Excerpt from

Neighborhoods and Health:

Building Evidence for Local Policy

By Kathryn L.S. Pettit, G. Thomas Kingsley, and Claudia J. Coulton
With Jessica Cigna

May 2003

SUBMITTED TO
The Office of the Assistant Secretary for Planning and Evaluation
U.S. Department of Health and Human Services

SUBMITTED BY
The Urban Institute
2100 M Street NW, Washington DC

Delivery Order 19, Contract No. HHS-100-99-0003.

Section 10

ECOLOGICAL CORRELATIONS

INTRODUCTION

As noted, this section formally analyzes the relationships between the contextual indicators (selections from the neighborhood conditions introduced in section 8) and health-related indicators (reviewed in section 9). We first test the hypotheses set out in section 7 by examining the bivariate correlations of contextual characteristics with four of the health indicators from section 9: teen birth rates, early prenatal care rates, low birth weight birth rates, and age-adjusted death rates.[38] This analysis is based on three-year averages for these indicators as in the last section, but the indicators are calculated for individual census tracts as the geographic unit of analysis instead of the high-poverty/nonpoor groupings used previously. We review the results for all the cities together and for each city separately.

We next move to multivariate analysis with a select set of contextual variables to test the independent relationship between a selected set of contextual indicators and the four health indicators listed above. The multivariate analysis reveals three things: (1) the independent association between the dependent health variable and the tract condition; (2) how much of the variation of each health indicator is explained by the five racial and socioeconomic factors versus other conditions in the city and time period; and (3) whether some of the key shifts in health indicators are statistically significant or due to random fluctuation of small numbers of events.

BIVARIATE CORRELATION ANALYSIS

The bivariate correlation analysis is organized around the hypotheses that were introduced in section 7. Under each static hypothesis, we discuss the aggregate and site-specific results from the data. As noted, we use the three-year rolling average approach to calculate values for the health variables, but we calculated them for all individual census tracts

[38] Infant mortality rates are not included because of very small numbers at the census tract level.

in all study sites. Each overlapping three-year average health indicator rate for each census tract is an observation. This method yielded a total of over 8,400 observations across the five sites. [39]

For the contextual indicators derived from the census, we interpolated annual estimates for all tracts based on 1990 to 2000 trends. The census indicator for a given year was related to the value of the health indicators as of the mid-point of the three-year period employed for them. For example, each 1991/1993 health indicator value is paired with the 1992 estimated poverty rate, percentage female-headed households, etc. To test our dynamic hypotheses, we compare the correlation coefficients for two time periods: 1990 to 1995 and 1995 to 2000.

STATIC HYPOTHESES

Socioeconomic conditions

Hypothesis: Census tracts with a majority non-white population and higher levels of immigrants will have higher levels of mortality and poor maternal and infant health outcomes than majority white census tracts.

Our first hypothesis addresses race, ethnicity, and nativity. As shown in table 10.1, the aggregate results are all significant in the expected directions for the overall percentage minority and the African-American and Hispanic percentages of the population. However, when looking at the individual city correlations, the share of the population that is Hispanic significantly correlates for low-birth weight rates only in Cleveland. The remaining results are mostly consistent for all cities and indicators, with a few notable exceptions. (For the complete list of city-specific correlations, see annex table C.23.) In Providence, the relationships between low-birth weight rates and percentages minority and African-American are slightly positive but not significantly correlated. In Oakland, teen births and the percentage African-American were not significantly associated.

Relationships with the percentage foreign born are more complicated. Overall, high percentages of foreign-born population are associated with worse prenatal care and teen birth rate outcomes as expected, but with *better* low-birth-rates. [40] The correlation with age-adjusted death rates was not significant. On closer examination of the city-specific correlations,

---

[39] Census tract codes for Cleveland mortality data from 1990 to 1996 were truncated to four digits. We aggregated the Cleveland death data from 1997 to 2000 to the 4-digit tracts in order to compare rates across time.

[40] This is consistent with previous research documented in Vega 2001 and Weigners 2001.

Table 10.1: Cross Site Analysis
Correlations between Health Indicators and Socioeconomic Conditions by Census Tract

	Pct. low birth weight		Pct. prenatal care in first trimester		Teen birth rate		Age-adjusted death rate	
Census tracts with a majority non-white population and higher levels of immigrants will have higher levels of mortality and poor maternal and infant health outcomes than majority white census tracts.								
Percent minority	0.53	*	-0.56	*	0.26	*	0.36	*
Percent African-American	0.54	*	-0.45	*	0.19	*	0.32	*
Pct. Hispanic	0.05	*	-0.34	*	0.18	*	0.13	*
Percent foreign-born population	-0.09	*	-0.02	*	0.08	*	0.00	
Low-income census tracts as measured by poverty rate, median income,and public assistance and AFDC/TANF recipiency will be associated with poorer scores on the mortality and the maternal and infant health measures than higher income tracts.								
Poverty rate	0.52	*	-0.66	*	0.33	*	0.59	*
Average family income	-0.37	*	0.56	*	-0.27	*	-0.45	*
Pct. pop. receiving public assistance	0.52	*	-0.53	*	0.28	*	0.52	*
Pct. pop. receiving AFDC/TANF**	0.52	*	-0.62	*	0.39	*	0.43	*
Census tracts with higher social risk factors, as measured by lower education, low employment rates, and higher shares of female headed families, will have worse scores on the mortality and the maternal and infant health measures than those with lower values.								
Pct. pop. age 25 and over with no HS degree	0.39	*	-0.66	*	0.37	*	0.55	*
Percent population age 16 and over not employed	0.42	*	-0.45	*	0.24	*	0.48	*
Percent fem-headed HH of HH w/kids	0.58	*	-0.63	*	0.28	*	0.48	*
Percent mothers w/HS education	-0.31	*	0.67	*	-0.35	*	-0.48	*

* Significant at the .05 level.

**AFDC/TANF correlations only include Cleveland, Denver, and Indianapolis because administrative data was not available for Oakland and Providence data reflected households instead of individual level data.

the relationships with health indicators are split. Cleveland and Indianapolis show *better* outcomes in tracts with higher immigrant levels, but the immigrant areas in the three more Latino cities have *worse* birth and death indicators.

Hypothesis: Low-income census tracts, as measured by poverty rate, median income, and public assistance and AFDC/TANF recipiency, will be associated with poorer scores on the mortality and the maternal and infant health measures than higher income tracts.

The next group of characteristics describes the economic conditions of the neighborhood. As many studies have found personal economic status to be linked with birth and mortality outcomes, it is not surprising that the aggregate correlations show the same relationships at the census tract level. These correlations are overwhelmingly strong and consistent across cities and measures. As with the racial variables, low birth weight in Providence and teen birth rates in Oakland are the two exceptions.

Hypothesis: Census tracts with higher social risk factors, as measured by lower education, low employment rates, and higher shares of female-headed families, will have worse scores on the mortality and the maternal and infant health measures than those with lower values.

In aggregate, all of the social risk factors—low levels of education and employment, and higher percentages of female-headed households with children—were significant and in the hypothesized direction. At the city level, the indicators were not always significantly related to low birth weight in Providence or teen birth rates in Oakland.

Physical stressors

Census tracts with poor housing quality, as measured by older housing, overcrowded units, and lower home values, will have higher levels of mortality and worse maternal and infant health outcomes than stronger housing markets.

For the next group of contextual variables, we use age of housing, overcrowded conditions, and home values as proxies for physical housing quality. Table 10.2 shows that indicators of better physical housing conditions are generally associated with better maternal and mortality outcomes as hypothesized, though the relationship is weaker than we saw with the socioeconomic indicators. Looking at city-by-city correlations (annex C.23) the only notable findings inconsistent with the hypothesis were that older housing was significantly related to *better* maternal and infant outcomes in Oakland and Providence.

Table 10.2: Cross Site Analysis
Correlations between Health Indicators and Physical Stressors by Census Tract

	Pct. low birth weight		Pct. prenatal care in first trimester		Teen birth rate		Age-adjusted death rate	
Census tracts with poor housing quality, as measured by age of the housing, overcrowded units, and home values, will have higher levels of mortality and poor maternal and infant health outcomes than stronger and more stable markets.								
Pct. housing units built before 1960	0.16	*	-0.17	*	0.17	*	0.40	*
Pct. overcrowded units	0.14	*	-0.31	*	0.25	*	0.22	*
Avg. owner-occupied home values	-0.26	*	0.44	*	-0.22	*	-0.33	*
Avg. amount of home purchase mortgage ($)	-0.23	*	0.45	*	-0.28	*	-0.06	*

* Significant at the .05 level.

Table 10.3: Cross Site Analysis
Correlations between Health Indicators and Social Stressors by Census Tract**

	Pct. low birth weight		Pct. prenatal care in first trimester		Teen birth rate		Age-adjusted death rate	
Hypothesis: Census tracts with high total, violent or property crime rates will have poorer scores on the mortality and the maternal and infant health measures than safer communities.								
Total Part I Crimes per 1000 population	0.07	*	0.02		0.30	*	0.23	*
Property Crimes per 1000 population	0.07	*	0.02		0.29	*	0.22	*
Violent Crimes per 1000 population	0.07	*	0.00		0.35	*	0.27	*

* Significant at the .05 level.
** Crime correlations do not include Providence, and reflect different years for each city, depending on data availability. See Table 7.2 for details.

Social stressors

Hypothesis: Census tracts with high total, violent, or property crime rates will have poorer scores on the mortality and the maternal and infant health measures than safer communities.

As in section 8, we used data on crime rates provided by our local partners as a proxy for social stressors in neighborhoods. Table 10.3 shows that census tracts with higher levels of crime generally have worse maternal and mortality outcomes as hypothesized, but that relationship did not hold for early prenatal care rates. The correlations are much stronger for teen birth rates and age-adjusted death rates than for low-birth weight rates. In Oakland, there is a remarkable 93 percent correlation between teen birth rates and violent crime rates, although the results for low-birth weight rates are insignificant. Although the aggregate calculation showed no association between crime and prenatal care rates, higher crime rates do have a substantial and negative association with prenatal care rates in Indianapolis.

Social networks

Hypothesis: Census tracts with less stable populations, as measured by renter occupancy, vacancy rate, and mobility rate, will have higher levels of mortality and worse maternal and infant health outcomes than stronger and more stable markets.

In this grouping, we select crude proxies for social networks, with the idea that less stable neighborhoods will have less opportunity for neighborhood cohesion than more stable ones. Overall, the correlations between the health indicators and our mobility variables confirm the hypothesis (see table 10.4). City-by-city results are not always consistent, however (annex C.3). The expected inverse relationship between renter occupancy and positive health outcomes holds generally, but the results for vacancy rate and percentage moved vary across cities and health indicators. Only the vacancy rates and percentage moved in Cleveland and Denver are consistently associated with worse health effects.

Hypothesis: Places with less change in total or minority population or a higher rate of home improvement or refinancing loans will have better mortality and birth outcomes.

The correlations for all three of these contextual indicators representing the social network category are in the opposite direction of the hypothesis, with more population change and fewer home improvement loans associated with better outcomes (see table 10.4). Gentrification is one possible explanation for the better maternal outcomes associated with greater population change, though more research would need to be done to confirm this. The separate city correlations are often inconsistent for the percentage population change and the rate of home improvement, with some positive and some negative values. However, larger percentage change in the minority population relates to better indicators for all cities except Providence.

Table 10.4: Cross Site Analysis
Correlations between Health Indicators and Social Network Indicators by Census Tract

	Pct. low birth weight		Pct. prenatal care in first trimester		Teen birth rate		Age-adjusted death rate	
Census tracts with less stable populations, as measured by renter-occupancy, vacancy rate, and mobility rate, will have higher levels of mortality and worse maternal and infant health outcomes than stronger and more stable markets.								
Pct. renter-occupancy	0.35	*	-0.51	*	0.25	*	0.42	*
Rental vacancy rate	0.20	*	-0.26	*	0.15	*	0.21	*
Pct population age 5 and over in different house 5 yrs ago	0.08	*	-0.29	*	0.10	*	0.11	*
Places with less change in total or minority population or a higher rate of home improvement or refinancing loans will have better mortality and birth outcomes.								
Pct. change in total population, 1990-2000	-0.11	*	0.06	*	-0.01		-0.16	*
Pct. change in minority population, 1990-2000	-0.19	*	0.17	*	-0.08	*	-0.14	*
Rate of home improvement loans	0.04	*	-0.07	*	0.06	*	0.23	*

* Significant at the .05 level.

DYNAMIC HYPOTHESES

Hypothesis: The correlation between high-minority tracts and poor birth and mortality outcomes will remain positive, but will have decreased over the 1990s.

Consistent with the hypothesis, table 10.5 shows that the relationships between percent minority and early prenatal care levels and low birth weight rates weakened from the 1990–1995 to the 1995–2000 period but still continued to be significant and substantial. We see that for low-birth- weight rates, the strength of the relationships fell for both percentage African American and percentage Hispanic, while the declines in association with low early prenatal care rates occurred only for the percentage African American. For overall percentage minority and percentage African American, the correlation coefficients for age-adjusted deaths changed very little from the beginning to the end of the decade. There was a decrease, however, in the magnitude of the association with the percentage Hispanic.

Contrary to our hypotheses, however, in the late 1990s teen births appear to be *more concentrated* in minority (both black and Hispanic) areas than in the early 1990s. The pattern of much higher correlations occurs with the vast majority of our indicators and for all cities. After the tremendous progress seen over the decade in this indicator, we suspect that the places remaining with high teen births tend to be in the more segregated and distressed areas.

Hypothesis: The correlation between low-income tracts and births with late or no prenatal care will remain positive, but will have decreased over the 1990s. The correlation between low-income tracts and high rates of low-birth weight births will remain positive, but will have decreased over the 1990s.

Consistent with the hypothesis, lower early prenatal care rates are slightly less correlated with the income variables in 1990–1995 than in 1995–2000. More striking, the association between low-birth weight rates and tract economic conditions has diminished considerably over the 1990s. This is likely due to several interacting factors, including (1) advanced medical technology enabling an increasing number of low-birth weight infants born in high-poverty and nonpoor areas to survive, (2) the increase of low-birth weight infants due to women having children at later ages, and (3) reductions in low-birth weight infants in poor areas due to better and earlier prenatal care.

Table 10.5: Cross Site Analysis
Dynamic Correlations between Health Indicators and Racial and Economic Conditions by Census Tract

	Pct. low birth weight		Pct. prenatal care in first trimester		Teen birth rate		Age-adjusted death rate	
Dynamic Hypothesis: The correlation between high-minority tracts and poor birth and mortality outcomes will remain positive, but has decreased over the 1990's.								
Percent minority								
1990/1992 - 1993/1995	0.57	*	-0.62	*	0.24	*	0.36	*
1995/1997 - 1998/2000	0.47	*	-0.49	*	0.40	*	0.36	*
Percent African-American								
1990/1992 - 1993/1995	0.57	*	-0.53	*	0.19	*	0.32	*
1995/1997 - 1998/2000	0.49	*	-0.35	*	0.27	*	0.32	*
Pct. Hispanic								
1990/1992 - 1993/1995	0.06	*	-0.33	*	0.17	*	0.15	*
1995/1997 - 1998/2000	0.03	*	-0.37	*	0.33	*	0.10	*
Dynamic Hypothesis: The correlation between low-income tracts and births with late or no prenatal care will remain positive, but has decreased over the 1990's. The correlation between low-income tracts and high rates of low birth weight births will remain positive, but has decreased over the 1990's.								
Poverty rate								
1990/1992 - 1993/1995	0.57	*	-0.67	*	0.32	*	0.59	*
1995/1997 - 1998/2000	0.44	*	-0.64	*	0.48	*	0.58	*
Avg family income								
1990/1992 - 1993/1995	-0.41	*	0.58	*	-0.25	*	-0.43	*
1995/1997 - 1998/2000	-0.31	*	0.55	*	-0.42	*	-0.48	*

* Significant at the .05 level.

MULTIVARIATE ANALYSIS

Purposes

Once the bivariate analysis confirmed most of our hypothesized relationships, the next step was to develop a multivariate model. The multivariate regression allows us to go beyond the findings of the bivariate correlations in three ways. First, it enables us to look at the influences of several variables simultaneously. As stated before, many of the census tract contextual conditions are correlated with each other. Lower income tracts also tend to have higher shares of minority population and higher social risk factors, like female-headed families and welfare recipiency. Multivariate analysis allows us to identify the independent association between a health condition and each contextual variable, holding all other variables constant.

Second, in addition to estimating the strength and direction of the independent relationships, the model used in this analysis allows us to separate out how much of the variation among cities can be explained by underlying differences in the contextual variables. As we viewed trends in section 9, we sometimes speculated that rate or trend differences between cities were due to demographic and racial/ethnic differences. For example, we stated that Denver and Providence (the more Hispanic cities) might not be experiencing the same magnitude of teen birth reductions as Cleveland and Indianapolis (the more African-American cities). In this section, we will quantify how much of the rate differences are accounted for by the five contextual indicators we specify. The remaining unexplained difference relates to other factors particular to each city and not specified in our model, including health services, policy initiatives, or census tract characteristics we omitted.

The third benefit of the multivariate models is to test shifts over time in our dependent variables for statistical significance. We know from the discussion of rare events in section 7 that rates based on a small number of events can fluctuate widely from year to year, so any given annual rate may not represent the underlying true rate. Statistical tests allow us to sort out changes due to this “white noise” from true underlying trends. For example, in section 9, we saw that early prenatal care rates in Providence fell from 1997/1999 to 1998/2000. From looking at simple line graphs, it is impossible to tell whether this is a disturbing change due to lower shares of women receiving timely prenatal care or a random shift caused by few events that may well bounce back the next year. In this section, we identify key shifts in indicators and test whether they represent statistically significant changes or not.

Methodology

We implemented four regression models, each with one of our four health indicators as the continuous dependent variable. As with the bivariate analysis, each overlapping three-year average health indicator rate for each census tract is an observation. Each observation was weighted by the total number of births to account for the varying degree of precision in tracts of different sizes

As to the independent variables, we employed only a subset of the contextual indicators used in the bivariate analysis, primarily because most of the latter were correlated with one another and using them all would have biased the models. The final selection was based on a combination of the strength of correlation, how well the variable represented the concept, and the independence of the variable in relation to the others. Unfortunately, we had to exclude all of our indicators for physical conditions because they were all too closely correlated with the other factors.[41] Crime rates were also omitted because they were not available for all of the tracts in the cities for all years. We finally selected five tract-level variables to serve as the independent variables in all four models: percentage African American, percentage Hispanic, average family income, percentage not employed, and percentage of population that moved in the past five years.

Table 10.6: Coefficients for Contextual Independent Variables in the Multivariate Regressions

	Dependent Variable			
	Teen birth rate	Early prenatal care rates	Low birth weight rates	Age-adjusted death rates
R-squared	0.45	0.77	0.56	0.46
Percent African-American population	0.04 *	-0.13 *	0.06 *	1.45 *
Percent Hispanic population	0.09 *	-0.23 *	-0.02 *	-0.44
Average Family Income (000)	-0.06 *	0.11 *	-0.01 *	-2.58 *
Pct of population age 16 and over that is not employed	0.20 *	-0.22 *	0.09 *	10.35 *
Pct pop. age 5 and over who moved in past five years	0.07 *	-0.11 *	0.03 *	8.02 *

* Indicates significance at the .001 level.
Note: For full model results, see Annex Tables C.24 - C.27.

[41] Average home value and average mortgage origination were 80 percent correlated with average family income, and percentage overcrowded was 60 percent correlated with the percentage of the population that is Hispanic.

In addition to the five variables describing tract characteristics, we include three additional series of variables: city dummies, year dummies, and city-year interactions. The first set of city dummy variables controls for differences in the health-related rates solely due to conditions in the each city (other than the five we specified), while the second controls differences due to the time period of the rate. Each dummy variable is coded 1 to indicate the presence of specific attributes for a case and 0 to indicate their absence. So, for Denver observations the variable d_denv = 1; for observations in the other four cities, it would equal zero. All the possibilities within a set of dummy variables cannot be included in the model, since information about all but one of the dummies determines the value of the last category. For example, if you know the values of four of the city dummies are zero, you can figure out the observation is in the fifth city. This means the fifth city variable would not be independent of the other four, and bias the model. To account for this, one city and one year need to be left out of the series. For this analysis, we chose Cleveland and 1999 as the omitted choices.

The third set of variables consists of interaction variables between the city and year of the rate, calculated by multiplying the city and year dummy variables. These variables control for differences in the rates due to particular conditions in a city in a given time period.

The paragraphs below present the results for our four independent variables. Under each, we offer findings in three areas paralleling the three purposes of the multivariate analyses described above. First, we discuss the strength of the overall model, noting the R-squared and the level and significance of individual coefficients (table 10.6).

Second, we examine the overall explanatory power of the contextual variables using data in table 10.7. The values in the table present results for four cities in reference to how much they differ from the Cleveland value in 1999 (since as stated above Cleveland was the reference city and 1999 the reference year for the dummy variables in the model). These differences are averaged across all years for which data were available.

For each indicator, the first line ("average difference in city rates") is calculated by averaging the difference in a city's overall rates from Cleveland's rates. The second line under each indicator is the estimated difference in a city rate from the Cleveland rate that is due to differences in the five contextual characteristics. In section 9, for example, we referred to the fact that low birth weight rates are higher for more African-American areas like in Cleveland. The third line is the estimated difference in rates that is due to "unobserved characteristics" (i.e., not explained by the contextual variables), and that amount is important in interpreting results in each city[42]. These unobserved factors could be the characteristics of the individuals in the city or characteristics of the census tracts not included as independent variables in the model. They could also include influences of programs aimed at improving the particular health indicator (like Healthy Start) or barriers to healthy outcomes, such as lack of insurance or appropriate care facilities

Finally, we examine shifts in the indicator trends over the decade, charting year-by-year changes in the “due to unobserved characteristics” variable (using differences from the Cleveland 1999 values to standardize). This is also important to interpretation. For example, if the trend for teen birth rates in a city goes down on this chart it means that a total decline that might have been observed in section 8 was not totally due the contextual variables (e.g., changes in the race/income indicators) but was also partially explained by something else (e.g., local programs, broader changes in attitudes). We also note whether these trends are statistically significant. (Full data on these results is found in annex tables C.27-C.30.)

Teen Birth rates

Overall strength of model. In the first regression model, the teen birth rate for mothers age 15–19 is the dependent variable. The R-squared is moderately strong, with the independent variables accounting for 45 percent of the variation in the model. The percentage of the population not employed has the highest coefficient, with a 1 percentage point change relating to a 0.20 point change in teen birth rates (table 10.6). All of the city dummies are significant, indicating that the levels of teen birth rates vary by city characteristics not captured by the model. Interestingly, only the year dummy variables through 1994 are significant, reflecting unique conditions in those years. From section 9, we know that these were the years of the largest decreases in teen births. The only significant interaction variables are for three early years in Oakland, signaling the effect of unmeasured city conditions in Oakland in those particular years above and beyond the overall city and year influences.

Explanatory power of contextual variables. Over the 1990s, Denver's teen birth rate averaged about 4 points above the rate of Cuyahoga County (table 10.7). Summing of the regression coefficients for the city, year, and interaction dummies gives us the percentage points’ difference that is not explained by the contextual characteristics included in the model. In this case, 2 percentage points’ difference is explained by the contextual variables (percentage non-Hispanic black, percentage Hispanic, etc.) and another 2 percentage points are due to factors that are unique to Denver during this period that the model does not measure. For

[42] The difference in rates due to unobserved characteristics calculated for each city and year by first adding the regression coefficients of city dummy, the year dummy, and the city-year interaction term. This value by city and year is then subtracted from the value for Cleveland (the reference site) in that year.

Indianapolis, the census tract characteristics would predict that the average teen birth rate over the decade would be about 0.8 points *lower* than Cleveland's rate, but, in fact, the rate was 1.4 points *higher*. This indicates that there are conditions in Indianapolis that increase the teen birth rate to 2.2 points above the predicted level.

Both Oakland and Providence show the opposite situation. The model predicts that with the contextual characteristics of their census tracts (both poorer and more Hispanic than Cleveland), their rates would both be above that of Cleveland. However, there must be some unobserved beneficial factors that bring the expected higher rates down for both cities (1.6 points lower in Oakland and 2.4 points lower in Providence).

Shifts in indicator trends. Figure 10.1 is the graphic illustration of the year-to-year change in Thus, the differences in the chart have already taken into account that Providence has generally lower incomes than Cleveland (Cuyahoga County), and that Denver has a greater share of Hispanic population, and so on. Again, the rate differences are all expressed in relation to the 1999 Cleveland rate (the reference city and year for the dummy variables). The figure looks very similar to the graph of teen birth rates from section 9, with all cities showing downward trends. We can see that Oakland's rate is declining at a faster pace. For the four cities for which we have complete data series (all except Providence), the changes from 1991 to 1999 are all highly significant (at the .001 level). The Providence trend is not statistically significant despite the fact that it parallels Oakland's decline.

Early Prenatal Care

Overall strength of model and contextual relationships. The census tract-level early prenatal care rate was the dependent variable in the second model. The overall model has the most predictive power of all four models tested, with 77 percent of the variation in the rates explained by the independent variables. Our five contextual variables are all highly significant (table 10.6 and annex table C.25). The percentage Hispanic population and the percentage of the population not employed were the strongest of the set, with a 1 percent increase corresponding to a .23 and .22 percentage point respective decrease in early prenatal care rates.

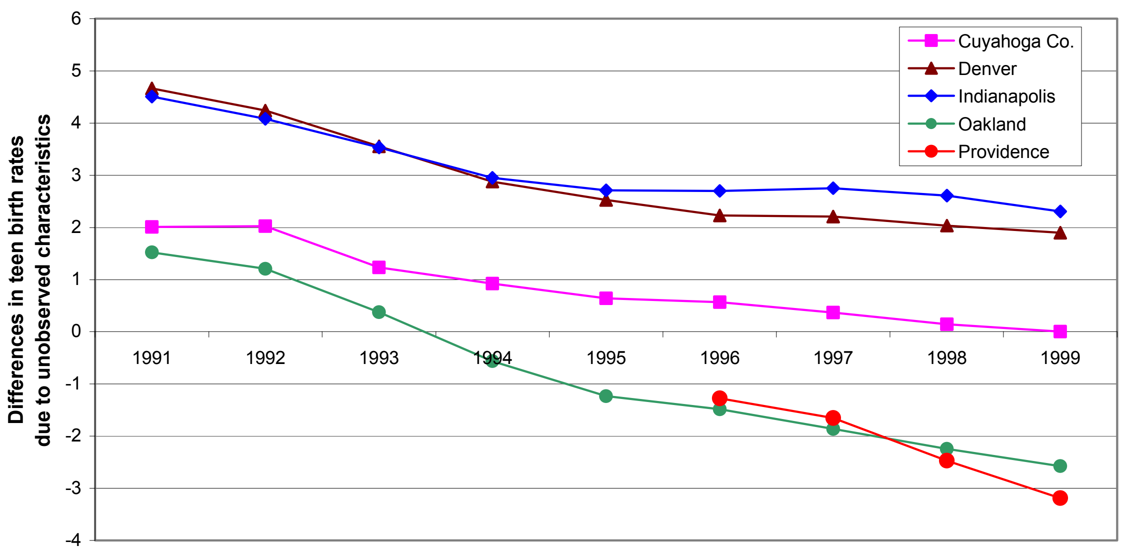
Table 10.7: Decomposition of Differences between City Health Indicators

	Denver	Indianapolis	Oakland	Providence*
*Note: Differences are all relative to Cuyahoga County 1999 rates.***				
Births to teens (age 15-19) as percent of females age 15-19				
Average difference in city rates, 1991 - 1999	4.0	1.4	2.4	-0.2
Difference due to five contextual variables in model				
Percentage points	1.9	-0.8	4.0	2.2
Difference due to unobserved characteristics				
Percentage points	2.0	2.2	-1.6	-2.4
Percent of births to mothers receiving prenatal care in first trimester				
Average difference in city rates, 1991 - 1999	-11.0	-6.5	-1.2	-16.2
Difference due to five contextual variables in model				
Percentage points	-5.4	1.9	-7.3	-8.4
Difference due to unobserved characteristics				
Percentage points	-5.5	-8.4	6.1	-7.8
Percent births with low birth weight				
Average difference in city rates, 1991 - 1999	0.9	-1.1	-0.2	-0.9
Difference due to five contextual variables in model				
Percentage points	-2.0	-1.1	0.8	0.3
Difference due to unobserved characteristics				
Percentage points	2.9	-0.1	-1.0	-0.3
Age-Adjusted Death Rates per 100,000 population				
Average difference in city rates, 1991 - 1999	91	-114	-20	NA
Difference due to five contextual variables in model				
Rate difference	90	-81	9	NA
Difference due to unobserved characteristics				
Rate difference	1	-34	-29	NA

*The averages for Providence are only for 1996 to 1999.

** Differences are all relative to Cuyahoga County 1999 rates because they are the reference site and year for the dummy variables.

Figure 10.1: Differences in Teen Birth Rates Due to Unobserved Characteristics



Explanatory power of contextual variables. On average over the 1990s, Denver's early prenatal care rate was about 11 points below that of Cleveland (table 10.7). The explanation for the difference is equally split between the unobservable conditions in Denver during this time and the contextual variables we include in the model. In other words, with the demographic, economic, and social status of Denver's tracts in the 1990s, the model predicts that the early prenatal care rate would be about 5.4 points lower than Cuyahoga County's rate. But, in addition to this, other factors particular to Denver that are not measured in this model are associated with another 5.5 point drop in the average rate. In addition to omitted census tract attributes, some potential policy factors were already listed in section 9—fewer facilities accepting Medicaid and immigrants choosing nontraditional health providers (either because of preference or belief that those using public health services will risk deportation).

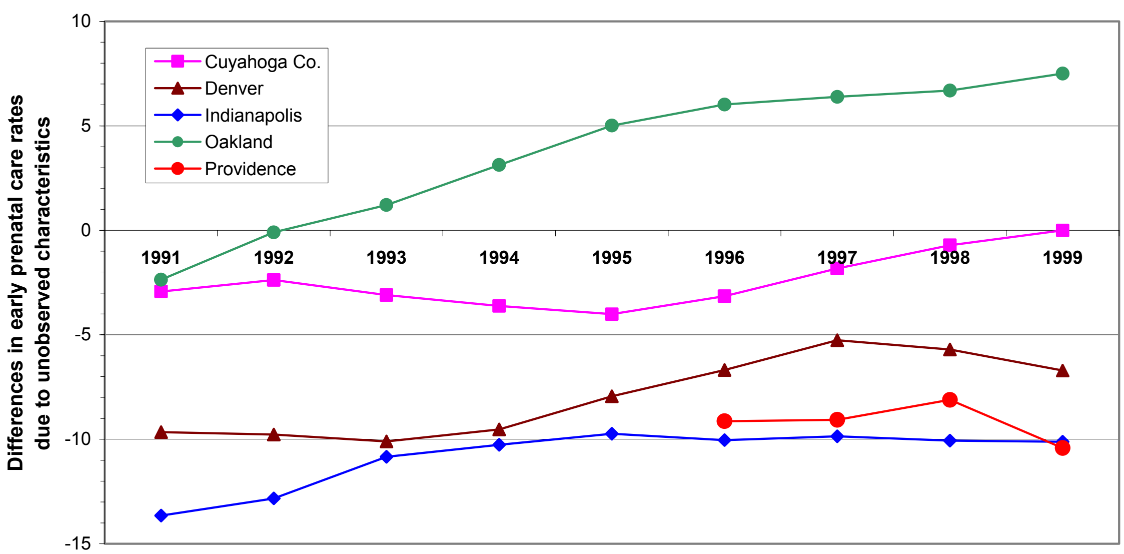
The differences in Providence appear similar to Denver, with about half of its lower rate (8.4 points) explained by our chosen contextual indicators and the other half (7.8) by other factors not specified in the model. Like Denver, Providence had a growing immigrant population during the 1990s, which may result in higher teen birth rates.

In Oakland, we see a different situation. The socioeconomic contextual indicators in Oakland should place its rate more than 7 points lower than Cuyahoga County, but other unmeasured circumstances there raise the expected rate by 6 percentage points, narrowing the gap in the overall rates to 1 percentage point. While not conclusive, this finding is consistent with the positive impact of the Oakland Healthy Start initiative described in section 7.

Indianapolis is the only city where the variation in the five contextual variables corresponds to a higher early prenatal care rate than Cleveland (+1.9 percentage points). The unmeasured factors, however, reduce the average rate by 8 points, more than offsetting the influence of the included tract attributes.

Shifts in indicator trends. Figure 10.2 displays the regression-adjusted differences in the early prenatal care rate among the cities. The chart shows the sharp increases in Oakland that we described in section 9, along with the increases in Indianapolis in the early decade. The decreases in Denver in 1997 and Providence in 1998 stand out amid the general improvement in prenatal care in the 1990s. With our model, we can perform a joint test of significance to see if these downturns are significant or just due to random variations. The regression shows that both changes are significant, at the 0.04 level for Providence and the .01 level for Denver. Given that the prenatal care rates in the high-poverty areas in these two cities were already lowest among the five cities, these further declines are troubling.

Figure 10.2: Differences in Early Prenatal Care Rates Due to Unobserved Characteristics



Low-Birth Weight Rates

Overall strength of model and contextual relationships. In the third model, our independent variables explained 56 percent of the variation in the low-birth weight rates. Our five contextual variables are all highly significant, but the coefficients are quite small (table 10.6 and annex table C.26.) The percentage of the population not employed had the highest correlation, with a 1 percent increase corresponding to a .09 percentage point decrease.

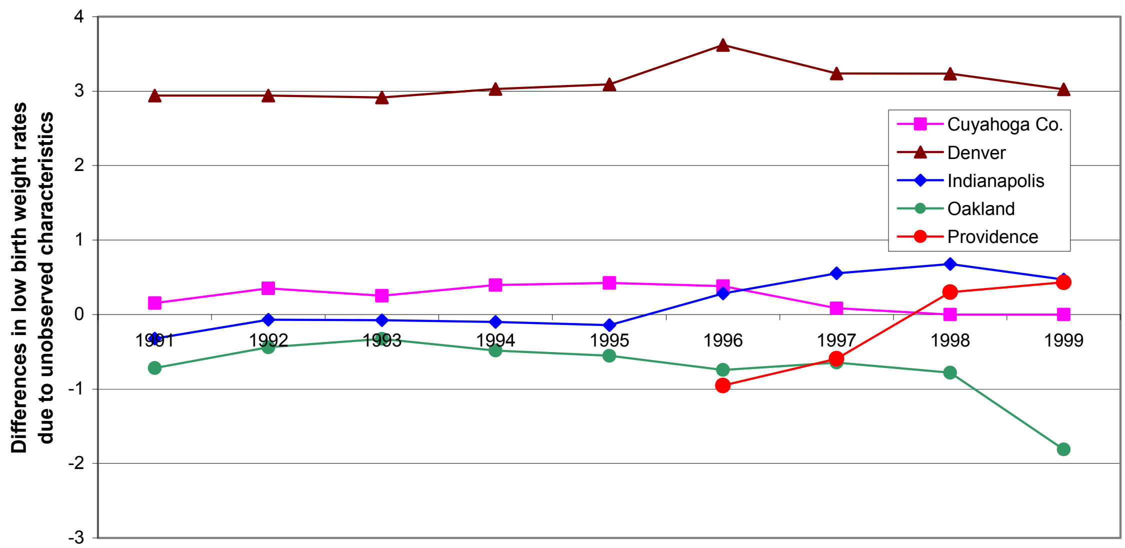
Explanatory power of contextual variables. The model predicts that the low-birth weight rate in Denver would be 2 points lower than the Cleveland rate, but the characteristics of Denver during this period raise the rate by 2.9 percentage points, placing the end rate above Cleveland's. Only in Denver does this portion of the difference increase the rate. In both Indianapolis and Providence, our specified contextual variables explain most of the difference between the city rates, with unspecified city influences lowering the rate slightly more. The power of this methodology is apparent when looking at the case of Oakland. Just looking at the difference in overall low birth weight rates between Oakland and Cleveland, it appears that the two are quite similar, only .2 percentage points different. However, decomposing the difference reveals that, as with the previous two models, the levels of the contextual factors in Oakland's should be associated with a rate worse than Cleveland's (0.8 points higher), but the city must have some other protective factors that compensate for the contextual conditions.

Significance of shifts in indicator trends. As stated in section 9, the trends in low-birth weight rates were the least consistent of all the maternal and infant health indicators, and the regression-adjusted means confirm this (figure 10.3). Both Denver (beginning in 1996) and Oakland (beginning in 1997) show statistically significant improvements in low-birth weight rates, at the .04 and .002 significance level, respectively. The increases in Indianapolis from 1995 and Providence in 1996 are also significant.

Age-Adjusted Death Rates per 100,000 Population

Overall strength of model and contextual relationships. The final model has the age-adjusted death rates as the dependent variable. The explanatory power of the independent variables is moderately strong, with a .45 R-squared value. This is the one model where not all of the tract descriptive variables are significant—the share of the census tract population who are Hispanic is not correlated with the age-adjusted death rates. Of the remaining three variables, the percentage of population not employed is the most highly associated, with a 1 percentage point increase linked with a 10.4 change in the death rates. The percentage of the population living in a different house five years ago has the second highest coefficient (8.0) (table 10.6 and annex table C.27.)

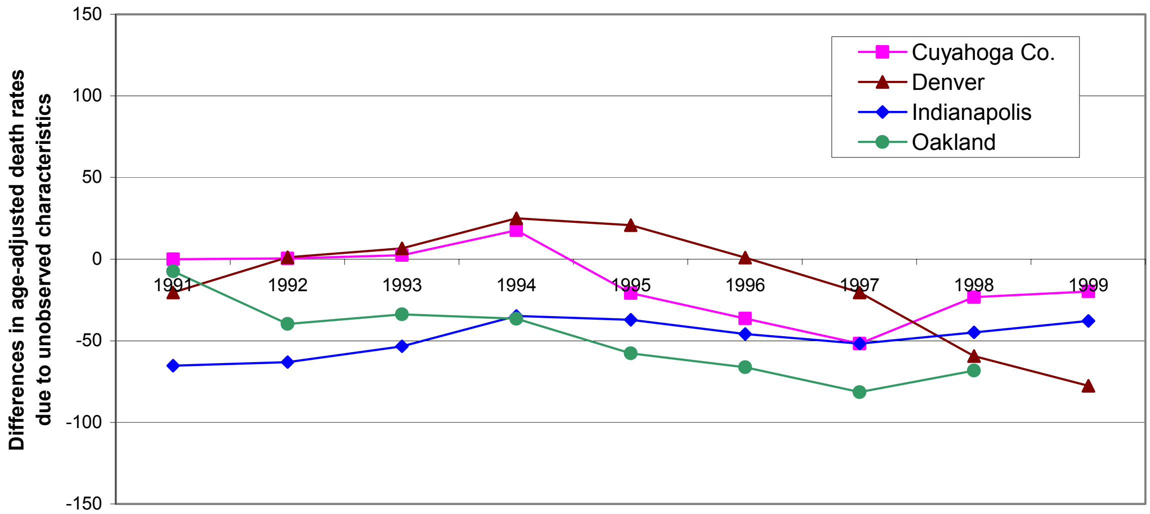
Figure 10.3: Differences in Low Birth Weight Rates Due to Unobserved Characteristics



Explanatory power of contextual variables. The three cities show contrasting situations for the differences in age-adjusted death rates among cities. For Denver, almost all of the difference between its and Cleveland's rates can be explained by the contextual indicators. For Indianapolis, the census tract contextual indicators account for the majority of the lower death rate compared to Cleveland, but the unmeasured conditions do increase the total magnitude of the difference. Finally, Oakland's city-specific influence is much stronger than that of the census descriptors.

Significance of shifts in indicator trends. Three of the cities (Cleveland, Indianapolis, and Oakland) show upward trends in the differences between the regression-adjusted means of the age-adjusted death rates in the late 1990s (see figure 10.4). When tested, however, only the Cleveland trend is statistically significant.

Figure 10.4: Differences in Age-adjusted Death Rates Due to Unobserved Characteristics



SUMMARY

The following points summarize the key findings of this section:

1. Using the bivariate methodology, we find that most of our hypotheses about the relationships between neighborhood and health conditions proved correct, with a few occasions of site differences. The correlations confirmed that higher rates of minority population, lower socioeconomic status, and lower quality housing are correlated with lower early prenatal care rates and higher rates of low birth weights, teen births, and age-adjusted deaths.

2. Two of our hypotheses were not completely verified with the proxy measures we used. First, higher levels of social stressors (measured by crime rates) were significantly related only to higher rates of low birth weight, teen births, and age-adjusted deaths (not to early prenatal care rates). Second, the hypothesis about stronger social networks correlating with lower levels of mortality and better maternal and infant outcomes was confirmed by one set of proxy variables (rates of renter occupancy, rental vacancy, and mobility) but was related to worse health outcomes with the second set (change in total and minority population and rate of home improvement loans).

3. The multivariate analysis demonstrated that much of the variation among the health indicators is explained by five selected independent variables (percentage African-American, percentage Hispanic, average family income, percentage not employed, and percentage of population that moved in the past five years). The most predictive model was the one with early prenatal care rates as the dependent variable, though the remaining models also have substantial explanatory power (R-squares range from 0.45 to 0.77). Of the five independent variables in the model, the percentage of population that is not employed was the variable most highly correlated with three of the health indicators (with early prenatal care rates as the exception).

4. The models also show that a portion of the variation is not explained by the five selected contextual indicators but, rather, by conditions particular to the city and time period of the measure. For example, Oakland's rates for maternal, infant, and mortality outcomes were consistently better than the model predicts given the contextual conditions in Oakland’s census tracts. Finally, we identify which of the trends represent significant changes versus random fluctuations. For example, the results of the model enable us to confirm that the early prenatal care rates in both Denver and Providence have fallen by a statistically significant amount in the recent years—going against positive trends in the United States and the other three cities.