The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing[1]

Jonathan Crane
University of Illinois at Chicago

Why are the social problems of ghettos so bad? This article proposes that ghettos are communities that have experienced epidemics of social problems. One important implication of this theory is that the pattern of neighborhood effects on social problems should be nonlinear in large cities. As neighborhood quality decreases, there should be a sharp increase in the probability that an individual will develop a social problem. The jump should occur somewhere near the bottom of the distribution of neighborhood quality. This hypothesis is tested by analyzing the pattern of neighborhood effects on dropping out and teenage childbearing. The analysis strongly supports the hypothesis, with exceptions for certain subgroups. Even after controlling for individual characteristics, black and white adolescents are exposed to sharp increases in the risk of dropping out and having a child in the worst neighborhoods in large cities.

I. THE EPIDEMIC THEORY OF GHETTOS

The word "epidemic" is commonly used to describe the high incidence of social problems in ghettos. The news is filled with feature stories on "crack epidemics," "epidemics of gang violence," and "epidemics of teenage childbearing." The term is used loosely in popular parlance but

[1] I wish to thank David T. Ellwood, Christopher Jencks, Mary Jo Bane, Naomi Goldstein, Lisa Whittemore, Glenn Loury, and Lee Rainwater for their invaluable advice and helpful comments on earlier drafts. The data utilized in this paper were made available in part by the Inter-university Consortium for Political and Social Research. The data for the Public Use Samples of basic records from the 1970 census were originally collected by the U.S. Department of Commerce, Bureau of the Census. Neither the collector of the original data nor the consortium bears any responsibility for the analyses or interpretations presented here. Requests for reprints should be addressed to Jonathan Crane, University of Illinois at Chicago, Institute for Government and Public Affairs, 921 West Van Buren, Chicago, Illinois 60607.

0002-9602/91/9605-0006$01.50

turns out to be remarkably apt. In fact, I propose that ghettos are neighborhoods that have experienced epidemics of social problems.

In my earlier work (Crane 1989) I used a mathematical model designed to describe the spread of infectious diseases in order to characterize how social problems spread. This "contagion model" is essentially a more general version of Schelling's (1971; 1978, pp. 101–10) well-known "tipping model," which was formalized by Granovetter (1978) and Granovetter and Soong (1983).[2]

The basic assumption of my model is that social problems are contagious and are spread through peer influence. The large body of empirical work on delinquency and differential association supports this assumption, at least for several types of adolescent social problems (e.g., Kandel 1980).[3]

The key implication of the contagion model is that there may be critical levels of incidence of social problems in populations. The incidence of problems tends to move toward equilibrium levels. If the incidence stays below a critical point, the frequency or prevalence of the problem tends to gravitate toward some relatively low-level equilibrium. But if the incidence reaches a critical point, the process of spread will explode. In other words, an epidemic may occur, raising the incidence to an equilibrium at a much higher level.[4]

Two basic conditions determine a community's susceptibility to epidemics: (1) the residents' risks of developing social problems and (2) their susceptibility to peer influence. Steinberg (1987) found that children from single-parent homes are more susceptible to peer pressure to engage in antisocial behavior. Liebow (1967), Fischer (1977), and MacLeod (1987) found that poor urban males of various ages often develop strong peer subcultures that value behaviors that generate social problems. Adoles-

[2] Schelling (1978) used the term "critical mass model" to describe his finding that examples with multiple equilibria exhibit critical mass properties. Granovetter (1978) used the term "threshold model" because an important assumption is that individuals will develop a certain behavior when the proportion of the relevant population engaging in that behavior reaches a particular threshold. However, I prefer the term "contagion model" to emphasize the medical analogy, which I have found provides a quick intuitive understanding of the model.

[3] Social problems may also be contagious among children and adults, but the mechanisms by which they spread might be different. For example, the parent-child relationship may be relatively more important for children than for teenagers. Adults may influence each other less directly through social norms. The analysis here is confined to adolescents because of limitations of the data, described below.

[4] Similar kinds of epidemic results can be generated by a number of different models with different mathematical approaches. All of these models fall under the general rubric of catastrophe theory. For examples of alternative approaches that could be applied to this problem, see Varaiya and Wiseman (1984).

cent street gangs are a good example of this. (Recent media reports of a sudden explosion of gang membership in Los Angeles suggest that there may have been an epidemic of joining there.) Rainwater (1970) found that poor black teenage females often begin having sex because it is valued as a sign of maturity in their peer group, despite the fact that few have an "autonomous interest" in it.

Thus, we would expect epidemics of social problems to be most common in poor, minority neighborhoods, particularly in cities. But not all such neighborhoods experience epidemics; ghettos are the ones that do.[5] Poor neighborhoods that do not become ghettos have higher than average rates of social problems, but they do not experience the epidemic interaction that generates a whole much greater than the sum of its parts.

This theory of epidemics implies that the pattern of incidence of a particular social problem in the neighborhoods of a city should take a specific form. There should be two separate distributions.[6] One should include all the neighborhoods that have not experienced an epidemic of the problem. Poor neighborhoods should be concentrated at the high end of this distribution, but still at a much lower level than the low end of the second distribution. The second distribution, much smaller than the first, should include all the ghetto neighborhoods. I have used this implication to test the theory and found that various indices of juvenile delinquency in Chicago and murder rates in Los Angeles conformed to this pattern (Crane 1989).

Another testable implication of the theory is that neighborhood effects on social problems should follow a very specific form. The relationships between neighborhood quality and the incidence of particular social problems should be nonlinear. Social problems should increase as neighborhood quality declines, but not at a constant rate. Somewhere near the bottom of the distribution of neighborhood quality, there should be a jump in the rate of increase. This is because the prevalence of problems should be much higher in those neighborhoods that have experienced an epidemic than in those that have not. Thus, the epidemic theory implies that there are very strong neighborhood effects at at least one point near the bottom of the distribution of neighborhood quality.

This pattern of neighborhood effects enables us to distinguish the epidemic hypothesis from other theories. Jencks and Mayer (1990) organize

[5] In everyday usage, the term "ghetto" sometimes refers to any poor and/or ethnically homogeneous neighborhood. But here I apply its other usage, which refers specifically to "bad" neighborhoods, i.e., those rife with social problems. There are at least two synonyms that share this meaning, but "slum" seems to have gone out of use and "underclass neighborhood" (Ricketts and Sawhill 1988) is not yet widely used.

[6] If we take into account dynamic aspects of the contagion process, it could also be a single bimodal distribution (Crane 1989). But the basic idea is the same.

theories of neighborhood effects into four general categories. (1) Contagion theories (of which the epidemic theory is one example) focus on peer influence as the mechanism of neighborhood effects. (2) Collective socialization theories argue that adults in the neighborhood serve as role models and sources of social control. (3) Institutional theories emphasize the role that schools, businesses, political organizations, social service agencies, and the police play in the community. (4) Social competition theories see neighbors as competitors for scarce resources.

The first three theories imply that bad neighborhoods increase social problems, while the fourth predicts that good neighborhoods increase them. Some versions of the first three predict that social problems increase at an accelerating rate as neighborhood quality declines. But the acceleration is smooth and steady. None of them, except the epidemic version of the contagion theory, imply an extremely large and very sharp increase in the incidence of social problems. It is not that the others are inherently inconsistent with the existence of such a pattern. It is just that they do not explain it. Each of these other theories needs some kind of intermediate mechanism with critical mass–type properties embedded within it to account for a large jump. The epidemic process could provide that intermediate mechanism.

To say that the existence of this pattern of neighborhood effects would thus prove the epidemic theory is certainly going too far. There may be other mechanisms that could generate a critical mass phenomenon in social problems. But, in the absence of alternatives, determining whether or not there is an extremely large and very sharp increase in social problems at the bottom of the distribution of neighborhood quality is a sufficient test of the epidemic hypothesis.

In the following sections, neighborhood effects on dropping out of high school and teenage childbearing are analyzed to determine whether there was such a large jump in the incidence of these social problems in the worst neighborhoods of large cities and in other areas as well.

II. THE LITERATURE ON NEIGHBORHOOD EFFECTS

In their comprehensive literature review, Jencks and Mayer (1990) found that strong conclusions about the existence and strength of neighborhood effects could not be drawn from the existing body of empirical work. There are a number of reasons for this, but the main one is simply that the literature is extremely small.[7] While there is no need to repeat their

[7] There is, however, a large literature on the effects of the social composition of schools (Jencks and Mayer 1990). The relationship between neighborhood effects and school effects is discussed in Sec. V below.

review, the studies of neighborhood effects on the two outcomes examined here can be summarized briefly.

Datcher (1982) used the Panel Study of Income Dynamics (PSID) to do a longitudinal study of urban males, aged 13–22, who lived with their parents in 1968. She used zip codes to create geographical boundaries and then examined the effect that area-averaged income had on individual education attainment in 1978. She controlled for parents' educational attainment and income, family size, region, community size, age, and the head of household's educational aspirations for his or her children. She found that an increase of $1,000 (10%) in zip code–area income raised the educational attainment of the men by approximately one-tenth of a school year for both blacks and whites.

Corcoran et al. (1987) expanded Datcher's analysis. They looked at all the individuals in the PSID who were between the ages of 10 and 17 in 1968. They examined the effects of four different 1968 zip code–area characteristics (median income, percentage of female-headed homes, male unemployment rate, percentage of people receiving public assistance) on the individual educational attainment in 1983. They controlled for race, region, city size, religion, family structure, family income in 1968, welfare receipt, the educational attainment of the head of household and spouse, and the work hours of the head of household and spouse.

For men, a 2-SD increase (eight percentage points) in the proportion of female-headed families in the zip code area decreased educational attainment by about one-fourth of a year. A 2-SD increase (10 percentage points) in the welfare-receipt rate reduced schooling by about half a year. Neither median income nor the male unemployment rate had an effect. For women, a 2-SD increase in the male unemployment rate of the zip code area decreased attainment by about half a year, a 2-SD increase in the proportion of female-headed families reduced schooling by about one-fourth of a year, and a 2-SD increase in the welfare-receipt rate lowered attainment by a little less than half a year. Median income had no effect.

There are no studies of neighborhood effects on teenage childbearing per se, but there is one each on pregnancy and contraception. Hogan and Kitagawa (1985) looked at Chicago census tracts to construct a 1979 sample of unmarried black females between the ages of 13 and 19. They used two different measures of neighborhood effects. In one, they constructed a composite measure of neighborhood quality using the tracts' poverty rate, median family income, male-female ratio, number of children per ever-married female, and several indices of juvenile delinquency among teenage males. They divided the tracts into three groups: the top quarter, the middle half, and the bottom quarter. The second neighborhood variable was a dummy indicating whether the female lived in the

West Side ghetto. They controlled for social class, parents' marital status, and number of siblings.

Hogan and Kitagawa (1985) found no significant difference in teen pregnancy rates between the high-quality neighborhoods and middle-quality neighborhoods. But the chances of becoming pregnant in a given month were a little more than a third higher in the low-quality neighborhoods. Living on the West Side increased pregnancy risk by almost two-fifths. Since two measures at very different geographic levels were used, the overall neighborhood effect cannot be determined precisely. It was probably larger than the effect of either one alone but less than the sum of both.

Using the same data set, Hogan, Astone, and Kitagawa (1985) examined the practice of contraception among black females in Chicago. They found that females in low-quality neighborhoods were about half as likely as those in better neighborhoods to use contraception when they first had sex.

In short, what little evidence we have suggests that neighborhood effects on educational attainment are quite small, while neighborhood effects on pregnancy are quite large. It is possible, however, that the two studies of educational attainment underestimated neighborhood effects because zip code areas are poor proxies for neighborhoods. Zip code areas are probably larger than most neighborhoods, and they are not designed to be socioeconomically or ethnically homogeneous. Zip code–area estimates might be biased downward by a large amount because the error in measuring neighborhoods introduces randomness.[8]

III. METHODOLOGY

One reason that so little work has been done on neighborhood effects is that there are so few data sets that provide information on both individuals and their neighborhoods. However, there is one extremely valuable data set that has gone virtually untouched, although it has been around for almost two decades. In 1970, the Census Bureau defined a geographic unit called a neighborhood and made data on individuals' neighborhoods publicly available in one of its Public Use Microdata Samples (PUMS). The neighborhoods themselves were not identified, but the values of 55 neighborhood indices were attached to individuals' records. Since this data set is nonlongitudinal, it has certain methodological weaknesses.

[8] Of course, it is also true that estimates of neighborhood effects in all three studies may have been biased in either direction for a number of other reasons. In Sec. V below, the potential sources of bias in the analysis done here are discussed, and most of them apply to these three studies as well.

However, its enormous size (over 2 million individuals) gives it some unique and extremely important advantages. In particular, it enables us to examine neighborhood effects on small subgroups, such as those living in urban ghettos.

In the following analysis, I examine the effects of neighborhood quality on dropping out and teenage childbearing. The data were drawn from the 1/100, 15% Neighborhood Characteristic File of the PUMS. These so-called neighborhoods were specially designed for this version of the PUMS. They were about the same size as census tracts, averaging 4,000–5,000 people. They were formed by computer, using geographic keys associated with each household record. They were normally contiguous and relatively compact. However, socioeconomic and demographic data were not used to define them (U.S. Bureau of the Census 1973). So, despite the nomenclature, these geographic units were not necessarily neighborhoods in any meaningful social sense.

I examined only teenagers living with their parents, because social problems almost certainly affected both the probability of moving out of their parent's home and the quality of the neighborhood into which they moved. In a model that assumed one-way causality and made no adjustment for this simultaneity effect, estimates of neighborhood effects would probably be biased upward.[9] No instrument could be found to distinguish the effects in each direction, so it was not possible to specify a simultaneous equation model. Teenagers were also excluded on the basis of two data-cleaning criteria. Those with inconsistencies in their records were left out. And individuals with certain allocated values for the outcome variables were omitted.[10]

There were a total of 113,997 16–19-year-olds (56,233 females) in the 1970 PUMS. After the various exclusions, the study samples consisted of 92,512 teenagers for the analysis of dropping out of high school and 44,466 females for the childbearing analysis. The ramifications of the exclusions are discussed in Section V below.

Recall that the epidemic theory of ghettos implies that there should be a sharp increase in the incidence of social problems among the worst neighborhoods in large cities. To determine whether there was such a

[9] This is also the reason this study focuses exclusively on teenagers. The simultaneity problem applies to all adults, so without longitudinal data or an identifying instrument, any estimates of neighborhood effects on their social problems are likely to be biased upward. The problem does not apply to children, but there was no information on their social problems in the data set.

[10] If a question is left unanswered or the response is ambiguous, the Census Bureau may allocate a response in the PUMS. One method of allocation was to substitute the response of the previous person of the same type. Those with dropout and childbearing responses allocated in this way were excluded.

jump in dropping out or childbearing, a piecewise linear logit model (Amemiya 1981) is used to estimate the pattern of neighborhood effects across the distribution of neighborhood quality. A large increase in the slope from one piece to the next would indicate a precipitous rise in these problems at a particular point in the neighborhood distribution.

In the model, the reduced-form equation for the probability that individual i drops out (or has a child), P_{Di}, is:

$$P_{Di} = \frac{1}{1 + e^{(\alpha + \mathbf{X}_i \boldsymbol{\beta} + \mathbf{N}_{ij} \boldsymbol{\gamma}_j)}}, \qquad (1)$$

where $\mathbf{X}_i$ is a vector of personal characteristics; $\boldsymbol{\beta}$ is a vector of coefficients; $\mathbf{N}_{ij}$ is a vector indicating neighborhood quality for each individual i within a particular interval j of the neighborhood quality distribution; $\boldsymbol{\gamma}_j$ is a vector of coefficients indicating the neighborhood effect of each interval j of the distribution; and α is an intercept. The parameters α, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}_j$ are generated by maximum-likelihood estimates.

For blacks and Hispanics, the neighborhood quality distribution is divided into seven intervals. A percentile scale is used to establish boundaries for the intervals.[11] For example, the bottom group includes the 5% of the black population living in the worst neighborhoods. In ascending order, the intervals are percentiles 0–5, 6–10, 11–25, 26–50, 51–75, 76–90, and 91–100. The same intervals are used for whites. But an eighth one is added, 0–.1, for the purpose of comparison, in the same range of the bottom interval for blacks. In other words, just .1% of white teenagers lived in neighborhoods as bad as those inhabited by 5% of black teens.

The index of neighborhood quality used in the models estimated below is the percentage of workers in the neighborhood who held professional or managerial jobs (%HIGH STATUS). It was chosen because it had a larger effect on both dropping out and childbearing than any of 15 other indices,[12] as well as several composites. Also, it dominated all of these other indices when they were run together in various combinations. And

[11] Approximations were necessary because the raw data were provided in discrete form.

[12] The 15 other measures were the poverty rate, the family poverty rate, the median income, the unemployment rate, the male labor-force participation rate, the female labor-force participation rate, the proportion of female-headed families, the percentage of the population that was black, the percentage of the population that was Hispanic, the proportion of the population between the ages of 16 and 21, the median level of educational attainment, the proportion of adults who had completed less than eight years of school, the proportion of families who had moved within the last five years, the proportion of households with more than one person per room, and a Gini coefficient of income distribution.

%HIGH STATUS alone explained almost as large a proportion of the variation in each outcome variable as did all 16 of the indices combined. (A two-way correlation matrix showing the relationship between 10 of these indices is presented in table A1 in the Appendix.)

Just why %HIGH STATUS had the strongest effects is not entirely clear. It could be that if kids see a lot of role models for traditional success in their neighborhood, it gives them more incentive to stay in school and avoid having children. Or it may be that high-status people use their affluence and influence in wider society to bring into the neighborhood resources that make local institutions and services better.

While these hypotheses are probably valid to some extent, my guess is that, for the most part, the relationship reflects a selection process. Affluent people can live wherever they want to. They choose to live in good neighborhoods. If few affluent individuals are around, it is probably because the neighborhood is undesirable. This hypothesis is essentially predicated on the assumption that individuals can gain better information about neighborhood quality in their everyday lives than can social scientists with statistical instruments.

The control variables were chosen largely on the basis of previous empirical work. There is a fairly large literature on the determinants of educational attainment. Parents' educational attainment, family income, and father's occupational status are important factors (Corcoran et al. 1987; Sewell, Hauser, and Wolf 1980; Jencks et al. 1972); cognitive ability, academic achievement, attitudes, and aspirations also have effects (Sewell et al. 1980; Jencks et al. 1972). Jencks et al. found that race had no effect on educational attainment after controlling for background variables, but Corcoran et al. found a significant black advantage over whites. Sewell et al. found that gender affects the pattern of attainment. They also found that family size, number of parents in the home, rural origin, and maternal employment had indirect effects, but no direct effects, on educational attainment.

Of these variables, eight are available in the 1970 PUMS data set, including family income, parents' educational status, family head's occupational status, household structure, family size, rural origin, gender, and race. The first six were included as controls in the models of dropping out. Gender and race (and also Hispanic ethnicity) are accounted for in the analysis by estimating separate equations for various subgroups. Measures of ability, achievement, attitudes, and aspirations are not available. Place size, region, and residential mobility have been included because it seems plausible that they might have effects. Two other variables included indicate whether the family head was in the military and whether the head lacked a census-designated occupation; in both of those cases their occupational status score was zero.

In general, demographic analyses of the determinants of teen childbearing have tended to yield equivocal results (Hogan and Kitagawa 1985). But Hogan and Kitagawa did find that a composite variable of social class, parents' marital status, the presence of a sister who was a teen mother herself, and parental control of dating had significant effects on the childbearing age of black teenagers in Chicago. Zelnick, Kantner, and Ford (1981) found that the number of parents in the home had a significant effect. For the sake of comparison, and since it seemed plausible that they might have an effect, the same measures of SES and family structure and all the other controls in the models of dropping out were also included in the childbearing models. The one exception was household size, which was excluded because it is endogenous.

IV. RESULTS

Figure 1 shows the effect of changes in the percentage of high-status workers in the neighborhood on the estimated dropout probabilities for black, white, and Hispanic teenagers.[13] Since the epidemic theory emphasizes the importance of changes at the low end of the distribution, figure 1 and all the figures included below are structured to highlight the effects of decreases in %HIGH STATUS. With vertical axes positioned at the right, the figure is read right to left, from higher percentage to lower ones, and demonstrates the movement from better neighborhoods to worse ones. The confidence intervals of the estimates presented in figure 1 and all other figures below are given in tables A2–A12 in the Appendix. The logit equations that generated the estimates shown in each figure are available from the author.

The estimates of dropout probabilities are calculated for teenagers with average background characteristics for their racial/ethnic group. In other words, the control variables are fixed at their mean levels.

The slope of the line between each set of points indicates the magnitude of the neighborhood effect in that portion of the neighborhood quality distribution. For example, from right to left in the figure, the estimated dropout probability for Hispanics increases from .083 to .121 as the percentage of high-status workers falls from 43.3% to 27%. The slope is −.0023. In other words, each one-percentage-point decrease in %HIGH STATUS increases by .0023 the dropout probability of an average Hispanic teenager living in the best neighborhood.

As figure 1 clearly shows, the pattern of neighborhood effects on drop-

[13] The "white" category contained all individuals who were neither blacks nor Hispanics, including Asians, Native Americans, and other minorities. Since whites constituted such a high percentage of this group, the term "white" is used as a shorthand.

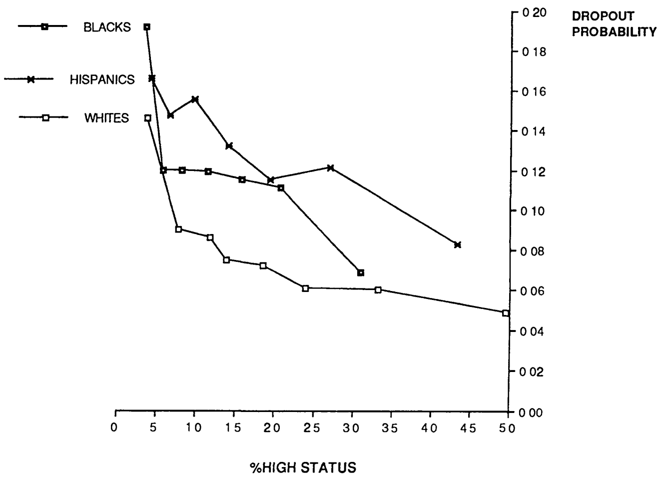

FIG. 1.—Estimated dropout probability as a function of the percentage of high-status workers in the neighborhood for blacks, whites, and Hispanics.

ping out is precisely the one predicted by the epidemic hypothesis for both blacks and whites. In each case, there is a sharp jump in dropout probabilities in the lowest range of %HIGH STATUS. The neighborhood effect is extremely large at the very bottom for both groups. For Hispanics, however, the pattern of neighborhood effects is essentially linear.

The extreme sharpness of the increase can be expressed by the ratio of the slopes around the point where it occurs. For blacks, there is virtually no neighborhood effect in the middle range of the distribution. As the percentage of high-status workers falls from 20.7% to 5.6%, the estimated dropout probability increases from just .111 to .120. The slope of the line fitted to the four observations in this range (using the least squares criterion) is just −.00065. But as %HIGH STATUS falls just two more percentage points, to 3.5%, the estimated dropout probability jumps up to .192. The slope at the bottom of the distribution is −.034. The ratio of the steep slope to the flat slope is 52.7, a very sharp increase indeed. In other words, the neighborhood effect among the very worst neighborhoods is more than 50 times greater than the effect in the middle. There is also a sharp decrease in the dropout probability at the top of the distribution for blacks. The slope in the range of %HIGH STATUS between the top two points is −.0041, more than six times greater than the slope in the middle (though still less than an eighth as large as the slope at the bottom).

Is it likely that the jump resulted from random variation in the estimates? We can address this question by testing a null hypothesis of linearity, that is, that the slope in the middle of the distribution was the true slope for the entire distribution. To do this, I extrapolate from the line fitted to the middle five points by extending it in each direction, down to 3.5% and up to 31%. I then check to see whether the dropout probabilities at these points are significantly different from the extrapolated values. The dropout probability in the worst neighborhoods is significantly higher than the corresponding extrapolated value. But the dropout probability in the best neighborhoods is not significantly lower.[14]

For whites, the estimated dropout probability rises from .049 to .090 as the percentage of high-status workers falls from 49.4% to 7.7%. The slope of the line fitted to these seven points is .00095. As the percentage of high-status workers falls from 7.7% to 3.6%, the dropout probability jumps to .146, which is significant under the same hypothesis test described above. The slope between these two points is $-.014$. Thus the neighborhood effect is almost 15 times greater below the key point. Recall that just .1% of white teenagers lived in those very worst neighborhoods. But these results suggest that the ones who did dwell in bad neighborhoods were not immune to epidemics of dropping out.

For Hispanics, the pattern of increase is approximately linear across the whole distribution. The slope of the line fitted to all the observations is $-.002$, so neighborhood effects are fairly large. But there was no sharp jump. However, the validity of these results is open to question because of problems with the definition of Hispanic ethnicity in the 1970 census. Different criteria were used in different regions of the country. If the estimates are actually composites in which the dropout probabilities of very different populations are averaged together, sharp increases within each separate population could be masked. Of course, there is no evidence that this is, in fact, the case.

In order to determine whether these precipitous increases are either concentrated in or confined to urban ghettos, separate equations were estimated for blacks living in the largest cities and those living in other places.[15] Figure 2 shows the patterns of neighborhood effects for the two groups.

There are sharp jumps for both groups. In these cities, there are nontrivial neighborhood effects across the entire distribution. The estimated dropout probability rises from .063 to .145 as the percentage of high-status workers falls from 28.1% to 5%. The slope of the line fitted to

[14] The criterion for significance is the .05 level of a one-tailed test. The value at 31% is significant at the level of .10, but not at .05. The value at 3.5% is significant at .01.

[15] The "largest cities" were central cities of urbanized areas, in which the urbanized area had a population of more than 1 million.

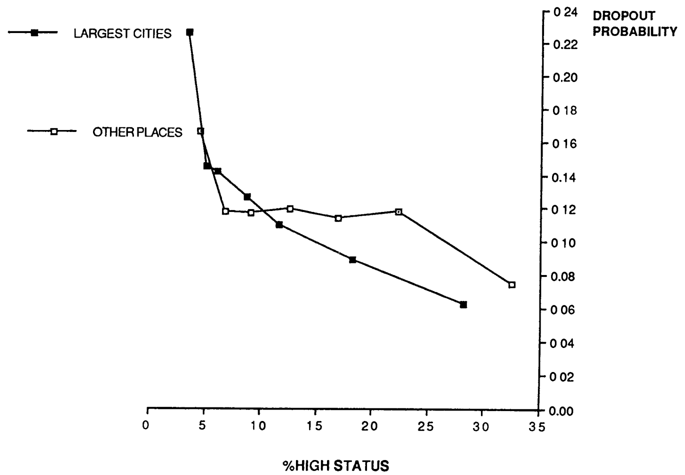

FIG. 2.—Estimated dropout probability as a function of the percentage of high-status workers in the neighborhood for blacks in the largest cities and other places.

these points is −.0036. But neighborhood effects are much larger at the very bottom. As %HIGH STATUS falls to 3.4%, the dropout probability leaps to .226, which is significantly greater than the value extrapolated from the fitted line. The slope below the key point is .051, more than 14 times greater than the slope above it.

The increase is even sharper outside the largest cities, but this was because there was virtually no neighborhood effect at all in the center of the distribution. The estimated dropout probability is almost constant around .118, between 22.2% and 9%. As the percentage of high-status workers falls to 4.5%, the probability rises to .116, which is significant. The slope below the key point is .022, over 300 times greater than the slope of the line fitted to the middle five points of the distribution. There is also a sharp decrease at the top of the distribution. The slope in that range is 60 times greater than the slope in the middle. But the estimated dropout probability is not significantly lower than the value extrapolated from the fitted line.

The level, the absolute increase, and the rate of increase in dropout probabilities below the key point are all greater within the largest cities than outside them. So these results suggest that epidemics may have occurred outside large cities but that they were worse in ghettos.

To ensure that these results are not just the result of the arbitrary grouping, several models were also estimated in which progressively smaller cities were included in the large-city category. As expected, the size of the jump diminishes as more cities are added, but it remains quite large. Equations were also estimated for whites inside and outside of the largest cities. The results were similar to those for blacks. There is a sharp increase at the bottom in each case, but it is significant only in the cities.

While all ghetto teenagers suffer from social problems, recently people have become particularly concerned about males. Separate models were estimated for black males and black females in the largest cities, to see if males were particularly prone to epidemics. Figure 3 shows the results.

There is a dramatic difference in the results, with an enormous increase in estimated dropout probabilities for black males in ghettos. As the percentage of high-status workers falls from just 5.6% to 3.4%, the probability explodes from .146 to .345, which is highly significant. The slope of the line between these two observations is huge, −.09. Neighborhood effects are almost 38 times greater below the key point than above it. In marked contrast, the estimated dropout probability actually decreases at the very bottom for black females.

It is possible that these results are accurate and that the special concern about males is more well placed than anyone ever imagined. But it is hard to believe that there could really have been a difference of this magnitude. Models were estimated for black males and females outside the largest cities and for white males and females within them. There are sharp increases at the bottom for all of these groups, and they are about the same size for both sexes in each case. This suggests that one or both of the observations at %HIGH STATUS = 3.4% may be a bad estimate. Random error could be responsible. The upper bound of the 95% confidence interval of the observation for females is .182, while the lower bound of the confidence interval for males is .229. Another possibility is bias. As discussed in Section V below, there is evidence that the observation may be biased downward for females.

There is evidence of epidemics of teenage childbearing too. Figure 4 presents the results for four groups: all black females, black females in the largest cities, all white females, and white females in the largest cities. None of the neighborhood dummies or the control variables are significant in the equation estimated for Hispanic females. This is probably because there are very few of them with children in the study sample (Crane 1989).

For all black females, the estimated probability of having a child rises from .082 to .124 as the percentage of high-status workers in the neighborhood falls from 31.2% to 5.6%. As %HIGH STATUS falls to 3.5%, the probability jumps to .161. The slope below the key point is 13 times greater than the slope above it, but the increase in the childbearing probability is not significant. Separate models were also estimated for black

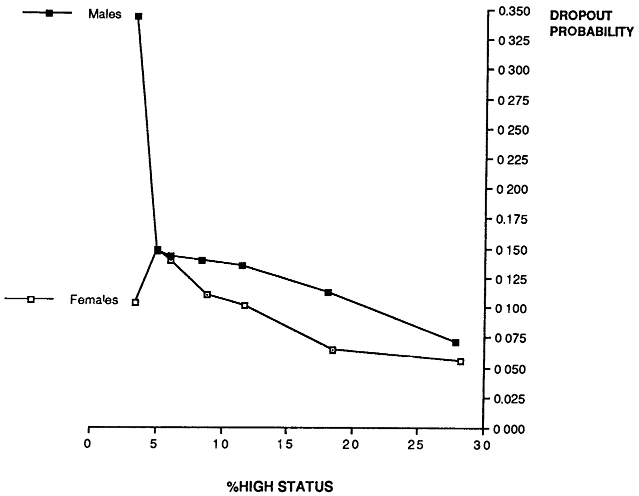

FIG. 3.—Estimated dropout probability as a function of the percentage of high-status workers in the neighborhood for black males and black females in the largest cities.

females in smaller cities, suburbs, towns, and rural areas. There are some nonlinearities, but none of the increases are significant.

For black females in the largest cities, the estimated probability of having a child rises from .074 to .120 as the percentage of high-status workers falls from 31.2% to 6%. As %HIGH STATUS falls to 3.5%, the probability jumps to .198. The sharp increase occurs at the third-to-the-last point rather than the second-to-the-last, as in all the other cases. Since these two observations are so close together, this aberration could have been generated by a small random error in either one. But the basic story is the same either way. The childbearing probability is highest by far in the very worst neighborhoods. Taking the third-to-last observation as the key point, the ratio of the slopes is 16. The increase in childbearing probabilities is significant for the last observation, though not for the second-to-last.

The results are similar for white females. There is a sharp increase at the bottom for all white females. But it is not significant for the group as a whole or for any subgroup outside the largest cities. Within these cities, however, neighborhood effects at the bottom of the distribution are huge. The estimated childbearing probability increases from .001 to .01, as the percentage of high-status workers falls from 46.3% to 7.5%. As %HIGH STATUS falls to 3.5%, the probability explodes to .102, which is easily significant. The slope below the key point is −.023,

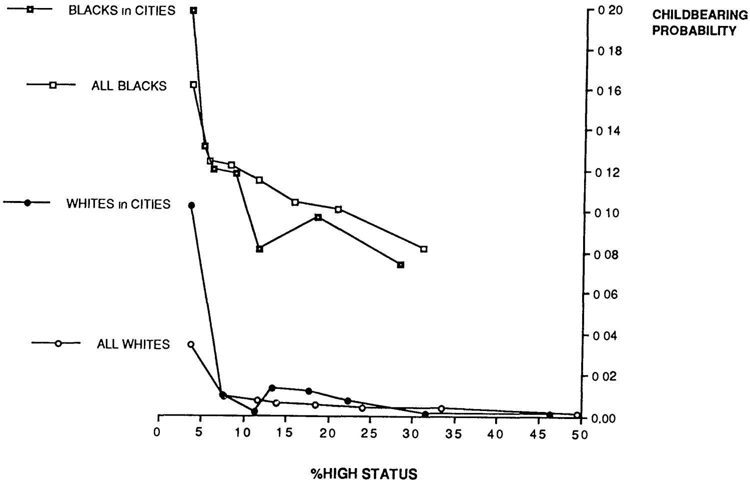

FIG. 4.—Estimated childbearing probability as a function of the percentage of high-status workers in the neighborhood for all blacks, blacks in the largest cities, all whites, and whites in the largest cities.

almost 100 times greater than the slope above it. Those few white females who lived in the worst neighborhoods of the largest cities were more like black females in terms of childbearing than other whites.

The pattern of results found here could be sensitive to the scale of measurement. The choice of scales were not completely arbitrary. The nonlinear pattern of effects in this linear scale suggests that there was a "social effect." If the pattern of effects had been linear, it could be argued that neighborhood effects were generated entirely by the high-status individuals directly and thus were directly proportional to the percentage of individuals in the community. The nonlinear results suggest that neighborhood effects were the product of social interaction, that the whole was greater than the sum of its parts, at least in the worst neighborhoods. Nevertheless, equations scaled in percentiles were also estimated. The results are essentially the same in each case, except that neighborhood effects tend to be a little larger at the very top of the distribution in the percentile scale.

Taken together, these results generally provide strong support for the hypothesis. Neighborhood effects on both dropping out and teenage childbearing were much larger in urban ghettos than anywhere else. Both the sharpness of the increases and the fact that they occur at virtually the same place in each distribution are striking. While any one of these jumps could have been generated by an unusually large random error, it is exceedingly unlikely that there would be so many large random errors in the same direction at the same point in the distributions. The absence of the pattern among Hispanics is not so troubling, given problems in the way the group was defined in the census. The absence of the pattern among black females in the largest cities is troubling, but that result may be biased.

V. BIAS

There are several different possible sources of bias in estimating neighborhood effects. These include sampling bias, measurement error, specification error, and endogeneity bias. Wherever possible, the analysis was structured so that estimates of both the overall neighborhood effect and the relative size of the effect at the bottom would be biased downward, in order to make these estimates as conservative as possible.

The single most important potential source of bias is the exclusion of certain teenagers from the study sample. Recall that teenagers not living with their parents had to be excluded because the direction of causality could not be identified for them. Individuals with inconsistencies or other problems in their records were also left out. Unfortunately, dropouts and childbearers were disproportionately excluded, and thus the total bias may be very large. Almost half the dropouts and about two-thirds of the childbearers are excluded from the study sample. Fortunately, the net bias is probably downward, especially in the worst neighborhoods. So, the estimates of neighborhood effects are probably conservative.

The estimates are biased if the correlation between the dropout (or childbearing) probabilities and neighborhood is different for the excluded group than for those in the study sample. If the partial correlation is more negative among the excluded teenagers than in the sample, then the estimates are biased downward. If the partial correlation is less negative or positive among the excluded teens, then the estimates are biased upward.

For the group that was excluded in the data-cleaning process, we have information on dropout rates, childbearing rates, and the percentage of high-status workers in the neighborhood, but we do not know how reliable this information is. The dropout and childbearing rates are known for those teenagers who were left out because they lived away from home. But we do not have any information about the neighborhoods they lived in when they dropped out or became pregnant. Their neighborhood at the time of the census may have been a good proxy for those who lived with relatives, but not for those living independently. A substantial proportion of teenagers living on their own probably moved out after (and often because) they left school and/or became pregnant. And the neighborhoods they moved to were almost surely poorer, on average, than the ones they moved from. So, any conclusions drawn for this group about the correlation between neighborhood quality and social problems, and thus about bias, have to be conditioned on assumptions about where they lived previously.

Table 1 compares the dropout and childbearing rates of various included and excluded groups. Teenagers left out of the sample were much more likely to have left school or had a child than those who were included. The dropout rates of the excluded teens are three-and-one-half times greater than those in the sample. Excluded females are more than eight times as likely to have given birth. This is not just a composition effect, resulting from the fact that blacks and Hispanics are disproportionately excluded. The incidence of these problems is also much higher among the excluded teenagers for minorities.

Teenagers excluded because of inconsistencies lived in very low quality neighborhoods. The mean of %HIGH STATUS is 23.0% for those included in the analysis but just 7.4% for those with inconsistencies. The dropout rate of this group is more than four percentage points higher than that of those from comparable neighborhoods in the sample. The childbearing rate is 12 percentage points higher.[16] So, if the data for this group are accurate, then the data cleaning may bias estimates of neighborhood effects downward, especially in the models of childbearing.

To analyze bias among teenagers living away from home, they were divided into three mutually exclusive groups: students in college dorms, teens living on their own or in institutions, and teens living with relatives other than their parents.

A little less than a quarter of those away from home were students in college dorms. By definition, their dropout rate is zero, while the childbearing rate of the females is .4%.[17] It is almost certainly true that these college students came from above-average neighborhoods. The incidence of these problems among teenagers from comparable neighborhoods in the sample was probably low, but not that low. Thus, excluding college students in dorms probably biases estimates of neighborhood effects downward.

Half of those who lived away from home were on their own or in institutions. As shown in table 1, their dropout and childbearing rates are extremely high. However, we do not know anything for certain about the neighborhoods where they lived when they dropped out or became pregnant. It seems likely that this group would have tended to originate

[16] For this comparison, I used teens in the sample who lived in neighborhoods in which the proportion of high-status workers was 7% or 8%.

[17] Note that, since the average age of 16–18-year-olds in dorms was 17.96, the relevant comparison group is 18-year-olds in the sample. Their childbearing rate was 5.8%.

TABLE 1

Dropout and Childbearing Rates of Groups Included in and Excluded from the Study Sample

Group	Total	Dropout Rate (%)	Females	Childbearing Rate (%)
Total PUMS	113,997	11.4	56,219	7.5
Study sample	92,512	7.7	44,466	3.0
Total excluded	21,485	27.4	11,753	24.6
Blacks in sample	10,459	12.7	5,280	12.4
Blacks excluded	3,952	32.8	1,942	35.7
Whites in sample	77,508	6.8	37,007	1.7
Whites excluded	16,236	25.0	9,092	21.0
Excluded by data cleaning	2,698	20.9	1,235	20.6
Not living with parent	18,787	28.3	10,518	25.1
Students in dorms	4,318	.0	2,221	.4
On own or in institution	9,368	43.5	5,569	38.6
Living with relatives	5,101	24.4	2,728	17.5

from low-quality neighborhoods with high rates of social problems that cause teenagers to strike out on their own or to become institutionalized. If so, excluding them probably biases estimates of neighborhood effects downward. However, it is conceivable that affluent teens from rich neighborhoods were more likely to live independently because they could get enough financial help from their parents. If that was the case, excluding this group could bias estimates upward.

A little more than a quarter of those away from home lived with relatives. The mean value of %HIGH STATUS for this group is 18.6%. Both the dropout and childbearing rates for these teens are considerably higher than for those in the sample from comparable neighborhoods.[18] If all of these teenagers actually grew up with their relatives, then leaving them out almost certainly biases the estimates downward. But these high rates of social problems probably reflect the fact that dropouts and teenage mothers were more likely to move in with relatives (which is precisely why they were excluded from the study sample). In that case, the size and direction of the bias depend on the difference in quality between the neighborhoods of the parents and the relatives.

If the parents lived in neighborhoods of quality similar to that of the relatives or better, then the correlation is more negative than in the sample and excluding these teenagers probably biases estimates down-

[18] For this comparison, I used teenagers in the sample who lived in neighborhoods in which the proportion of high-status workers was 18% or 19%.

ward. If the parents lived in neighborhoods that were a little bit or even a fair amount worse than the relatives, then there is probably very little bias. The estimates could only be biased upward if the parents lived in much worse neighborhoods than the relatives.

For the sake of comparison, models were estimated that included some or all of these groups. Since these models suffered from simultaneity bias, we would expect them to generate higher estimates of neighborhood effects, and they do. The estimates are consistently larger, by up to about three-fourths. There are sharp increases at the very bottom of the distributions. But in most cases the ratios of the slopes are a little smaller because neighborhood effects are relatively larger above the key point.

Analysis of biases on the overall neighborhood effects does not necessarily say much about whether the very large effects found at the bottom of the distributions are biased in any way. The dropout rate for excluded teenagers from neighborhoods with 4% or fewer high-status workers is 49.4%, compared to 19.9% for those in the sample. The childbearing rate for excluded females from these neighborhoods is 40.7%, compared to 16.5% for females in the sample. Of course, these differences probably result at least partly from endogeneity, that is, the tendency for teenagers with these problems to leave home and move to bad neighborhoods. But these differences are so large that they suggest that even the high dropout and childbearing probabilities in the worst neighborhoods may be biased downward by the exclusions.

Recall that there was not a sharp increase in the dropout probability for black females in the largest cities. One possible reason for this exception to the general pattern is that black females in ghettos were the most likely to be living away from home and thus excluded from the sample. If so, the estimate may be biased downward more for these females than for anyone else. It is conceivable that a sharp increase could be masked by such bias, though that bias would have to be quite large.

Among black females in the worst neighborhoods (i.e., %HIGH STATUS = 4) of the largest cities, the dropout rate for those in the sample is just 10.3%. But among the black females in these neighborhoods who were excluded from the sample, the rate is 50.0%. In marked contrast, the difference for black males in ghettos is only about 10 percentage points. This suggests that many more black females than black males who originated from these neighborhoods were excluded, probably because females with children were the type of teenager most likely to leave home. Thus, the estimate for black females in ghettos may very well be biased downward by a large amount. Whether the bias is in fact great enough to mask a sharp increase is impossible to say.

The census undercount of minorities is a second possible source of sampling bias. Supplemental surveys designed to identify people left out of the census suggest that those who were missed tend to lack family attachments (Bound 1986). So the census probably did not leave out many teenagers who lived at home or with relatives. But a nontrivial proportion of those living on their own may have been missed. It seems likely that those missed were disproportionately concentrated in bad neighborhoods and perhaps also in cities. The dropout rate was almost certainly extremely high among those who were left out. Thus, estimates of neighborhood effects on dropping out for blacks and Hispanics may be biased downward to some extent by the census undercount. The relative size of the childbearing rate among those missed is less certain because having a child may make a female less mobile and easier to find. But, in general, the teenagers left out of the census were probably the ones who were the most likely to develop social problems. Unfortunately, this is the catch-22 of statistical analyses of problems of the ghetto. It is hard to find the people who matter most.

Measurement error is another possible source of bias. The most fundamental problem in measuring neighborhood effects is defining the neighborhoods themselves. The concept of neighborhood is a little like the concept of obscenity. It is hard to define, but most people know it when they see it. In essence, a neighborhood is a geographic area with unbroken borders in which the density of social ties among residents is significantly greater than the density of ties between residents and nearby nonresidents.

In practice, neighborhoods are defined by using socioeconomic criteria, subjective conceptions, and/or geographic features to determine the boundaries. We really do not know how well these methods work. (We could get some idea of this by surveying social ties and measuring their correlation with operational boundaries.) Measurement error could very well be quite large even when all three methods are used together. This is probably the most fundamental problem in attempting to measure neighborhood effects. Recall that the census used only geographic criteria to define the neighborhoods in the sample used here. This aggravates the problem.

There is one saving grace, however. The bias that measurement error generates is unambiguously downward (i.e., toward zero), thus making estimates of neighborhood effects conservative. This is because overlaps between true boundaries and operational boundaries make the operationally defined neighborhoods weighted averages of actual separate neighborhoods. This adds an element of randomness to the measured association between neighborhood characteristics and the dependent variable.

There are two possible sources of specification error: structural misspecification of the model and omitted variable bias. Structural misspecification of the model may bias estimates of neighborhood effects downward. Since neighborhoods were not specifically identified in the census, it was not possible to specify a multilevel hierarchical model. In general, nonhierarchical specifications of multilevel data tend to bias estimates of group-level effects downward.

Omitted variables are a potential source of bias in any study, of course. In this particular case, some of the potentially important omissions are a historical record of the neighborhoods in which the teenagers grew up, an index of their parents' permanent income, measures of attitudes, achievements, and abilities, and an index of the social composition of the teenagers' schools.

It is quite plausible that neighborhoods begin to affect the probability that an individual will develop social problems very early in life; this influence probably continues at least through adolescence. If this is true, then the neighborhood measure used here is just a proxy for some duration and age-weighted index of all the neighborhoods a teenager ever lived in. To the extent that the proxy is imperfect, there is greater randomness in the association between neighborhood quality and the dependent variables, biasing estimates of neighborhood effects downward.

The same basic argument holds true for family income. Recall that family income in the census year had just a small effect on dropping out and no significant effect on childbearing for the whole sample—a surprising result. Among subgroups, family income made more difference for blacks, but its effects still were not very large. It is possible that what really matters is permanent income, and temporary income may not necessarily be a very good proxy for it. To the extent that this is true, the omission of a measure of permanent income may bias estimates of the neighborhood effect upward. However, measures of parental education and occupational status were included, and they may have been reasonably good proxies for permanent income.

There was a little historical information in the 1970 PUMS, and it does offer at least a little insight into the effect of omitting these two longitudinal variables. Recall that, while there was no measure of duration of residence in the neighborhood in the sample, there was a measure of duration of residence at the particular address. Surely this was highly correlated with length of stay in the neighborhood. All those who had been at the same address a long time had been in the neighborhood a long time. Given that census neighborhoods were quite small, it is almost certainly true that a nontrivial proportion of recent movers came from another neighborhood.

If long-term residence and long-term income were significantly better predictors than short-term residence and income, then there should have been some significant interaction effects between %HIGH STATUS and duration of residence and/or family income and duration of residence.

Several different models were estimated to test for interactions, but none were detected. While this examination was by no means definitive, it does suggest that the omission of longitudinal variables may not have been a serious problem.

If true, this result suggests some important things both about the way that neighborhood effects work and about the fundamental nature of dropping out and childbearing behavior.[19] It implies that these behaviors may not be determined by fundamental attitudes or deeply rooted personality traits, factors that affect a person for a long time or develop early in childhood. This result is consistent with epidemic theory, which is predicated on the assumption that the short-term dynamics of peer interaction are an important determinant of these problems. Finally, this result suggests some cause for optimism. If short-term phenomena are responsible for the problems, short-term policy interventions might be quite effective.

As noted in Section III above, cognitive ability, academic achievement, attitudes, and aspirations have been found to affect dropping out and/or childbearing. Omitting them may cause estimates of neighborhood effects on dropping out to have an upward bias. But these factors are probably important mechanisms of neighborhood effects. Their omission would only generate bias to the extent that they are correlated with differences in neighborhoods but unaffected by them.

The same basic argument applies to school effects. Part of the neighborhood effects found here may actually have been attributable to school effects. But it is also possible that schools were mechanisms of neighborhood effects. Although it is important to distinguish between neighborhood effects and school effects and to determine their relation, if any, this issue does not really affect the basic interpretation of the results here. Since the two effects cannot be distinguished, it might be more precise technically to call the overall effect found here a "social context" effect.[20] But whether the social processes that generated the sharp jumps occurred in neighborhoods, schools, or both, these sharp increases are no less striking. And, if anything, the epidemic theory makes even more sense when applied to schools because social networks are probably denser in schools than in neighborhoods.

There is no way to estimate the net effect of omitted variable bias. But there is one piece of evidence that suggests that it may not be that large. Most of the important omissions are intimately related to SES. To the extent that these factors affect dropping out and childbearing, it may be as proxies for unmeasured aspects of SES. If so, adding them to the

[19] I would like to thank Christopher Jencks for pointing this out.

[20] I would like to thank Anthony Bryk for pointing this out.

model might not reduce estimates of neighborhood effects by much. Among those measures of SES included in the model, there was a pattern of diminishing marginal effects as they were added sequentially.

Models were estimated without any measures of SES for several subgroups. Then the SES variables were added one at a time in different orders, and the models were reestimated. Adding either mother's education, father's education, or head's occupational status reduced the estimated dropout probability in the worst neighborhoods by approximately half for whites and a third for blacks. Adding the second of these reduced the estimates by about an additional fifth for both groups (with some variation depending on the order in which they were introduced). Adding the third of these reduced the estimates just a few percentage points more in each case. Adding family income as the fourth measure had almost no effect for whites and no effect at all for blacks.[21] The pattern of change in models of childbearing probability was quite similar. Of course, this pattern of diminishing marginal effects would not necessarily hold true for additional measures of SES. But it may indicate that the residual relationship between neighborhood quality and the two outcomes is unrelated to SES.

Another possible source of bias is the endogeneity of the control variables. The model here assumes that all of the control variables are unaffected by neighborhoods. But family income, occupational status, family structure, and residential tenure could each have been affected by neighborhood quality. If so, then estimates of neighborhood effects are biased downward. The total neighborhood effect would be the measured effect, which was direct, plus the sum of indirect effects through these other variables.

One final issue of bias needs to be addressed. Tienda (1989) argued that neighborhood effects in general may be impossible to distinguish statistically because any observed effect could be caused by selection bias. Individuals sort themselves in nonrandom ways among neighborhoods, and so there may be unobservable differences among them. (Of course, all models of behavior share this problem, and it begs the question of whether any kind of statistical inference about people is valid.) But one of the strengths of the epidemic theory is that it can be distinguished in such a way that this is not a big problem. To generate sharp increases in neighborhood effects, there would have to be some kind of "tipping"

[21] Family income had surprisingly small effects in general, especially for blacks. It did not have much of an effect on either dropping out or childbearing in any of the models. And when it was added as the first index of SES, it reduced estimates of the neighborhood effect in the worst neighborhoods by just a few percentage points in all the models. In those cases, the addition of the second index of SES most affected the estimates.

process (Schelling 1971) in the sorting along the lines of unobservable characteristics. While Schelling argued that neighborhoods do tip along racial lines, it seems highly implausible that they could tip along the lines of unobservable characteristics. For tipping to occur, people must be able to observe characteristics reliably, so that people generally agree on which individuals belong to which groups. In fact, the easy observability of race may be the very reason that it is such a powerful driving force in the development of in-group/out-group distinctions.

In sum, most of the potential biases are downward. I cannot be absolutely sure that the estimates are lower bound, because the magnitudes of the actual biases cannot be calculated. But, unless the net bias is upward and very large in ghettos specifically, the basic finding of sharp increases at the bottom of the neighborhood distributions in the largest cities is valid.

VI. CONCLUSION

The pattern of neighborhood effects on both dropping out and teenage childbearing was precisely the one implied by the epidemic hypothesis, for both blacks and whites. There were sharp jumps in the dropout probabilities of all blacks, black males, all whites, white males, and white females in the worst neighborhoods of the largest cities. There were also large jumps in childbearing probabilities for both black females and white females in these cities. All of these very large neighborhood effects were significant. Outside the largest cities, there were precipitous increases in the dropout probabilities of both blacks and whites and in the childbearing probabilities of black females. But these increases were not significant. In all the cases where there was a jump, it occurred at approximately the same point in the distribution, that is, in neighborhoods where about 4% of the workers held high-status jobs.

There was a huge jump in the dropout probability for black males in urban ghettos but not for black females. However, there was evidence that the estimate of the black females' dropout probability in these neighborhoods may have been biased downward, perhaps by a very large amount.

The pattern of neighborhood effects on dropping out was approximately linear among Hispanics. However, the validity of this finding is open to question since the Census Bureau used different definitions of Hispanic ethnicity in different regions of the country. There were too few Hispanic females with children in the study sample to distinguish any effects on childbearing.

In short, the evidence here does provide strong support for the epidemic theory. But, even though there is other evidence that supports the hypothesis (Crane 1989), more confirmation is needed. First of all, the data have to be updated. Second, additional ways to test the theory need to be found. Third, alternative explanations of the mechanisms of neighborhood effects that are consistent with these results should be developed, and then, of course, ways to distinguish between the competing theories would have to be worked out. Fourth, research that seeks to determine what types of teenagers leave home and where they go is needed to shed more light on the effects of bias in this study.

It would be premature to make policy recommendations on the basis of these results. The results are promising, however, so it makes sense to think about some policies that should be considered if this study is borne out by further work. The theory suggests that if we knew how to improve the "quality" of particular types of neighborhoods, such efforts could be very effective in reducing the incidence of social problems. Unfortunately, we do not know that much about how to improve neighborhoods. And, while the theory does suggest a couple of things about this, its clearest implications refer to targeting strategies. Thus, for the moment, let us assume that we do have effective policies for improving neighborhoods and consider issues of targeting. The substance of such policies is considered below.

The epidemic theory suggests that we should target two types of neighborhoods for policy interventions: neighborhoods that have already undergone epidemics and neighborhoods that are at high risk of doing so.

Recall that a key result of the model was that neighborhoods will tend to move toward equilibrium levels of social problems. Consider what this suggests about the effects of policies aimed at reducing the incidence of a social problem. Suppose that the incidence of a problem begins at some high equilibrium level and is then forced downward by a policy intervention. The incidence will naturally tend to move back to the high equilibrium unless the policy reduces it so much that it is drawn toward some low-level equilibrium. This implies that a policy intervention that is large but not quite large enough is little better than no intervention at all: if the policy fails to reduce the incidence of a problem enough, that problem will tend to revert back to a high level of incidence on its own. Therefore, a relatively large investment of resources would be required to reverse an epidemic in a community where it had already occurred, large enough to push the neighborhood down to the point where it moves naturally toward a low equilibrium.

Thus, in the context of limited resources, the optimal strategy would be one of "sequential saturation." In other words, if it were too expensive to reverse all epidemics at once, the best approach would be to concentrate resources in a fraction of places rather than to spread resources equally. An initial set of neighborhoods should be helped until they reach low equilibrium levels of social problems. Once attained, these low levels would be relatively inexpensive to maintain: neighborhoods would be pulled back naturally by the magnetic quality of the equilibrium if exogenous forces generated an increase in the incidence of a problem. Therefore, most of the resources could then be transferred to a second set of neighborhoods.

Another implication of the stability of low-level equilibria is that it would be relatively inexpensive to prevent epidemics in at-risk neighborhoods. Thus, we should take the opposite approach for resources aimed at the at-risk communities. These resources should be spread out rather than concentrated. One nice implication of the results of this study is that neighborhoods at risk might be relatively easy to identify, since all the epidemic effects were found at approximately the same point in the index of neighborhood quality.

The epidemic theory has less to say about how to improve neighborhood quality. But it is not totally silent on this question. Recall that it is predicated on the assumption that social problems spread through peer influence. This suggests that education programs aimed at reducing negative peer pressure and teenagers' susceptibility to it could be effective in preventing epidemics. In neighborhoods where epidemics have taken place, programs aimed at generating positive peer influence might help. It would be theoretically possible to generate a kind of reverse epidemic of positive behavior that would push the incidence of social problems down to some stable low-level equilibrium.

Whether or not the epidemic theory of ghettos is right, the empirical findings here are striking. And they may have important policy implications in and of themselves. They suggest that neighborhood improvement strategies might be very effective in reducing dropout and teen childbearing rates in ghettos and perhaps also in bad neighborhoods outside of large cities. If there are sharp increases in dropping out and early childbearing when neighborhoods descend to a particular threshold of deterioration, there might also be sharp decreases in these social problems if we could find a way of pushing neighborhoods back above these thresholds.

APPENDIX

Some Results from Correlation Analyses of Neighborhood Indices and Logit Models of Dropping Out and Teenage Childbearing

TABLE A1

THE TWO-WAY CORRELATION MATRIX FOR 10 INDICES OF NEIGHBORHOOD QUALITY

	%HIGH STATUS	Median Income	% Poor	% Female Headship	Median Education	% In Labor Force	% Unemployed	% Blacks	% Hispanics	Gini Coefficient
%HIGH STATUS	. . .	.767	−.515	−.391	.699	.291	−.363	−.325	−.131	−.137
Median income			−.744	−.450	.701	.552	−.396	−.373	−.167	−.494
% Poor				.595	−.675	−.639	443	.560	.223	.658
% Female headship					−.409	−.390	.404	698	.177	.375
Median education						.368	−.332	−.372	−.291	−.306
% In labor force							−.380	−.269	−.068	−.562
% Unemployed								.290	.202	.244
% Blacks									−.033	.318
% Hispanics										.095
Gini coefficient										. . .

NOTE —All of the correlations are significant at the level of .0001.

TABLE A2

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG BLACKS

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.5% HIGH STATUS	1.171 (.233)	.192	.140–.259
5.6% HIGH STATUS	.617 (.227)	.120	.086–.166
8.1% HIGH STATUS	.614 (.214)	.120	.088–.162
11.4% HIGH STATUS	.605 (.203)	.119	.088–.159
15.7% HIGH STATUS	.567 (.205)	.115	.085–.154
20.7% HIGH STATUS	.525 (.216)	.111	.080–.151
31.0% HIGH STATUS*	. . .	.069	.043–.109

NOTE.—N = 10,459.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A3

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG WHITES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.6% HIGH STATUS	1.198 (.279)	.146	.098–.213
7.7% HIGH STATUS	.650 (.099)	.090	.078–.104
11.7% HIGH STATUS	.596 (.099)	.086	.074–.099
13.8% HIGH STATUS	.455 (.099)	.075	.065–.087
18.3% HIGH STATUS	.407 (.089)	.072	.063–.083
23.9% HIGH STATUS	.233 (.089)	.061	.053–.070
33.2% HIGH STATUS	.215 (.096)	.060	.052–.070
49.4% HIGH STATUS*	. . .	.049	.040–.061

NOTE.—N = 77,508.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A4

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG HISPANICS

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
4.2% HIGH STATUS	.792 (.291)	.166	.110–.243
6.6% HIGH STATUS	.646 (.268)	.147	.100–.211
9.6% HIGH STATUS	.705 (.237)	.155	.110–.212
13.5% HIGH STATUS	.523 (.234)	.132	.094–.183
19.3% HIGH STATUS	.365 (.237)	.115	.081–.161
27.0% HIGH STATUS	.416 (.242)	.121	.084–.169
43.3% HIGH STATUS*	. . .	.083	.049–.136

NOTE.—N = 4,684.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A5

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG BLACKS IN THE LARGEST CITIES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.4% HIGH STATUS	1.466 (.275)	.226	.156–.314
5.0% HIGH STATUS	.925 (.271)	.145	.098–.209
6.0% HIGH STATUS	.887 (.245)	.142	.098–.196
8.6% HIGH STATUS	.769 (.224)	.127	.091–.173
11.6% HIGH STATUS	.610 (.236)	.110	.077–.154
18.2% HIGH STATUS	.372 (.248)	.089	.061–.128
28.1% HIGH STATUS*	. . .	.063	.033–.118

NOTE.—N = 3,498.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A6

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG BLACKS OUTSIDE OF THE LARGEST CITIES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
4.5% HIGH STATUS	.892 (.216)	.166	.122–.221
6.7% HIGH STATUS	.499 (.214)	.118	.086–.160
9.0% HIGH STATUS	.489 (.192)	.117	.088–.154
12.5% HIGH STATUS	.516 (.160)	.120	.095–.151
16.8% HIGH STATUS	.460 (.162)	.114	.090–.144
22.2% HIGH STATUS	.498 (.195)	.118	.089–.156
32.5% HIGH STATUS*	. . .	.075	.045–.123

NOTE.—N = 6,961.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A7

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG BLACK MALES IN THE LARGEST CITIES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.4% HIGH STATUS	1.920 (.346)	.344	.229–.481
5.0% HIGH STATUS	.816 (.345)	.147	.090–.235
6.0% HIGH STATUS	.775 (.309)	.143	.091–.217
8.4% HIGH STATUS	.749 (.295)	.140	.091–.209
11 5% HIGH STATUS	.708 (.295)	.135	.088–.203
18.0% HIGH STATUS	.493 (.312)	.112	.070–.174
27.9% HIGH STATUS*	. . .	.071	.029–.164

NOTE —N = 1,737.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A8

ESTIMATES AND CONFIDENCE INTERVALS FOR DROPPING-OUT RATES AMONG BLACK FEMALES IN THE LARGEST CITIES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.4% HIGH STATUS	.678 (.401)	.103	.056–.182
5.0% HIGH STATUS	1.090 (.394)	.148	.083–.249
6.0% HIGH STATUS	1.021 (.348)	.139	.084–.223
8.8% HIGH STATUS	.754 (.325)	.110	.068–.174
11.7% HIGH STATUS	.657 (.327)	.101	.062–.161
18.4% HIGH STATUS	.165 (.375)	.064	.036–.113
28.3% HIGH STATUS*	. . .	.055	.022–.132

NOTE —N = 1,761
* This is the omitted category Its confidence ınterval is calculated from the SE of the intercept.

TABLE A9

ESTIMATES AND CONFIDENCE INTERVALS FOR CHILDBEARING AMONG BLACKS

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.5% HIGH STATUS	.767 (.279)	.161	.108–.233
5.6% HIGH STATUS	.463 (.273)	.124	.083–.182
8.1% HIGH STATUS	.445 (.255)	.122	.084–.175
11.5% HIGH STATUS	.374 (.234)	.115	.081–.160
15.8% HIGH STATUS	.264 (.238)	.104	.073–.147
20.8% HIGH STATUS	.232 (.261)	.101	.068–.147
31.1% HIGH STATUS*	. . .	.082	.047–.139

NOTE.—N = 5,280.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A10

ESTIMATES AND CONFIDENCE INTERVALS FOR CHILDBEARING AMONG WHITE FEMALES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.6% HIGH STATUS	3.152 (.656)	.034	.012–.093
7.7% HIGH STATUS	1.851 (.256)	.009	.006–.014
11.7% HIGH STATUS	1.600 (.260)	.007	.005–.011
13.9% HIGH STATUS	1.467 (.240)	.006	.004–.009
18.4% HIGH STATUS	1.238 (.224)	.005	.004–.007
24.0% HIGH STATUS	.955 (.224)	.004	.003–.006
33.4% HIGH STATUS	.889 (.235)	.004	.002–.005
49.6% HIGH STATUS*	. . .	.001	.001–.003

NOTE.—N = 37,007.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A11

ESTIMATES AND CONFIDENCE INTERVALS FOR CHILDBEARING AMONG BLACKS IN THE LARGEST CITIES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.4% HIGH STATUS	1.126 (.357)	.198	.121–.308
5.0% HIGH STATUS	.632 (.354)	.131	.078–.213
6.0% HIGH STATUS	.530 (.309)	.120	.076–.185
8.8% HIGH STATUS	.516 (.292)	.118	.077–.178
11.7% HIGH STATUS	.093 (.294)	.081	.051–.125
18.4% HIGH STATUS	.296 (.336)	.097	.058–.158
28.3% HIGH STATUS*	. . .	.074	.034–.156

NOTE.—N = 1,761.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept.

TABLE A12

ESTIMATES AND CONFIDENCE INTERVALS FOR CHILDBEARING AMONG WHITES IN THE LARGEST CITIES

Variable	Coefficient (SE)	Estimated Dropout Probability	95% Confidence Interval
3.5% HIGH STATUS	4.513 (1.104)	.102	.018–.412
7.5% HIGH STATUS	2.093 (.917)	.010	.002–.044
11.3% HIGH STATUS	.362 (1.059)	.002	.000–.010
13.3% HIGH STATUS	2.377 (.828)	.013	.003–.05
17.7% HIGH STATUS	2.245 (.794)	.012	.003–.042
22.3% HIGH STATUS	1.689 (.798)	.007	.002–.024
31.5% HIGH STATUS	.108 (.846)	.001	.000–.006
46.3% HIGH STATUS*	. . .	.001	.000–.008

NOTE.—N = 4,084.
* This is the omitted category. Its confidence interval is calculated from the SE of the intercept

REFERENCES

Amemiya, T. 1981. "Qualitative Response Models: A Survey." *Journal of Economic Literature* 19:1483–1536.

Bound, J. 1986. "Appendix: NBER-Mathematics Survey of Inner-City Black Youth: An Analysis of the Undercount of Older Youths." Pp. 443–60 in *The Black Youth Employment Crisis,* edited by R. B. Freeman and H. J. Holzer. Chicago: University of Chicago Press.

Corcoran, M., R. Gordon, D. Laren, and G. Solon. 1987. "Intergenerational Transmission of Education, Income and Earnings." Unpublished manuscript. University of Michigan, Ann Arbor.

Crane, J. 1989. "The Epidemic Theory of Ghettos." Working paper, Center for Health and Human Resources Policy Discussion Paper Series, Cambridge, Mass.

Datcher, L. 1982. "Effects of Community and Family Background on Achievement." *Review of Economcs and Statistics* 64:32–41.

Fischer, C. S., et al. 1977. *Networks and Places: Social Relations in an Urban Setting.* New York: Free Press.

Granovetter, M. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83:1420–43.

Granovetter, M., and R. Soong. 1983. "Threshold Models of Diffusion and Collective Behavior." *Journal of Mathematical Sociology* 9:165–79.

Hogan, D. P., N. M. Astone, and E. M. Kitagawa. 1985. "Social and Environmental Factors Influencing Contraceptive Use among Black Adolescents." *Family Planning Perspectives* 17:165–69.

Hogan, D. P., and E. M. Kitagawa. 1985. "The Impact of Social Status, Family Structure, and Neighborhood on the Fertility of Black Adolescents." *American Journal of Sociology* 90:825–55.

Jencks, C., and S. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood: A Review." In *Concentrated Urban Poverty in America,* edited by M. McGeary and L. Lynn. Washington, D.C.: National Academy.

Jencks, C., M. Smith, H. Acland, et al. 1972. *Inequality.* New York: Harper & Row.

Kandel, D. B. 1980. "Drug and Drinking Behavior among Youth." *Annual Review of Sociology* 6:235–85.

Liebow, E. 1967. *Talley's Corner.* Boston: Little, Brown.

MacLeod, J. 1987. *Ain't No Makin' It.* Boulder, Colo.: Westview.

Rainwater, L. 1970. *Behind Ghetto Walls.* Chicago: Aldine.

Ricketts, E., and I. Sawhill. 1988. "Defining and Measuring the Underclass." *Journal of Policy Analysis and Management* 7:316–25.

Schelling, T. C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143–86.

———. 1978. *Micromotives and Macrobehavior.* New York: Norton.

Sewell, W. H., R. M. Hauser, and W. C. Wolf. 1980. "Sex, Schooling, and Occupational Status." *American Journal of Sociology* 86:551–83.

Steinberg, L. 1987. "Single Parents, Stepparents, and the Susceptibility of Adolescents to Antisocial Peer Pressure." *Child Development* 58:269–75.

Tienda, M. 1989. "Poor People and Poor Places: Deciphering Neighborhood Effects on Behavioral Outcomes." Paper presented at the annual meeting of the American Sociological Association, San Francisco. August 9–13.

U.S. Bureau of the Census. 1973. *Supplement no. 1 to Public Use Samples of Basic Records from the 1970 Census: Description and Technical Documentation.* Washington, D.C.: Government Printing Office.

Varaiya, P., and M. Wiseman. 1984. "Bifurcation Models of Urban Development." Pages 61–88 in *Regional and Industrial Development Theories: Models and Empirical Evidence,* edited by A. E. Anderson, W. Isard, and T. Puu. Amsterdam: Elsevier Science.

Zelnick, M., J. F. Kantner, and K. Ford. 1981. *Sex and Pregnancy in Adolescence.* Beverly Hills, Calif.: Sage.