The Interplay of Race/Ethnicity and Parental Education in Educational Attainment: A Cohort Analysis

Christopher Byrd

Criminology Major

SO-337

Dr. Jason Wollschleger

May 12, 2025

Author Note

Parts of this paper may be flagged for plagiarism, as I've heavily borrowed from an earlier literature review I completed on the subject. If you have any questions, please direct them to my email: cbyrd25@my.whitworth.edu or byrds12@outlook.com.

Abstract

Educational attainment is a key predictor of life outcomes. Despite this, access to opportunity within the American education system remains inequitable. This study examines how the intersectionality of race/ethnicity and parental education shape educational attainment among U.S. adults born between 1980 and 1984, using data from the National Longitudinal Survey of Youth 1997 (NLSY97). Prior research has highlighted racial disparities in school funding, disciplinary outcomes, and postsecondary access, as well as the effects of parental education on individual success. However, the interaction between the aforementioned variables remains under-explored in cohort-based analyses. This study applied multiple imputation and proportional odds ordinal logistic regression to model how educational attainment of residential parents, combined with race, influenced respondents' highest educational outcome completed by ages 33–37. Results indicate that while increased parental education is positively correlated with respondent educational attainment overall, the strength of this relationship varies significantly by race/ethnicity-parental interactions. Notably, Hispanic respondents demonstrated higher baseline odds of attainment but received smaller marginal benefits from increased parental education. These findings are at odds with current literature and statistics about intergenerational mobility and suggest that societal or cultural structures may moderate expected outcomes within certain racial groups. The results reinforce the importance of examining race and class as intersecting rather than additive influences in educational research. Implications for future research include the need to refine models of intergenerational achievement and to further investigate the structures that disrupt or facilitate educational mobility across different racial groups.

The Interplay of Race/Ethnicity and Parental Education in Educational Attainment: A Cohort Analysis

School serves as one of the most important institutions within an individual's life, affecting development and socialization (Conley, 2019, pp. 551–555). Throughout this experience, students of color are subjected to ethnic-racial discrimination. According to the UN Human Rights Office of the High Commissioner (OHCHR) (1965), ethnic-racial discrimination refers to negative behaviors experienced by individuals who identify as people of color. The effects of this discrimination are varied in reach and impact, from educational outcomes to adjustment within a school environment (Bottiani et al., 2017; Del Toro et al., 2024; Pearman, 2022).

Of particular interest to this paper is how societal forces and structures affect differing educational outcomes, particularly with regards to parental educational achievement and ethnic-racial discrimination. Specifically, this analysis looks to answer the following question: among U.S. adults born from 1980 to 1984, how does the intersection of race/ethnicity and parental education affect highest educational milestone completed (ages 33-37)? This question is of particular interest as educational inequality is rampant in the United States, and can affect all facets of life, from wealth (Lurie et al., 2021) to health (Zajacova & Lawrence, 2018). To answer this question, data from the National Longitudinal Survey of Youth 1997 (NLSY97) was analyzed.

The Literature

Availability of Resources

The degree to which students and educators have access to resources significantly impacts student outcomes. Condron and Roscigno (2003) found a positive correlation between school spending and test scores, with schools serving predominantly minority and low-income populations consistently underfunded. Looking to the National Center for Education Statistics (2023), over 30% of Hispanic, Black, and Native students attend low-income schools, compared to only 7% of White students. This stark disparity is often attributed to broader patterns of social stratification within U.S. society (Kozol, 1991), with concrete examples within California's education system (Reed, 2005).

Oakes (1990) argued that these inequities arise because inner-city schools lack the political influence needed to secure adequate funding. As a result, schools with predominantly minority student populations often face severe resource deficits specifically within science, limiting access to experienced teachers, adequate facilities, and technology necessary for academic success. It's notable that this article was preempted by a need for human capital - more workers within science and technology - and only partially driven with the intent to discover racial inequities.

Belonging Within Schools

Bottiani et al. (2017) completed a study which showed that gaps between White and Black students within out-of-school suspension (OSS) had a significant negative correlation on sense of belonging within a Black student's school. Alongside this, there was a positive correlation in adjustment problems with higher Black suspension rate. A low-SES population within a school also contributed to these findings. King (2011) found the opposite: school discipline had no effect on belonging and outcomes based upon race. This study is limited in effect, however, as it only examines whether there was discipline. There was no analysis of gaps based upon race.

Georgiades et al. (2013) provided an interesting set of results in their study on belonging in school, emotional issues, and adjustment issues related to racial congruence and immigration generational status. For Asian and White students, an increase in congruence resulted in a negative correlation on behavioral problems, while the opposite was true for Black students (all groups being 3rd-generation immigrants). For emotional problems, we see a negative correlation for White and Black students when congruence increases, but a positive correlation for Asian students. The overall effects of congruence are very mixed. However, they are attenuated significantly by sense of belonging, as this contributes directly to emotional and behavioral adjustment. That is to say a sense of not belonging explains emotional and behavioral issues at a statistically significant level when a positive correlation is observed, rather than congruence.

Both studies (Bottiani et al., 2017; Georgiades et al., 2013) contain parallels to critical race theory (CRT) (Delgado & Stefancic, 2017) and strain theory. Looking upon CRT, systemic inequalities, whether it be unequal representation in the student body or school punishment, increase racially divided outcomes in school performance, perpetuating negative stereotypes, and increasing inequalities outside of schooling environments. Within strain theory, the pressures of belonging and increased punishment creates strain upon students’ psyche, resulting in the findings of Georgiades et al. (2013).

School Achievement

Testing in Primary and Secondary Settings

The Black-White score gap is a numerical manifestation of systemic inequities within education and society. Pearman (2022) found that Black students are performing upwards of 0.5 standard deviations (SD) worse than their White peers, with this increasing modestly (0.1-0.2 SDs) with a SD of 1 or more with regards to county-level implicit and explicit bias. Del Toro et al. (2024) found similarly that racist encounters with an educator, specifically in math, lead to lower test scores in both victimized students and students aware of the aggression. Students experiencing vicarious discrimination were subject to even lower scores than students directly affected, though the converse is true for class grades.

The score gap is also directly related to resources, where every additional ~$1,000 in funding per student accounts for a subject-dependent increase of 6%-10% on state-administered standardized tests (Condron & Roscigno, 2003; Reed, 2005). Reed (2005) found that clearance rates for the 10th grade exit test in California were appreciably lower in every race except Asians when compared to White students. Furthermore, over half of all Hispanic students were in low-performing schools, where there were 7% less properly credentialed educators on staff.

Post-Secondary Achievement and Family

Reed (2005) found data supporting the concept that generational achievement affects outcomes for minority students. There is a correlation between completion of secondary and post-secondary education by a mother and their four-year-olds likelihood to attend early childhood education (ECE) programs. According to Jones (2024), there are significant benefits associated with attendance of ECE programs, amounting to higher wages and higher completion of education. As this is a capstone article, there are some issues, but it has been accepted by CSUMB into their digital commons, hence it’s inclusion.

Moving back towards Reed (2005), she finds that attenuating factors include general SES level, though SES is predicated often upon academic achievement. The general trend shows a vicious cycle of lower educational outcomes begetting lower educational outcomes. Assari (2018) undertook a similar study to the current one, finding that between Black and non-

Hispanic White individuals, educational attainment was greater for White individuals, but educational mobility was higher for Black individuals.

Theoretical Framework

Racial/Ethnic Generational Achievement

A preeminent theory which attempts to explain segmentation in immigrant class outcomes and mobility is segmented assimilation theory, proposed by Portes and Zhou (1993). In their paper, Portes and Zhou examined second generation immigrant households, arguing that not all immigrant groups follow the same upward trajectory of class mobility. Instead, assimilation pathways are shaped by numerous structural factors, such as race and community support structures. This often leads to outcomes that diverge in effects, from downward class mobility to assimilation of cultural ideas. This theory provides a valuable framework from which to observe the impact and effects of generational achievement. In particular, it helps explain how generational achievement may impact educational attainment differently depending on the social and institutional context in which families are embedded.

Hypothesis

From the available research and theory, the expected outcome of the current study is that higher levels of parental educational and being White will lead to increased educational outcomes, whilst the converse is true for lower parental educational and individuals who are a race/ethnicity other than White. Given previous literature, it’s expected that an increase in residential mother’s highest grade completed (HGC) will have a greater impact than father’s HGC.

Methods

NLSY97 Survey Design, Sampling, and Participants

As previously described, the sample used was the NLSY97 conducted by the US Bureau of Labor Statistics (BLS). This is a nationally representative sample carried out every year pre-2011 and every two years thereafter, with selection finalized in 1997 and round 1 of questions carried out in 1998. This database was chosen due to its large sample size (n = ~8,000 respondents), robust and comprehensive question selection, representative nature, and ease of access.

Individuals were chosen who were born between 1980 and 1984. At the time of the first interview, respondents were ages 12 to 18. The sampling strategy, as described by the BLS, is as follows:

> The NLSY97 cohort was selected in two phases... In the first phase, a list of housing units for the cross-sectional sample and the over-sample was derived from two independently selected, stratified multistage area probability samples. This ensured an accurate representation of different sections of the population defined by race, income, region, and other factors. In the second phase, sub-samples of the eligible persons identified in the first phase were selected for interview.

The cross-sectional sample was used for all data analysis. Further information about the NLSY97 is available on the NLS website[1].

Within the NLSY97, the data was subset to specify participants in round 18 who responded with an answer to the following question: [What is] the highest degree received as of the survey date? This was on a scale of 0 to 7, with zero representing no high school completion and 7 representing a doctorates degree. Important to note is the difference in scales between

[1]https://www.nlsinfo.org/content/cohorts/nlsy97

parental and respondent education levels (see appendix A, fig. A & B for the complete codebook). All responses falling outside of this range (non-interviews or invalid skips) were removed.

Data Analysis

A unique aspect of the data which made analysis more complex was the number of answers missing from the residential parent's educational variable (round 1) (see appendix for codebook). Residential parent (separated into mother and father - coded as mom and dad) was chosen, as this study is attempting to determine how parents living in the same household affect their children's future education. Figure 4 visualizing represents the missing data using an upset plot. To determine whether the pattern was random, Jamshidian and Jalal's (2010) non-parametric missing completely at random (MCAR) test[2] was used. Following this, data was imputed using the `mice` package, which utilizes Gibbs sampling. Imputation has been shown to produce better analyses than casewise deletion, especially when the data is MCAR or missing at random (MAR) (Afghari et al., 2019). The imputation method chosen was predictive mean matching (PMM), as PMM provides plausible data that is more robust than other methods (Buuren, 2018).

Prior to imputation, three dummy variables were coded (figure B) for race/ethnicity, with non-Black/non-Hispanic as the reference variable. Once the imputation was completed, an proportional ordinal logistic regression was ran on each imputation (a total of 5 imputations were created) and pooled together to create table 2 and figures 6 and 7. ggplot2 was used to construct all plots within the document due to its wealth of documentation, plethora of packages expanding on its functionality, and it's flexibility. The following regression equation was used:

---

[2]This and all other tests utilize a 95% CI.

$$HighestDegree = (HGCMother * Black) + (HGCMother * Hispanic) + (HGCMother * Mixed) + (HGCFather * Black) + (HGCFather * Hispanic) + (HGCFather * Mixed)$$

The last step taken was to fit a weighted general linear model (GLM) on an imputation using the `survey` package. The above equation was utilized in this regression. This GLM helps illustrate possible outcomes and variations within the imputations, alongside allowing weighting, which `mice` lacks support for.

Limitations

As previously noted, imputations can introduce error into analyses from inconsistent, variable data, possibly affecting results. Compounding this, mice does not support weighting, meaning it is hard to generalize results to broader society. Additionally, running a POLR off a single imputation to allow weighting increases the risk of introducing variability from the imputations.

Furthermore, the results from the GLM have limited use, as GLMs are designed for univariate models, not multivariate. While an POLR function exists for the survey package, seemingly unfixable errors arose while attempting to code the POLR function. As POLR does not necessarily handle survey weights correctly, the decision to forgo using it was made in favor of a GLM, despite its limitations.

Lastly, GED and HS are interpreted as discrete ordinal categories, with GED placed below HS. While societally they may be seen as different, practically they have similar impact and weight. See Study Limitations.

Results

A total of 6,734 respondents answered with the highest degree they'd received was, with 6,707 of these responses being valid to the question. This is out of the original 8,984 respondents enrolled in the NLSY97 in round 1. Figures 1 through 3 provide a detailed breakdown of respondents and parental educational achievement.

A total of 2,686 responses lacked HGC responses from one or both parents, as illustrated in figure 4. Jamshidian and Jalal's (2010) MCAR test gives 4 different patterns of missingness within this data, with both tests of normality and homoscedasticity presenting significant results, those being a Hawkin's probability of $p < .01$ and a non-parametric test of $p < .01$ ($\alpha = 0.05$ for both tests). However, the multivariate normality test was inconclusive, so any methods requiring a normal distribution were generally avoided. Due to time, no additional normality tests were run.

After imputing the data, the quality of the imputations were inspected using density plots (figure 5). The graphs show that the imputed data is roughly the same probability as the observed data, with the expected smoothing from a PMM imputation. Graphs displaying the logarithmic odds estimates for predictors and thresholds are visible in figures 6 and 7, and table 2. Lastly, imputation 1 was selected as an example with weights applied along a GLM, which produced significant results for 7 of the 12 predictors.

Analysis

Given the data for parental education is not MCAR, and significantly so (Table 1), there is sufficient evidence to suggest the data is MAR and has a pattern predicted by an observed variable, which explains the missingness in the data, hence allowing imputations. It is important to note that any time an imputation is used, there is a chance bias and variability which may have not existed in the dataset prior to the imputation will be introduced (Buuren, 2018), though the benefits of an imputation in the present study from the number of missing cases ($n = 2686$) outweigh the drawbacks. The relative homogeneity between the imputed and observed values supports the quality of the imputations, likely leading to less effect on the study outcome when compared to casewise deletion.

The results of the pooled POLR reveals that five of the eleven predictors are statistically significant at CI = 95%, those being HGC of residential father and mother (both $p < .01$, α = 0.05), the interaction effect of residential mother's HGC and being Hispanic ($p < .01$, α = 0.05), the interaction effect of residential father's HGC and being Hispanic ($p < .01$, α = 0.05), and being Hispanic ($p < .01$, α = 0.05) (table 2). The effects of residential mother and father HGC are in line with the literature, with mildly positive log odd estimates. This means the effect is rather mild, but still noticeable and prevalent.

Interestingly, being Hispanic is indicative of increased log odds, which seems at odds with current data (Pew Research Center, 2023). Furthermore, Hispanic residential parents see a decreased effect of parental education increasingly respondents' educational attainment, with an estimate of -0.113792 for mother's ($p < .01$, α = 0.05) and -0.110221 for father's ($p < .01$, α = 0.05).

Looking at the non-significant predictors in the POLR model, most have an intercept within ± 0.05 of 0, suggesting that, even if these predictors (all race-parent variables with the exception of Hispanic) were significant, their effects compared to the reference group, non-White/non-Hispanic individuals, are negligible. Black and mixed-race show predictor estimates that have slightly stronger negative correlation, but each have extremely high *p*-values. This means there is no basis to support intersectionality within the effects of race and education, at least among Black and mixed-race individuals.

When observing the GLM data (Table 3), the results are identical in significance to the POLR model results, apart from the interaction effect between respondents being Black and their residential father's HGC. In this specific imputation, Black:Father's HGC is significant ($\alpha = 0.05$)

Unsurprisingly, the threshold estimates for education show that almost every jump between educational level has increasing log odds which are significant, except for *HS|AA* and *AA|BA*. This is possibly due to the strong emphasis on college completion in the US.

Discussion

Context

Overall, the data partially supports and refutes the test hypotheses. While it is true that higher grade completion from a residential parent positively affects educational attainment, when interacting with race, White individuals do not see significant benefits of HGC by their parent's over individuals of other races and ethnicities — with the exception of Hispanic individuals, who experience reduced benefits from both parent's HGC. However, when not accounting for HGC, Hispanic individuals have increased log-odds of achieving higher education.

Moving back to parental effects, while the expected results was that a residential mother's HGC would have greater log-odd effects on educational achievement than that of the father, in this cohort the opposite is true. One possible explanation is errors introduced by the imputation influencing this result. Another is that, proportionally, residential fathers have attained marginally higher degrees in percentage. Also plausible is the effects of labor market alignment — the need of the market for a particular skill — or role modeling effects.

Using segmented assimilation theory (Portes & Zhou, 1993), it is possible that the results we see from Hispanic individuals come from a multitude of societal and familial structures. Previous research shows Hispanic educational attainment (AA or above) to be severely below that of many other racial groups in the US, particularly White (Kim et al., 2024). We are beginning to see increases in the rates at which Hispanic individuals attain higher education, and while these rates may not be high compared to the general population (48.1% versus 29.5%), segmented outcomes in the individuals studied can possibly explain the results. These segmentations could arise in multiple fashions. For example, this could be due to the citizenship status of individuals surveyed and whether this proportion is generalizable to the US population of Hispanic individuals. This may also align with findings on immigrant optimism, where high educational aspirations persist notwithstanding socioeconomic disadvantages (Kao & Tienda, 2022), potentially driving attainment independently of parental HGC. Furthermore, Kao and Tienda (2022) found that the immigration status of Hispanic individuals is less important to educational achievement. This may be due in part from societal or cultural barriers which prevent the full relation of educational capital from parent to child in Hispanic families. The findings suggest that simply improving parental education levels may not equitably improve intergenerational mobility for all racial groups. Educational policy must then address these societal and cultural barriers, should they be present.

Study Limitations

This study is not without limitations. As previously discussed in the limitations subsection of methods, imputations have a risk of introducing error into the study design. Additionally, certain algorithms used, primarily the GLM, are not suited for this type of data, categorical, and their results are purely illustrative. Furthermore, this study does not control for factors such as SES, citizenship status, location, racial composition of location, clustering, etc. and this isn’t generalizable to the United States population despite the representative sample.

Lastly, certain variable answers may create better results when collapsed into a single answer, namely completion of high school or a GED. Functionally, these certificates are identical in usage, and thus future studies should consider collapsing the two outcomes.

Implications for Future Research

Future research should expand upon the present study by including confounders which may create more generalizable results, such as the aforementioned factors above. Additionally, a longitudinal study tailored specifically to researching the intersectionality of race and parental education may yield better results. Segmenting race and ethnicity further may allow a more nuanced view of racial-ethnic divides in attainment. Furthermore, future research may seek to define educational success like educational achievement and job success, as education is only one small part in a person's success. With regards to ethnic and racial groups, future studies should attempt to identify societal and cultural factors which moderate gains or decreases in attainment.

References

Afghari, A. P., Washington, S., Prato, C., & Haque, M. M. (2019). Contrasting case-wise deletion with multiple imputation and latent variable approaches to dealing with missing observations in count regression models. *Analytic Methods in Accident Research*, *24*, 100104. https://doi.org/10.1016/j.amar.2019.100104

Arneback, E., & Jämte, J. (2022). How to counteract racism in education – A typology of teachers' anti-racist actions. *Race Ethnicity and Education*, *25*(2), 192–211. https://doi.org/10.1080/13613324.2021.1890566

Assari, S. (2018). Parental Education Attainment and Educational Upward Mobility; Role of Race and Gender. *Behavioral Sciences*, *8*(11), 107. https://doi.org/10.3390/bs8110107

Bartos, O. J., & Wehr, P. (2002). *Using Conflict Theory* (1st ed.). Cambridge University Press.

Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and white adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, *109*(4), 532–545. https://doi.org/10.1037/edu0000155

Buuren, S. van. (2018). *Flexible Imputation of Missing Data, Second Edition*. Taylor & Francis Group.

Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the Air: A Nationwide Exploration of Teachers' Implicit Racial Attitudes, Aggregate Bias, and Student Outcomes. *Educational Researcher*, *49*(8), 566–578. https://doi.org/10.3102/0013189x20937240

Condron, D. J., & Roscigno, V. J. (2003). Disparities within: Unequal Spending and Achievement in an Urban School District. *Sociology of Education*, *76*(1), 18. https://doi.org/10.2307/3090259

Conley, D. (2019). *You May Ask Yourself: An Introduction to Thinking like a Sociologist* (J. Cahill, Ed.; Core Sixth). W. W. Norton & Company.

Del Toro, J., Legette, K., Christophe, N. K., Pasco, M., Miller-Cotto, D., & Wang, M.-T. (2024). When ethnic–racial discrimination from math teachers spills over and predicts the math adjustment of nondiscriminated adolescents: The mediating role of math classroom climate perceptions. *Developmental Psychology*, *60*(12), 2242–2257. https://doi.org/10.1037/dev0001833

Delgado, R., & Stefancic, J. (2017). *Critical Race Theory: An Introduction*. Dev Publishers & Distributors.

Georgiades, K., Boyle, M. H., & Fife, K. A. (2013). Emotional and Behavioral Problems Among Adolescent Students: The Role of Immigrant, Racial/Ethnic Congruence and Belongingness in Schools. *Journal of Youth and Adolescence*, *42*(9), 1473–1492. https://doi.org/10.1007/s10964-012-9868-2

Jamshidian, M., & Jalal, S. (2010). Tests of Homoscedasticity, Normality, and Missing Completely at Random for Incomplete Multivariate Data. *Psychometrika*, *75*(4), 649–674. https://doi.org/10.1007/s11336-010-9175-3

Jones, C. N. (2024). *Benefits of Early Childhood Education on Later Adult Life*.

Kao, G., & Tienda, M. (2022). Optimism and Achievement: The Educational Performance of Immigrant Youth*. *Interdisciplinary Perspectives on the New Immigration*, 83–101. https://doi.org/10.4324/9781315054216-4

Kim, J. H. "Jane," Claudia, M., Zhao, Z., & Swirksy, E. (2024). *Race and Ethnicity in Higher Education: 2024 Status Report*. American Council on Education.

King, L. D. (2011). *An examination of the influence of school discipline, race/ethnicity, and gender on students' attitudes toward school*. University of South Carolina.

Kozol, J. (1991). *Savage Inequalities*. The Trumpet Club.

Lurie, L. A., Hagen, M. P., McLaughlin, K. A., Sheridan, M. A., Meltzoff, A. N., & Rosen, M. L. (2021). Mechanisms linking socioeconomic status and academic achievement in early childhood: Cognitive stimulation and language. *Cognitive Development*, *58*, 101045. https://doi.org/10.1016/j.cogdev.2021.101045

National Center for Education Statistics. (2023). *Concentration of Public School Students Eligible for Free or Reduced-Price Lunch*. NCES.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science.* https://www.rand.org/content/dam/rand/pubs/reports/2006/R3928.pdf

Pearman, F. A. (2022). Collective Racial Bias and the Black-White Test Score Gap. *Race and Social Problems*, *14*(4), 283–292. https://doi.org/10.1007/s12552-021-09347-y

Pew Research Center. (2023, August 16). *Educational Attainment of Hispanic Population in the U.S., 2021*. Pew Research Center. www.pewresearch.org/chart/us-hispanics-education/

Portes, A., & Zhou, M. (1993). The New Second Generation: Segmented Assimilation and its Variants. *The Annals of the American Academy of Political and Social Science*, *530*(1), 74–96. https://doi.org/10.1177/0002716293530001006

Reed, D. (2005). Educational resources and outcomes in California, by race and ethnicity. *California Counts: Public Trends and Profiles*, *6*(3). https://www.ppic.org/wp-content/uploads/CC_205DRCC.pdf

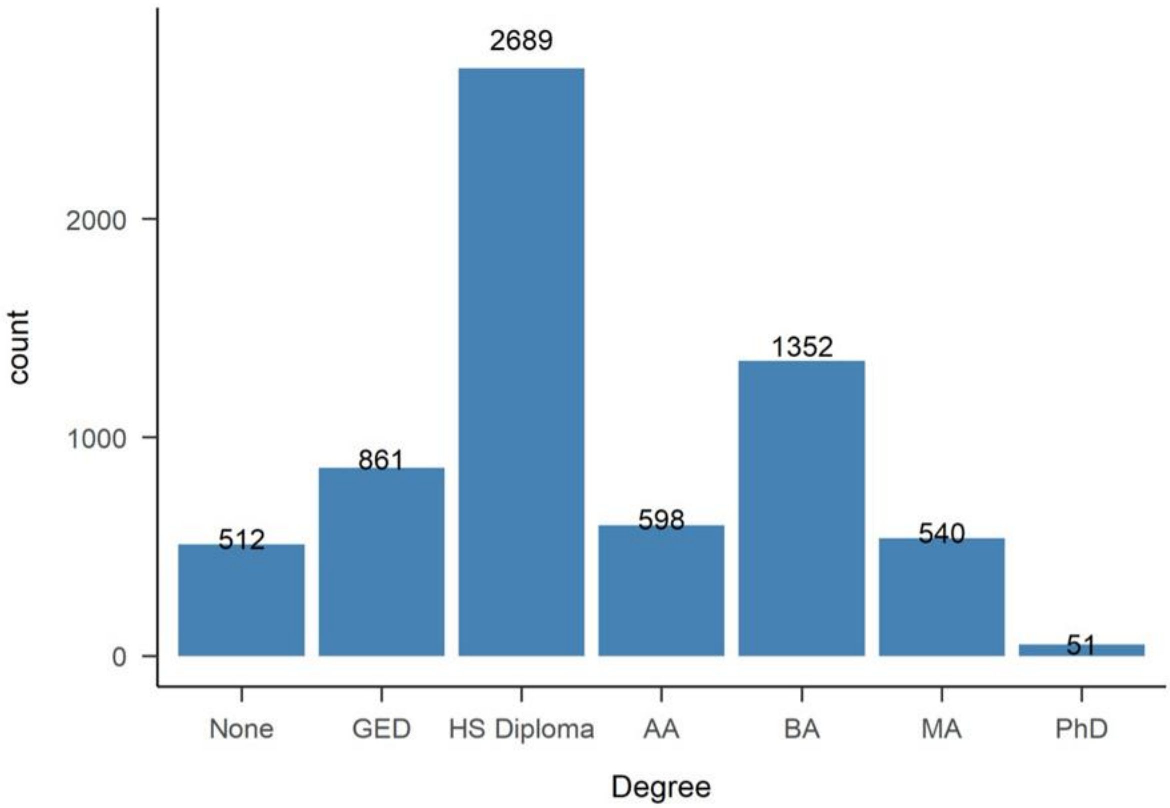
United Nations Human Rights Office of the High Commissioner. (1965, December 21). *International Convention on the Elimination of All Forms of Racial Discrimination*. UN OHCHR. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-convention-elimination-all-forms-racial

Zajacova, A., & Lawrence, E. M. (2018). The Relationship Between Education and Health: Reducing Disparities Through a Contextual Approach. *Annual Review of Public Health*, *39*(1), 273–289. https://doi.org/10.1146/annurev-publhealth-031816-044628

Figures

Figure 1

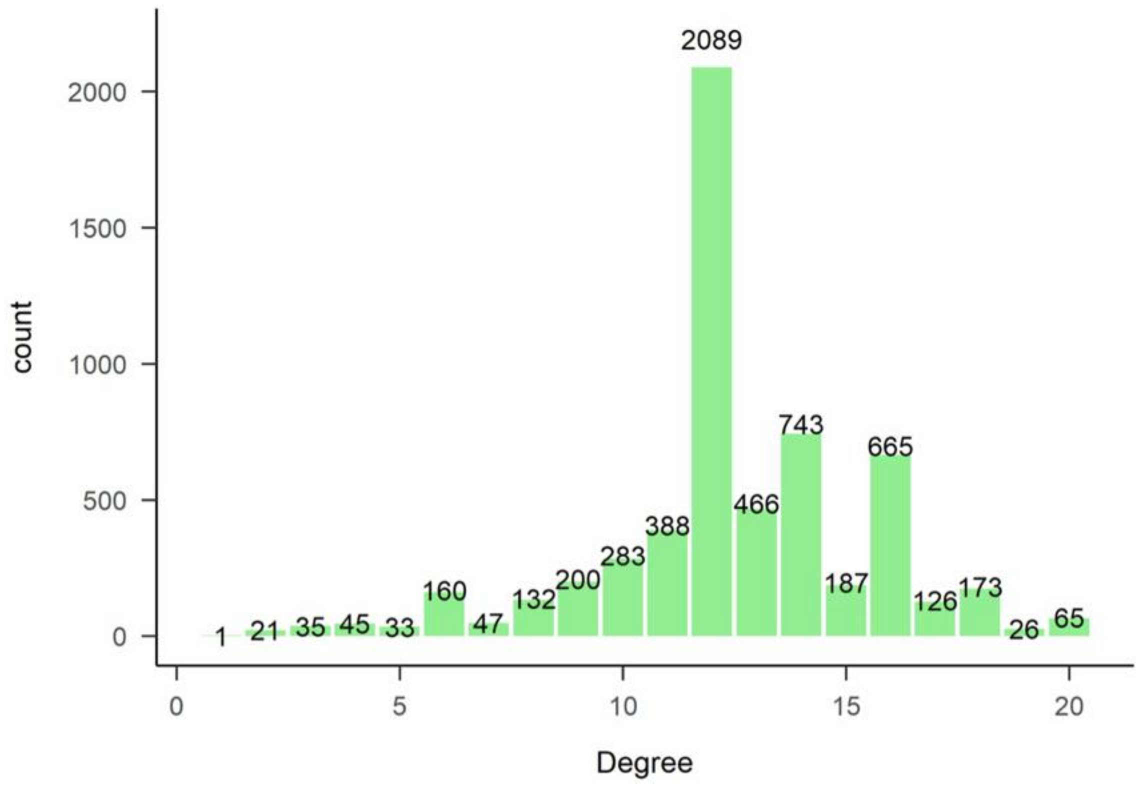
Educational Attainment of Respondents



Note. Collected during round 18 of the NLSY97.

Figure 2

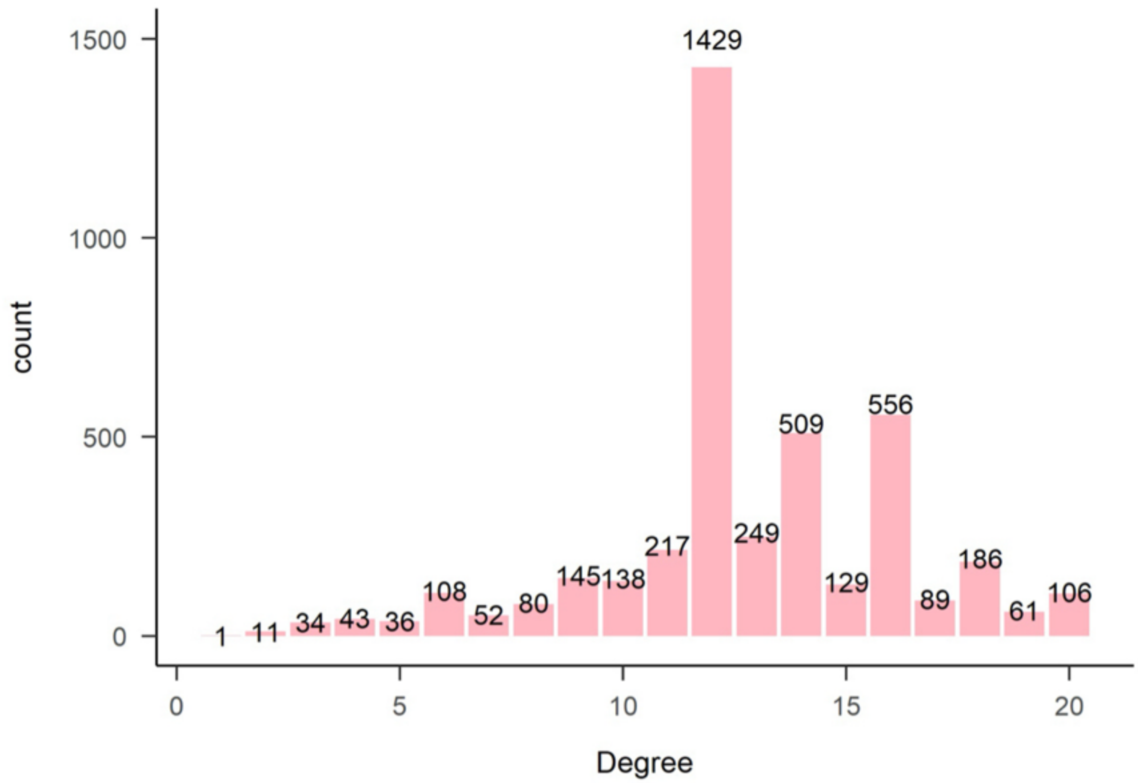
Highest Grade Completion of Residential Mom



Note. N/A responses were removed for this graph.

Figure 3

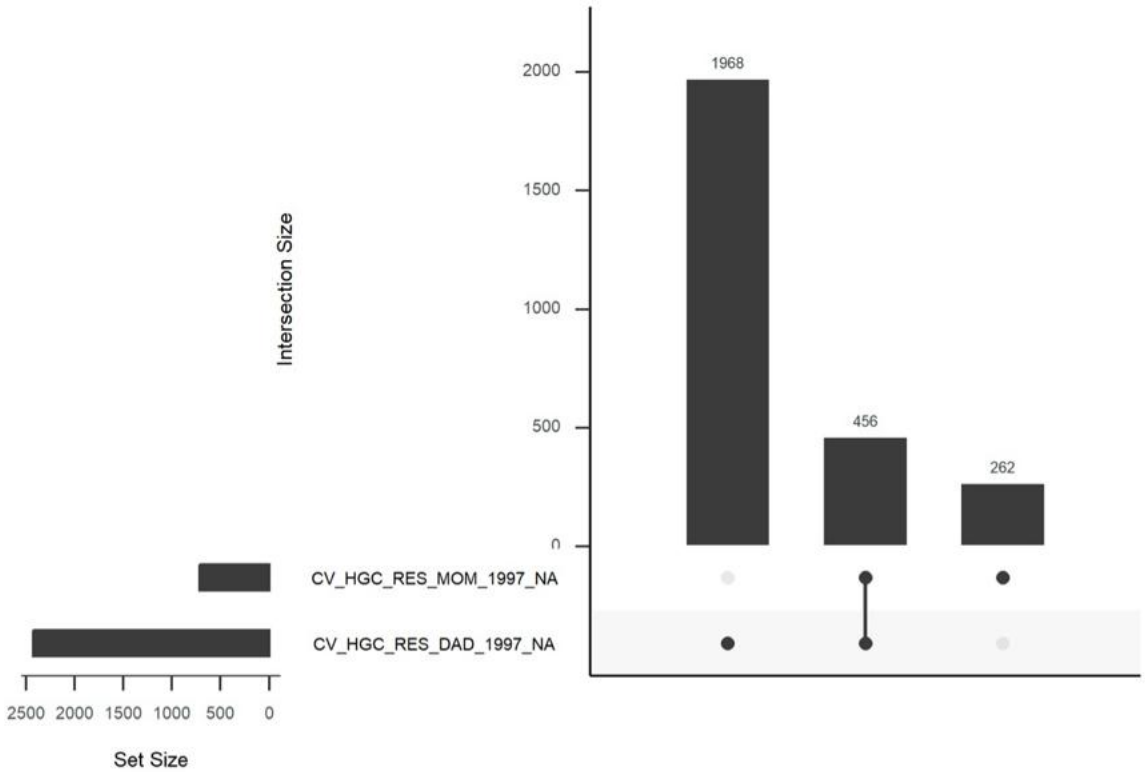
Highest Grade Completion of Residential Dad



Note. N/A responses were removed for this graph.

Figure 4

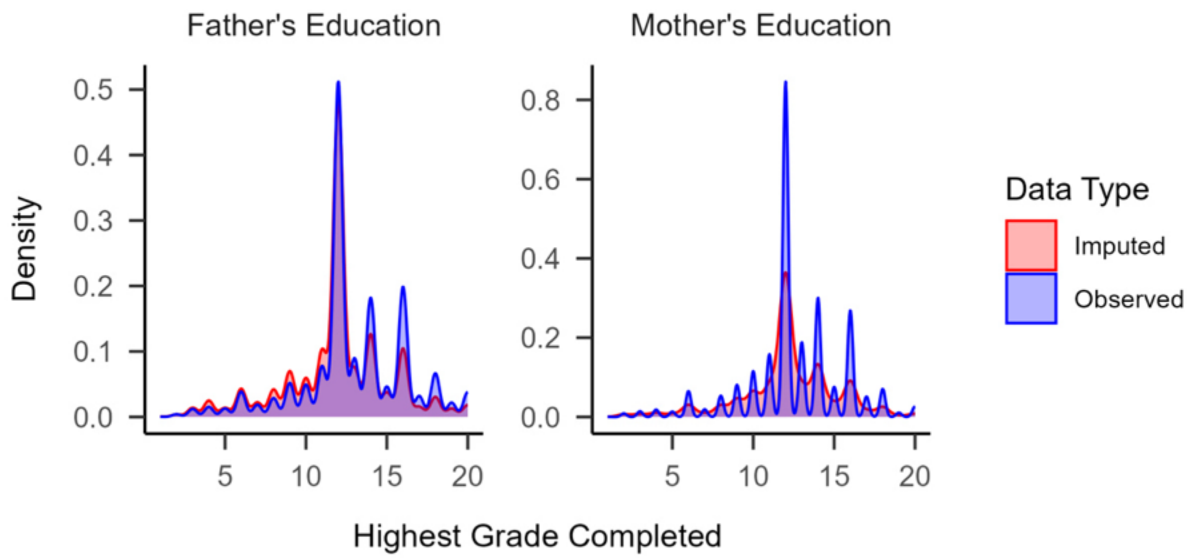
Occurrence of Missing Residential Parental Responses



Note. Upset plot of missing responses to residential parent’s highest grade completed.

Figure 5

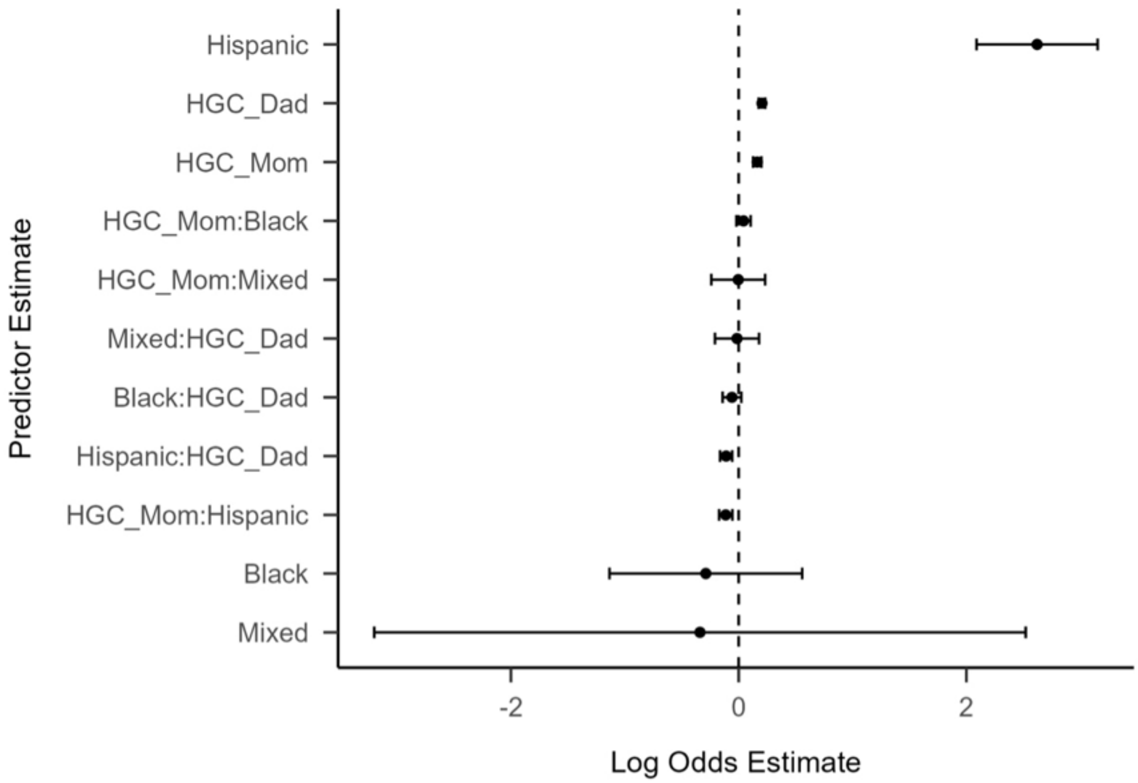
Density Plot of Imputed Versus Observed Values



Note. Pooled density plot values comparing educational attainment between observed and imputed values.

Figure 6

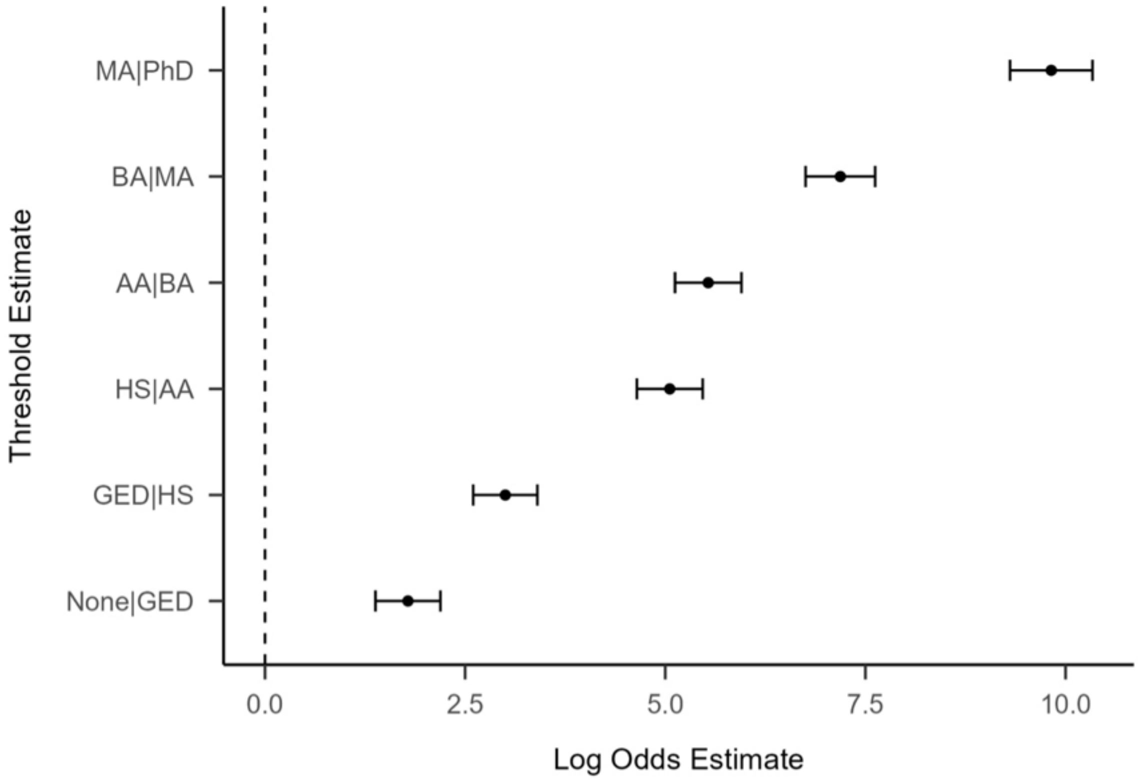
Pooled Coefficients Estimates of Imputed polr Model



Note. Unweighted ordinal logistic regression model of imputed data from NLSY97.

Figure 7

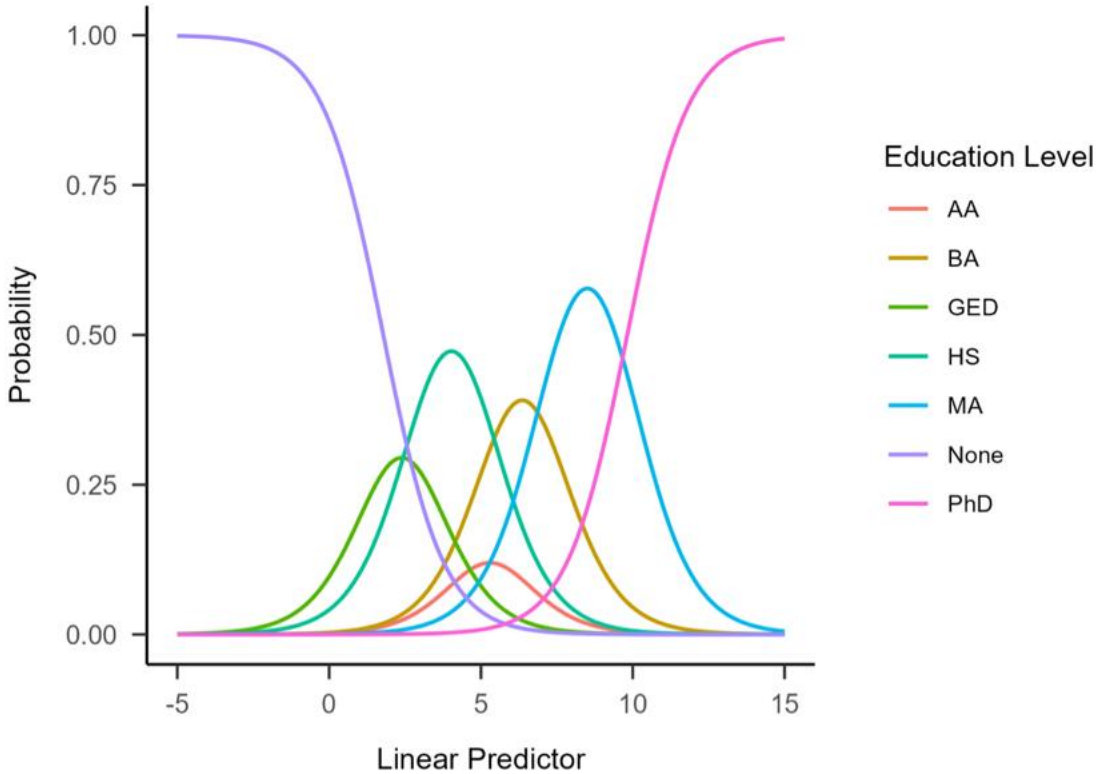
Pooled Coefficient Estimates of Imputed polr Model



Note. Imputed polr model of NLSY97 data.

Figure 8

Density Graph of polr Model



Note. Utilizes pooled OLR results to display probability of educational attainment based upon predictors.

Tables

Table 1

MCAR Test of Parental Education Missingness

Pattern	Highest Degree	Highest Grade Completed		*n*
		Mom	Dad	
1	1	1	1	3917
2	1	1	N/A	1968
3	1	N/A	N/A	456
4	1	N/A	1	262
Total	4	2	2	6603

Note. The specific MCAR test being ran utilizes Jamshidian and Jalal's (2010) proposed methodology. Furthermore, test information can be found within the R package `MissMech`.

Table 2

Ordinal Logistic Regression Results for Educational Attainment

Term	Estimate	*SE*	*F*	*df*	*p*
			Predictor Estimate		
HGC Mom	0.1631633	0.0180142	9.0575029	56.98497	<0.01
Black	-0.288742	0.4160886	-0.693945	33.10176	0.4925585
Hispanic	2.6211007	0.2696701	9.7196573	185.7731	<0.01
Mixed Race*	-0.339786	1.4578296	-0.233077	975.1149	0.8157508
HGC Dad	0.2047639	0.0146422	13.984501	144.9801	<0.01
HGC Mom:Black	0.0426151	0.0315178	1.3520959	85.62714	0.1799051
HGC Mom:Hispanic	-0.113792	0.0288421	-3.945341	42.09671	<0.01
HGC Mom:Mixed	-0.003684	0.1189891	-0.030964	91.32734	0.9753661
Black:HGC Dad	-0.059038	0.0371335	-1.589899	10.50609	0.1447160
Hispanic:HGC Dad	-0.110221	0.026111	-4.221232	41.14578	<0.01
Mixed:HGC Dad	-0.014649	0.0986168	-0.148549	456.7391	0.8819751
			Threshold Estimate		
None|GED	-0.785746	0.2062901	8.6564781	313.0477	<0.01
GED|HS	3.0015602	0.2038075	14.727426	321.1901	<0.01
HS|AA	5.0560201	0.208676	24.22904	313.7444	<0.01
AA|BA	5.5361057	0.210876	26.252892	320.1433	<0.01
BA|MA	7.1869071	0.2209116	32.532963	378.2244	<0.01
MA|PhD	9.8210439	0.2622659	37.446899	724.3489	<0.01

Note. Uses polr on pooled imputed dataset. See *methods* for regression formula.

*Non-Hispanic.

Table 3

Survey GLM of Weighted Imputation

Coefficients	Estimate	*SE*	*T*	*Pr(>|t|)*
(Intercept)	0.278254	0.134809	2.064	0.03905
HGC Mom	0.117171	0.011554	10.141	<0.01
Black	0.125328	0.252801	0.496	0.62008
Hispanic	1.764051	0.174008	10.138	<0.01
Mixed Race*	-0.275646	0.922629	-0.299	0.76513
HGC Dad	0.142678	0.009553	14.935	<0.01
HGC Mom:Black	0.028286	0.022246	1.272	0.20359
HGC Mom:Hispanic	-0.070874	0.018047	-3.927	<0.01
HGC Mom:Mixed	0.020633	0.068754	0.3	0.76411
Black:HGC Dad	-0.064504	0.016949	-3.806	<0.01
Hispanic:HGC Dad	-0.078902	0.015649	-5.042	<0.01
Mixed:HGC Dad	-0.029156	0.053379	-0.546	0.58494

Note. This general linear model (GLM) uses imputation 1 (default value) to model a GLM of the highest level of education obtained. Weights are from round 18 (2017) of the NLSY97.

*Non-Hispanic.

Appendix

Figure A

Codebook

Var	Key
CV_HGC_RES_DAD	0 None
Comment	1 1st
Highest grade completed by respondent's residential father (includes both biological and non-biological fathers).	2 2nd
	3 3rd
	4 4th
	5 5th
	6 6th
	7 7th
	8 8th
	9 9th
	10 10th
	11 11th
	12 12th
	13 1st Year College
	14 2nd Year College
	15 3rd Year College
	16 4th Year College
	17 5th Year College
	18 6th Year College
	19 7th Year College
	20 8th Year College or more
	95 Ungraded
	-3 Invalid Skip
	-4 Valid Skip

Var	Key
KEY!RACE_ETHNICITY	1 Black
Comment	2 Hispanic
Combined race and ethnicity variable.	3 Mixed Race (Non-Hispanic)
	4 Non-Black/Non-Hispanic

Var	Key
CV_HIGHEST_DEGREE_EVER_EDT	0 None
	1 GED
Comment	2 High School Diploma
The highest degree received as of the survey date	3 Associate's Degree
	4 Bachelor's Degree
	5 Master's Degree
	6 PhD
	7 Professional Degree
	-3 Invalid Skip
	-4 Valid Skip
	-5 Non-Interview

Figure B

Codebook (Cont.)

Var	Key	
CV_HGC_RES_Mom	0	None
Comment	1	1st
Highest grade completed by respondent's residential mother (includes both biological and non-biological mothers).	2	2nd
	3	3rd
	4	4th
	5	5th
	6	6th
	7	7th
	8	8th
	9	9th
	10	10th
	11	11th
	12	12th
	13	1st Year College
	14	2nd Year College
	15	3rd Year College
	16	4th Year College
	17	5th Year College
	18	6th Year College
	19	7th Year College
	20	8th Year College or more
	95	Ungraded
	-3	Invalid Skip
	-4	Valid Skip

Var	Key	
DV_RACE_BLACK	0	Not black
Comment	1	Is black
Dummy varriable for \`polr\` model		

Var	Key	
DV_RACE_HISPANIC	0	Not hispanic
Comment	1	Is hispanic
Dummy varriable for \`polr\` model		

Var	Key	
DV_RACE_MIXED	0	Not mixed
Comment	1	Is mixed
Dummy varriable for \`polr\` model		

Note. Mixed refers to non-Hispanic individuals of mixed race.