Reliability of Student and Parent Reports

of Socioeconomic Status in NELS-88

Robert M. Hauser

Megan Andrew

University of Wisconsin-Madison

July 6, 2007

Executive Summary

This report outlines an analysis assessing the relative reliability of student and parent reports of key socioeconomic characteristics of the family for three race-ethnic groups: non-Hispanic Whites, Blacks, and Latinos. The analysis uses a nationally representative sample of 8^{th} grade students in 1988. Key findings include:

1. Neither students nor their parents are very reliable or precise in reporting key socioeconomic characteristics of the parents.
2. The quality of such reports by students (and by parents) is lower among Hispanics than Blacks and lower among Blacks than among Whites. ***These differentials in the quality of socioeconomic reports have important implications for inferences about race-ethnic differences in the academic correlates of socioeconomic status.***
3. Eighth-grade students are not notably worse than their parents in reporting the parents' occupations in NELS-88, nor are Black and White students substantially worse at reporting parents' occupations than their educational attainments.
4. While parents are more reliable in reporting status characteristics than their 8^{th} grade students, their reports are also imprecise. ***Thus, we should not privilege reports from any one source as a "gold standard."***

The implications of these findings are two-fold:

1. We should not set too high a standard for the validity of Census-linked data that may be used to measure the socioeconomic status of students. A reasonable standard might be that such indicators be no worse than those obtained from direct reports by students or their parents.
2. Even if Census-linked data fail to meet that standard, it might be better to combine them with direct student reports of parental status characteristics than to choose only one such indicator.

In conclusion, our analysis suggests that students can provide reports of socioeconomic background characteristics, though their reports are often less reliable than that of their parent(s). However, neither student nor parent reports are infallible, so measurement error is a key concern. By linking with Census data, the problem of error could be easily mitigated, enhancing the utility and accuracy of the NAEP data and our understanding some of our nation's most serious educational inequalities.

Introduction

This brief report addresses the *reliability* and *precision* of student and parent reports of socioeconomic variables in the first round of the National Educational Longitudinal Study of 1988 (NELS-88), that is, when the students were in the 8th grade.[1] *Reliability* refers to the proportion of total variance in a measured variable that is attributable to true variance in the construct that it represents. *Precision* is inverse to the (absolute) size of the error variance in a measured variable. That is, the larger the error variance, the less precise is the measure. It is necessary to consider *both* reliability and precision. This may occur because the size of true variances may vary across populations; for example, a variable could be less reliable in one population than another even if it is more precise. This report does not address other aspects of the quality of student and parent reports, e.g., whether the distributions of student and parent reports have similar shape or location. We generally find that no one report of a socioeconomic measure is a panacea for the daunting task of accurately describing the relationship between an individual's background and his or her achievement. Therefore, the question becomes more a matter of the extent to which analysts and others are willing to sacrifice accurate measurement. We begin with a brief discussion of reliability and precision in the context of latent factor models in order to orient the reader. We then describe the data used and our main findings in this report. Readers who are familiar with statistical reliability and precision and latent factor models can skip the background section of this report.

[1] McLaughlin and Cohen (1997) report parent-student correlations of maternal and paternal educational attainment, but they do not estimate structural models nor investigate variations in the quality of reports by race-ethnicity.

Background

In this report, we utilize latent factor measurement models to assess the relative reliability and precision of student and parent reports of key socioeconomic background measures. In the title of this statistical model, “latent factor” refers to some underlying “true” construct that generates what we observe and “measurement” means simply that we use this general type of statistical model to look at how reliably and precisely a given quantity is measured. By quantifying reliability and precision in a statistical model, we can begin to speak to issues of accuracy and whether we are measuring what we think we are measuring. In the case of socioeconomic background measures more specifically, we can begin to speak to issues of whether students and their parents provide reliable information about their socioeconomic status, if they do so to an equivalent degree, and the precision in those reports.

To clarify, we use a simple example of measuring height for a group of students. Suppose two individuals separately measure each student with a yard stick and record that measurement. It is likely that each individual would not provide the same exact measurement of each student’s height. Therefore, each of the two individuals measure student height imperfectly and any calculations we make with these data on student height would include this error. But how much error is included in a given measure, whether it is height or something else? Is one individual more reliable in reporting the true student height?

In this scenario, we would be able to use a statistical model, a latent factor model, to quantify student height into a two parts: “true” student height and error.[2] The error represents the difference from a given measurement of a given student’s height and the student’s “true” height and can be thought of as the precision of a given measurement. We can define precision

[2] We denote “true” with quotation marks to underscore that the true or real quantity of interest is not directly observed. It is derived from the data an analyst uses and, therefore, it and the reliability of a given quantity of interest is not fixed and may vary from data set to data set.

and reliability in terms of this model. The less absolute error in our measure of height relative to the total variability in the data, the greater is the precision of the height measure. The more absolute error in the measure of height relative to the total variability in the data, the lower is the precision of the measure.

In statistical terms, reliability is defined as the proportion of the variance of the "true" height to the overall variance of height in the data. Again, in order to quantify the statistical reliability of a given measure, one would need multiple measures of the same quantity of interest. Therefore, our latent factor model can also provide information about which individual in our height example provides more accurate information on average when the two are compared. A latent factor measurement model can be summarized using the simple path diagram in Figure 1:

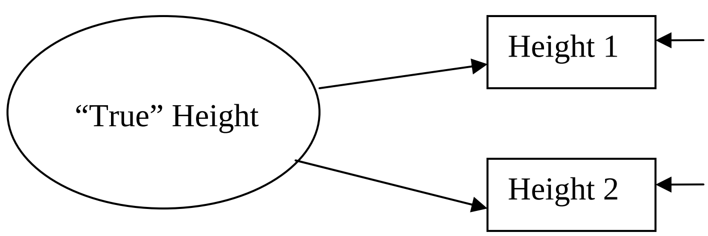

Figure 1. Schematic Latent Factor Measurement Model

In this path diagram, a "true" underlying quantity, height, generates the two resulting measure of height. These two measures of height are not measured perfectly, a characteristic denoted in short-hand by the two shorter arrows feeding into each. In order to estimate a latent factor measurement model, a number of standard statistical assumptions must be made. Noted above, the underlying factor described as "true" height is not observed. Therefore, in order to estimate the paths from "true" height to height 1 and height 2, one must assume one of these paths is equal to 1 and estimate the other path relative to the path set to 1. For example, if we set the path from "true" height to height 1 to 1, we can estimate a number quantifying the statistical relationship between "true" height and height 2 that will denote the relationship between the metrics of the two sets of reports. If the path from "true" height and height 2 is greater than 1, then the second person is reporting fewer inches of height relative to those reported by the first person. Conversely, if that path is smaller than 1 then the relationship between the metrics of the first and second reports is reversed. One can easily extrapolate from this example to the task at hand: determining the reliability of student and parent reports of socioeconomic background characteristics. Substituting "student report" for height 1 and "parent report" for height 2, one can easily see that the relative paths provide similar information about the relative metrics of each report in a latent factor measurement model.[3] There is technically no limit to the number of factors that may be included simultaneously in a model so long as there are multiple measures of each underlying construct of interest.

Data

NELS-88 is a national probability survey administered by the National Center for Educational Statistics. Students were initially contacted in the 8th grade in 1988 and regularly interviewed until 2000 when the study concluded. Additional data were collected from principals, teachers, and parents at certain waves of the study. In particular, a parent of the student was interviewed in 1988 and 1992, the first and third waves of the study. Post-secondary

[3] In fact, with only two indicators of a single construct, the error variances in the model of Figure 1 cannot be identified unless, as is sometimes done for convenience, they are assumed to be equal. However, as soon as there is at least one more variable in the model that is correlated with the two indicator variables, either a measure of the original construct or some other construct, all of the parameters of the model of Figure 1 are identified.

transcript data were included in the study in 2000. A total of 12,144 students completed all waves of the study, and we limit the base sample to these students. We delete Asian and Native American respondents in our analytic sample, leaving 9968 observations. We choose to focus on White, Black, and Latino respondents in order to draw comparisons between white youth and traditionally more disadvantaged race-ethnic groups. While Native American groups often face severe socioeconomic disadvantages, we do not consider this group due to insufficient numbers. Following these restrictions we are left with a total sample of 9968 students. For this remaining sample, we impute using a Markov chain Monte Carlo (MCMC) method implemented in Stata. The MCMC method is based on a small number of draws from existing data from a predictive distribution which are then used to impute the data. The imputations were based on a large number of variables including additional socioeconomic characteristics, family variables, academic achievements, educational expectations, and educational attainment measures rather than merely variables used in the present analysis. In this paper, we simply use one imputation of the data, but supplementary analyses suggest that our findings hold once we incorporate all imputations of the data and adjust standard errors for the effects of multiple imputation.

In addition to the race-ethnic groups described above, we consider a number of social background variables. For a number of these variables, we have a student and parent report in 1988 when students were in the 8^{th} grade. For parent education, occupation, and labor force status measures, we use both parent and student reports based on the 1988 wave of the study. Years of education measures for each parent are derived from a categorical variable and equivalent years were assigned for each category. Occupation measures for each parent are similarly based on a categorical variable. Rather than use the original occupational coding, we assign the appropriate occupational educational logit value, based on Hauser and Warren (1997), to each of the possible 15 categories. If either parent was not known to be in the regular, civilian labor force, we assigned the mean rather than imputing a value.

We were concerned that the extensive use of imputation might distort our findings, and for that reason we repeated all of the analyses reported below with two other treatments of the data. First, we deleted observations where the student reported that their parent was not in the regular labor force. Second, we also deleted observations where occupational data were missing. The analyses of these two substantially reduced samples yielded findings that are essentially the same as those reported below. The only notable change is that, in the two reduced minority samples, the students' reports are *more* reliable and *more* similar in reliability to parents' reports than in the full samples.

Table 1 presents data on the percentage of missing data for each analytic variable (described in detail below) by race-ethnic group. Most variables have moderate amounts of missing data (around 10 percent), but there are notable exceptions. For example, about a quarter of data for father's education and occupation is missing for black students. About 50 percent of data is missing for parent reports of father's occupation and education. The large discrepancy in missing data between student and parent reports largely is a function of instructions to the surveyed parent to disregard questions about the student's other parent if they were a single parent. The high rate of missing reports mirrors the higher portion of African American families in the data headed by a single mother.

Table 2 presents means and standard deviations of analytic variables by race-ethnicity. We note that, as in other comparisons of student and parent reports, students report higher average levels of parents' schooling than parents. The disparity in mean levels of schooling is especially large among Latinos. However, with a single exception, parents report higher levels of occupational status than do students. Thus, there is no *global* tendency for students to report higher parental status than their parents.

Model

The analyses of reliability reported herein are based on a simple structural equation model of the five pairs of reports. We consider five underlying, latent constructs corresponding to key social background measures: father's education, mother's education, father's occupation, mother's occupation, and family income. These constructs can be understood to generate the observed student and parent reports for each construct. The idea not only is that each indicator reflects the construct it was intended to represent but that it is also affected by a random error term. There are no correlations between error terms *within* the five latent constructs. For example, the errors of parent's and student's reports of educational attainment and occupational status are never correlated. However, there may be correlations between error terms of different variables reported by the same person on the same occasion. For example, there may be correlations between errors in a student's reports of her mother's and her father's educational attainments.

Within each race-ethnic group we estimate a latent factor measurement model with five latent factors: father's years of education, mother's years of education, father's occupation, mother's occupation, and family income. These factors are included in the same model and are correlated with one another. Each factor generates two observed reports. For father's education, mother's education, father's occupation, and mother's occupation, we observe a parent and a student report. For family income, we observe a parent report in 1988 and a parent report in 1992. A portion of this model is replicated schematically in Figure 2 using father's education as an example. The remaining four factors and the associated reports could be represented similarly, along with the correlations between factors, and all parameters are estimated simultaneously.

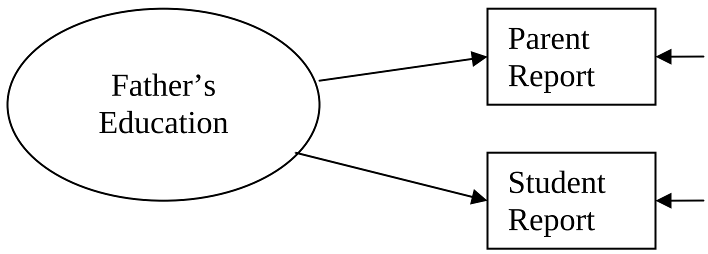

Figure 2. Partial Schematic Latent Factor Measurement Model, Father's Years of Education

Similar to the model shown schematically in Figure 1, the five underlying factors in the model partially represented in Figure 2 are not directly observed but their properties are estimated from the data. Therefore, in order to estimate the paths from each factor to the two reports, one must assume one of these paths is equal to 1 and estimate the other path relative to the path set to 1. We choose to fix the loadings of parent reports and the second income report. In statistical terms, the metric of the unobservable true values of the socioeconomic variables is the same as the metric of the parents' reports of those variables.

Findings

Table 3 shows the fit of several models, estimated separately for the three large race-ethnic groups. Models were estimated in LISREL 8.80 by maximum likelihood (Jöreskog and Sörbom 1996). In Model 1 (shown schematically in Figure 3), there are no correlations between errors in survey reports. Model 2 (shown in Figure 4) adds a correlation between errors in the student's reports of their parents' educational attainments. Model 3 (shown in Figure 5) adds a correlation between errors in the parent's report of parents' educational attainments. In all three groups, a fit diagnostic, the modification index, suggests adding free parameters in this order. In general, Model 3 fits better than Model 2, which fits better than Model 1. The key fit statistic in Table 3 is BIC (the Bayesian Information Criterion), which is useful in assessing fit when large sample sizes yield so much statistical power that no parsimonious model fits by conventional criteria (Raftery 1993; 1995). A negative value of BIC indicates satisfactory model fit. By this criterion, all three models fit satisfactorily in the Black sample, but fit still improves in Models 2 and 3. In the White sample, fit is essentially the same in Models 2 and 3.

Table 4 shows the slopes in each population group, as estimated in Model 3. In the Black sample, except in the case of mother's education, the free slopes are very close to 1.0, implying that the metric of students' reports is essentially the same as that of parents' reports. In the White sample, the metrics of students' reports of mother's educational attainment and father's occupational status are very close to those of the parent reports, but the slopes of student reports of father's educational attainment and of mother's occupational status are significantly lower than those of the parents' reports. That is, the metric of students' reports is compressed, relative to that of parents' reports. The pattern holds for all of the Latino students' reports of parents' statuses. As a check on these findings, Model 4 (in Table 3) constrains all of the slopes of observables on true scores to equal 1.0 within each group. Among Latinos and Whites – but not among Blacks – these constraints lead to a significant deterioration of fit.

The final two lines of Table 1 (labeled "combined") report a test of differences in the freely estimated slopes among Blacks, Latinos, and Whites. The line for Model 3 simply estimates Model 3 for all three samples combined, but with no cross-group constraints on parameters. Note that the chi-square test statistic and degrees of freedom for this model are just the sum of the test statistics and degrees of freedom in the runs of Model 3 in the three samples.

The line for Model 4 here pertains to a model in which the freely estimated slopes are specified to be equal in the three samples. Since there were originally five free slopes in each sample, these constraints add 10 degrees of freedom to the model. The chi-square test statistic increases substantially relative to its degrees of freedom, by 125.9, so we find that there are significant differences in the free slopes among the three populations.

Regardless of the slight differences in the best fitting models among Blacks, Latinos, and Whites, we have chosen to focus on a common specification, Model 3, to report standardized loadings and error correlations. Standardized loadings and error correlations allow us to make direct comparisons across the three samples of students, each with different variances. The standardized loadings and error correlations are shown in Figures 6, 7, and 8 for Blacks, Latinos, and Whites, respectively. Among Blacks, there are few differences in the loadings between students' and parents' reports and the standardized loadings center on 0.80, meaning that about 64 percent of the variance in students' and parents' reports is attributable to the true values of the socioeconomic variables.[4] There are moderate correlations—both for students and parents—between the reports of father's and of mother's education. In the Latino sample, by contrast, parent's reports are of similar or higher reliability than those of black parents but the reliability of student reports is substantially lower, except in the case of father's educational attainment. Again, there are moderate correlations between the students' and between the parents' reports of paternal and maternal educational attainment. Among Whites, the reliability of parents' reports is the same or higher than that of students' reports, but both are at least as high as in the Black or Latino samples.

[4] We make this calculation by squaring the coefficient. That is, 0.8 to the second power is 0.64, which is 64 percent of the variance.

Reliabilities may be misleading with regard to the *precision* of the reports in the three populations because they depend on the size of the true variances as well as the size of the error variance. Thus, Table 5 displays the estimated error variances under Model 3 in each sample. These are in the metrics of the original reports by students and by parents. Comparisons among the error variances pertain directly to the precision of reports relative to their average values (but not to upward or downward biases in the reports).

Variances are always larger among Blacks than among Whites, regardless of the variable and the person reporting it. The same holds among Latinos relative to Whites, with two exceptions, the parental reports of occupational statuses of fathers and of mothers. That is, with these two exceptions, minority reports of status variables are less precise than those of Whites.

Comparisons of the precision of reporting the same variable by students and parents are less clearly patterned. In the case of father's educational attainment, reports by parents and students are similar in precision among Blacks and among Latinos, but the parents' reports are more precise among Whites, that is, have lower variance. In the case of mother's educational attainment, parents' reports are more precise in all three population groups. In the case of father's occupational status, student reports are more precise among Blacks, parent reports are more precise among Latinos, and precision is essentially the same among White parents and students. Finally, in the case of mother's occupational status, parents are more precise among Latinos and Whites, but there is no difference in precision among Blacks.

In all three populations, students are more precise in reporting father's education than mother's education while parents are more precise in reporting mother's education than father's education. One might want to explain the latter finding by the greater likelihood that the reporting parent is a mother than a father. However, parents in all three groups are more precise in reporting father's occupational status than mother's occupational status. Among Blacks and Whites, but not among Latinos, students are more precise in reporting father's occupational status than mother's occupational status.

Conclusion

What are the lessons from these findings? First, eighth-grade students are not very reliable or precise in reporting key socioeconomic characteristics of their parents. Second, the quality of such reports by students (and by parents) is lower among Hispanics than Blacks and lower among Blacks than among Whites. These differentials in the quality of socioeconomic reports have important implications for inferences about race-ethnic differences in the academic correlates of socioeconomic status. Third, students are not notably worse than parents in reporting their parents' occupations, using a scheme like that in NELS-88, nor are Black and White students substantially worse at reporting parents' occupations than their educational attainments. For students as young as eighth-graders, the major problems in reporting occupation are not quality-related, but pertain to respondent burden and the cost of coding detailed occupational protocols.

Last, while parents are more accurate in reporting status characteristics than their children in the eighth grade, their reports are also substantially unreliable and imprecise. Thus, we should not privilege reports from any one source as a "gold standard." Rather, it would be better to obtain such reports from more than one source and either to construct measures based on all such reports or – where possible – to model response variability when analyzing the correlates of socioeconomic variables.

These findings appear to have two operational implications in the context of efforts to improve the measurement of socioeconomic status in NAEP. The first is that we should not set too high a standard for the validity of Census-linked data that may be used to measure the socioeconomic status of students. A reasonable standard might be that such indicators be no worse than those obtained from direct reports by students or their parents. The second implication is that, even if Census-linked data fail to meet that standard, it might be better to combine them with direct student reports of parental status characteristics than to choose only one such indicator.

References

Hauser, Robert M. and John Robert Warren. 1997. "Socioeconomic Indexes for Occupations: A Review, Update, and Critique." Pp. 177-298 in *Sociological Methodology 1997*, edited by A. E. Raftery. Cambridge: Basil Blackwell.

Jöreskog, K. G. and Dag Sörbom. 1996. *Lisrel 8 User's Reference Guide*. Chicago, IL :: Scientific Software International.

McLaughlin, Donald H. and Jon Cohen. 1997. "Nels:88 Survey Item Evaluation Report." National Center for Education Statistics, U.S. Department of Education, Washington, DC.

Raftery, Adrian E. 1993. "Bayesian Model Selection in Structural Equation Models." Pp. 163-80, *Testing Structural Equation Models*, edited by K. A. Bollen and J. S. Long. Newbury Park: SAGE Publications.

—. 1995. "Bayesian Model Selection in Social Research." Pp. 111-63 in *Sociological Methodology 1995*, edited by P. V. Marsden. Cambridge: Basil Blackwell.

Table 1. Missing Data by Variable by Race-Ethnicity

	White	Black	Latino
Father's Years Education - Student	0.12	0.26	0.20
Mother's Years Education - Student	0.09	0.13	0.16
Father's Occupational Status - Student	0.08	0.25	0.16
Mother's Occupational Status - Student	0.04	0.07	0.08
Father's Education - Parent	0.18	0.49	0.28
Mother's Education - Parent	0.08	0.14	0.14
Father's Occupation - Parent	0.19	0.53	0.32
Mother's Occupation - Parent	0.14	0.23	0.29
1988 Family Income - Parent	0.05	0.07	0.10
1992 Family Income - Parent	0.13	0.12	0.24

Table 2. Weighted Means and Standard Deviations for Model Variables, Imputed Sample

Variable	White	Black	Latino
Father's Years Education - Student	14.02	13.08	12.72
	(2.64)	(2.3)	(2.34)
Mother's Years Education - Student	13.69	13.23	12.44
	(2.3)	(2.28)	(2.09)
Father's Occupational Status - Student	0.03	-0.53	-0.49
	(1.19)	(1.13)	(1.18)
Mother's Occupational Status - Student	0.14	-0.23	-0.14
	(.14)	(1.11)	(.89)
Father's Education - Parent	13.62	12.81	11.61
	(2.74)	(2.25)	(2.8)
Mother's Education - Parent	13.07	12.63	11.11
	(2.1)	(1.96)	(2.5)
Father's Occupational Status - Parent	0.27	-0.30	-0.27
	(1.2)	(1.16)	(1.03)
Mother's Occupational Status - Parent	0.39	0.09	-0.26
	(1.07)	(1.17)	(1.06)
Logged 1988 Family Income - Parent	4.04	3.37	3.37
	(.85)	(1.11)	(1.05)
Logged 1992 Family Income - Parent	3.99	3.34	3.27
	(.91)	(1.06)	(1.09)
	N=7632	N=974	N=1362

Table 3. Fit of Selected Models of Socioeconomic Measurement

	Model	Chi-Sq	df	BIC
Black	1	128.4	25	-43.6
	2	56.0	24	-109.2
	3	43.0	23	-115.3
	4	47.2	28	-145.5
Latino	1	300.5	25	120.1
	2	158.7	24	-14.5
	3	100.6	23	-65.4
	4	297.8	28	95.7
White	1	569.7	25	346.2
	2	151.5	24	-63.1
	3	151.0	23	-54.6
	4	412.0	28	161.7
Combined	3	294.6	69	-340.7
	5	420.5	79	-306.9

Table 4. Estimated Slopes of Student's Reports of Socioeconomic Variables

	Black	Latino	White
Father's education	1.030	0.781	0.918
	(0.046)	(0.021)	(0.008)
Mother's education	1.016	0.711	0.999
	(0.043)	(0.025)	(0.011)
Father's occupational status	1.001	0.825	0.992
	(0.042)	(0.037)	(0.011)
Mother's occupational status	0.907	0.624	0.819
	(0.048)	(0.029)	(0.014)
Family income	1.030	0.890	0.934
	(0.062)	(0.048)	(0.018)

Table 5. Error Variances (standard errors) in Socioeconomic Variables

	Blacks	Latinos	Whites
Father's education, student	1.775	1.778	1.392
	(0.144)	(0.096)	(0.038)
Father's education, parent	1.865	1.782	0.870
	(0.143)	(0.136)	(0.039)
Mother's education, student	2.169	2.066	1.530
	(0.137)	(0.094)	(0.038)
Mother's education, parent	0.959	1.601	0.662
	(0.102)	(0.135)	(0.031)
Father's occupational status, student	0.377	0.855	0.305
	(0.034)	(0.034)	(0.009)
Father's occupational status, parent	0.456	0.272	0.309
	(0.036)	(0.029)	(0.010)
Mother's occupational status, student	0.566	0.485	0.441
	(0.038)	(0.020)	(0.010)
Mother's occupational status, parent	0.565	0.305	0.329
	(0.043)	(0.032)	(0.012)
Family income, 1988, parent	0.470	0.568	0.259
	(0.045)	(0.032)	(0.009)
Family income, 1992, parent	0.418	0.505	0.297
	(0.042)	(0.037)	(0.010)

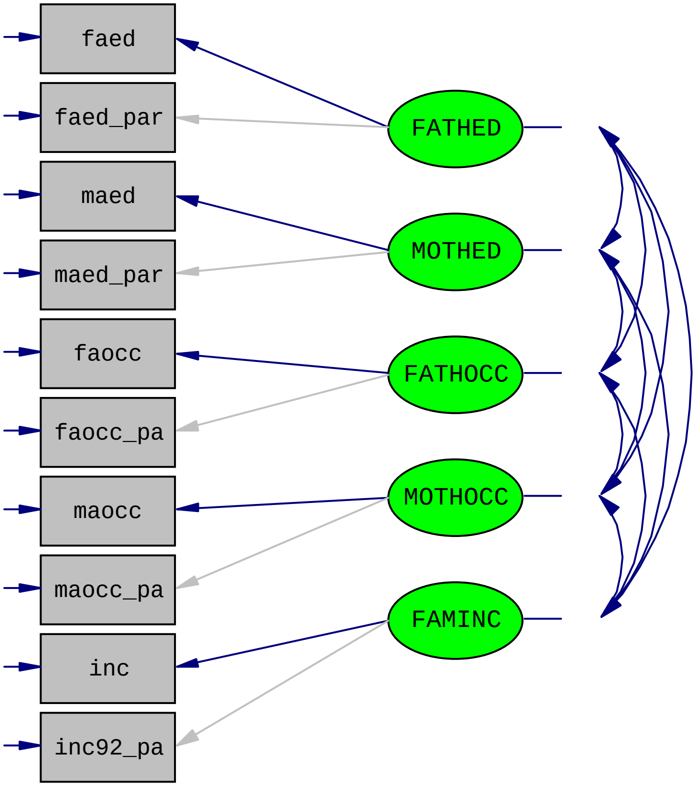

Figure 3. Model 1 of SES Measurement

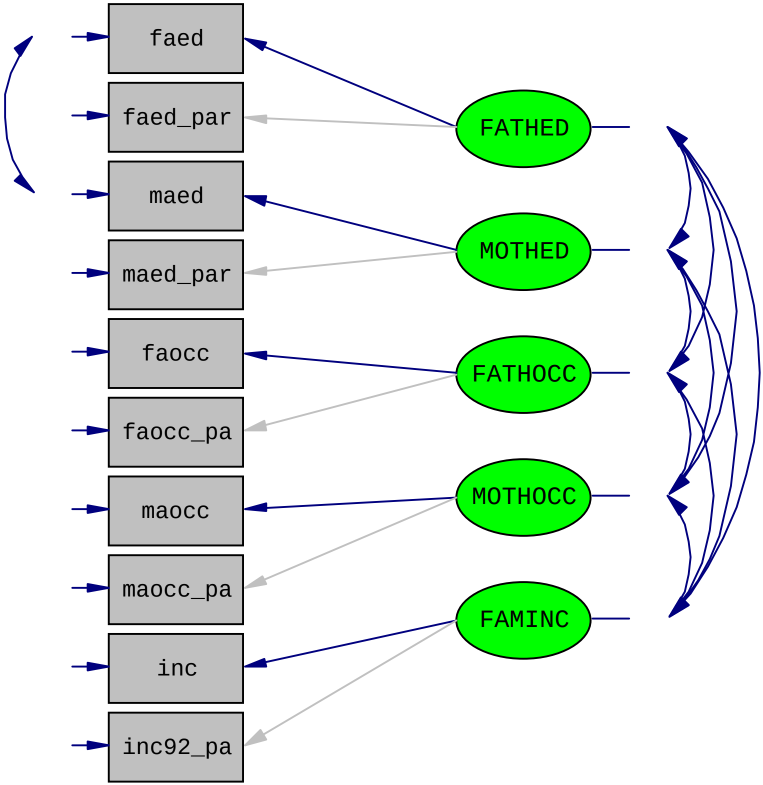

Figure 4. Model 2 of SES Measurement

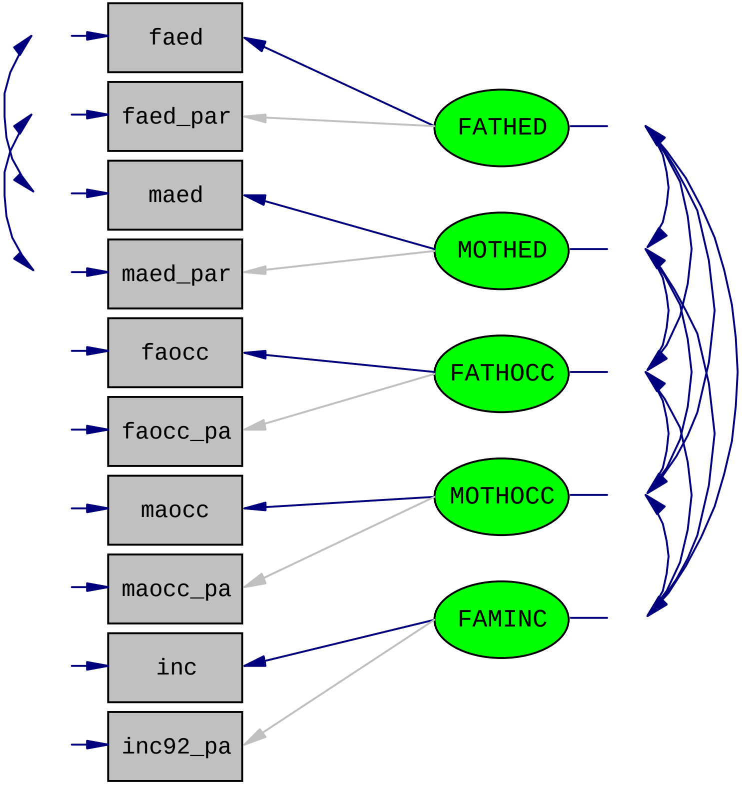

Figure 5. Model 3 of SES Measurement

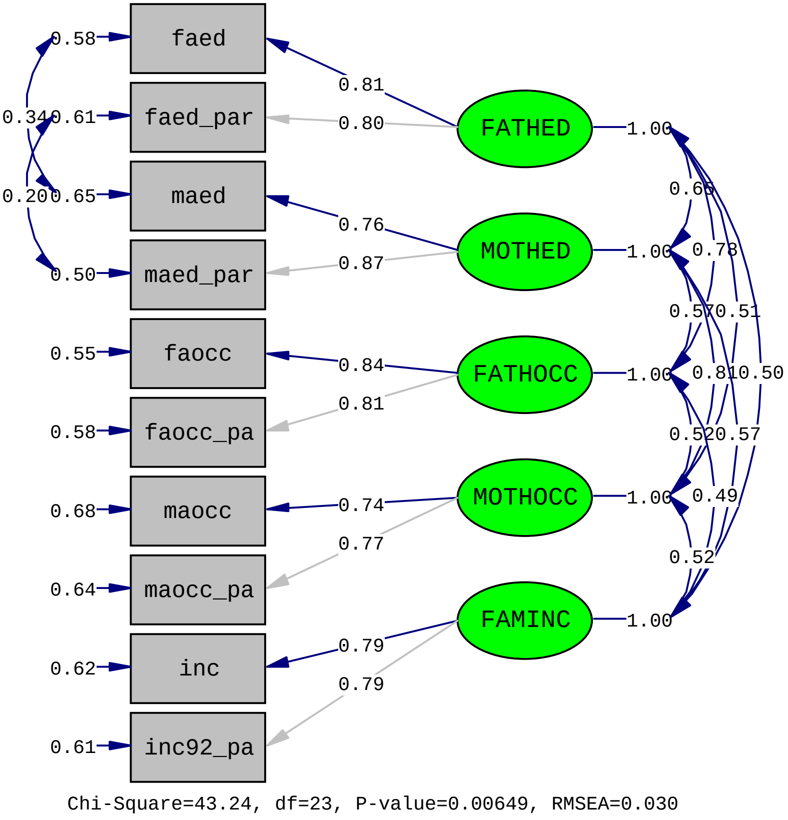

Figure 6. Standardized Estimates, Model 3, Blacks

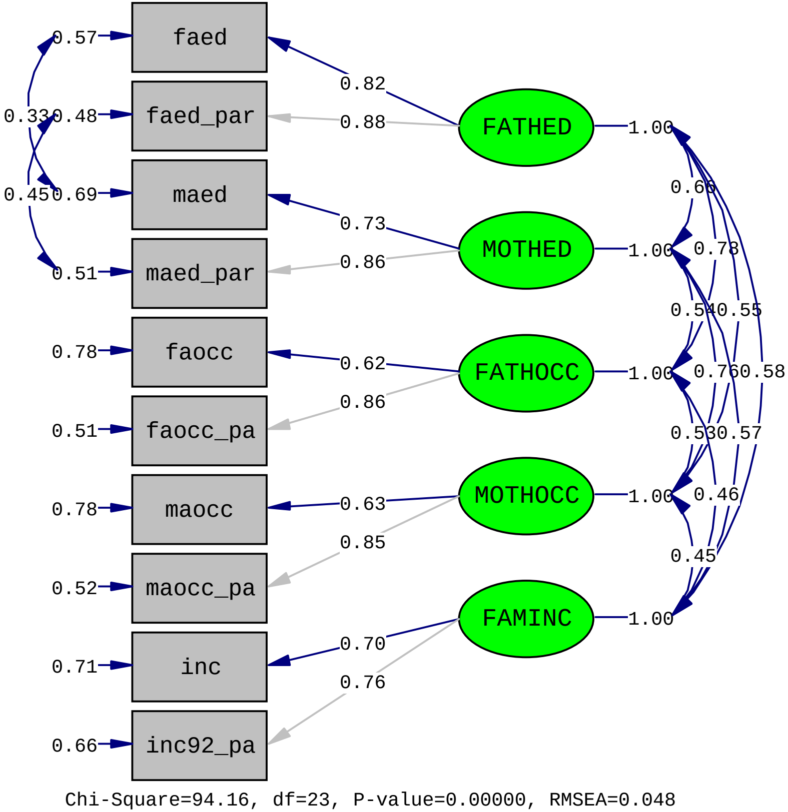

Figure 7. Standardized Estimates, Model 3, Latinos

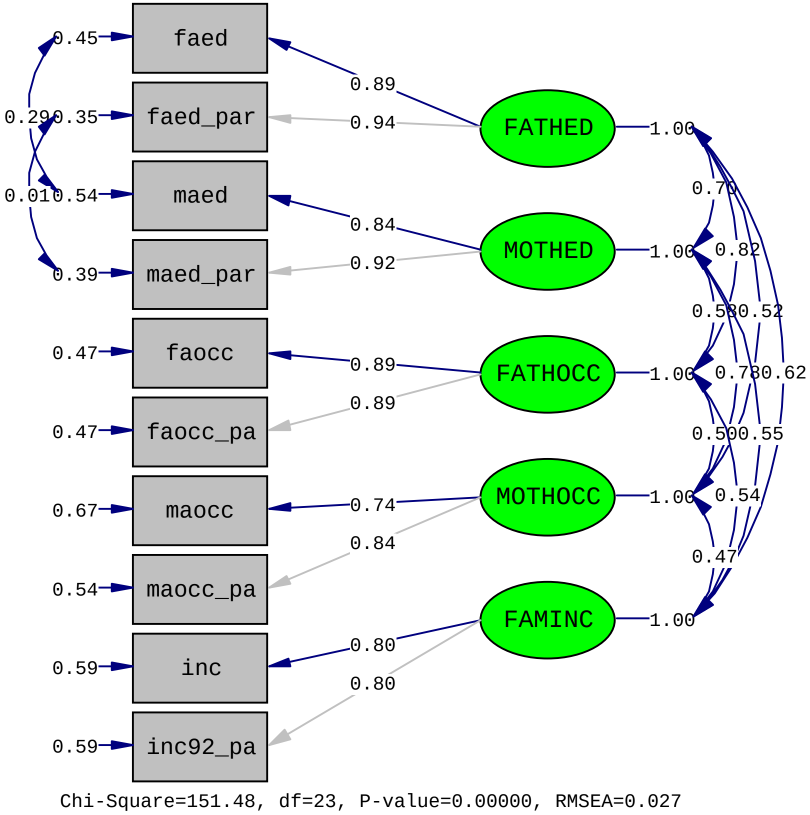

Figure 8. Standardized Estimates,
Model 3, Whites