sup> Wilson's (1987, p. 137) proposition also is consistent with the empirical results of Matsueda and Heimer (1987). Using individual-level data on delinquency, these authors show that an excess of cultural definitions favorable to delinquent behavior is a critical intervening variable mediating the influence of structural characteristics and social bonding on delinquency. Because the cultural definitions variable itself also is shown to depend substantially on structural characteristics and social bonding (these factors jointly explain about two-thirds of the variance in the cultural definitions variable), however, a clean interpretation of the Matsueda-Heimer model as an instance of cultural deviance theory cannot be given. For, according to Kornhauser (1978, p. 26), if a delinquent/criminal subculture is a dependent variable, surviving only in the process of social disorganization and dependent on the supply of recruits through exogenous factors, the theory is classified as a social disorganization model. It is beyond the scope of this paper to evaluate the veracity of Kornhauser's classification system. Even if it is not accepted, however, some additional theoretical articulation of similarities and differences between the social disorganization and differential association/differential social organization models appears called for, at a minimum.

Blau and Blau 1982: Sample consists of SMSAs with populations  $\geq$  250,000. The log base 10 transformation was applied to the dependent variable. Variable not included in the above table: (1) SES inequality in race measure, which had a positive effect.

Blau and Golden 1986: See Blau and Blau for sample description. The log base 10 transformation was applied to the dependent variable. Variables not included in above table: (1) SES inequality in race measure, which had a null effect; (2) arrest rates, which had a null effect; and (3) "product" terms or interaction terms between all independent variables and the percentage-southern-born variable, none of which had statistically significant effects except the interaction with percentage poor.

Crutchfield et al. 1982: Sample consists of SMSAs with populations  $\geq$  500,000. Variables not included in the above table: (1) median education, which had a null effect, and (2) mobility, which had a positive effect.

DeFronzo 1983: Sample consists of SMSAs for which cost-of-living data were available. Variables not included in above table: (1) adjusted AFDC per aided family member, which had a positive effect, and (2) household activity index, which had a null effect.

Gastil 1971: Sample consists of 48 contiguous states. Gastil's Southernness Index is based on the persistence of southern cultural traditions that emerged in the pre-Civil War era. The index reflect the southernness of a state as influenced by migration patterns of southerners. Variables not included in above table: (1) percentage urban, which was included only in the state-level model and had a null effect; (2) physicians per 1,000 population, which had a null effect in the state and SMSA models and had a positive effect in the city-level model; (3) median years school completed, which had null effects in the state- and SMSA-level models and a negative effect at the city level; (4) hospital beds per 1,000 population, with null effects in all three levels of aggregation; and (5) percentage cities with populations  $\geq$  300,000, which was used in the state-level only and had a null effect in the model.

Huff-Corzine et al. 1986: Sample consists of 48 contiguous states. Variables not included in the above table: (1) Balance of Structural Poverty Index (see Loftin and Hill below) not in the above table, which had a positive effect in the model; (2) percentage rural, which had a positive effect; (3) hospital beds per 100,000 population, which had a null effect in the Southernness Index model and had a positive effect in the percentage born in the South model; (4) percentage born in South, which had a null effect.

Jackson 1984: Sample consists of cities with populations  $\geq$  25,000 in model 1, cities with population  $\geq$  50,000 in model 2, and cities with populations  $<$  50,000 in model 3. Relative deprivation was measured

using the ratio of black to white median income. Variable not included in above table: household activity ratio, which had null effects in models 1 and 3 but had a positive effect in model 2.

**Loftin and Hill 1974:** Sample consists of 48 contiguous states. Introduced the Structural Poverty Index, which includes: (1) infant mortality rates, (2) percentage of persons 25 years old and over with less than five years of school, (3) percentage of population illiterate, (4) percentage of families with income under \$1,000, (5) Armed Forces Mental Test failures, (6) percentage of children living with one parent. Variables not included in above table: (1) Balance of Structural Poverty Index, which had a positive effect; (2) percentage rural, which had a null effect; (3) hospital beds per 100,000 population, which had a positive effect in the Confederate South model but had a null effect in the Southernness Index model.

**Loftin and Parker 1985:** Sample consists of the 49 largest cities (part of another working data set from which data were available for these cases). Variable not included in above table: (1) infant mortality used as instrumental variable (proxy replacing poverty line variable), which had a positive effect.

**Messner 1982:** Sample consists of SMSAs for which FBI crime data were available.

**Messner 1983a:** See sample for Messner (1982) above. Variables not included in above table: (1) Balance of Structural Poverty Index (see Loftin and Hill above), which had positive effects; and (2) percentage black squared, which had negative effects.

**Messner 1983b:** Sample consists of cities for which FBI crime data were available for 1969, 1970, and 1971.

**Parker and Smith 1979:** Sample consists of 48 contiguous states. Variables not included in the above table: (1) severity measure (median months served for homicide sentence for those released in 1970), which had a null effect; (2) certainty measure (number of admissions to prison sentenced for homicide in 1970 per total number of homicides in 1970), which had a null effect; (3) percentage urban, which had a positive effect; (4) Balance of Structural Poverty Index (see Loftin and Hill above) except percentage illiterate for which they had no data, which had a positive effect in the model.

**Rosenfeld 1986:** Sample consists of SMSAs with populations  $\geq 250,000$ . Relative deprivation measure is the product of (*a*) the mean income of families below the poverty level, (*b*) percentage families below the poverty level, and (*c*) ratio of median years of school completed by heads of poverty families to median years of school completed of all family heads, which had positive effects in both models.

**Sampson 1985:** Sample consists of cities with populations  $\geq 250,000$ . The

natural logarithm transformation has been applied to the dependent variable. Relative deprivation is measured using the ratio of white to black median income. Variable not included in above table: ratio of the number of homicide arrests to the number of homicides, which had a negative effect.

Sampson 1986: Sample consists of cities with populations  $\geq 100,000$ . The dependent and the relative deprivation variables are measured as those defined in Sampson (1985) above. Variables not included in the above table: (1) West regional indicator, which had a positive effect; (2) occupational status (proportion of civilian work force in professional and managerial positions), which had a positive effect; (3) police aggressiveness, which had a null effect; (4) local incarceration risk, which had a null effect; and (5) state prison risk, which had a positive effect.

Simpson 1985: Sample consists of SMSAs with populations  $\geq 250,000$ . The log base 10 transformation has been applied to the dependent variable. Variables not included in the above table: (1) SES difference in income (see Blau and Blau 1982), which had positive effects in both models; and (2) the logit form of the age structural variable (white males ages 15–29) squared.

Smith and Parker 1980: Sample consists of 48 contiguous states. Variables not included in above table: (1) Balance of the Structural Poverty Index (see Loftin and Hill above), which had a positive effect; (2) percentage urban, which had a null effect; and (3) hospital beds per 100,000 population, which had a null effect.

Williams 1984: Sample consists of SMSAs with populations  $\geq 250,000$ . The log base 10 has been applied to the dependent variable. Model 1 is based on Blau and Blau's (1982) equation and model 2 is based on Messner's (1982) equation—both studies outlined above. Variables not included in the above table: (1) racial inequality (Blau and Blau's model), which had a positive effect.

Williams and Flewelling 1988: Sample consists of cities with populations  $\geq 100,000$ . The log base 10 has been applied to the dependent variable. Variable not included in the above table: (1) justifiable homicide ratio, which had a positive effect.

## APPENDIX B

### Definitions

Homicide rate: (Number of murder and nonnegligent manslaughter offenses/total resident population)  $\times 100,000$ .

Population size: Number total resident population.

Population per square mile: (Total population/land area square miles).

Percentage black population: (Number of black population/total resident population)  $\times$  100.

Percentage of the population 15–29 years of age: (Number of 15–29 year olds/total resident population)  $\times$  100.

Percentage divorced males: (Number divorced males/number males 14 [15 in 1980] years old and over)  $\times$  100.

Percentage of children not living with both parents: (Number of children 18 years old and under not living with both parents/number of children 18 years old and under)  $\times$  100.

#### Income Measures (in current dollars): Income Distributions for Families

Median family income.

Percentage of families living below the official poverty level.

Gini index of income concentration for families.

Percentage unemployed: (Number employed in civilian labor force/number in civilian labor force)  $\times$  100.

South: South dummy variable defined in the text.

#### Sources

U.S. Department of Justice, Federal Bureau of Investigation. Various years. *Uniform Crime Reports*. Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census. Various years. *Census of Population*. Vols. 1 and 2. *Characteristics of the Population*. Washington, D.C.: Government Printing Office. (State volumes.)

U.S. Bureau of the Census. Various years. *County and City Data Book: A Statistical Abstract*. Washington, D.C.: Government Printing Office.

U.S. Bureau of the Census. 1978 and 1984. *City and County Data Book Consolidated File, County Data 1947–1977 and 1983 (MRDF)*. Washington, D.C.: Bureau of the Census.

#### REFERENCES

- Bailey, William C. 1984. "Poverty, Inequality, and City Homicide Rates: Some Not So Unexpected Findings." *Criminology* 22:531–50.
- Baldwin, John D. 1985. "Thrill and Adventure Seeking and the Age Distribution of Crime: Comment on Hirschi and Gottfredson." *American Journal of Sociology* 90:1326–30.
- Ball-Rokeach, Sandra J. 1973. "Values and Violence: A Test of the Sub-culture of Violence Thesis." *American Sociological Review* 38:736–49.

- Berry, B. J. L. 1972. "Latent Structure of the American Urban System." Pp. 11-60 in *City Classification Handbook*, edited by B. J. L. Berry. New York: Wiley.
- Blalock, Hubert M., Jr. 1964. *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- Blau, Judith R., and Peter M. Blau. 1982. "The Cost of Inequality: Metropolitan Structure and Violent Crime." *American Sociological Review* 47:114-29.
- Blau, Peter M., and Reid M. Golden. 1986. "Metropolitan Structure and Criminal Violence." *Sociological Quarterly* 27:15-26.
- Breault, Kevin D. 1987. "Beyond the Pale: A Large Sample Structural Analysis of U.S. Crime Rates." Working paper. University of Chicago, Department of Sociology.
- Cohen, Lawrence E. 1987. "Throwing Down the Gauntlet: A Challenge to the Relevance of Sociology for the Etiology of Criminal Behavior." Review essay on *Crime and Human Nature*, by James Q. Wilson and Richard J. Herrnstein. *Contemporary Sociology* 16:202-05.
- Cohen, Lawrence, and Marcus Felson. 1979. "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review* 44:588-608.
- Cohen, Lawrence, Marcus Felson, and Kenneth C. Land. 1980. "Property Crime Rates in the United States: A Macrodynamical Analysis, 1947-77; with Ex Ante Forecasts for the Mid-1980s." *American Journal of Sociology* 86:90-118.
- Cohen, Lawrence, James R. Kluegel, and Kenneth C. Land. 1981. "Social Inequality and Predatory Criminal Victimization: An Exposition and Test of a Formal Theory." *American Sociological Review* 46:505-24.
- Cohen, Lawrence, and Kenneth C. Land. 1987. "Age Structure and Crime: Symmetry vs. Asymmetry, and the Projection of Crime Rates through the 1990s." *American Sociological Review* 52:170-83.
- Cohen, Lawrence, and Richard Machalek. 1988. "A General Theory of Expropriative Crime: An Evolutionary Ecological Approach." *American Journal of Sociology* 94:465-501.
- Cook, Philip J. 1986. "The Demand and Supply of Criminal Opportunities." Pp. 1-27 in *Crime and Justice: An Annual Review of Research*, vol. 7. Edited by Michael Tonry and Norval Morris. Chicago: University of Chicago Press.
- Cook, R. Dennis, and Sanford Weisberg. 1982. *Residuals and Influence in Regression*. New York: Chapman & Hall.
- Crutchfield, Robert D., Michael R. Geerken, and Walter R. Gove. 1982. "Crime Rate and Social Integration." *Criminology* 20:467-78.
- Curtis, Lynn. 1975. *Violence, Race, and Culture*. Lexington, Mass.: Heath.
- DeFronzo, James. 1983. "Economic Assistance to Impoverished Americans." *Criminology* 21:119-36.
- Dixon, Jo, and Alan J. Lizotte. 1987. "Gun Ownership and the 'Southern Subculture of Violence.'" *American Journal of Sociology* 93:383-405.
- Doerner, William G. 1975. "A Regional Analysis of Homicide Rates in the United States." *Criminology* 13:90-101.
- Erlanger, Howard S. 1976. "Is There a 'Subculture of Violence' in the South?" *Journal of Criminal Law and Criminology* 66:483-90.
- Farley, John E. 1987. "Suburbanization and Central-City Crime Rates: New Evidence and a Reinterpretation." *American Journal of Sociology* 93:688-700.
- Flango, Victor Eugene, and Edgar L. Sherbenou. 1976. "Poverty, Urbanization, and Crime." *Criminology* 14:331-46.
- Gastil, Raymond D. 1971. "Homicide and a Regional Culture of Violence." *American Sociological Review* 36:412-27.
- Gordon, Robert A. 1967. "Issues in the Ecological Study of Delinquency." *American Sociological Review* 32:927-44.



- . 1968. "Issues in Multiple Regression." *American Journal of Sociology* 73:592–616.
- Greenberg, David F. 1985. "Age, Crime, and Social Explanation." *American Journal of Sociology* 91:1–21.
- Hackney, Sheldon. 1969. "Southern Violence." *American Historical Review* 74:906–25.
- Hanushek, Eric A., and John E. Jackson. 1977. *Statistical Methods for Social Scientists*. New York: Academic.
- Hindelang, Michael J., Michael R. Gottfredson, and James Garofalo. 1978. *Victims of Personal Crime*. Cambridge, Mass.: Ballinger.
- Hirschi, Travis, and Michael Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89:552–84.
- . 1985a. "Age and Crime, Logic and Scholarship: Comment on Greenberg." *American Journal of Sociology* 91:22–27.
- . 1985b. "All Wise after the Fact Learning Theory, Again: Reply to Baldwin." *American Journal of Sociology* 90:1330–33.
- Huff-Corzine, Lin, Jay Corzine, and David C. Moore. 1986. "Southern Exposure: Deciphering the South's Influence on Homicide Rates." *Social Forces* 64:906–24.
- Jackson, Pamela Irving. 1984. "Opportunity and Crime: A Function of City Size." *Social Science Review* 68:172–93.
- Janson, Carl-Gunnar. 1980. "Factorial Social Ecology: An Attempt at Summary and Evaluation." *Annual Review of Sociology* 6:433–56.
- Johnston, J. 1984. *Econometric Methods*, 3d ed. New York: McGraw-Hill.
- Kim, Jae-On, and Charles W. Mueller. 1978. *Factor Analysis: Statistical Methods and Practical Issues*. Beverly Hills, Calif.: Sage.
- Klein, L. R. 1962. *An Introduction to Econometrics*. Englewood Cliffs, N.J.: Prentice-Hall.
- Kornhauser, Ruth Rosner. 1978. *Social Sources of Delinquency: An Appraisal of Analytic Models*. Chicago: University of Chicago Press.
- Land, Kenneth C., Patricia L. McCall, and Lawrence E. Cohen. 1989. "Age Structure and Crime Rates: Is There a Connection?" Paper presented at the annual meeting of the American Sociological Association, San Francisco.
- Loftin, Colin, and Robert H. Hill. 1974. "Regional Subculture and Homicide: An Examination of the Gastil-Hackney Thesis." *American Sociological Review* 39:714–24.
- Loftin, Colin, and Robert Nash Parker. 1985. "An Errors-in-Variable Model of the Effect of Poverty on Urban Homicide Rates." *Criminology* 23:269–87.
- Lundsgaarde, Henry Peder. 1977. *Murder in Space City*. New York: Oxford University Press.
- Maddala, G. S. 1977. *Econometrics*. New York: McGraw-Hill.
- Matsueda, Ross L., and Karen Heimer. 1987. "Race, Family Structure, and Delinquency: A Test of Differential Association and Social Control Theories." *American Sociological Review* 52:826–40.
- McCall, Patricia L., Kenneth C. Land, and Lawrence E. Cohen. 1988. "Violent Criminal Behavior: Is There a General and Continuing Influence of the South?" Working paper. Duke University, Department of Sociology.
- Merton, Robert K. 1938. "Social Structure and Anomie." *American Sociological Review* 3:672–82.
- Messner, Steven F. 1982. "Poverty, Inequality, and the Urban Homicide Rate." *Criminology* 20:103–14.
- . 1983a. "Regional and Racial Effects on the Urban Homicide Rate: The Subculture of Violence Revisited." *American Journal of Sociology* 88:997–1007.
- . 1983b. "Regional Differences in the Economic Correlates of the Urban

- Homicide Rate: Some Evidence on the Importance of Cultural Context." *Criminology* 21:477-88.
- Miller, S. M., and P. Roby. 1970. *Future of Inequality*. New York: Basic.
- Neter, John, William Wasserman, and Michael H. Kutner. 1983. *Applied Linear Regression Models*. Homewood, Ill. Irwin.
- Parker, Robert Nash, and M. Dwayne Smith. 1979. "Deterrence, Poverty, and Type of Homicide." *American Journal of Sociology* 85:614-24.
- Reed, John S. 1971. "To Live—and Die—in Dixie: A Contribution to the Study of Southern Violence." *Political Science Quarterly* 86:429-43.
- . 1972. *The Enduring South: Subcultural Persistence in Mass Society*. Lexington, Mass.: Lexington.
- Rosenfeld, Richard. 1986. "Urban Crime Rates: Effects of Inequality, Welfare Dependency, Region, and Race." Pp. 116-30 in *The Social Ecology of Crime*, edited by James M. Byrne and Robert J. Sampson. New York: Springer-Verlag.
- Sampson, Robert J. 1985. "Race and Criminal Violence: A Demographically Disaggregated Analysis of Urban Homicide." *Crime and Delinquency* 31:47-82.
- . 1986. "Crime in Cities: The Effects of Formal and Informal Social Control." Pp. 271-312 in *Communities and Crime*, edited by Albert J. Reiss, Jr., and Michael Tonry. Chicago: University of Chicago Press.
- . 1987a. "Communities and Crime." Pp. 91-114 in *Positive Criminology*, edited by Michael R. Gottfredson and Travis Hirschi. Newbury Park, Calif.: Sage.
- . 1987b. "Urban Black Violence: The Effect of Male Joblessness and Family Disruption." *American Journal of Sociology* 93:348-82.
- Schurman, Leo, and Solomon Kobrin. 1986. "Community Careers in Crime." Pp. 67-100 in *Communities and Crime*, edited by Albert J. Reiss, Jr., and Michael Tonry. Chicago: University of Chicago Press.
- Schuessler, K., and G. Slatin. 1964. "Sources of Variation in U.S. City Crime, 1950 and 1960." *Journal of Research in Crime and Delinquency* 1:127-48.
- Shaw, Clifford R., and Henry McKay. 1942. *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.
- Silberman, Charles E. 1978. *Criminal Violence, Criminal Justice*. New York: Random House.
- Simpson, Miles E. 1985. "Violent Crime, Income Inequality, and Regional Culture: Another Look." *Sociological Focus* 18:199-208.
- Smith, M. Dwayne, and Robert Nash Parker. 1980. "Type of Homicide and Variation in Regional Rates." *Social Forces* 59:137-47.
- Stahura, John M., and John J. Sloan III. 1988. "Urban Stratification of Places, Routine Activities and Suburban Crime Rates." *Social Forces* 66:1102-18.
- Sutherland, Edwin H. 1947. *Principles of Criminology*, 4th ed. Philadelphia: Lippincott.
- Townsend, Peter. 1974. "Poverty as Relative Deprivation: Resources and Style of Living." Pp. 15-41 in *Poverty, Inequality, and Class Structure*, edited by D. Wedderburn. Cambridge: Cambridge University Press.
- Theil, Henri. 1971. *Principles of Econometrics*. New York: Wiley.
- Wilks, J. A. 1967. "Ecological Correlates of Crime and Delinquency." Pp. 138-56 in *Crime and Its Impact*. Task Force report. Washington, D.C.: Government Printing Office.
- Williams, Kirk R. 1984. "Economic Sources of Homicide: Reestimating the Effects of Poverty and Inequality." *American Sociological Review* 49:283-89.
- Williams, Kirk R., and Robert L. Flewelling. 1988. "The Social Production of Criminal Homicide: A Comparative Study of Disaggregated Rates in American Cities." *American Sociological Review* 53:421-31.

- Wilson, James Q., and Richard J. Herrnstein. 1985. *Crime and Human Nature*. New York: Simon & Schuster.
- Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Wirth, Louis. 1938. "Urbanism as a Way of Life." *American Journal of Sociology* 44:1-24.
- Wolfgang, Marvin E., and Franco Ferracuti. 1967. *The Subculture of Violence*. London: Tavistock.