# 9

# Applying Generalized AI to Conventional Quantitative Data

One of the strengths of Viterna's study (chapter 8) is her integration of qualitative interview data with the examination of cross-case evidence. The qualitative evidence, usually personal narratives, reinforces and enlivens her analysis of cross-case patterns. Rarely do most social scientists have the opportunity to join and triangulate different types of evidence in a single study. The most common situation is for the researcher to have a set of quantitative data on multiple cases, most often at the individual level, and nothing more. Furthermore, the analyst often does not participate in the collection of the data and thus has little opportunity to enrich the quantitative analysis with qualitative evidence.

The purpose of this chapter is to demonstrate that generalized AI can be usefully applied to conventional quantitative data. Because generalized AI is fundamentally descriptive in nature, it can be used to complement findings derived using conventional quantitative methods. Applying different analytic techniques to the same data does not make the research multimethod; however, using multiple analytic techniques allows the researcher to observe the impact of different underlying assumptions on findings, especially when the techniques make contrasting assumptions regarding the nature of causation.

The demonstration of generalized AI presented in this chapter uses data from the National Longitudinal Survey of Youth (NLSY), 1979 sample. The analysis is restricted to Black females and focuses on membership in the set of respondents in poverty as the outcome. Before presenting the application of generalized AI to the NLSY data, I offer two quantitative analyses. The first applies logistic regression techniques to a binary dependent variable, in poverty versus not in poverty. The second analysis parallels the first, except that the dependent and independent variables are operationalized as *fuzzy sets* (see appendix B). The second quantitative analysis uses ordinary least squares (OLS) regression to build a bridge between the logistic regression analysis and the application of generalized

AI to the fuzzy-set data. As discussed previously, AI is fundamentally set-analytic in nature. To utilize the truth table techniques presented in this work, causal conditions must be operationalized as crisp or fuzzy sets. To ensure comparability of results, I use the fuzzy sets that were prepared for the generalized AI application as my dependent and independent variables in the second quantitative analysis.

## LOGISTIC REGRESSION ANALYSIS

The first quantitative analysis regresses "poverty status" on three interval/ratio-scale variables (respondent's parents' income-to-poverty ratio, respondent's years of education, and respondent's Armed Forces Qualifying Test percentile score) and two dichotomous variables (married vs. not married, and having one or more children vs. having no children). Details regarding the measures used in the logistic regression are provided in appendix E.

Poverty status is a dichotomy, with 1 indicating that household income is less than or equal to the "poverty level" for households of that type (determined by the number of adults, the number of children, and so on), and 0 indicating that household income is greater than the poverty level. For example, if the respondent's household income is $14,000 for a family of four (two adults and two children), and the poverty level for households of that type is $15,000, the income-to-poverty ratio is 14,000/15,000 = 0.93, which would translate to a score of 1 on poverty status. An income-to-poverty ratio of 1.0 or less indicates that the respondent is in poverty.

Parents' income-to-poverty ratio is constructed in the same manner, as a ratio of household income to poverty level. However, it is entered into the logistic regression analysis as a ratio-scale independent variable, not as a dichotomy. Respondent's years of education is linked to educational degrees—such that, for example, a score of 12 indicates that the respondent completed high school. The Armed Forces Qualifying Test (AFQT) is mistakenly treated as a generic test of intelligence by some researchers (e.g., Herrstein and Murray in *The Bell Curve*). However, it is best viewed as a test of the respondent's trainability, which is how it is used by the military. Basically, it is a measure of the respondent's degree of retention of school-based learning. Thus, it is indirectly a measure of school performance, as well as a measure of the respondent's degree of acquiescence to authority.

Table 9-1 reports the results of the logistic regression of poverty status on the five independent variables. All five have statistically significant effects on the odds of being in poverty for Black females. Having children more than doubles the odds of poverty (odds ratio = 2.171), while being married dramatically reduces the odds (odds ratio = 0.125). Parents' income-to-poverty ratio, respondent's years of education, and respondent's AFQT percentile score all reduce the odds of poverty. Overall, these results are consistent with those reported in Ragin and Fiss (2017) and, more generally, with findings reported in the research literature on poverty.

TABLE 9-1 Logistic regression analysis of poverty status, Black female sample

| | Coefficient (standard error) | Odds ratio |
|---|---|---|
| Children (1 = yes) | 0.775 *** (0.225) | 2.171 |
| Married (1 = yes) | −2.083*** (0.244) | 0.125 |
| Parents' income-to-poverty ratio | −0.112* (0.045) | 0.894 |
| Respondent's years of education | −0.468*** (0.074) | 0.627 |
| AFQT percentile score | −0.020** (0.007) | 0.980 |
| Constant | 5.703*** (0.906) | 299.853 |

NOTES: $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; pseudo-$r^2 = 0.285$; likelihood-ratio $\chi^2 = 274.47$ (df = 5); $N = 775$.

## OLS REGRESSION ANALYSIS USING FUZZY SETS

The OLS regression analysis that follows serves as a bridge between the logistic regression analysis, just presented, and the application of generalized AI, still to come. The regression analysis is unconventional in that it uses fuzzy sets in place of the more familiar variables used in the logistic regression analysis. Before presenting the results of the OLS regression, I provide an overview of the construction and calibration of the relevant fuzzy sets.

The dependent variable is degree of membership in the set of households in poverty. This fuzzy set uses the following benchmarks to convert a respondent's income-to-poverty ratio into degree of membership in the set of households in poverty:

| Income-to-poverty ratio | Poverty membership score |
|---|---|
| 0 to 1 | 1 to 0.95 |
| 1 to 2 | 0.95 to 0.5 |
| 2 to 3 | 0.5 to 0.05 |
| 3+ | 0.05 to 0 |

The use of a ratio of three times the poverty level for full membership in the set of cases not in poverty is a conservative cutoff value, but also one that is anchored in substantive knowledge regarding what it means to be out of poverty. For example, in 1989, the weighted average poverty threshold for a family of two adults and two children was about $12,500 (Social Security Bulletin, Annual Statistical Supplement

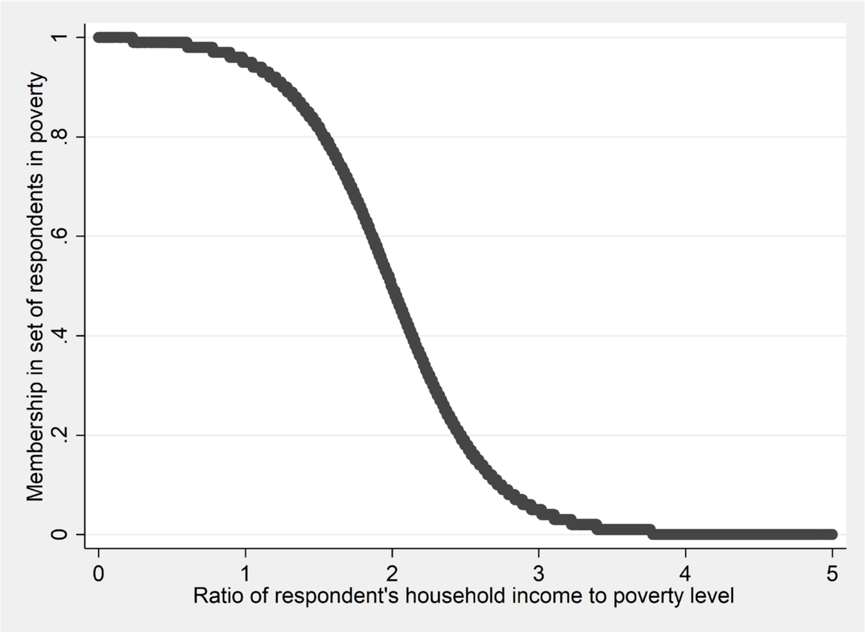


FIGURE 9-1. Calibration of degree of membership in poverty.

1998: tbl. 3.E). Three times this poverty level corresponds to $37,500 for a family of four, a value that lies just slightly above the median family income of $35,353 in 1990 (U.S. Census Bureau, Historical Income Tables—Families, tbl. F-7).

Figure 9-1 illustrates the translation of income ratio values to fuzzy membership scores. For presentation purposes, the *x*-axis has been truncated at an income-to-poverty ratio of 5, consistent with the fact that the threshold for non-membership in the outcome set (an income-to-poverty ratio of 3) has been surpassed by a substantial margin. Note that the calibration of degree of membership in poverty is much more nuanced than the dichotomous dependent variable used in the logistic regression analysis. The dichotomy treats respondents who are barely out of the set of households in poverty (e.g., with an income-to-poverty ratio of 1.01) the same as respondents who are well-off (e.g., with an income-to-poverty ratio of 5 or even greater). The crossover point of the fuzzy set, which separates respondents who are more in versus more out of the set in poverty, is an income-to-poverty ratio of 2.

In place of the two dichotomous variables used in the logistic regression analysis, married versus not married and having children versus not having children, the OLS regression analysis uses a single fuzzy set, *favorable domestic situation*, coded as follows:

| Family status combination | Membership in favorable domestic situation |
|---|---|
| married, no children | 1.0 |
| married with children | 0.6 |
| unmarried, no children | 0.4 |
| unmarried with children | 0.0 |

The membership scores are arrayed according to the association of the categories with poverty. Marriage tends to offer a degree of insulation from poverty, while having children makes poverty more likely. Thus, the highest membership score in the fuzzy set *favorable domestic situation* is for respondents who are married without children (1.0); the lowest membership score is for unmarried respondents with children (0.0). The two middle combinations, *married with children* and *unmarried without children*, both entail domestic situations that are equivocal with respect to poverty avoidance, earning them membership scores close to the crossover point (0.5). However, respondents who are married with children are coded as slightly more in than out of *favorable domestic situation* (0.6), while respondents who are *unmarried without children* are coded as slightly more out than in (0.4).

In place of parents' income-to-poverty ratio, the OLS regression analysis uses *not-low parental income*, a specific calibration of parents' income-to-poverty ratio. The numerator of this measure is based on the average of the reported 1978 and 1979 total net family income in 1990 dollars. The denominator is the household-adjusted poverty level for that household. As explained in appendix B, fuzzy sets use adjectives to specify the range of relevant variation in a source variable. While "parental income" does not make sense as a fuzzy set, "high parental income" and "low parental income" can both be calibrated as fuzzy sets, using data on parental income-to-poverty ratio as the source variable. It is important to note that "low parental income" is not the simple mathematical reverse (i.e., set negation) of "high parental income." A middle-income respondent registers relatively low membership in both "low income" and "high income." The negation of "high income" is "not-high income"; the negation of low income is "not-low income."[1] The benchmarks for degree of membership in *not-low-income parents* are as follows:

| Parents' income-to-poverty ratio | Membership in not-low income |
|---|---|
| 0 to 2 | 0 to 0.05 |
| 2 to 3 | 0.05 to 0.5 |
| 3 to 5.5 | 0.5 to 0.95 |
| 5.5+ | 0.95 to 1 |

Degree of membership in the set of respondents with *not-low AFQT scores* is based on categories used by the Department of Defense to place enlistees. The military divides the AFQT scale into five categories based on percentiles. Persons in categories I (93rd to 99th percentiles) and II (65th to 92nd percentiles) are considered above average in trainability; those in category III (31st to 64th percentiles) are considered about average; those in category IV (10th to 30th percentiles) are designated as below average in trainability; and those in category V (1st to 9th percentiles) are designated as well below average. To construct the fuzzy set of respondents with *not-low AFQT scores*, I use respondents' AFQT percentile scores. The threshold for full membership (0.95) in the set of respondents with *not-low AFQT scores* was placed at the 30th percentile, in line with its usage by the military; respondents who scored greater than the 30th percentile received fuzzy membership scores greater than 0.95. The crossover point (0.5) was set at the 20th percentile, and the threshold for non-membership was set at the 10th percentile, again reflecting the practical application of AFQT scores by the military. Respondents who scored worse than the 10th percentile received fuzzy scores less than 0.05 in degree of membership in the set of respondents with *not-low AFQT scores*.

| AFQT percentile score | Membership in not-low AFQT score |
|---|---|
| 1st to 10th | 0 to 0.05 |
| 10th to 20th | 0.05 to 0.5 |
| 20th to 30th | 0.5 to 0.95 |
| 30th+ | 0.95 to 1 |

Respondent's years of education serves as the source variable for the fuzzy set, degree of membership in the set of *educated* respondents. The translation of years of education to fuzzy membership scores is detailed below. Respondents with twelve or more years of schooling are more in than out of the set of educated respondents (fuzzy score > 0.5). Those with fewer than nine years of education are treated as fully out of the set of educated respondents (fuzzy score of 0.0), and those with sixteen or more years of education are treated as fully in the set of educated respondents.

| Years of education | Membership in educated |
|---|---|
| 0–8 | 0.0 |
| 9 | 0.1 |
| 10 | 0.2 |
| 11 | 0.4 |
| 12 | 0.6 |
| 13 | 0.7 |
| 14 | 0.8 |
| 15 | 0.9 |
| 16 (max.) | 1.0 |

Table 9-2 reports the results of the OLS regression analysis using fuzzy-set membership scores for the dependent and independent variables. Overall, the results are entirely consistent with the logistic regression analysis reported in table 9-1.

TABLE 9-2 OLS regression analysis of degree of membership in poverty, Black female sample

| | Coefficient (standard error) | Standardized coefficient |
|---|---|---|
| Favorable domestic situation | −0.487*** (0.036) | −0.366 |
| Not-low parental income | −0.188*** (0.030) | −0.180 |
| Educated | −0.440*** (0.059) | −0.237 |
| Not-low AFQT score | −0.215*** (0.032 | −0.216 |
| Constant | 1.151*** (0.035) | – |

NOTES: ***$p < 0.001$; $r2 = 0.466$; $F = 167.91$ (df = 4 and 770); $N = 775$.

All four independent variables have negative effects on degree of membership in poverty, and all four are statistically significant at $p < 001$.[2] The metric regression coefficients indicate the decrease in membership in poverty associated with full membership in each of the condition sets. Thus, the four independent variables utilize the same metric. For example, a respondent with full membership in *favorable domestic situation* (i.e., respondent is married and childless) registers a 0.487 decrease in the outcome, degree of membership in poverty. Full membership in the set of *educated* respondents also has a very strong metric effect on membership in poverty, a reduction of 0.440. The $r^2$ value of this analysis, 0.466, is impressive for individual-level data.[3]

## APPLICATION OF GENERALIZED AI TO NLSY DATA

The first step in applying generalized AI to conventional quantitative evidence is to reconceptualize the dependent variable. Instead of being viewed as a raw quantity that simply varies across cases, the dependent variable must be reformulated as one or more qualitative outcomes. Fortunately, this focus on qualitative outcomes is consistent with the logic of the calibration procedure used to create fuzzy sets from conventional ratio- and interval-scale variables. To create a fuzzy set, the researcher specifies numerical values for the two main qualitative breakpoints—the threshold for full membership in the set and the threshold for full non-membership.[4] For example, the calibration of membership in the set of respondents in poverty, described above, uses an income-to-poverty ratio of 1.0 as the threshold for full membership in the set. Respondents with a ratio of 1.0 or less are classified as in poverty. Likewise, the qualitative threshold for non-membership in the set is an income-to-poverty ratio of 3.0. Respondents with a ratio of 3.0 or greater are classified as fully out of poverty. Thus, there is a direct link between generalized AI's focus on qualitative outcomes and the interpretive work that is central to the construction and calibration of fuzzy sets.

From the perspective of generalized AI, there are two key questions addressed by the analysis: (1) What causally relevant conditions are shared by respondents with full membership in the set in poverty? (2) What causally relevant conditions are shared by respondents with full non-membership in this set? Note that these two analyses are independent of each other. In other words, the "negative" cases (i.e., those with non-membership in the outcome set) do not serve as analytic foils for the examination of the positive cases, as they do in the two quantitative analyses presented above. Rather, these "negative" cases are accorded equal analytic attention and are treated as instances of a separate outcome. This feature of generalized AI contrasts sharply with the two quantitative analyses.

The causally relevant conditions under consideration are the four fuzzy sets used in the OLS regression analysis: *favorable domestic situation, not-low-income parents, educated respondent*, and *not-low test score*. Note, however, that it is the absence (or negation) of these conditions that should be linked to membership in the set of respondents in poverty, while their presence should be linked to non-membership in this set. In other words, the interpretive inferences (see chapter 6) that shape the coding of conditions in the two truth tables are opposite.

Table 9-3 presents the results of the application of generalized AI to the set of respondents in poverty (outcome set membership ≥0.95). There are three main steps. First, respondents are sorted into truth table rows based on their profiles. Membership scores greater than 0.5 (the crossover point) are treated as present (1); membership scores less than 0.5 are treated as absent (0). For example, the first row of the table summarizes the eighty-one respondents in poverty who have less than 0.5 membership in three conditions (not-low parental income, not-low AFQT scores, and favorable domestic situation), and greater than 0.5 membership in one—the set of educated respondents. Second, the four conditions are transformed from "present versus absent" codings (panel A) into "contributing versus irrelevant" codings (panel B). The revised codings are based on substantive and theoretical knowledge. For example, the absence of a *favorable domestic situation* is clearly linked to poverty, while its presence is not. Dashes are used in panel B to indicate irrelevance (see chapter 6). Third, low-frequency combinations ($N <$ 20) have been dropped from the table, which is motivated by the focus on the most widely shared combinations of contributing conditions (i.e., "modal configurations"). The three listed rows together embrace 67 percent of the respondents experiencing poverty.

The next step is to simplify the panel B results. In fact, the first and second rows (~nlpinc•~nlafqt•~fdomsit and ~nlpinc•~educ•~nlafqt•~fdomsit) are both

TABLE 9-3 Conditions linked to poverty (frequency cutoff: $N \geq 20$)

| Panel A | | | | |
|---|---|---|---|---|
| Not-low parental income (nlpinc) | Educated (educ) | Not-low AFQT (nlafqt) | Favorable domestic situation (fdomsit) | *N* |
| 0 | 1 | 0 | 0 | 81 |
| 0 | 0 | 0 | 0 | 59 |
| 0 | 1 | 1 | 0 | 23 |

| Panel B | | | | |
|---|---|---|---|---|
| Not-low parental income (nlpinc) | Educated (educ) | Not-low AFQT (nlafqt) | Favorable domestic situation (fdomsit) | *N* |
| 0 | – | 0 | 0 | 81 |
| 0 | 0 | 0 | 0 | 59 |
| 0 | – | – | 0 | 23 |

logical subsets of the third row (~nlpinc•~fdomsit). Thus, table 9-3 panel B reduces to a single modal configuration:

$$\sim\text{nlpinc}\bullet\sim\text{fdomsit} \rightarrow \text{poverty}$$

Here and below, an arrow indicates the superset/subset relation, a multiplication sign indicates the logical term *and* (combined conditions), and a tilde indicates *not* (set negation). In short, poverty is strongly linked to the combination of low parental income and an unfavorable domestic situation. The other two conditions, being educated and having not-low AFQT scores, are not consistently absent among respondents in poverty. It is important to note, in regard to low parental income and unfavorable domestic situation, that (1) they are conjunctural in the modal configuration, meaning that it is their combination that matters; and (2) both concern family characteristics, in the current household and in the family of origin.

The application of generalized AI to the avoidance of poverty (using the qualitative breakpoint of an income-to-poverty ratio of 3.0 or greater) follows the same general pattern. Table 9-4 panel A shows the high-frequency combinations among the respondents who avoid poverty, along with conventional presence/absence coding of their conditions. Panel B shows the results of the application of interpretive inferences to panel A. Conditions that do not contribute to the outcome are converted into dashes, indicating irrelevance. For example, respondents in the second row of panel A have unfavorable domestic situations, which is not linked to avoiding poverty. Accordingly, this condition is converted into a dash in panel B. Finally, this analysis, like the one preceding it, uses a frequency threshold of twenty respondents, and in so doing embraces 75 percent of the cases avoiding poverty.

TABLE 9-4 Conditions linked to avoiding poverty (frequency cutoff: $N \geq 20$)

Panel A

| Not-low parental income (nlpinc) | Educated (educ) | Not-low AFQT (nlafqt) | Favorable domestic situation (fdomsit) | *N* |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 54 |
| 1 | 1 | 1 | 0 | 45 |
| 0 | 1 | 1 | 1 | 33 |
| 0 | 1 | 1 | 0 | 22 |

Panel B

| Not-low parental income (nlpinc) | Educated (educ) | Not-low AFQT (nlafqt) | Favorable domestic situation (fdomsit) | *N* |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 54 |
| 1 | 1 | 1 | – | 45 |
| – | 1 | 1 | 1 | 33 |
| – | 1 | 1 | – | 22 |

Simplifying the results reported in table 9-4 panel B is straightforward. The first three rows are all logical subsets of the fourth row, which leads to a single modal configuration:

$$\text{educ}\bullet\text{nlaftq} \rightarrow \text{avoiding poverty}$$

In other words, avoiding poverty is strongly linked to the combination of being in the set of educated respondents and having not-low AFQT scores. The other two conditions, not-low-income parents and a favorable domestic situation, are not consistently present among respondents avoiding poverty. The results indicate that being educated and retaining school-based learning, the basis for a not-low AFQT score, together offer a degree of protection from poverty, regardless of domestic situation and parental income. The fact that they are conjunctural is consistent with the interpretation that one without the other would not be as effective.

These findings contrast dramatically with the generalized AI results for respondents in poverty. The conditions linked to being in poverty are having an unfavorable domestic situation and low-income parents; low education and low AFQT scores are not consistently linked to poverty. However, as just demonstrated, being educated and not having low AFQT scores are both strongly linked to avoiding poverty. These contrasting findings are not accessible using techniques that merge the two outcomes into a single analysis (i.e., almost all forms of conventional quantitative analysis; see Lieberson 1985). With generalized AI, it is not necessary to use "negative" cases as a foil for the positive cases. The two analyses are separate and equal.

## CLARIFYING THE TWO MODAL CONFIGURATIONS

It is important to point out that the two generalized AI solutions, while dramatically different in substance, overlap. This can be verified simply by deriving their intersection. If their intersection produces anything other than a null set, there is logical overlap:

$$(\sim\text{nlpinc}\bullet\sim\text{fdomsit}) \bullet (\text{educ}\bullet\text{nlaftq}) = \sim\text{nlpinc}\bullet\text{educ}\bullet\text{nlaftq}\bullet\sim\text{fdomsit}$$

The overlap occurs in part because the process of applying interpretive inferences eliminates non-contributing conditions on the basis of theoretical and/or substantive knowledge, not on the basis of empirical analysis. Overlap might be acceptable if there were no respondents in the intersection of the two modal configurations (i.e., in the four-way combination just derived). However, as is clear from tables 9-3 and 9-4, there is a nontrivial number of such respondents.

It is a straightforward matter to resolve the overlap, either by awarding it to one of the two modal configurations or by removing it from both, a more conservative strategy. For example, to assign the overlap to the modal configuration for poverty, it is necessary to remove the overlap from the modal configuration for avoiding poverty. The removal can be accomplished by intersecting the modal configuration for avoiding poverty with the *negation* of the modal configuration for poverty. This restricts the modal configuration for avoiding poverty to the combinations of conditions *not covered* by the modal configuration for poverty:

$$\begin{aligned}\text{avoiding poverty} &= (\text{educ}\bullet\text{nlaftq}) \bullet \sim(\sim\text{nlpinc}\bullet\sim\text{fdomsit})\\ &= (\text{educ}\bullet\text{nlaftq}) \bullet (\text{nlpinc} + \text{fdomsit})\\ &= \text{educ}\bullet\text{nlaftq}\bullet\text{nlpinc} + \text{educ}\bullet\text{nlaftq}\bullet\text{fdomsit}\end{aligned}$$

Here and below, a plus sign indicates the logical term *or* (alternate conditions or alternate combinations of conditions). Using De Morgan's theorem, the negation of (~nlpinc•~fdomsit) is (nlpinc + fdomsit). In essence, the scope of the modal configuration for avoiding poverty has been narrowed, while the scope of the modal configuration for poverty (~nlpinc•~fdomsit) is unchanged.

Alternatively, the overlap can be assigned to the modal configuration for avoiding poverty. In this scenario the overlap must be removed from the modal configuration for poverty, which can be accomplished by intersecting it with the negation of the modal configuration for avoiding poverty, as follows:

$$\begin{aligned}\text{in poverty} &= (\sim\text{nlpinc}\bullet\sim\text{fdomsit}) \bullet \sim(\text{educ}\bullet\text{nlaftq})\\ &= (\sim\text{nlpinc}\bullet\sim\text{fdomsit}) \bullet (\sim\text{educ} + \sim\text{nlafqt})\\ &= \sim\text{nlpinc}\bullet\sim\text{fdomsit}\bullet\sim\text{educ} + \sim\text{nlpinc}\bullet\sim\text{fdomsit}\bullet\sim\text{nlaftq}\end{aligned}$$

De Morgan's theorem is applied to educ•nlaftq to produce ~educ + ~nlafqt. The scope of the modal configuration for poverty has been narrowed, while the scope of the modal configuration for avoiding poverty (educ•nlafqt) is unchanged.

Finally, the most conservative strategy is to remove the overlap from both modal configurations, which yields

$$\text{in poverty} = \sim\text{nlpinc}\bullet\sim\text{fdomsit}\bullet\sim\text{educ} + \sim\text{nlpinc}\bullet\sim\text{fdomsit}\bullet\sim\text{nlaftq}$$
$$\text{avoiding poverty} = \text{educ}\bullet\text{nlaftq}\bullet\text{nlpinc} + \text{educ}\bullet\text{nlaftq}\bullet\text{fdomsit}$$

In this version of the results, not being educated or having low AFQT scores accompanies the core conditions linked to poverty (low-income parents combined with an unfavorable domestic situation), and not-low parental income or a favorable domestic situation accompanies the core conditions linked to avoiding poverty (being educated combined with having not-low test scores).

All three solutions to the problem of overlapping solutions are valid. The choice of strategies for addressing the overlap must be based on substantive and theoretical knowledge and interests. For example, if the researcher in this example wanted to emphasize the challenge of avoiding poverty for Black females, she might favor the more restrictive modal configuration for that outcome, and leave intact the less restrictive, two-condition configuration for being in poverty.

## DISCUSSION

The findings of the application of generalized AI to the NLSY data on Black females add depth to the results of the two regression analyses. In both regression analyses, independent variables are evaluated with respect to their separate contributions to the explanation of variation in the dependent variable. Variation in the dependent variable is key; without variation, there is nothing to explain. Both analyses confirm that the independent variables all have significant net effects on their respective dependent variables. The application of generalized AI, by contrast, separates the dependent variable into two qualitative outcomes and two separate analyses—full membership in the set of respondents in poverty and full membership in the set of respondents avoiding poverty. The conditions strongly linked to these two outcomes differ: having low-income parents combined with an unfavorable domestic situation is linked to being in poverty; being educated combined with having not-low test scores is linked to avoiding poverty. These are not simple net effects; both solutions involve combinations of conditions.

The two regression analyses and the generalized AI analysis provide convergent results. However, greater nuance is offered by the generalized AI application.[5] So-called independent variables with generic net effects are recast as modal configurations that differ by outcome. The application of generalized AI reveals subtle differences among the four causal conditions. The two conditions that are consistently linked to poverty are inconsistently linked to avoiding poverty, and vice versa. These contrasting effects are masked in the regression analyses.