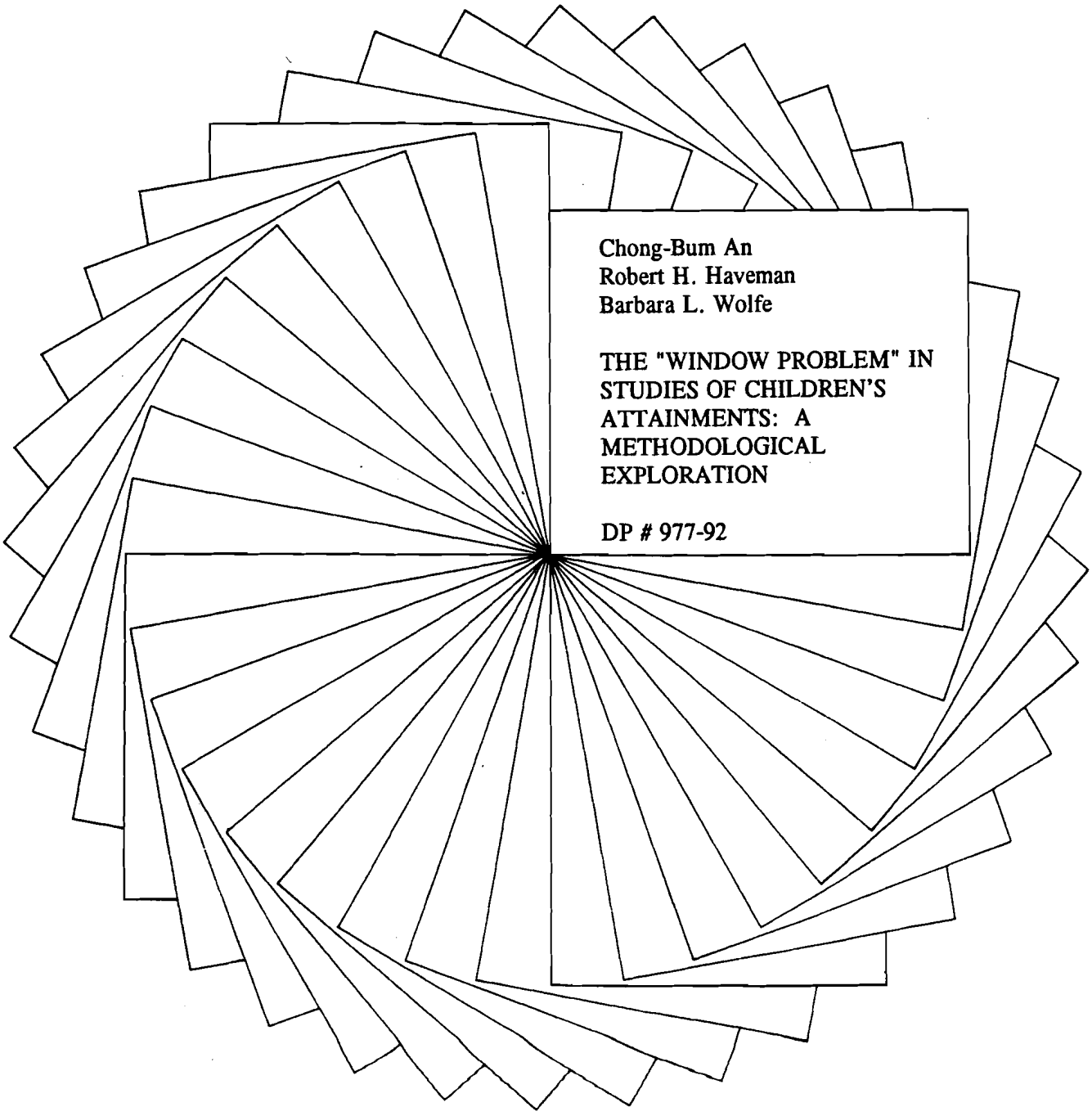
University of Wisconsin-Madison

Institute for Research on Poverty

Discussion Papers



Chong-Bum An
Robert H. Haveman
Barbara L. Wolfe

THE "WINDOW PROBLEM" IN STUDIES OF CHILDREN'S ATTAINMENTS: A METHODOLOGICAL EXPLORATION

DP # 977-92

Institute for Research on Poverty
Discussion Paper no. 977-92

The "Window Problem" in Studies of Children's Attainments: A Methodological Exploration

Chong-Bum An
Daewoo Research Institute
Seoul, Korea

Robert H. Haveman
Department of Economics
La Follette Institute of Public Affairs
Institute for Research on Poverty
University of Wisconsin-Madison

Barbara L. Wolfe
Department of Economics
Department of Preventive Medicine
Institute for Research on Poverty
University of Wisconsin-Madison

July 1992

The authors wish to thank the participants at a seminar at the Graduate School of CUNY, Rebecca Blank, Robert Hauser, Christopher Jencks, Charles Manski, and Robert Mare, for their helpful comments on this paper, and the Russell Sage Foundation and the Institute for Research on Poverty for their financial support.

Abstract

Numerous studies of the determinants of children's attainments rely on observations of circumstances and events at age 14 as proxies for information over the entire childhood period. These age-14 observations are often referred to as "window" observations. Using twenty-one years of panel data from the Michigan Panel Study of Income Dynamics on 825 children who were 14 to 16 years old in 1979, we evaluate the effects of using window information in models of the determinants of the attainments (e.g., education and nonmarital fertility) of young adults. Correlations between window and full-childhood variables are presented, along with five tests of the reliability of estimates based on window measurements. The tests are designed to evaluate the differential effects of data accuracy, multiple occurrence of events, duration of circumstances, and the timing of events or circumstances on the reliability of window and full-childhood information. We conclude that most of the standard window variables serve as weak proxies for multiyear information in such models, and draw the implications of these findings for future data collection and research.

The "Window Problem" in Studies of Children's Attainments: A Methodological Exploration

I. INTRODUCTION

Numerous studies in the 1980s employed recently available longitudinal microdata on families and individuals to estimate the effects of family circumstances and events early in an individual's life on his/her attainments later in life. The attainments analyzed included schooling, fertility behavior (especially, teen nonmarital births), welfare recipiency, and labor market success. Hypotheses drawn from economics and sociology concerning the potential effect of various circumstances or events experienced by a child while growing up on the child's potential for later success or failure were tested. The circumstances or events included parental occupation and education; growing up in a mother-only family, a poor family, or a family receiving welfare; and experiencing a parental divorce or a geographic move.

In an ideal study of this sort, longitudinal information on a rich set of circumstances and events spanning the entire childhood period would be available for testing these hypotheses. However, many of the published studies used longitudinal data that do not contain such long-duration childhood information. In fact, some prominent longitudinal data sets (e.g., the National Longitudinal Survey of Youth) do not begin collecting information on individuals until they are 14 years old; researchers using these data sets are therefore constrained from observing events and circumstances in the preadolescent period. In other cases, researchers desire to study outcomes later in life (e.g., attainments among 20 to 30 year olds). When longitudinal data are limited in the length of time over which they have been collected, choosing to focus on the attainments of older individuals means sacrificing information on circumstances earlier in their lives. Because of either data limitations or researcher choice, then, information on these family, school, and neighborhood variables during a

brief observation "window," often a single year, is used as a proxy for information spanning the entire childhood period.[1]

Several questions are pertinent to evaluating the reliability of studies of children's attainments that use variables based on window observations. The primary questions pertain to the accuracy of data reported, the multiple occurrence of events, the duration of circumstances, and the timing of events or circumstances during childhood.

> -Is a single-year report on a family background or economic status (income) variable as accurate as, say, an average of multiple reports?
>
> -Can a variable based on an observation window reliably capture the effect of events that might occur intermittently or with a low frequency throughout the childhood years (e.g., parental separations)?
>
> -Can a window observation reliably measure the effects of circumstances that may be present over longer or shorter durations during childhood, and for which duration of occurrence may matter (e.g., living in poverty)?
>
> -Can a window observation reliably measure the effects of events or circumstances that may be present during particular periods of childhood, and for which the timing of occurrence matters (e.g., geographic moves in, say, early childhood)?[2]

By definition, variables specified by a window observation sacrifice information relevant for assessing the effect of events and circumstances for which multiple occurrence, timing, and duration are relevant. The question, then, is one of statistical reliability and accuracy in measurement: How reliable (or accurate) are estimates of effects based on variables constructed from window information relative to estimates based on variables constructed from more complete longitudinal information?[3]

In this paper, we provide evidence on the reliability and accuracy of variables based on truncated observations of important childhood circumstances and events in empirical estimates of the determinants of children's success. Section II describes our approach to answering this question and the criteria that we use in assessing the accuracy of estimates that rely on variables based on window observations. Section III provides our estimates and assessment; Section IV concludes.

II. METHODS AND CRITERIA

Our assessment of the reliability of the estimates of effects in attainment models based on snapshot observations of parental situations rests on a series of "tests," all of which employ longitudinal data on a sample of nearly two thousand children. These data are taken from the Michigan Panel Study of Income Dynamics (PSID), and include children who were both aged 0 to 6 in 1968 and in the sample in 1989. For each child, detailed annual information on family background (age, race, one/two parents present, location, number of siblings), resources (family income, income source, adult labor supply, home ownership), and events (parental separation, remarriage, change in geographic location) is recorded and made specific to the age of the child. The attainments of these children in young adulthood are also known.

We employ a variety of comparisons and tests designed to estimate reliability in each of the dimensions noted in the questions above; no single test can answer the questions in all dimensions. Our strategy is to present tests based on a variety of viewpoints regarding the definition and interpretation of reliability.

First, we present a set of correlations between variables based on a window observation and variables constructed from multiple years of longitudinal information. The correlations indicate the extent to which the snapshot variables serve as accurate measures of the diverse and/or changing character of family situations at a point in, or over a period of, time.

Second, we conduct a series of tests based on estimates of six single-equation probit models relating family background and family circumstance/event variables to each of two limited dependent variable outcomes--high school graduation and teen out-of-wedlock birth; three models are estimated for each of the two dependent variables.[4] The three models specify a variety of family circumstances and events[5] in different ways, allowing us to test whether the estimates based on the window

variables are comparable to estimates based on variables reflecting multiple occurrences, duration, and timing of the circumstance or event. The circumstance/event variables subject to the tests for window-reliability are specified as follows:

-number of occurrences (or years) of a circumstance or event for each of three childhood periods[6] (three variables);
-number of occurrences (or years) of a circumstance or event for the age 6-14 period (one variable);
-occurrence of a circumstance or event at age 14 (the window variable).

Using these models, we undertake five tests or comparisons designed to assess the reliability of the single-year, age-14 measurement of these circumstance/event variables:

-Test 1: A likelihood ratio test of the null hypothesis that adding information from the age 6-14 period to a specification including the age-14 window variable does not significantly improve the fit of the estimated model.

-Test 2: A sign-and-significance comparison in which the estimated coefficients on variables based on information from the age-14 window are compared with the estimated coefficients on the same variables measured over periods of varying length during the childhood period.

-Test 3: A comparison of the magnitude of the effects of simulated changes between those window and the multiyear variables which conform in terms of sign and statistical significance.

-Test 4: A test of the conformance between the implicit policy advice of the estimates based on window and multiyear variables.

-Test 5: A test of the ability of models using window variables to identify successful outcomes, relative to that of models relying on multiyear information.

III. RESULTS

Correlations

Table 1 presents simple correlations between pairs of variables based on (1) occurrence at age 14; (2) occurrence at any time during the age 6-15 period; and (3) number of occurrences during the age 6-15 period. The three measures reflect differences in multiple occurrences and duration. The variables are designated as either circumstance or event variables; the former represent economic or

TABLE 1

Correlations among Age-14 Variables and Variables Reflecting Age 6-15 Observations (N=825)

EVENTS		
	Ever Separated, Ages 6-15	Total Separations, Ages 6-15
Parental separation at age 14	.32	.40
	Ever Remarried, Ages 6-15	Total Remarriages, Ages 6-15
Parental remarriage at age 14	.27	.25
	Ever Moved, Ages 6-15	Total Moves, Ages 6-15
Moved at age 14	.27	.28
CIRCUMSTANCES		
	Ever Poor, Ages 6-15	Years Poor, Ages 6-15
Poor at age 14	.51	.74
	Average Income-to-Needs Ratio, Ages 6-15	
Income-to-needs ratio at age 14	.88	
	Ever Lived in SMSA, Ages 6-15	Years Lived in SMSA, Ages 6-15
Lived in SMSA at age 14	.83	.96
	Mother Ever Worked, Ages 6-15	Years Mother Worked, Ages 6-15
Mother worked when child was age 14	.48	.71

Source: Authors' computations based on Michigan Panel Study of Income Dynamics.

parental circumstances at different points during the childhood period, and the latter identify discrete changes occurring at different times. The correlations are estimated for the entire sample of observations.

Several comparisons are relevant. First, of the thirteen correlations shown in Table 1, only three exceed .75; six are at or below .40, suggesting that a variable constructed from a one-year observation is generally a weak proxy for variables reflecting multiple, relatively rare, or transitory occurrences over the childhood period, and that such one-year reports are also less accurate than an average of multiple-year reports.

Substantial differences in correlation coefficients are observed between the event and the circumstance variables. The age-14 window variable appears to capture circumstances more readily than events, which is not surprising. Economic and family circumstances are rather persistent variables; observed family income relative to needs when a child is 14 is not likely to be greatly different than income/needs at other points during childhood.[7] Conversely, the age-14 variable appears to be a rather poor proxy for events that do not occur regularly during the childhood years; the correlations for the event variables range from .25 to .40.

Moreover, the age-14 window variable tends to be more highly correlated with both event and circumstance variables that are measured in terms of multiple occurrences, rather than as dummy variables. Again, this is not surprising; the more often during childhood that an event or circumstance occurs, the more likely that an observation at a particular age will capture it. Finally, variables that tend to change slowly or infrequently over time (e.g., family income/needs or living in a metropolitan area) are better proxied by an age-14 observation than those that may change with some frequency (e.g., being in poverty).

Tests of Comparability

Test 1--Likelihood Ratio Test. Table 2 presents the tests necessary to determine if estimation results from models that use variables reflecting the occurrence of circumstances and events over the childhood period provide significantly more information compared to models that use variables based on age 14-information. Test results are shown for both the high school graduation and the teen out-of-wedlock birth outcomes.

For both of the outcomes, the tests indicate that models that include information over the childhood period yield estimates that are significantly different than those from models that include only the age-14 variables. This is the case both for models that add data over the entire 6-14 period and for those which add data for the three time periods separately.[8]

For the education outcome, the null hypothesis that there is no significant improvement from the addition of information over the childhood period to that observed at age 14 is rejected at the .10 (entire age 6 to 14 period) and .05 (three periods over age 6 to 14) significance levels. For the out-of-wedlock birth outcome, the null hypothesis is rejected at the .05 and .01 significance levels, respectively.

Conversely, the likelihood ratio test indicates that estimates obtained from models that add the age-14 variables to the multiyear childhood information are not significantly different from those based only on the age 6-14 variables. These results suggest that information on duration and timing do matter, but that adding information from the age-14 measurement to information measured over ages 6-14 does not.

The final test, comparing the results using variables from the three time periods to those using variables reflecting the entire age 6-14 period, indicates a significant difference (at the .10 level) in the teen out-of-wedlock birth estimates, but not in the high school graduation estimates.

TABLE 2

Likelihood Ratio Tests of Null Hypothesis that Variables Based on Multiyear Information Add No Relevant Information to Those Based on Age-14 Window Information

	Likelihood Ratio	Degrees of Freedom	Significance Level
HIGH SCHOOL GRADUATION MODELS			
A. Model based on age-14 window variables	-389.66	24	
B. Model based on age 6-14 variables	-386.62	24	
C. Model based on three age-period variables	-382.27	34	
D. Model based on age-14 window variables and age 6-14 variables	-383.15	31	
E. Model based on age-14 window variables and three age-period variables	-376.96	39	
Test 1: Model A vs. model D			< .10
Test 2: Model B vs. model D			< .50
Test 3: Model A vs. model E			< .05
Test 4: Model B vs. model E			< .25
Test 5: Model B vs. model C			< .50
TEEN OUT-OF-WEDLOCK BIRTH MODELS			
A. Model based on age-14 window variables	-136.48	14	
B. Model based on age 6-14 variables	-132.51	14	
C. Model based on three age-period variables	-123.17	26	
D. Model based on age-14 window variables and age 6-14 variables	-128.96	21	
E. Model based on age-14 window variables and three age-period variables	-119.55	31	
Test 1: Model A vs. model D			< .05
Test 2: Model B vs. model D			< .50
Test 3: Model A vs. model E			< .01
Test 4: Model B vs. model E			< .25
Test 5: Model B vs. model C			< .10

Source: Authors' computations based on Michigan Panel Study of Income Dynamics.

Test 2--Sign-and-Significance Comparison. In this comparison, the estimated coefficients on the age-14 event/circumstance variables are compared to coefficients on variables measured (1) during early childhood, ages 6-8; (2) during adolescence, ages 12-14; and (3) over the entire age 6-14 period.[9] Hence, there are fifteen comparisons for the education models (three comparisons for each of five variables) and eighteen for the out-of-wedlock birth models, for a total of thirty-three comparisons.

We conclude that the window and the multiyear variables do not convey the same information regarding "effects" if, at the .2 level of significance, either the two coefficients have different signs or the coefficients have the same sign but only one of them is statistically significant.[10]

Table 3 presents the thirty-three possible comparisons and indicates the age-14 and multiyear coefficients that are comparable to each other. In eleven of the fifteen possible comparisons for the high school graduation models, the age-14 variable is judged to yield statistically comparable information to that of the multiyear variables. In the out-of-wedlock birth models, however, comparability among the pairwise comparisons is far less frequent. In only seven of the eighteen comparisons is our sign-and-significance test met. Overall, the test is passed in only slightly more than one-half of the possible cases (eighteen of thirty-three).

In six of the eighteen cases in which comparability is observed, neither coefficient has statistical significance at the .2 level or less. In but twelve of the thirty-three cases do both the age-14 and multiyear variables have statistical significance at the .2 level. Eight of these twelve same-signed and significant comparisons are in the education estimates.

Test 3--Magnitude of Simulated Effect Test. This comparison concerns the implications for policy of the results from the thirty-three age-14 vs. multiyear pairwise comparisons. For those twelve cases in which the coefficients have the same sign and are statistically significant at the .2 level, we compare the magnitude of the effect on the dependent variable of equivalent,

TABLE 3

Comparability of Age-14 and Multiyear Coefficients

	Education		Out-of-Wedlock Birth
Years in Poverty		Average Ratio of Income to Needs	
-Ages 6-8	NC		NC
-Ages 12-14	NC		–
-Ages 6-14	–		–
Years Living in SMSA		Years Living in SMSA	
-Ages 6-8	–		+
-Ages 12-14	NC		NC
-Ages 6-14	–		+
Number of Parental Separations		Number of Parental Separations	
-Ages 6-8	0		NC
-Ages 12-14	0		NC
-Ages 6-14	0		NC
Years Mother Worked		Years Receiving Welfare	
-Ages 6-8	NC		NC
-Ages 12-14	+		0
-Ages 6-14	+		0
Number of Geographic Moves		Number of Geographic Moves	
-Ages 6-8	–		NC
-Ages 12-14	–		NC
-Ages 6-14	–		0
		Years Living with One Parent	
-Ages 6-8			NC
-Ages 12-14			NC
-Ages 6-14			NC

Source: Authors' computations based on Michigan Panel Study of Income Dynamics.

NC = Not comparable.
– = Coefficients both negative and significant < .2.
+ = Coefficients both positive and significant < .2.
0 = Neither coefficient significant < .2.

one-standard-deviation simulated changes in the age-14 and the multiyear independent variables. The results are presented in Table 4.

If we accept as equivalent simulated effects that are within 2 percentage points of each other, we find that nine of the twelve same-signed and significant cases indicate similar quantitative effects. A more demanding criterion of a difference of no more than 1 percentage point indicates comparable quantitative effects in but six of the twelve cases.

Test 4--Conformance of Policy Advice Test. Here, we take the coefficient on the multiyear variable as providing the correct implicit advice regarding policy, and ask if the one-year, age-14 variable yields the same advice. We interpret a significance level of less than .2 as providing a weak basis for policy advice. Table 3 and Appendix B serve as the basis for this discussion.

Using this standard, eighteen of the thirty-three coefficients based on multiyear variables indicate that policy intervention would affect the outcome with .8 confidence or more. Only twelve of the thirty-three age-14 coefficients have the same sign as the multiyear variables and meet the .8 confidence test. Stated alternatively, in one-third (six of eighteen) of the cases in which policy action would seem warranted, the window observation fails to provide this advice.

Test 5--Identification of Successful Outcomes Test. An important criterion in appraising an estimated limited dependent variable model is its ability to accurately identify the occurrences of an event that are observed in the data. The models that we have estimated using variables constructed from age 6-14 information yield substantially more accurate identifications than do the models using the age-14 variables. For the high school graduation model, those who do not graduate are correctly identified only 10 percent of the time in the models using the age-14 variable; models relying on information over the age 6-14 period correctly identify nearly 21 percent of the dropouts. In the teen out-of-wedlock birth model, the comparable percentages are 7 and 25, respectively.

TABLE 4

Simulated Changes in Coefficients of Age-14 and Multiyear Variables, Given an Increase of One Standard Deviation in the Measures of Multiyear Variables

Multiyear Variable	Change in Coefficients of: Age-14 Variables[a]	Multiyear Variables[b]
HIGH SCHOOL GRADUATION MODELS (Base Probability of Graduating = .770)		
Years in Poverty, Ages 6-14	-2.09	-4.10
Years Living in SMSA, Ages 6-8	-2.66	-8.09
Years Living in SMSA, Ages 6-14	-2.66	-2.44
Years Mother Worked, Ages 12-14	+3.47	+2.40
Years Mother Worked, Ages 6-14	+3.47	+2.61
Number of Geographic Moves, Ages 6-8	-3.88	-3.52
Number of Geographic Moves, Ages 12-14	-3.88	-2.26
Number of Geographic Moves, Ages 6-14	-3.88	-5.10
OUT-OF-WEDLOCK BIRTH MODELS (Base Probability of Out-of-Wedlock Birth = .137)		
Average Ratio of Income to Needs, Ages 12-14	-1.50	-1.46
Average Ratio of Income to Needs, Ages 6-14	-1.50	-1.49
Years Living in SMSA, Ages 6-8	+2.80	+7.15
Years Living in SMSA, Ages 6-14	+2.80	+2.76

Source: Authors' computations based on Michigan Panel Study of Income Dynamics.
Note: As a guide to reading the table, consider the following: The first row indicates that increasing the poverty variable as measured at the age-14 window by one standard deviation would reduce the probability of graduating high school by .021 (-2.09 ÷ 100); thus, the base probability of .770 would be reduced to .749, or by a little more than 2 percentage points. Increasing the same variable as recorded over ages 6 to 14 would reduce the probability of graduating high school by .041 (-4.10 ÷ 100); thus, the base probability of .770 would be reduced to .729, or by more than 4 percentage points. See Appendix B for coefficients of variables and Appendix C for their means and standard deviations.

[a]In each instance, the age-14 variable is the measure of the multiyear variable at age 14; for example, in row 1, the age-14 variable is In Poverty at Age 14.

[b]In each instance, the multiyear variable in this column is the same as the multiyear variable in the stub.

IV. CONCLUSION

This exploration has yielded rather discouraging results regarding the accuracy of single-year window observations (relative to averages of multiyear observations) and the reliability of empirical estimates based on such window variables. We conclude that, in general, single-year "window" variables serve as weak proxies for multiyear information recorded over the life of the child.[11]

Clearly, then, important information is lost when circumstances or events observed in but a single year are used to represent the complex and changing environment in either a different developmental period or over a longer period of time. Those estimates in the published literature based on one-year window observations should be interpreted very carefully. They may inaccurately reflect the effect of circumstances and events in but the single year of observation; they often provide biased and misleading estimates of the effects of a child's environment over a longer, or for a different, period of time.

These results also highlight a basic issue of data collection in the social sciences. Our results suggest a high priority for the collection of longitudinal information on individuals and families extending over the entire period of childhood. Such efforts are extremely costly, however, and require long periods of observation before data are ripe for analysis. An alternative might be the compilation of retrospective information on parental situations at various points during childhood from respondents who are children. While the costs of the latter strategy are smaller, the ability to accurately capture correctly timed information on important aspects of parental circumstances and events through this strategy is weaker. In any case, the importance of the process which determines whether children succeed or fail in later life--and of the role that family resources, stressful events, and general environment play in this process--suggests that careful attention should be paid to developing data that will permit reliable estimates of the effect of important environmental, family, and individual variables on attainments.

APPENDIX A

Recent Studies of the Determinants of Children's Attainments Relying on a Limited Window of Observation of Childhood Circumstances and Events

The following is a selection of studies that analyze the influence of family events and circumstances on children's achievements; the measurement of those events and circumstances is for a truncated period (often one year) sometime during the child's adolescent years. The studies are limited to those that have been published since 1980.

1. Brooks-Gunn, Jeanne, Greg J. Duncan, Pam Kato, and Naomi Sealand. 1991. "Do Neighborhoods Influence Child and Adolescent Behavior?" Draft mimeo.

   The PSID sample used in this study focuses on eighteen hundred black and white women who were observed between ages 14 and 18. Hence, the neighborhood variables, the welfare ratio, and whether or not the family was headed by the mother were all observed when the women were aged 14. The outcomes were dropping out of school and experiencing a teen out-of-wedlock birth.

2. Crane, Jonathan. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." American Journal of Sociology, 96(5): 1226-59.

   Has a sample of 113,997 16 to 19 year olds and observes family income, head's occupational status, household structure, and family size at the age of the child when teen childbearing and dropping out of high school were observed. The data are cross-sectional data.

3. Datcher, Linda. 1982. "Effects of Community and Family Background on Achievement." Review of Economics and Statistics, 64: 32-41.

Has a sample of males aged 13 to 22 in 1968 and used family data as of 1968. Family income, family size, parental expectations, receipt of transfer income, zipcode variables, and some psychosocial variables are all subject to the window problem. These variables were measured for males who were ages 13 to 22 in 1968, and then these males were observed in 1978.

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

The same sort of data set as Datcher, but the individuals on the PSID were aged 10 to 17 in 1968. Again, family background and neighborhood variables were observed for the observations at these ages.

5. Mayer, Susan. 1991. "The Effect of Schools' Racial and Socioeconomic Mix on High School Students' Chances of Dropping Out." Unpublished paper. Northwestern University.

Her data are for 26,321 students in the tenth grade and hence about 15 years old. Parental characteristics included father's occupation, whether they owned their home, had two or more cars and a dishwasher, and whether or not the family was headed by the mother.

6. Mare, Robert. 1980. "Social Background and School Continuation Decisions." Journal of the American Statistical Association, 75(370): 295-305.

Control variables were family income, father's occupational status, and living on a farm, all when the youth was 16. (On page 296 he has some interesting language regarding the limitations of these age-16 variables and their problems.)

7. McLanahan, Sara. 1985. "Family Structure and the Reproduction of Poverty." American Journal of Sociology, 90(4): 873-901.

The data are from the PSID, 1978 wave, and consist of respondents who were 17 to 27 in 1978. A wide variety of family variables were measured for the individual as of age 17, including parents' marital status, years since marital disruption, region of residence, city size, family welfare ratio, mother's employment status, and family welfare receipt.

8. Astone, Nan Marie, and Sara McLanahan. 1991. "Family Structure and High School Completion: The Role of Parental Practices." American Sociological Review, 56: 309-320.

They use a sample of fifty-eight thousand students in the High School and Beyond Study who were sophomores in 1980. The parental psychosocial variables, family structure, change in family structure, father's occupation, family income, household possessions, region, and urban-rural location were all observed for the sophomore year, or when the children were about 16 years old.

9. Case, Anne, and Lawrence Katz. 1991. "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths." Unpublished paper. Harvard University.

They use data from the 1989 NBER Boston Youth Survey, containing information on twelve hundred youths aged 17 to 24 in high-poverty areas in inner-city Boston. Some of the family background variables were measured at age 14 (father or mother present, parents not married) but others were measured at the time of the survey so that the youths would have been 17 to 24. These include living arrangements, in public housing, and family income of others in the household.

10. Hauser, Robert, and William Sewell. 1986. "Family Effects in Simple Models of Education, Occupational Status, and Earnings: Findings from the Wisconsin and Kalamazoo Studies." Journal of Labor Economics, 4(3): S83-S115.

Uses the Wisconsin High School data, a random sample of about ten thousand high school seniors in 1957. The data on family income were for the year of the survey, as was information on students' aspirations, number of siblings, father's occupation, and marital status--hence, measured when the children were about 18 years old. The same must be true of all of the previous studies using the Wisconsin data.

11. Sewell, William, Robert Hauser, and Wendy Wolf. 1980. "Sex, Schooling and Occupational Status." American Journal of Sociology, 86(3): 551-583.

See comments on #10. This paper also contains numerous references to other studies that use the Wisconsin data.

12. Manski, Charles, Gary Sandefur, Sara McLanahan, and Daniel Powers. 1990. "Alternative Estimates of the Effect of Family Structure During Adolescence on High School Graduation." Journal of the American Statistical Association, 87(417): 25-37.

This study uses the NLSY, and uses individuals aged 14 to 17 in 1979. Variables include family structure at age 14 and region of residence at age 14. The authors are clear that they are trying to measure the effects of these background characteristics when the children were adolescents. Still, there is but one year of observation designed to capture the entire period.

13. Duncan, Greg, and Saul Hoffman. 1990. "Welfare Benefits, Economic Opportunities, and Out-of-Wedlock Births Among Black Teenage Girls." Demography, 27(4): 519-535.

PSID data used on 874 black women beginning at age 14. Family background statistics measured at age 14 include region, city size, family welfare recipiency, family income, single-parent family, and number of persons in the household.

14. Duncan, Greg, and Saul Hoffman. 1990. "Teenage Welfare Receipt and Subsequent Dependence Among Black Adolescent Mothers." Family Planning Perspectives, 22(1): 16-20, 35.

A wide variety of family background variables, including family income, region, city size, and welfare recipiency, were measured when the girls were age **14**.

15. Hogan, Dennis, and Evelyn Kitagawa. 1985. "The Impact of Social Status, Family Structure, and Neighborhood on the Fertility of Black Adolescents." American Journal of Sociology, 90(4): 825-55.

Study uses data on about one thousand black girls in Chicago aged 13 to 19 in 1979. Many independent variables, including parental occupational status, parental labor force and employment status, family income, and housing characteristics, were measured as of the time of the survey, hence the girls were aged **13** to **19**.

16. McLanahan, Sara, and Larry Bumpass. 1988. "Intergenerational Consequences of Marital Disruption." American Journal of Sociology, 94: 130-152.

Uses the National Survey of Family Growth of 1982 with interviews of about eight thousand women aged 15 to 44. Family structure and region were measured when the girls were age **14**, but a number of other variables were based on recall of events and circumstances earlier in the girls' childhoods, allowing the authors to address timing issues to some extent.

17. Antel, John. 1988. "Mother's Welfare Dependency Effects on Daughters' Early Fertility and Fertility Out of Wedlock." Unpublished paper. University of Houston.

Sample included girls 18 years or younger in 1979, and the data from the NLSY were collected in 1978. A variety of family variables, including welfare receipt, state welfare benefit levels, geographic move in 1978, local unemployment rate, and urban-rural, were used in the analysis.

18. Lundberg, Shelley, and Robert Plotnick. 1990. "Effects of State Welfare, Abortion, and Family Planning Policies on Premarital Childbearing Among White Adolescents." Family Planning Perspectives, 22(6): 246-51, 275.

This study uses the NLSY, and while the main emphasis is on the effects of state policies, a variety of family background variables are used as controls, including living with a single parent, number of siblings, mother's work, region, and religiosity, and all of these are measured at age 14.

19. Ribar, David. 1991. "A Multinomial Logit Analysis of Teenage Fertility and High School Completion." Unpublished paper. Pennsylvania State University.

Used the NLSY to analyze this outcome, and hence measured family background and events variables at age 14. These included family structure, number of siblings, mother's working, father's working, region, urban-rural, religiosity, and magazines, newspapers, and library card.

20. Krein, Sheila. 1986. "Growing up in a Single Parent Family: The Effect on Education and Earnings of Young Men." Family Relations, 35: 161-168.

Uses the NLS, but makes efforts to record certain family structure events over the lifetime. However, family income is measured over the high school years and region is recorded at the date of the interview, hence at age 14 at the earliest.

21. Greenberg, D., and D. Wolf. 1982. "The Economic Consequences of Experiencing Parental Marital Disruption." Children and Youth Services Review, 4: 141-62.

Family structure and other background variables measured at ages 15 to 17.

22. Duncan, Greg, Martha Hill, and Saul Hoffman. 1988. "Welfare Dependence Within and Across Generations." Science, 239(January): 467-471.

Used data from the PSID on 1085 daughters aged 13 to 15, in which the economic status of the parents and their welfare recipiency were observed at the time the girls were 13 to 15.

APPENDIX B

Alternative Probit Estimates of the Effects of Family Background, Circumstances, and Events on High School Graduation and Teen Out-of-Wedlock Birth: Varying Periods of Circumstances/Events Observation

Variable	Coefficient	T-Statistic
HIGH SCHOOL GRADUATION MODELS		
Constant	-0.64E-1	-0.16
Non-Time-Varying Variables		
Nonwhite = 1	0.23	1.27
Female = 1	-0.32	-1.98
Nonwhite x Female	0.28	1.31
Catholic	0.85	3.51
Jewish	5.14	0.02
Protestant	0.57	2.75
Head Foreign Born	-0.34	-1.25
Father High School Graduate	0.33	2.17
Father Some College	0.35	2.26
Father College Graduate	0.62	2.81
Mother High School Graduate	0.29	2.00
Mother Some College	0.57	1.98
Mother College Graduate	1.18	2.35
One Parent in 1968	0.90E-1	0.51
No Parents in 1968	0.28	0.92
Number of Siblings	0.28	0.71
Preschool Child Care Time	0.16	0.14
Time-Varying Variables		
Years in Poverty, Ages 6-8	-0.75E-1	-1.10
Years in Poverty, Ages 9-11	-0.25E-1	-0.33
Years in Poverty, Ages 12-14	-0.45E-1	-0.66
Years in Poverty, Ages 6-14	[-0.52E-1]	[-2.12]
In Poverty at Age 14	[[-0.19]]	[[-1.33]]
Years in SMSA, Ages 6-8	-0.22	-2.21
Years in SMSA, Ages 9-11	0.26	1.71
Years in SMSA, Ages 12-14	-0.12	-0.98
Years in SMSA, Ages 6-14	[-0.23]	[-1.54]
In SMSA at Age 14	[[-0.21]]	[[-1.73]]

(appendix continues)

APPENDIX B (continued)

Variable	Coefficient	T-Statistic
Number of Parental Separations, Ages 6-8	-0.20	-1.14
Number of Parental Separations, Ages 9-11	0.76E-1	0.40
Number of Parental Separations, Ages 12-14	0.51	0.24
Number of Parental Separations, Ages 6-14	[-0.33E-1]	[-0.29]
Parents Separated at Age 14	[[0.19]]	[[0.58]]
Years Mother Worked, Ages 6-8	-0.23E-3	-0.00
Years Mother Worked, Ages 9-11	0.15E-1	0.23
Years Mother Worked, Ages 12-14	0.73E-1	1.31
Years Mother Worked, Ages 6-14	[0.31]	[1.74]
Mother Worked at Age 14	[[0.28]]	[[2.42]]
Number of Location Moves, Ages 6-8	-0.17	-2.35
Number of Location Moves, Ages 9-11	-0.34E-1	-0.44
Number of Location Moves, Ages 12-14	-0.12	-1.53
Number of Location Moves, Ages 6-14	[-0.11]	[-3.41]
Moved Location at Age 14	[[-0.41]]	[[-2.87]]
Log-Likelihood = -382.27		
N = 825		
TEEN OUT-OF-WEDLOCK BIRTH MODELS		
Constant	-0.83	-1.10
<u>Non-Time-Varying Variables</u>		
Nonwhite = 1	0.29	1.20
Any Religion = 1	-0.24	-0.64
Number of Siblings	-0.19	-0.29
Mother's Age at First Birth	-0.37E-2	-0.17
Mother a High School Graduate	-0.71	-3.13
Mother Had Out-of-Wedlock Birth	0.35	1.28
Bad Neighborhood in 1976	-0.49	-0.71

(appendix continues)

APPENDIX B (continued)

Variable	Coefficient	T-Statistic
Time-Varying Variables		
Average Income-to-Needs Ratio, Ages 6-8	-0.20	-0.90
Average Income-to-Needs Ratio, Ages 9-11	0.15	0.59
Average Income-to-Needs Ratio, Ages 12-14	-0.333	-1.56
Average Income-to-Needs Ratio, Ages 6-14	[-0.33]	[-2.26]
Income-to-Needs Ratio at Age 14	[[-0.33]]	[[-2.94]]
Number of Location Moves, Ages 6-8	0.28	2.21
Number of Location Moves, Ages 9-11	0.13	0.93
Number of Location Moves, Ages 12-14	-0.24	-1.51
Number of Location Moves, Ages 6-14	[0.62E-1]	[1.14]
Moved Location at Age 14	[[0.26]]	[[1.07]]
Number of Parental Separations, Ages 6-8	0.43	1.43
Number of Parental Separations, Ages 9-11	0.22	0.68
Number of Parental Separations, Ages 12-14	0.74	2.05
Number of Parental Separations, Ages 6-14	[0.50]	[3.01]
Parents Separated at Age 14	[[-0.36]]	[[-.01]]
Years Family Received Welfare, Ages 6-8	0.37	2.62
Years Family Received Welfare, Ages 9-11	-0.24	-1.45
Years Family Received Welfare, Ages 12-14	-0.73	-0.05
Years Family Received Welfare, Ages 6-14	[0.13E-1]	[0.34]
Family Received Welfare at Age 14	[[-0.16]]	[[-0.69]]
Years in SMSA, Ages 6-8	0.31	1.71
Years in SMSA, Ages 9-11	-0.37	-1.23
Years in SMSA, Ages 12-14	0.23	0.89
Years in SMSA, Ages 6-14	[0.41]	[1.55]
In SMSA at Age 14	[[0.34]]	[[1.60]]
Years Lived with One Parent, Ages 6-8	-0.28	-1.88
Years Lived with One Parent, Ages 9-11	0.21	1.23
Years Lived with One Parent, Ages 12-14	0.72E-1	0.52
Years Lived with One Parent, Ages 6-14	[0.34]	[0.99]
Lived with One Parent at Age 14	[[0.59]]	[[2.55]]
Log-Likelihood = -122.93		
N = 431		

Source: Authors' computations based on Michigan Panel Study of Income Dynamics.

Note: The coefficients and t-statistics shown for the non-time-varying variables are from the probit with time-varying variables measured in period-specific form.

[ ] = Coefficients and t-statistics from probit with time-varying variables measured over years 6-14, in place of period-specific variables.

[[ ]] = Coefficients and t-statistics from probit with variables measured at age 14, in place of period-specific variables.

APPENDIX C

Means and Standard Deviations of Variables Used in Probit Equations: Models of High School Graduation and Teen Out-of-Wedlock Birth

Variable	High School Graduation Model		Teen Out-of-Wedlock Birth Model	
	Mean	Standard Deviation	Mean	Standard Deviation
NON-TIME-VARYING VARIABLES				
Nonwhite = 1	0.50	0.50	0.30	1.20
Female = 1	0.52	0.50		
Nonwhite x Female	0.28	0.45		
Any religion = 1			-0.24	-0.64
Catholic	0.15E-1	0.12		
Jewish	0.72	0.45		
Protestant	0.21E-1	0.14		
Mother's Age at First Birth			-0.37	-0.17
Head Foreign Born	0.21	0.41		
Father High School Graduate	0.22	0.41		
Father Some College	0.93E-1	0.29		
Father College Graduate	0.93E-1	0.29		
Mother High School Graduate = 1			-0.71	-3.13
Mother High School Graduate	0.38	0.48		
Mother Some College	0.70E-1	0.26		
Mother College Graduate	0.02	0.23		
One Parent in 1968[a]	0.19	0.40		
No Parents in 1968[b]	0.37E-1	0.19		
Mother Out-of-Wedlock Birth = 1			0.35	1.28
Number of Siblings	2.66	1.63		
Preschool Child Care Time[c]	2146.90	676.98		
Bad Neighborhood in 1976[d]			-0.49E-1	-0.71

(appendix continues)

APPENDIX C (continued)

Variable	High School Graduation Model Mean	High School Graduation Model Standard Deviation	Teen Out-of-Wedlock Birth Model Mean	Teen Out-of-Wedlock Birth Model Standard Deviation
TIME-VARYING VARIABLES				
Years in Poverty (Child Lived in Family Whose Income Was below the Matched Poverty Line in Year of Observation = 1)				
Years in Poverty, Ages 6-8	0.78	1.13		
Years in Poverty, Ages 9-11	0.64	1.06		
Years in Poverty, Ages 12-14	0.64	1.05		
Years in Poverty, Ages 6-14	2.06	2.86		
In Poverty at Age 14	0.20	0.40		
Average Income-to-Needs Ratio (Average over Specified Ages of the Ratio of Family Income to the Matched Poverty Line)				
Average Income-to-Needs Ratio, Ages 6-8			2.06	1.56
Average Income-to-Needs Ratio, Ages 9-11			2.31	1.73
Average Income-to-Needs Ratio, Ages 12-14			2.56	1.99
Average Income-to-Needs Ratio, Ages 6-14			2.31	1.69
Income-to-Needs Ratio at Age 14			2.72	2.34
Years in SMSA (Child Lived in SMSA in Year of Observation = 1)				
Years in SMSA, Ages 6-8	2.18	1.28	2.21	1.27
Years in SMSA, Ages 9-11	2.15	1.32	2.19	1.30
Years in SMSA, Ages 12-14	2.15	1.34	2.18	1.32
Years in SMSA, Ages 6-14	6.47	3.82	6.58	3.77
In SMSA at Age 14	0.71	0.45	0.72	0.45
Number of Parental Separations (Parents of Child Separated or Divorced in Year of Observation = 1)				
Number of Parental Separations, Ages 6-8	0.08	0.28	0.09	0.30
Number of Parental Separations, Ages 9-11	0.09	0.29	0.09	0.30
Number of Parental Separations, Ages 12-14	0.06	0.25	0.06	0.25
Number of Parental Separations, Ages 6-14	0.23	0.47	0.25	0.49
Parents Separated at Age 14	0.03	0.16	0.03	0.16

(appendix continues)

APPENDIX C (continued)

Variable	High School Graduation Model Mean	High School Graduation Model Standard Deviation	Teen Out-of-Wedlock Birth Model Mean	Teen Out-of-Wedlock Birth Model Standard Deviation
Years Mother Worked (Mother Worked Outside the Home in Year of Observation =1)				
Years Mother Worked, Ages 6-8	1.58	1.29		
Years Mother Worked, Ages 9-11	1.63	1.33		
Years Mother Worked, Ages 12-14	1.77	1.30		
Years Mother Worked, Ages 6-14	4.98	3.37		
Mother Worked at Age 14	0.59	0.49		
Number of Location Moves (Change in Household Location of the Family in Year of Observation=1)				
Number of Location Moves, Ages 6-8	0.56	0.77	0.50	0.75
Number of Location Moves, Ages 9-11	0.52	0.76	0.51	0.74
Number of Location Moves, Ages 12-14	0.38	0.69	0.37	0.66
Number of Location Moves, Ages 6-14	1.46	1.64	1.38	1.57
Moved Location at Age 14	0.13	0.34	0.13	0.34
Number of Years Lived with One Parent (Lived with One Parent in Year of Observation = 1)				
Years Lived with One Parent, Ages 6-8			0.68	1.20
Years Lived with One Parent, Ages 9-11			0.80	1.27
Years Lived with One Parent, Ages 12-14			0.89	1.32
Years Lived with One Parent, Ages 6-14			2.35	3.51
Lived with One Parent at Age 14			0.31	0.46
Number of Years Family Received Welfare (Family Received Welfare in Year of Observation = 1)				
Years Family Received Welfare, Ages 6-8			0.35	0.86
Years Family Received Welfare, Ages 9-11			0.41	0.96
Years Family Received Welfare, Ages 12-14			0.46	0.96
Years Family Received Welfare, Ages 6-14			1.22	2.50
Family Received Welfare at Age 14			0.15	0.36

Source: Authors' computations based on Michigan Panel Study of Income Dynamics.
[a]Hence education variable is available only for that parent.
[b]Hence no education variable is available for either parent.
[c]Total number of hours allocated to child care in preschool years, ages 4 and 5.
[d]The sum of positive responses to (1) burglaries and robberies, (2) muggings, rapes, pushers, junkies, or too few police, (3) crowded area with too many people, too much noise, and bad traffic, (4) a poor neighborhood for kids, or (5) unkept yards, grounds, houses poorly kept up, or infrequent or sloppy garbage pickups being a problem in the neighborhood.

Endnotes

[1]A survey of such research studies since 1980 (Appendix A) reveals that most employ observations on the individuals studied at age 14, though in some cases even later.

[2]The importance of these "timing" effects has been emphasized by both developmental psychologists and sociologists. See, for example, Alwin and Thornton (1984), Krein (1986), and Wallerstein and Kelly (1986). Alwin and Thornton find it difficult to distinguish differential effects of early childhood and later childhood family influences on schooling experiences. The high degree of intertemporal correlation among many of their explanatory variables (e.g., parental education and occupation, family assets, and income) constrains their efforts. For variables with less intertemporal stability, they find greater differences between early and later family influences.

[3]This issue is also addressed by Cherlin and Horiuchi (1980) and more recently by Wu and Martinson (1990). Wu and Martinson document the diversity in family situations among children from parental histories of respondents, concluding that "snapshot measures understate greatly the complexity of parental situation."

[4]The sample used for the estimates includes the children who were aged 14 to 16 in 1979 (aged 3 to 5 in 1968). There are 825 children in the sample used for the high school graduation estimate, of whom 635 graduated. The sample used for the out-of-wedlock birth estimate includes 431 females, of whom 59 gave birth as a teen. The specification of the models follows that of Haveman, Wolfe, and Spaulding (1991) and An, Haveman, and Wolfe (1992). The estimated models are shown in Appendix B. Definitions of the variables and their means and standard deviations are shown in Appendix C. The family background variables are invariant across the education and out-of-wedlock birth models, and include race, gender, mother's age at first birth, whether or not the mother had an out-of-wedlock birth, neighborhood quality, religion, father foreign born, father and mother education, number of siblings, and child care time received over childhood years.

[5]The circumstance/event variables are:

High School Graduation	Teen Out-of-Wedlock Birth
-Years in Poverty	-Average Ratio of Income to Needs
-Years Living in SMSA	-Years Living in SMSA
-Number of Parental Separations	-Number of Parental Separations
-Years Mother Worked	-Years Receiving Welfare
-Number of Geographic Moves	-Number of Geographic Moves
	-Years with One Parent

[6]The three observation periods are (1) the early childhood period, ages 6 to 8; (2) the middle childhood period, ages 9 to 11; and (3) the adolescent period, ages 12 to 14.

[7]The correlation between an income variable measured at a single point in time with its average over nine years can be viewed as a measure of accuracy of data reported at a point in time. The .88 correlation for the income-to-needs ratio fits within the range observed in other studies. Other researchers have found correlations of from .61 to .94 comparing responses to a single question on circumstance variables asked at widely separate occasions. See Bielby, Hauser, and Featherman (1977), who used 1973 OCG data, and Hauser and Sewell (1986), who used Wisconsin and Kalamazoo data. In the 1977 study, correlations of .87 and .94 were reported for repeated questions of parental schooling and occupational prestige among white men; among black men, .64 and .92 were reported for those same questions. In the 1986 study, correlations of .73 to .78 for replies to questions on father's schooling, and .61 to .75 to questions regarding father's occupational status, were reported.

[8]The former is a somewhat constrained version of the latter, in which the effects of duration are equated over the three age periods.

[9]The estimates (and subsequent simulations) of the effects of the early childhood and adolescent variables are from probit equations including observations for all three of the childhood time periods. The estimates of the effects of the non-time-varying variables in Appendix B are from the specification including the three period-specific, time-varying variables. In virtually no case did the

significant (at the .05 level) coefficients on the variables shown become insignificant in the alternative specifications; similarly, none of the insignificant coefficients shown became significant.

[10]The .2 level of significance is an arbitrarily chosen level; however, use of an alternative level such as .25 does not substantially change the results.

[11]For example, as mentioned in the discussion of Table 3, in only eighteen of the thirty-three cases did the pairwise comparisons pass our sign/significance test. In only four of the eleven cases in which the age-14 variable substitutes for the multiyear variable during the early childhood (ages 6 to 8) period, and in but five of the eleven cases in which the age-14 variable proxies for information during the adolescent (ages 12 to 14) period, is this test passed. However, when the window variable serves as a proxy for full information over the entire age 6-14 childhood experience, the test is passed in nine of eleven cases.

References

Alwin, Duane F., and Arland Thornton. 1984. "Family Origins and the Schooling Process: Early Versus Late Influence of Parental Characteristics." American Sociological Review, 49: 784-802.

An, Chong Bum, Robert Haveman, and Barbara Wolfe. 1992. "Teen Out-of-Wedlock Births and Welfare Receipt: The Role of Childhood Events and Economic Circumstances." Review of Economics and Statistics, forthcoming.

Bielby, William T., Robert Hauser, and David Featherman. 1977. "Response Errors of Black and Nonblack Males in Models of Intergenerational Transmission of Socioeconomic Status." American Journal of Sociology, 82(May): 1242-1288.

Cherlin, Andrew, and S. Horiuchi. 1980. "Retrospective Reports of Family Structure: A Methodological Assessment." Sociological Methods and Research, 8(4): 454-469.

Hauser, Robert and William Sewell. 1986. "Family Effects in simple Models of Education, Occupational Status, and Earnings: Findings from the Wisconsin and Kalamazoo Studies." Journal of Labor Economics, 4(July, Part 2): S83-S115.

Haveman, Robert, Barbara Wolfe, and James Spaulding. 1991. "Childhood Events and Circumstances Influencing High School Completion." Demography, 28(1): 133-157.

Krein, Sheila Fitzgerald. 1986. "Growing Up in a Single Parent Family: The Effect on Education and Earnings of Young Men." Family Relations, 35: 161-168.

Wallerstein, J. S., and J. B. Kelly. 1986. Surviving the Breakup: How Parents and Children Cope with Divorce. New York: Basic Books.

Wu, Lawrence L., and Brian C. Martinson. 1990. "Parent Histories: Patterns of Change in Early Life." NSFH Working Paper No. 37, Center for Demography and Ecology, University of Wisconsin-Madison.