THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

USING SIPP TO ANALYZE BLACK-WHITE DIFFERENCES IN YOUTH EMPLOYMENT

No. 143

G. C. Cain and P. M. Gleason
University of Wisconsin

U. S. Department of Commerce BUREAU OF THE CENSUS

Using SIPP to Analyze Black-White Differences in Youth Employment

Glen G. Cain and Philip M. Gleason
University of Wisconsin

ABSTRACT

The lag in employment and earnings of black men aged 16-24 relative to white men of that age is analyzed, mainly using probit models to explain employment-to-population ratios and using hazard rate models to estimate or lose jobs. We find that most of the rate at which young men gain smaller likelihood gap between white and black employment lies in the that blacks had any job during a given month. A second focus is on the methods of using SIPP. One innovation is that we appended market variables to the SIPP records, based on the SMSA of the respondent's residence, and estimated the effects of these market variables on the employment of young men.

KEYWORDS

labor force, youth employment, probit analysis, hazard rate analysis

THE PROBLEM, MOTIVATION, AND SUMMARY OF RESULTS

Our study uses the 1985 Survey of Income and Program Participation (SIPP), covering the 32-month period from the end of 1984 until the first part of 1987,-to analyze the gap in employment and earnings of black youth compared-to white youth. The gap is manifest in the time-series trends that span the last 30 years or so, and it is the long-run trend that provides the context of this serious problem in American society. Beginning in the late 1950s, the labor force participation rates of black youth, ages 16-24, which at that time were about the same as those for white youth, began to decline. not only relative to whites, but in the case of young black men the decline was absolute. Table 1 and Figure I show the trends of the employment to population ratios, E/P, for the period 1954-55 to 1988 for white and blacks, males and females, for the age groups 16-19 and 20-24. The largest gains in youth employment are by white women; the largest declines are by black men.

---

[1]This research was mostly carried out while we were research follows of the American Statistical Association/National Science Foundation at the U.S. Bureau of the Census. We are also grateful to the Institute for Research on Poverty and the LaFollette institute for Public Policy at the University of Wisconsin for support. we thank Robert Fay, David McMillen. Daniel Rasprzyk, and Arnold Reznek of the Census Bureau for their advice and assistance.

That the employment experiences of young people is likely to be an important determinant of their success or failure in the labor market as adults is a reasonable supposition with some research support (Ellwood, 1982). Part of the long-run context of the labor market experience of black youth is the relative stagnation in the last 20 years of the previous gains in black family incomes relative to whites that had characterized the period from 1940 to 1970. The lag in black youth employment portends a continuation of this stagnation.

The SIPP data do not permit an analysis of long-term trends, but they do provide a detailed distributional picture of the lag in black youth employment relative to whites for points in time over several years in the 1980s. SIPP has several advantages over other essentially cross-sectional surveys: a large size that compares with the Current Population Survey, a true longitudinal design that follows individuals and permits analyses of changes in labor market performance, better information about the incomes and wages of the respondents, and information about the labor market locations of the respondents. One of our purposes in this paper is to discuss some special features of SIPP to inform future users of its strengths and weaknesses. In our substantive work, we exploit the improved descriptive information available from SIPP, and we seek evidence that will help us understand what has occurred over time and what to do about the problem. An important motivation for our work with SIPP is to find relationships among variables that have significance for policies.

The results of our research to date, which has focused on young men, may be summarized briefly. In a static framework the supply characteristics of black youth that explain (or are associated with) their low employment relative to whites are low levels of schooling attainment and several variables that designate low family incomes. On the demand side two characteristics of the young person's labor market have practical significance. Standard Metropolitan Statistical Areas (SMSAS) that have low unemployment and a service-oriented industrial structure are favorable to youth employment, especially to black youth employment.

The impact on youth employment of the above-mentioned variables measuring personal and market characteristics depends both on the mean levels of their incidence by race and on the racial differential "affect" (or partial derivative) of the variables. Thus, educational attainment has a larger effect for blacks than whites in increasing employment, which implies that the lower educational attainment of black youth compared to whites is doubly disadvantageous to their employment relative to whites. The effect of educational attainment is, however, one result from these essentially cross-sectional findings with SIPP that does not contribute to explaining the time-series trends, because the black gains in years of schooling have been somewhat more rapid than white gains during recent decades. In contrast, our findings from SIPP about labor market characteristics support the hypothesis that the relative concentration of blacks in the northeast and north central cities that experienced a lose of jobs and increases in unemployment during the last 10 to 20 years is part of the explanation for the downward time-series trend for black youth.

A striking finding of our analysis of both the average employment over the full panel period and of the changes in employment from month to month is that most of the gap in black youth employment relative to whites is attributable to a much larger proportion of blacks who have no or almost no employment at any time during the young person's months of record, which is a maximum of 32 months, although the average is around 15 months. Black young men who do obtain jobs have nearly the same total hours at work and earnings as obtained by whites, but there are proportionately more jobless months among black youth. Our analysis of employment spell lengths supports this result. Given an employment spell during the panel period, the average spell length or the percentage of months that the young man is employed is lower for blacks than whites, but the main source of the white-black gap is in the lower probability that blacks have any spell of employment during the survey period.

In another paper (Cain and Gleason, 1990) we report that the employment of black young men in poor families, defining family income without including the earnings of the young person, is lower than that of black young man in non-poor families and lower than the employment of white young men in poor families. In this paper we report that family income has a positive, but very small, association with the employment of both blacks and whites. The findings in this and the preceding paragraph suggest that among families with young men living at home. there is a sharper economic disparity between the "haves" and the "have nots" among black families than among white families.

DESCRIPTIVE STATISTICS OF THE EMPLOYMENT PERFORMANCE

Table 2 shows the employment performance for white and black men, ages 16-19 and 20-24. using the proportion of months worked, average hours worked per week.- and earnings per week (or wages per hour). The SIPP sample of some 30,000 households provides a large sample of white young men: 1,665 16-19 year-olds and 2,437 20-24 year-olds. Unfortunately, the samples of black youth, 260 and 270, respectively, are small. Using person-months as observations inflates the sample sizes, of course, but this introduces a problem of non-independence of the observations that compounds the otherwise minor problems of non-independence from the clustered sampling design. Somewhat offsetting the problem of non-independence of the residuals is that SIPP has a relatively abundant set of characteristics of the young men to use as explanatory variables that control for person-specific factors.

As shown in Table 2, the black-to-white employment ratios are low, closely matching the ratios in the later period of the time series in Table 1 that was based on data from the Current Population Survey. Table 2 also shows that the employment-to-population ratio captures most of the gap between blacks and whites. The more comprehensive measures of employment performance, hours and earnings, do show declines in the black-to-whits ratios for the 20-24 group, .69 and .64 compared to the E/P ratio of .73, but the declines are small, If we restrict the sample to young men with jobs, then the black-to-whits ratios of hours worked and wages are between .9 and 1.0, which reinforces the point that the main reason for the employment problem of black youth is not having any job in a given month.

ESTIMATION MODELS

Tables 3 and 4 present results from probit models that estimate the probability of working in a given month for the two color and two age groups. The young men included in the samples for these estimates are restricted to those who live in one of the 300 or so largest SMSAs for which we have supplementary data from the COUNTY AND CITY DATA BOOK, 1988, of the U. S. Bureau of the Census. (Access to the city and county residences of the SIPP respondents is available to Census Bureau employees, which we were during the 1989-1990 academic year.) Before discussing these estimates, a methodological comment on our selection of variables may be helpful.

A problem in using survey data for individuals to estimate determinants of an endogenous outcome such as the person's employment, is that the determinants used are often characteristics of the individual that are themselves endogenous outcomes. Thus, while a few explanatory variables like age are purely exogenous, many others may be endogenous in the sense of reflecting behavioral choices of the person, such as school enrollment., marital status, and one's living arrangements. Or, those characteristics may be endogenous in the sense of reflecting the effects of omitted exogenous factors. An example is receiving public assistance, which may represent the effect of living in a destitute neighborhood. Clearly, causality is difficult to establish with these data. Most of the variables that are person-specific embody some mix of endogenous and exogenous forces; educational attainment and being "on welfare" are examples.

In this light. we are especially interested in estimating the effects of variables that measure characteristics of the labor market, because for young people, especially those aged 16 to 19, the SMSA in which they live is exogenous. Moreover, such market characteristics as the level of unemployment, prevailing wages, the industrial structure, and, to a lesser extent, the degree of racial segregation are all either policy variables or represent variables that can be affected by policies. Unfortunately, the number of persons in the SIPP sample who live in any one SMSA is very few, especially among blacks, and it is difficult to obtain reliable estimates of the relations between their employment outcomes and the SMSA characteristics.

With these caveats, we turn to Tables 3 and 4. An asterisk designates statistical significance after allowing for our adjustment for the non-independence of the observations.[2] The effects of many of the variables are similar in sign for both color groups and both ago groups. Among the 20-24 age group the partial derivatives of the variables are likely to be larger for blacks than for whites because we measure the point of response at the mean employment proportion, which is .58 for

---

[2] We are indebted to Robert Fay of the Bureau of the Census for his help in giving us corrected standard errors of the coefficients of the predictor variables in a logit model. which is similar to our probit model. Fay's procedure (1989), the CPLX program, uses a jackknife technique to adjust for sources of non-independence in complex sample designs. We determined the orders of magnitude for adjustments from Fay's procedure for several logit models and then applied these adjustments to our estimates from various models, including standard regression models and Tobit models as wall as the probit models that we report here.

blacks and .80 for whites. Thus, the response point for whites is closer to the asymptote of the probit curve---at the value of 1.0. The functional form of the probit dictates relatively smaller changes in the dependent variable per unit of change in the independent variable when the function is evaluated at probabilities much larger (or much smaller) than .5.

Among the 16-19 age groups, the opposite result occurs, because the black moan employment rate is .29, well below .5. and the white mean is .5'2.

In this paper we discuss only several noteworthy findings. To report the effect on employment of a specific variable, we evaluate the probit function for different values of that variable while using the sample moan characteristics of all other variables. The effect (or, to use more cautious wording. the association) between employment and the educational attainment of the young men shows that blacks aged 20-24 who have less than a high school education are predicted to be employed in 44 percent of the months, which is far less than the predicted employment rate of 71 percent for black men of this age with 12 years Of schooling completed (and no college). The comparable estimations for 20-24 year-old white men are 73 percent for those with less than a high school education and 84 percent for those with a 12 years of schooling. The partial derivatives shown in Table 3 are the basis for these estimates.

Table 4 shows similar striking results for education among 16 to 19 year-old males. Black youth who have graduated from high school are estimated to have a 36 percent employment rate; those who are "on track" in high school (based on their age and their level of schooling) have a 28 percent employment rate; and those who are "lagging behind in high school" have a predicted employment rate of 12 percent. The corresponding predicted employment rates for white youth are 54 percent. 56 percent, and 43 percent. (In these calculations, the proportion of young men who are enrolled in school is hold constant, but interacting the educational attainments with enrollment status show similar racial differences in the effects of education.) For both age groups we show that some part of the gap in black youth employment is attributable both to their lower level of schooling attainments and to the larger effects of education for blacks. If taken at face value, these results suggest a policy of preventing school drop-outs to increase youth employment and narrow the white-black gap. However. our estimated large effects of educational attainment for blacks may reflect in part the selection of employable types of young men into those educational attainments. That is, the same types who are likely to attain more education are likely to fare better in the labor market, so both outcome variables may be causally affected by omitted ability variables.

To examine the relation between the economic status of the family and the employment of the young man, we focus on the sample of 16 to 19 year-olds, most of whom live with their parents. The overall average employment rate for the white teenagers is 52 percent, and the income of the family, excluding the income of the young person, has only a small positive effect on this employment percentage. Of three other indicators of economic distress --- receiving public assistance (being on welfare), living in public housing, and having only one parent present --- each has an incrementally negative effect on the white youth's employment rate, especially welfare status. Assuming each of these disadvantaged states and a low level of parental income of $10,000 per year, the white young man is predicted to have been employed in 29 percent of the months of record. Not being on welfare raises the employment rate to 41 percent -and having two parents present raises the employment rate from 41 percent to 43 percent. Not being in public housing raises the employment rate to 50 percent, which is close to the overall mean employment rate. The incremental negative effects of welfare and public housing are sizeable, but the proportions of white families in these categories are quite low: 8 percent and I percent, respectively, in our sample.

The same examination of black young men, aged 16-19, reveals several important differences. The overall employment rate is predicted to be 23 percent, evaluating the probit function with mean values of the variables. (Note that the employment rate of 23 percent that is predicted using means for all the variables is smaller than the original employment rate, 29 percent, for the black sample. Because the probit function is nonlinear, there I 3 no necessity for the two rates to be equal --- as in fact they are up to two decimal places for whites.) Assigning a family income of $10,000 and the statuses of being on welfare, in public housing, and in a one-parent family serve to lower the predicted employment rate by only one percentage point, from 23 to 22 percent. However, welfare status is positively associated with the youth's employment, given that the youth is in public housing and in a one-parent family. Indeed, the employment rate is predicted to be 16 percent if the only change under these conditions were that the family was not on welfare. More than off-setting the effect of welfare status is that of one-parent status, because the predicted employment rate rises from 16 to 25 percent it a two-parent family is assumed. Moreover, fully 51 percent of the black young men are in ono-parent families compared to 22 percent of the white young men. This is an even larger racial difference than that for welfare incidence: 28 percent for black families and 8 percent for white families.

We turn next to examining several market variables that refer to the young men's SMSA of residence and which we appended to the SIPP file. The unemployment rate (for the civilian labor force) in the SMSA has a negative and statistically significant effect on youth employment for both age groups of blacks and for the white 20-24 year-olds. The partial derivatives are much larger for blacks than whites, which shows up in the following predicted employment rates in response to an assigned change in unemployment. If blacks aged 20-24 live in a SMSA with an unemployment rate that is one standard deviation below the mean, about 4.5 instead of 6.5, their employment rate is predicted to increase to 68 percent from a base of 61 percent. (Again, the employment rate for blacks that is predicted using means, 61 percent, is somewhat different from the original employment rate, which is 58 percent.) The predicted white employment rate if the unemployment rate is decreased by one standard deviation (also about 4.6 instead of 6.5) is an increase to 85 percent from a base of 84 percent. Among blacks aged 16-19 the comparable change is an increase in the employment rate to 31 percent from a base level of 23 percent. Among whites aged 16-19 the corresponding change in from 52 percent to 53 percent.

A second market variable, also representing a demand-side influence, is the level of per capita retail sales, which serves as an indicator of an industrial structure in the SMSA that is relatively favorable to youth and employment. Like the unemployment rate, retail sales is, we believe, essentially exogenous to the labor supply decisions of young people. As indicators of demand, both variables are suggestive of what might be expected from policies to increase the employment opportunities of youth. As shown in Tables 3 and 4, the logarithm of per capita retail sales has a positive and statistically significant effect on the employment of all four age-race groups, and somewhat larger for blacks than whites. Using the same technique of evaluating the probit function at a level of retail sales that is one standard deviation above the mean, we show an increase in the employment rate of the blacks aged 20-24 from a base of 61 percent to a predicted rate of 70 percent; in the employment rate of blacks aged 16-19 from 23 percent to 34 percent; in the employment rate of whites aged 20-24 from 84 percent to 85 percent; and in the employment rate of whites aged 16-19 from 52 percent to 53 percent.

The role of the young person's residence in the central city of the SMSA along with the** SMSA characteristics --- its size, percentage .black, and its extent of residential segregation --- provide another set of interesting results. (The segregation index is the ratio of the SMSA's percent of blacks who live in the central city relative to the percent of whites who live in the central city. This index has boon found to be highly correlated with more refined segregation indexes, and it has the advantage of being easily calculated for all SMSAs. See Cain and Finnie. 1990, for a discussion of this point.) These market variables have essentially no effect on the employment rates of white young men, and indeed no effect was expected. Population size has no effect on the employment of blacks, but by holding size constant we are better able to measure the not effects of the percent black and of the residential segregation variable. Holding constant all other variables, the effect on employment of the percentage black in the SMSA is essentially zero for black 20-24 year-olds but significantly negative for black 16-19 year-olds.

Neither segregation nor central city residence, by itself, affects the employment rate of black youth, but the interaction of the two is quite substantial. To illustrate, if blacks aged 20-24 live in a SMSA that has a segregation index that is one standard deviation above the mean, the predicted employment percent for blacks who live in the central city is 57 percent, and the predicted employment percent for blacks who live in the suburban areas is 73 percent. These percents should be compared to the predicted overall mean employment rate of 61 percent. A similar strong interaction effect between central city residence and the extent of segregation is found for blacks aged 16-19. Compared to the overall predicted mean employment rate of 23 percent, a 17 percent employment rate is predicted for blacks in a central city residence in a segregated SMSA, and a 34 percent employment rate for a suburban residence in a segregated SMSA.

We do not wish to over interpret these interaction results, preferring to wait to see if other sources of data, including other panels of SIPP data. confirm our findings. The general findings about lower employment rates among black youth in SMSAs that are relatively segregated and have large black populations are found elsewhere, however. (See Cain and Finnie, 1990.) The interpretation of these findings in terms of their indication of supply- or demand-side forces is not clear. Our impression is that segregated black neighborhoods tend to have fewer sources of employment opportunities for youth. However, supply-side factors, such as poorer quality schooling and a larger proportion of families on welfare, are also likely to reduce youth employment.

We now adopt a conventional calculation to decompose the results of the probit functions to indicate the fraction of the white-black gap in employment that is attributable to the different characteristics of the two races. The white mean values of the variables are multiplied by the coefficients obtained from the black probit function. The resulting predicted black employment rates are 75 percent for blacks aged 20-24 and 36 percent for blacks aged 16-19, which compare with the original employment rates of 58 percent and 29 percent respectively. (Note that we have reverted to using the original employment rate in our sample rather than the employment rate that is predicted using the mean values of the sample characteristics.) Thus, the original gap between white and black employment rates for 20-24 year-olds, which is 23 percentage points (81 - 58). is reduced to 6 percentage points (81 - 75). The original gap between white and black employment rates for 16-19 year-olds, which is coincidentally also 23 percentage points (52 - 29) is reduced to 16 percentage points (52 - 36). Among the 20-24 year-olds almost all of the closing of the gap is attributable to differences in personal characteristics. Among 16-19 year-olds, however, the labor market characteristics. including residing in a central city, contributes over half of the closing of the gap. Again, these results should not be over interpreted. It is not the case, for example, that assigning black means of tho variables to the white functions results in a significant closing of the gap. The predicted white employment rates are not much reduced from their original levels. Thus, assigning the higher educational attainments of whites to the black function sharply increases the predicted employment rate for blacks, because the coefficients (or partial derivatives) of the black educational variables are large.

However, assigning the lower educational attainments of blacks to the white function only slightly decreases the predicted employment rate for whites, because the coefficients of white educational variables are small.

HAZARD RATE ESTIMATIONS OF EMPLOYMENT SPELLS

We now focus on employment continuity, which is defined using monthly units. Continuity from one month to the next merely requires that the young man reports being employed at any time in each month. One must be employed to be at risk of leaving that state, and the rate of such transitions, defined as the number of exits divided by the number at risk, is our measure of the hazard rate with these discrete data. Strictly speaking, our measure understates the rate of ending an employment spell, because within-month transitions are omitted, but we may view such short periods of being between jobs as a minor problem.

Our analysis of the hazard rates of employment exits is intended to complement the previous static analysis of the worker's average performance over the full period of the panel. Only one of the following three objectives of our analysis has been satisfactorily achieved, however, so our discussion will be brief.

(1) Do the same factors that affect the levels of employment performance also affect the transitions? We find that the hazard rate estimates are too unreliably estimated to answer the question. Few of the determinants of employment levels wore statistically significantly related to employment transitions.

(2) Another objective, the estimation of the durations of the employment spells, is a joint product of estimating hazard rate models, so the unreliable estimates of the latter imply weak estimates of the former.

(3) We have more confidence in our answer to the question of whether the gap in the white-black employment performance is explained by the problem of many short spells of being without a job or by the problem of a few long spells of being without a job. We find that it is the latter, and, in fact, around half or more of the gap is attributed to the relatively large fraction of black youth who had no employment spell during the panel period. Because hazard rate models of employment spells are restricted to the sample of those who have at least on* spell, some of the variables that determine employment performance may primarily influence whether a person has been able to obtain any job, and these variables are apparently less strongly related to continuity in an obtained job.

Several conceptual and data problems arise with hazard rate analysis that ether do not arise or are less severe with static analysis. The Proportion of time employed over an extended period of time, the focus of our previous analysis, is a clearer indicator of employment success than is continuity in employment. Thus, a job loss that in followed by a now job at a higher wage may constitute success, and this result may be positively associated with a favorable characteristic of the worker. Another source of ambiguity in interpreting a job loss is that a transition to full-time schooling would not have the same normative implications as a transition to unemployment. To handle this issue we estimate two types of models, one that defines a transition simply as a job loss and another that defines a transition as an exit to a "non-productive" state, defined as not being either employed or in school.

We have already mentioned the ambiguity of interpretation that arises in restricting the sample to persons with a job, because continuity in employment may be less crucial to a young person's overall employment performance than obtaining a first job in the period under analysis. In examining this issue we have experimented with hazard rate models that are restricted to young men who are all in a "non-productive” state, defined as above, and then estimate the (presumed) successful transition to either employment or schooling.

Ideally, hazard rate models require precise dates of the respondent's statuses over extended periods of time. Errors in reporting the timing of transitions are inevitable. With suitable assumptions hazard rate models permit the estimation of the spell length, even when the data include uncompleted (or right-censored) spells, but unknown starting dates --- left-censored spells

--- create serious difficulties. Employment and other labor force data are often left-censored in SIPP. Another problem in SIPP is that the four-month interval of survey interviews creates the so-called "seam effect," wherein respondents tend to assign their current monthly state to all four months of that survey period. The result is-that a change from the previous four-month period's status is recorded as having occurred in the first month of the current four-month survey period, even though the actual change may have occurred in the second or third month of the current survey period. That first month of the current survey period is called the seam month. This systematic error in the timing of transitions, which all users of SIPP have found, adds to the general problem of errors in the data. As discussed below, we find a pronounced seam effect in determining the report of a transition from employment to joblessness.

A basic organization of the data underlying the hazard rate model is the demographer's life table, which here begins with the number of men in the particular age-race group who have a first month of employment. For each succeeding month the number who exit is recorded and used to compute the percent (or rate) of exits. The employment life table may be restricted to only first-time exits or it may add reentries into employment, which become now spells that face the risk of a second (or higher order) exit. In either case, there are 32 rows of computed monthly exit rates from the SIPP panel, which are the basis for two main statistical findings. One is the total number of exits as a proportion of the total number of person-months at risk, which measures the average monthly exit rate for the panel period. For the four age-race groups of young men, the average monthly employment exit rates are: .033 for whites aged 20-24, .056 for blacks aged 20-24, .072 for whites aged 16-19. and .089 for blacks aged 16-19. Clearly, blacks have higher rates of job loss than whites, and the teenagers have higher rates than the 20-24 group.

We should point out that most of the spells under discussion were avoid this left-censoring. on-going at the beginning of the survey. To we experimented with using only newly starting spells of employment to construct the life table. We will not report these results except to point out that they are selective of young people with a propensity for short duration spells of employment, and each of these samples of beginning spells shows a higher average exit rate than those reported above for the entire sample.

A second statistic of interest from the life table is the trend in the monthly exit rate. We expect the trend to decline on grounds that tenure both implies the stability of a successful match between the worker and his job and reflects increasing on-the-job training and wage growth --all of which should reduce the probability of a job loss. Whites have a more rapidly decreasing trend in the monthly rate of job loss than do blacks for each of the two age groups. More information on the trends is provided below.

The employment life table suggests that blacks have shorter average durations of employment spells, although the statistics are derived from a table that does not use any of the explanatory variables that we have used in the probit functions above or in the hazard rate models presented below. However, by accounting for the difference in proportions of the samples who have no employment spells, and who are therefore not included in the employment life table, we see that most of the overall gap in employment between white and black youth is attributable to the never employed group. Recall that the overall racial gap in employment for 16-19 year-olds was 23 percentage points, a 52 percent employment rate for whites and a 29 percent employment rate for blacks. The gap is considerably less, 17 percentage points, for those who had at least one spell of employment. For this group the proportion of months employed out of the number of months at risk is .59 for whites and .42 for blacks. The comparable figures for the groups aged 20-24 are: total employment rates of 81 percent for whites and 58 percent for blacks; among those with at least one spell of employment, the proportion in .85 for whites and 70 for blacks.

A direct measure of the difference in the proportion of blacks and whites who had no employment spell is obtained as a by-product of our estimated logit functions of the dummy dependent variable, 1 if the person had an employment spell; 0 otherwise. Among 16-19 year-olds 30 percent of the blacks and 12 percent of the whites reported no employment spell of one month or more. Among the 20-24 group, 17 percent of the blacks and 5 percent of the whites reported no employment spell.

Table 5 shows an estimated model for the hazard of leaving an employed state for the 20-24 year-old groups of white and black young men. For this age group, compared to the 16-19 year-olds, the exit is more likely to be to unemployment rather than to school enrollment. In all groups the employment exit rate is extremely high in the fourth month of the employment spell --- a reflection of the seam effect. The second and third months are also more likely to be seam months than the first month, because an employment spell can start at any given month. Because most of the respondents were employed at the beginning of the survey, however, the fourth month is the usual seam month. The coefficient of "Time," defined an the observed tenure month of the worker's employment spell, is negative and, for whites only, large and statistically significant. A negative sign is evidence that the hazard of leaving employment declines as the worker's tenure increases. This is an expected result, but we cannot be sure it the effect of tenure is truly negative for homogeneous workers or it the negative sign is the result of unobserved heterogeneity that leads to the selective departure of workers who have a higher propensity for mobility.

Few of the socioeconomic and demographic variables have large, significant, and, as between the two color groups, consistent effects on :he transition from employment to not being employed. For both color groups the number of hours worked on the job has a negative effect, implying. as expected. that full-time jobs are more secure than part-time jobs. The wage earned on the job has a strong negative effect on job losses for black workers but. surprisingly, not for whites. White workers having a health limitation on working and having less than a high school attainment show significantly higher rates of job loss. The market demand variable, retail sales per capita, has a significant negative effect on exits from employment of black workers, but it has, unexpectedly, a positive effect for white workers.

We tried a variety of models to estimate hazard functions of various transitions regarding labor force statuses and school enrollment statuses. Up to now we do not believe we have much improved upon our estimations of static models of employment performance, at least with respect to findings that have substantive, as distinct from methodological. content.

COMMENTS ON METHODS AND PROBLEMS IN USING SIPP

We organize our comments around three aspects of our work with SIPP: the static analysis of employment outcomes, the analysis of transitions in labor force status, and the issues of the sample size and the complex sample design of SIPP. Our research used the administratively compiled 32-month longitudinal panel of core interviews, which means that we benefitted from an internal edit that somewhat reconciles non-response items and other problems that appear in the successive waves of interviews. A disadvantage in using the longitudinal data is that we did not have the information on special topics from the non-repeated modules of each individual survey. The module information could be appended to the longitudinal survey, adding to its comprehensiveness.

The quality of the labor force and wage data appears high. The distributions of wage rates for young workers was more credible than the distribution that one of us obtained from the 1980 decennial census, showing fewer outliers and fewer anomalies. (Sao Cain, 1987, for a discussion of wage distributions for young workers in 1980.) Job amputations were recorded for only 7 percent of the white workers and 12 percent of black workers aged 20-24. Missing wage values appeared for less than 2 percent of all the race-ago groups. The data for years of schooling and for school enrollment status showed anomalous discontinuities that we decided to correct (or adjust) on the basis of neighboring values within the regular academic year.

The income data, which appears to be an especially strong feature of SIPP, played little role in our analysis, partly because the effects of family income on the labor supply behavior of youth were minimal and partly because young people had so little non-labor income that we received no benefit from the thoroughness of coverage of this type of income in SIPP. Any analysis of the labor supply of prime-age workers would find these data quite valuable, however.

Most reports of labor force data for young people were proxy reports: 70 percent for the 16-19 age group and 50 percent for the 20-24 age group. The percents were the same for both color groups. We can report that a dummy variable for a proxy report tended to show a negative coefficient for employment and a positive coefficient for a job loss transition, but we do not know whether these coefficients represent a downward bias in the proxy report of employment performance or, alternatively, are an accurate report of the behavior of those young people who are not around to answer the interviewer's questions. The negative partial derivatives measuring the relation between proxy status and employment were larger for blacks than whites, so if-.-.a proxy report gives a downward bias to the employment success of the young person, then some part of the white-black gap can be attributed to proxy reporting.

Our most ambitious undertaking with SIPP data was to add information about the labor market of residence of the young person, using those cases where his residence was identified as in an SMSA. As stated earlier, our methodological motivation was our belief that market variables are largely exogenous to the behavior we are analyzing, and that market variables have implications for policy interventions. We do not have a clear verdict on the success of this method with SIPP. The sample size for any SMSA is small, sometimes zero, and the properties of sampling variability need to be investigated. At the same time, the results for the two plausible demand-side variables that we used in our analysis are encouraging.

Our analysis of labor force transitions, as noted above, was relatively unsuccessful in terms of yielding substantive findings of interest, oven for the white sample, which seems large enough for reliable estimates. There are two major explanations for these "non-results," assuming our models wore reasonably well-specified. (1) Employment transitions of young people may be truly unsystematically related to the standard explanatory variables used, oven though the state (or incidence) of the young person's employment is systematically related to those variables. (2) The SIPP data are not sufficiently precise with respect to the timing and reporting of employment states.

In our analysis of transitions we used the general device of controlling for the particular problem (or source of the problem) with specific variables available in the SIPP records. Thus, we dealt with the errors that might be introduced by proxy reports, missing data, left-censoring, and the seam problem by the straightforward device of adding dummy variables that identified the presence of that particular problem. In the case of left-censored observations, we estimated the hazard model with and without a dummy variable for left-censoring. In the case of the seam problem, we claim to estimate successfully a negative duration dependence of the hazard of a job loss in a model in which the effect of a seam month is "held constant."

Turning to the issue of estimation with person-month observations and the problem of non-independence of residuals, we can report the results of a trial estimation of a model that used Robert Fay's CPLX procedure (1989) for computing standard errors that allow for such non-independence. In two logit models to estimate the probability of being employed or not employed in a given month among white and black men aged 20-24, t-statistics were calculated for the coefficients on 10 independent variables, all entered as dummy variables: summer month, living alone, living in a metropolitan area, living in the South, less than a high-school attainment, greater than a high-school attainment, and the ages 21, 22, 23, and 24.

Table 6 shown the comparison of the unadjusted t-statistics, the “conservative” adjustment in which the standard error was calculated assuming that the sample size. was effectively only as large as the number of persons, and, finally, t-statistics from the CPLX procedure. Three conclusions stand out. Assuming the CPLX procedure provides correct estimates, the unadjusted t-statistics are overstated by a factor of approximately two. Second, the conservative procedure yields t-statistics that are approximately 70 percent of their true value. Third, the CPLX adjustment varies from variable to variable, depending on the extent of clustering bias. Without access to a proper adjustment procedure, therefore, the conservative adjustment is better than no adjustment.

We have mentioned several times the problem of a small sample size concerning the black sample of young men and concerning our effort to measure market effects on behavior. We believe that the device of merging several SIPP panels should be tried an a way of doubling or tripling the sample size. The data for labor force measurements are consistent across panels, and the panel identification can be used as an additive or interactive control variable if there is reason to believe the panels differ in some administrative way.

Just a few years ago SIPP had the reputation of being forbiddingly complex and difficult to use. There have since been many editing improvements, reformatting of the data in more accessible forms, and experience in using it, so this reputation is no longer correct. Our work may be taken as illustrative of the capacity of SIPP to yield useful research output and as indicative of ways to improve upon our methods and upon the mechanical manipulations of SIPP data.

REFERENCES

CAI.N, GLEN G. (1987), "Black-White Differences in Employment of Young People: An Analysis of 1980 Census Data," Institute for Research on Poverty, Discussion Paper no. 844-87. Madison: University of Wisconsin.

CAIN, GLEN G., and FINNIE, ROSS E. (1990), "The Black-White Difference in Youth Employment: Evidence for Demand-Side Factors," Journal of Labor Economics, 8, S364-395.

CAIN, GLEN G. and GLEASON, PHILIP M. (1990), "The Earnings of Black and White Youth and Its Relation to Poverty," Presented at the Meetings of the Association for Public Policy Analysis and Management, San Francisco, available from the LaFollette Institute of Public Affairs, University of Wisconsin.

ELLWOOD, DAVID T. (1982), "Teenage Unemployment: Permanent Scare or Temporary Blemishes?," in The Youth Labor market Problem: its Nature, Causes and Consequences, edited by Richard B. Freeman and David B. Wise, Chicago: The University of Chicago Press, pp. 349-385.

FAY, ROBERT E. (1989), "CPLX: Contingency Table Analysis for Complex Sample Designing Program Documentation, Center for Demographic Studies, U.S. Bureau of the Census.

Table 1

Employment-to-Population Ratios (E/p) for youth Ages 16-19 and 20-24 by Race and Gender Selected Years, 1954 - 1988

	Men				Women			
	16 - 19		20 - 24		16 - 19		20 - 24	
Year	White	Black	White	Black	White	Black	White	Black
1954-55	.31	.53	.79	.77	.37	.26	.43	.42
1964-63	.46	.39	.80	.80	.33	.21	.46	.45
1974-75	.52	.29	.77	.65	.43	.21	.58	.45
1984-85	.49	.25	.78	.59	.47	.22	.67	.46
1988	.52	.29	.80	.64	.50	.26	.70	.51

Source: U.S. Department of Labor, Bureau of Labor Statistics, Handbook of Labor Statistics, Bulletin 2340, August 1989, Compiled from Tables: 3, 15, and 16; pp. 12-18, 63-75.

TABLE 2

1985 SIPP, 32-Month Period, 1984-1987, Means of Employment Performance Measures, Black and White Men

Performance	Age 16-19			Age 20-24		
Measure	Black	White	Ratio	Black	White	Ratio
E/P	.28	.52	.54	.58	.79	.73
Hours Worked (per week)	7.7	13.8	.56	21.6	31.3	.69
Hours Worked (workers only)	27.6	27.0	1.02	37.0	39.7	.93
Earnings (per week)	$34	$61	.56	$132	$205	.64
Wage (workers only)	$4.26	$4.24	1.00	$5.87	$6.47	.91
Number of Person-Months	4,033	24,795			3,280	35,462
Persons	260	1,665		270	2,437	

TABLE 3

Probit Coefficients and Estimated "Effect" of Independent Variables on the Probability of Being Employed, Young Men, 20-24 (1985 SIPP)

	WHITE (E/P=.80)		BLACK (E/P=.58)	
VARIABLES	COEFE	PARTIAL DERIV	COEFF	PARTIAL DERIV
HISPANIC	.002	.000	.392*	.150
YEAR = 1984	-.097	-.024	.200	.077
1985	-.007	-.002	.155	.059
1986	-.010	-.003	.088	.034
SUMMER	.200*	.049	.045	.017
AGE = 21	.035	.009	-.124	-.048
22	.076	.019	.015	.006
23	.212*	.052	.135	.051
24	.135*	.033	.065	.025
VETERAN	.089	.022	-.284	-.109
HEALTH LIM.	-.601*	-.147	-.375*	-.143
EDUC, < H.S.	-.366*	-.090	-.693*	-.265
EDUC, SOME COLL	.138*	.034	-.287*	-.110
EDUC, COLL DEGREE	.048	.012	.168	.064
PROXY RESPONSE	-.097*	-.024	-.374*	-.143
ONE PARENT HH	-.361*	-.088	-.282	-.108
TWO PARENT HH	-.311*	-.076	-.159	-.061
MARRIED	.376*	.057	1.870*	.503
KIDS	.056	.014	-.120	-.046
NONLAB INC (DUMMY)	.029	.007	-.206	-.079
FAM INC ($500/Mo)	-.003	-.002	.007	-.004
INC*MARRIED ($500)	-.034*	- - -	-.168*	- - -
ENROLLED	-.777*	-.190	-.544*	-.208
LOW RENT HOUSING	-.799*	-.195	.443	.169
PUBLIC HOUSING	-.132	-.032	-.268	-.102
CENTRAL CITY	.044	.026	.330*	-.105
SEG. INDEX	.016	.006	.124*	.002
CENT CITY * SEG	-.024	- - -	-.160*	- - -
LOG POP (1 STD DEV)	.007	.002	.033	.013
% BLACK (1 STD DEV)	.040*	.010	.037	.014
UNEMP RATE (1986)	-.022*	-.005	-.082*	-.031
LOG RET SALES (1 SD)	.053*	.086	.245*	.094
N (PERSON-MONTHS)	23,215		2,578	

Note - 1) Asterisk represents statistical significance at 10% (two-tailed test)
2) Partial derivatives of FAN INC, LOG POP, t BLACK, and LOG RET SALES represents the effect on employment probability of a change in the independent variable of a magnitude noted in the parentheses.

TABLE 4

Probit Coefficients and Estimated "Effect" Of Independent Variables on the Probability of Being Employed, Young Men, 16-19 (1985 SIPP)

VARIABLES		WHITE (E/P = .52) COEFF	PARTIAL DERIV	BLACK (E/P= .29) COEFF	PARTIAL DERIV
HISPANIC		-.263*	-.105	-1.105*	-.333
YEAR =	1984	.074	.030	.091	.027
	1985	.056	.022	-.172	-.052
	1986	.083	.033	-.067	-.020
SUMMER		.285*	.113	.163	.049
AGE = 17		.381*	.152	.572*	.173
18		.534*	.213	.725*	.219
19		.859*	.342	.976*	.294
PARENTS ON WELF		-.330*	-.132	.255*	.077
HEALTH LIM.		-.190*	-.076	.228	.069
EDUC, BEHIND		-.281*	-.112	-.818*	-.247
EDUC, ON TRACK		.062	.025	-.240	-.072
EDUC, SOME COLL		-.315*	-.126	-.857*	-.259
PROXY RESPONSE		-.236*	-.094	-.606*	-.183
ONE PARENT HH		-.045	-.018	-.333*	-.101
LIVING ALONE		.140*	.056	-.205	-.062
NONLAB INC (DUMMY)		-.095*	-.038	.547*	.165
FAMILY INC ($500/MO)		-.014*	.006	-.021	.006
ENROLLED		.419*	-.167	.430*	-.130
LOW RENT HOUSING		.355*	.141	-.354	-.107
PUBLIC HOUSING		-.178	-.071	-.037	-.011
CENTRAL CITY		-.012	-.037	.962	.003
SEG. INDEX		.029*	.007	.223*	-.010
CENT CITY * SEG		-.030	- - -	.353*	- - -
LOG POP (1 STD DEV)		.034	.013	.064	.019
% BLACK (1 STD DEV)		-.028	-.011	.149*	-.045
UNEMP RATE (1986)		-.009	-.003	.116*	-.035
LOG RET SALES (1 SD)		.028	.011	.348*	.105
N (PERSON-MONTHS)		16,293		2,931	

Note - 1) Asterisk represents statistical significance at 10% (two-tailed test)
2) Partial derivatives of FAM INC, LOG POP, % BLACK, and LOG RET SALES represents the effect-on employment probability of a change in the independent variable of a magnitude noted in the parentheses

TABLE 5
Employment Hazard Logit Model
Coefficient Estimates for White and Black Males, 20-24

		WHITE	BLACK
INTERCEPT		-6.137*	20.819
TIME		-.083*	-.009
WAVE MONTH #2		.375*	1.168
WAVE MONTH #3		.578*	1.343
WAVE MONTH #4		1.354*	3.023*
WAGE		-.010	-.184
WAGE MISSING		.511*	.585
HOURS PER WEEK		-.010*	-.022
HISPANIC		-.043	-.477
YEAR =	1984	-.310	.245
	1986	.452*	.376
	1987	.360*	-.682
SUMMER MONTH		.020	.059
AGE =	21	-.147	.081
	22	-.115	-.904
	23	-.342	-.455
	24	-.299	-.069
	25	-.336	-.485
VETERAN STATUS		.079	-.167
HEALTH LIMITATION		.467*	.191
EDUC, < HIGH SCHOOL		.387*	-.211
EDUC, SOME COLLEGE		.057	-.659
EDUC, COLL DEGREE		-.163	-.422
PROXY RESPONSE		-.224*	-.217
ONE PARENT HH		.534*	.249
TWO PARENT HH		.387*	-.082
MARRIED		-.199	-1.075
KIDS		.057	.197
NONLAB INC (DUMMY)		-.074	-.100
FAM INC ($500/Mo)		.029	-.043
INC*MARRIED ($500)		-.015	.023
LOW RENT HOUSING		-.029	1.220
PUBLIC HOUSING		-.004	.533*
CENTRAL CITY		-.301	-.439
SEGREGATION INDEX		-.066	.038
CENT CITY * SEG		.107	.104
LOG POP (1 STD DEV)		-.100	.034
% BLACK (1 STD DEV)		.092	-.296
UNEMP RATE (1986)		.015	-.065
LOG RET SALES (1 SD)		.480	-.394
N (PERSON-MONTHS)		20,877	1,589

Note - 1) Asterisk represents statistical significance at 10% (two-tailed test)
2 Coefficients of FAM INC, INC*MARRIED, LOG POP, % BLACK, and LOG RET SALES have been multiplied

by the number noted in parentheses.

TABLE 6
CPLX Correction for Non-independence Across observations

T-Statistics for Logit Estimates of White Males, Aged 20-24

Variable	Orig	Cons	CPLX	Cons/Orig	CPLX/Orig	CPLX/Cons
Summer	3.19	1.26	3.16	.39	.99	2.51
Alone	7.11	2.80	3.95	.39	.56	1.41
Metro	4.91	1.93	2.85	.39	.58	1.48
South	8.90	3.51	5.11	.39	.57	1.46
ED<HS	14.24	5.61	7.26	.39	.51	1.29
ED>HS	5.52	2.18	2.52	.39	.46	1.16
Age, 21	5.03	1.98	2.80	.39	.56	1.41
Age, 22	5.30	2.09	2.87	.39	.54	1.37
Age, 23	5.17	2.04	2.96	.39	.57	1.45
Age, 24	7.64	3.01	5.13	.39	.67	1.70
			T-Statistics for Logit Estimates of Black Males, Aged 20-24			
Summer	0.75	0.29	0.80	-.39	1.07	2.76
Alone	6.20	2.41	3.59	.39	0.58	1.49
Metro	2.16	0.84	0.62	.39	0.29	0.74
South	2.08	0.81	1.00	.39	0.48	1.23
ED<HS	8.14	3.16	3.70	.39	0.44	1.17
ED>HS	5.19	2.02	2.77	.39	0.53	1.37
Age, 21	1.95	0.76	1.15	.39	0.59	1.51
Age, 22	0.86	0.33	0.49	.39	0.57	1.48
Age, 23	2.57	1.00	1.23	.39	0.48	1.23
Age, 24	5.79	2.25	2.64	.39	0.46	1.17

Notes: Orig = the originally reported T-statistics, unadjusted
Cons = "conservative" T-statistics computed by deflating the reported T-statistic by the factor of the square root of the ratio of the number of persons in the sample to the number of person-months