Technical Appendix

This technical appendix summarizes the data and methods used to study the wage and benefits trajectories experienced by young adults from disadvantaged backgrounds found in the essay series, Diverging employment pathways among young adults.

DATA

National Longitudinal Survey of Youth 1997

We used the National Longitudinal Survey of Youth 1997 (NLSY97) for these analyses. The NLSY97 is an ongoing nationally representative survey of people born between 1980 and 1984 conducted by the Bureau of Labor Statistics. The NLSY97 collects detailed education, employment, and training information as respondents transition from school to work and adulthood. Interviews were conducted annually from 1997 to 2011 and biennially since then. The first survey took place in 1997 and 1998 when respondents were between the ages of 12 and 18 The cohort has been interviewed 18 times, and respondents were between 32 and 38 in the most recent interview round (2017/2018). The survey has a high retention rate compared to similar longitudinal studies: 8,984 individuals were interviewed in the first interview round, of whom 6,734 were interviewed in Round 18, the last round from which we pulled data.

TABLE 1

Demographic characteristics of the full NLSY97 sample (N=8,984)

	Unweighted N	Unweighted percentage	Weighted N	Weighted percentage
Gender				
Male	4,599	51.2%	9,945,147	51.3%
Female	4,385	48.8%	9,433,306	48.7%
Race/ethnicity				
White, non-Latino or Hispanic	4,406	49.2%	12,893,738	66.8%
Black, non-Latino or Hispanic	2,333	26.1%	2,980,438	15.4%
Latino or Hispanic	1,899	21.2%	2,485,722	12.9%
Other, non-Latino or Hispanic	318	3.6%	943,439	4.9%
Highest degree earned by age 27				
Drop out or GED	1,865	23.0%	3,497,083	20.2%
High school diploma	3,754	46.3%	7,721,281	44.6%
Post-secondary degree	2,491	30.7%	6,089,856	35.2%

Source: Child Trends analysis of NLSY97 data
Note 1: Missing data on a characteristic will result in Ns that do not add up to 8,984
Note 2: Percentages may not add to 100 due to rounding

Analytic sample

Our primary populations of interest are those who experienced disadvantage as teens, following our previous work.[1] We include in this sample those who meet any of the following criteria during the first interview round, or the second interview round if Round 1 data was missing:

- Family income equal to or less than 200% of the Federal Poverty Line (FPL)
- Neither parent has any post-secondary education
- Respondent was born to a teen mother (mother was aged 19 years old or younger when her first child was born)
- Family receives public assistance

Due to missing data on the disadvantage indicators in Rounds 1 and 2 of the survey, 327 survey respondents could not be assessed for inclusion in our sample. Results from a chi-square difference test indicate that individuals who were missing data on these variables (and thus excluded from our sample) are more likely to be white and Latino or Hispanic than the disadvantaged youth comprising our analytical sample, as well as more to likely have higher levels of education (see Table 2). For instance, 19.3% of the analytical sample has a post-secondary degree, whereas 28.1% of the sample excluded due to missing data has a post-secondary degree.

TABLE 2

Results of chi-square test of differences between the analytic sample and those missing data

	Sample excluded due to missing data (N= 327)		Analytic sample (N= 5,839)			
	N	%	N	%	Chi-square value	P-value
Gender						
Male	178	54.43%	2,978	51.00%	1.460	0.227
Female	149	45.57%	2,861	49.00%		
Race/ethnicity						
White, non-Hispanic	132	41.64%	2,268	38.90%	34.040	0.000
Black, non-Hispanic	88	27.76%	1,853	31.78%		
Latino or Hispanic	27	8.52%	169	2.90%		
Other, non-Hispanic	70	22.08%	1,540	26.42%		
Highest degree earned by age 27						
Drop out or GED	43	21.61%	1,346	29.89%	11.919	0.003
High school diploma	100	50.25%	2,286	50.77%		
Post-secondary degree	56	28.14%	871	19.34%		

Source: Child Trends analysis of NLSY97 data
Note 1: Ns and percentages in this table are unweighted
Note 2: Missing data on a characteristic will result in Ns that do not add up to 327 or 5,839
Note 3: Percentages may not add to 100 due to rounding
Note 4: Those missing data include those missing data on all of our disadvantaged criteria, earnings, or benefits

We provide demographic information on two populations in Table 3: those who experienced disadvantage in adolescence and those who did not experience disadvantage in adolescence. The disadvantaged sample is much more likely to be Black or Latino or Hispanic, have less parental wealth, and less education by age 27 than those who did not experience any of our adolescent disadvantage criteria.

TABLE 3

Demographic characteristics of disadvantaged and non-disadvantaged samples

	Disadvantaged sample		Non-disadvantaged sample	
	Number of observations	Weighted percentage	Number of observations	Weighted percentage
Disadvantaged criteria at first interview				
Family lives at or below 200% FPL	4,576	60.6%	2,175	0.0%
Neither parent has a post-secondary degree	5,627	75.2%	2,780	0.0%
Mother 19 years old or younger at first birth	5,387	40.1%	2,742	0.0%
Family receives public assistance	5,676	3.3%	2,763	0.0%
Female	5,839	48.7%	2,818	49.0%
Race/ethnicity				
White, non-Latino or Hispanic	5,830	58.3%	2,809	81.1%
Black, non-Latino or Hispanic	5,830	20.3%	2,809	7.5%
Latino or Hispanic	5,830	17.0%	2,809	6.4%
Other, non-Latino or Hispanic	5,830	4.4%	2,809	5.0%
Parental wealth 1997 ($2019)	4,486	$96,488	2,024	$304,910
Highest degree earned by age 27				
No degree	4,503	13.7%	2,120	1.8%
GED	4,503	14.2%	2,120	5.3%
High school	4,503	50.6%	2,120	35.5%
Associate degree	4,503	6.9%	2,120	7.2%
Bachelor's degree	4,503	12.4%	2,120	40.6%
Graduate degree	4,503	2.2%	2,120	9.4%

Source: Child Trends analysis of NLSY97 data
Note: Missing data on a characteristic will result in Ns that do not add up to 5,839

Data Transformation

Given that the NLSY97 is structured by survey round, we transformed the dataset so that variables could be examined by respondents' age at interview rather than by survey year. This transformation enabled us to examine how income and benefits change as youth age, and to observe how age salient employment, training, and educational experiences during key developmental periods were associated with youths' income and benefits.

MEASURES

Measurement of income and benefits

Using the transformed dataset, the analyses focus on patterns in respondents' annual earnings and benefits offered by their employers between ages 18 and age 31.

Annual income from earnings: Respondents reported their gross annual income from wages, salary, commissions, and/or tips from all jobs. Earnings income was adjusted via the Consumer Price Index (CPI) to the year 2019. Those not reporting any income were coded as having zero income. Respondents report their annual income from the previous year instead of their annual income in the year of the survey. To ensure that this variable aligned with all other variables computed at the time of the survey interview, the income variable was coded as income at the respondent's age during the year prior to the survey year.[2]

Fringe benefit index: The fringe benefits index measures whether a person received health insurance, retirement benefits, or paid time off from any of their employers at a given age. The NLSY97 gathers information about the fringe benefits offered to respondents from any job lasting 13 weeks or more (age 16 and above). Respondents were asked to provide this information for each week in the past year, which allowed us to first compute an index equal to 0, 1, 2 or 3, measuring whether the respondent received none, some, or all of the benefits from all jobs at all weeks for all ages available. Those not working or self-employed were coded as having zero benefits. Given that the Annual Income from Earnings is a yearly measure, we computed the mode of the weekly benefits index, to calculate the most common situation in terms of benefits received for each youth at each age.

METHODS

Trajectory Analysis

We employed trajectory analysis (see below) to identify patterns in respondents' annual earnings income and benefits offered by their employers between ages 18 and 31. We chose to end the analysis at age 31 since all youth had turned 31 in the last survey round. However, we cannot capture information for all ages for all youth because of attrition, and because in 2011 the NLSY97 started to interview youth every other year. As a consequence, some youth are not interviewed at certain ages (e.g., if a youth was 28 in 2011, for the next interview in 2013 he or she will be 30, so we do not have information for age 29). For our sample, the younger the respondent, the more age years they will miss in the data. For instance, for the youngest survey respondents, who were 26 in 2011, we do not have data at ages 27, 29, and 31, whereas for the oldest youth who were 31 in 2011 we have data at all ages from 18 to 31. The implication of this change in the timing of data collection is that the sample of older respondents gets overrepresented for later ages.

Trajectory analysis models identify a set of predicted earnings and benefits trajectories and the probabilities that individuals from the sample are a member of each of the trajectories.[3] This is a person-based analysis method, rather than a variable-based method. That is, the trajectory analysis identifies unique groups of respondents based on both the income and fringe benefit variables simultaneously. Respondents are assigned to the trajectory group that best approximates their annual earnings and benefits pattern using finite mixture models (FMM). Recently, many researchers have turned to finite mixture models to model change within people over time.[4] There are different types of FMM that can be applied to study group trajectories, but all of them identify latent trajectory groups, based on a series of observed and measured characteristics, and assign individuals to one of these groups. We used group-based trajectory modeling (GBTM), also known as latent class growth analysis.[5]

This method identified groups of individuals with similar earnings and benefits trajectories and assigned NLSY97 individuals to the group that best approximated their trajectory pattern. We ran this analysis for respondents who experienced disadvantage in adolescence. Because outcomes so often differ by race and gender, we also ran the same analysis separated by white respondents who experienced disadvantage as teens, Black respondents who experienced disadvantage as teens, Latino or Hispanic respondents who experienced disadvantage as teens, and male and female respondents who experienced disadvantage as teens.

Given that data from all rounds were combined, sampling weights from Round 1 were used to adjust for over-sampling.

Group-based trajectory models can fit any number of groups and functional forms (e.g., cubic, quadratic), and we used an empirical approach to select the models with the best statistical fit. The selection of the number of groups and functional forms is based on three criteria, following Nagin (2009): the best model fit (Bayesian Information Criterion, or BIC), the averaged predicted probabilities of group membership (at least 70%), and the size of the smallest group (minimum of 5% of the sample).[6] We first tested different functional forms (i.e., linear, quadratic, and cubic terms) and found that models with cubic functional forms produced a better fit. We then tested cubic models with two to eight groups and found that four-group models with cubic functional form had the best results (see Table 4).[7]

TABLE 4

Model fit statistics for group-based trajectory analysis with 2-8 groups

	Average probability assigned to each group										
Number of groups	1	2	3	4	5	6	7	8	BIC (persons)	BIC (persons-years)	% in smallest group
2	95.5%	95.0%							-874502.4	-874475.1	46.0%
3	94.3%	89.7%	93.1%						-866812.2	-866771.3	20.7%
4	94.2%	81.8%	90.2%	92.8%					-862620.8	-862566.2	14.2%
5	92.3%	84.9%	89.9%	83.6%	94.1%				-858951.6	-858883.4	9.1%
6	91.1%	76.7%	88.4%	86.4%	86.3%	93.4%			-856844.2	-856762.4	7.1%
7	90.1%	79.9%	84.9%	87.6%	76.6%	87.4%	94.0%		-855454.2	-855358.7	4.9%
8	90.2%	86.5%	95.5%	86.0%	83.0%	76.0%	76.2%	89.0%	-854318.6	-854209.5	3.3%

Source: Child Trends analysis of NLSY97 data
Note: Models with 5, 7, and 8 groups had a highly singular variance matrix

Descriptive analysis

Once we identified these clusters, or trajectory groups, we conducted descriptive analyses to determine young adults' characteristics for each of the wage and benefits trajectories. Table 5 explains the coding of the variables used in the descriptive analysis. For the descriptive analysis, we chose to report the variables around key developmental periods.

TABLE 5

Coding of variables used in the descriptive statistics and multivariate models

	Coding of the variable	Ages used in the descriptive analysis	Ages used in the multivariate analysis
BASIC DEMOGRAHICS			
Gender	Gender was dummy coded (Female = 1)	N/A	N/A
Race/ethnicity	Race/ethnicity was coded as white non-Latino or Hispanic, Black non- Latino or Hispanic, other non-Latino Hispanic, and Latino or Hispanic. For the multivariate analysis, the omitted category was Black.	N/A	N/A
ASVAB score	Scores from the Armed Services Vocational Aptitude Battery (ASVAB) in 1999 were used to measure respondents' cognitive ability. The variable accounts for participant scores on four of the ASVAB subtests (Mathematical Knowledge, Arithmetic Reasoning, Word Knowledge, and Paragraph Comprehension). This variable was standardized for the multivariate analysis.	Not included in descriptive data	N/A
Parental wealth in 1997	Parental wealth in 2019 dollars. For the descriptive analysis it was divided into 5 quantiles (10%, 25%, 50%, 75%, and 90%) and the mean for each quantile is represented. For the multivariate analysis the actual value was used in dollars.	N/A	N/A
Ever incarcerated	A cross-round variable in the NLSY97 presents the first age at which a respondent was incarcerated, if ever. Using this, a dichotomous variable was created to indicate whether the respondent had ever been incarcerated as of age 23, 27, or 31. The multivariate analysis includes a dichotomous variable that captures if respondent was ever incarcerated by age 27.	By age 23, 27, or 31	By age 27
Age at time of first interview	Age at time of first interview	Not included in descriptive data	N/A
EDUCATION			
Educational attainment	The date the respondent received a degree (GED, high school diploma, associate degree, bachelor's degree, or graduate degree) was used to identify the highest degree attained from ages 18 to 31. These variables were then recoded into five dummy coded variables: Less than a high school diploma, GED, high school diploma, associate degree, bachelor's degree, and graduate degree. In the multivariate analysis, "No degree" was used as the reference group.	Computed for each age from 18 to 31	Age 27. If missing closest age with data between 25 to 29.
Enrollment status	Variables were dummy coded to indicate in what level of education, if any, a respondent was enrolled at the time of the interview. Five variables were created for ages 20 to 24: Not enrolled in school; enrolled in high school; enrolled in 2-year college; enrolled in 4-year college; enrolled in graduate program.	Computed for each age from 20 to 24	Not used in the multivariate analysis
Mixing school and work	To determine whether a respondent mixed school and work, we determined whether the individual worked during the school year (roughly determined to be weeks 1-23 and 36-53 of the year), based on weekly work arrays, during ages 16 to 23. We determined if a respondent was enrolled in some type of education that year. Combining these two factors, we created the following 5 categories: Disconnected from school and work; did not work during school year, enrolled; worked fewer than 20 weeks during school year, enrolled; worked at least 20 weeks during school year, enrolled; worked during school year, not enrolled.	Computed for each age from 16 to 23	Not used in the multivariate analysis

TABLE 5 CONTINUED

Coding of variables used in the descriptive statistics and multivariate models

	Coding of the variable	Ages used in the descriptive analysis	Ages used in the multivariate analysis
FAMILY CHARACTERISTICS			
Married or cohabiting	We created variables using a self-report item to indicate whether the respondent was married and/or cohabitating at different ages.	Ages 23 and 27	Age 23
Has a child	We created variables that identify individuals who reported having their first child by ages 19, 23, and 27.	Ages 19, 23, and 27	By age 19 and between ages 20 and 27. Age 27 only identifies those who first became a parent after age 19 and by age 27.
Has had a child since last interview	We created a variable using a self-report item to indicate whether the respondent had a baby since the date of the previous interview.	Computed for each age from 19 to 31	Not used in the multivariate analysis
Ever received SNAP	Respondents were asked to report if they or their spouse had ever received food stamps in 1997 and asked if they or their spouse had received food stamps since the prior interview every subsequent year through 2005. We combined these variables to create a dichotomous variable to indicate whether a respondent had ever received food stamps by age 25.	By age 25	Not used in the multivariate analysis
Geography of residence	We created 6 dichotomous variables based on metro area (metropolitan statistical area) residence and one based on urban (as opposed to rural) residence for respondents for each year from age 18 onwards: lives in a metro area; lives in a metro area, city center; lives in a metro area, not sure; lives in a metro area, outside central city; lives in a non-metro area; lives outside of the United States; lives in an urban area.	Computed for each age from 18 to 31	Not used in the multivariate analysis
WEALTH, INCOME, AND POVERTY			
Poverty level	Family income relative to the federal poverty line. Below or equal to100% FPL; below or equal to 150% FPL; below or equal to 200% FPL; below or equal to 400% FPL.	Computed for each age from 18 to 31	Not used in the multivariate analysis
Household wealth	The NLSY97 dataset contains created variables for household net worth at 5-year intervals starting at age 20. Household wealth was adjusted for CPI to the year 2019 for ages 20, 25, and 30, and divided into the following quantiles: 10%, 25%, 50%, 75%, 90%.	Ages 20, 25, and 30	Not used in the multivariate analysis
Partner income	Respondents with partners were asked to report their partner's gross income from wages, salary, commissions, or tips from all jobs. Partner income was adjusted for CPI to the year 2019 and divided into the following quantiles: 10%, 25%, 50%, 75%, 90%.	Ages 20, 25, and 30	Not used in the multivariate analysis
TRAINING EXPERIENCES			
Ever had an internship	We created dichotomous variables to indicate whether the respondent self-reported having an internship between the ages of 16 to 25. These variables were based upon the employment history data, which indicated whether each job held was an internship.	Ages 16 to 25 combined	Not used in the multivariate analysis

TABLE 5 CONTINUED

Coding of variables used in the descriptive statistics and multivariate models

	Coding of the variable	Ages used in the descriptive analysis	Ages used in the multivariate analysis
TRAINING EXPERIENCES			
Participated in relationship-based CTE	During rounds in which respondents were enrolled in secondary school, respondents were asked to report whether they participated in a school-to-work training program. There were seven types of programs identified in the NLSY97 data. To create a more parsimonious measure of participation in school-to-work training programs, we created two summary variables. The first summary variable was a dummy coded variable that indicated whether respondents participated in any relationship-based training program between ages 16 to 19. School-to-work programs categorized as relationship-based include cooperative education, internship/apprenticeship, and mentoring. Relationship-based and non-relationship-based programs were distinguished given the importance of positive relationships in youth-focused programs.	Ages 16 to 19 combined	Ages 16 to 19 combined
Participated in other CTE	During rounds in which respondents were enrolled in secondary school, respondents were asked to report whether they participated in a school-to-work training program. This is a dummy coded variable that indicates whether respondents participated in any non-relationship-based training programs between ages 16 to 19. Non-relationship (other) school-to-work programs include career major, job shadowing, school-sponsored enterprise, and tech prep. Relationship-based and non-relationship-based programs were distinguished given the importance of positive relationships in youth-focused training programs.	Ages 16 to 19 combined	Not used in the multivariate analysis
Participated in a training program	This variable is based on a self-report item that asks whether the respondent participated in any of the following training programs: adult basic education (pre-GED); apprenticeship; business or secretarial school; community or junior college; correspondence course; formal company training run by employer; GED program; government training; nursing school (LPN or RN); seminar or training program outside of work; seminar or training program at work run by someone other than employer; vocational rehabilitation center; vocational, technical, or trade school; and K-12 school-based training, including ROTC (for Round 1 only).	Between ages 16 to 19 combined, 20 to 23 combined, and 24 to 27 combined	Between ages 20 to 23 combined and 24 to 27 combined
WORK PATTERNS AND EMPLOYER CHARACTERISTICS			
Ever in the military	NLSY97 employment arrays were used to determine if a respondent worked in the military at any time between ages 18 to 31.	Between ages 18 to 31 combined	Between ages 18 to 31 combined
Employer type	Dichotomous variables were created based on NLSY97 categories of employer type. Employer type was determined only for the main job for each year. The NLSY97 defines the main job of a respondent as the current or most recent job as of the interview date. The 4 employer type variables are: Main employer is government; main employer is non-profit; main employer is private; and worked without pay in a family business.	Computed for each age from 18 to 31	Not used in the multivariate analysis
Number of weeks spent self-employed	Using the employment arrays, we determined in which weeks a respondent reported working their main job if that job was classified as self-employment. Those weeks were then summed across ages 16 to 19 and 20 to 23. The average number of weeks for each age range is reported in the descriptive data.	Between ages 16 to 19 combined and 20 to 23 combined	Between ages 20 to 23 combined

TABLE 5 CONTINUED

Coding of variables used in the descriptive statistics and multivariate models

	Coding of the variable	Ages used in the descriptive analysis	Ages used in the multivariate analysis
WORK PATTERNS AND EMPLOYER CHARACTERISTICS			
Limited in work due to chronic illness	Starting in 2007, respondents were asked to report if they were limited in type or amount of work for health reasons. We created dummy variables for ages 23 (the first age at which all respondents would have been asked this question) to 27.	Age 23 and 27	Age 27. If missing, closest age with data between 23 to 26.
Held more than one job	We determined the number of jobs a respondent worked each week based on the employment arrays. We then created 6 variables based on how many weeks out of a given year the respondent worked more than one job: Did not hold more than one job at a time; had more than one job 1-4 weeks; had more than one job 5-12 weeks; had more than one job 13-24 weeks; had more than one job 25-49 weeks; had more than one job 50+ weeks.	Computed for each age from 18 to 31	Not used in the multivariate analysis
Worked 50+ hours per week (overwork)	Using the employment arrays, we determined the average hours worked per week in a given job and the weeks in a given year in which a respondent worked. By summing hours across jobs, we determined the number of weeks in each year in which a respondent would have worked 50 or more hours. We then created three dichotomous variables based on the number of weeks in each year that a respondent would have overworked: Did not overwork; overworked 1-24 weeks; overworked 25+ weeks.	Computed for each age from 18 to 31	Not used in the multivariate analysis
Full-time work	Using the employment arrays, we determined the average hours worked per week in a given job and the weeks in a given year in which a respondent worked. By summing hours across jobs, we determined the number of weeks in each year in which a respondent would have worked 35 or more hours. We then created six dichotomous variables based on the number of weeks in each year that a respondent would have worked full time: Did not work full time; worked full time 1-4 weeks; worked full time 5-12 weeks; worked full time 13-24 weeks; worked full time 25-49 weeks; worked full time 50+ weeks.	Computed for each age from 18 to 31	Not used in the multivariate analysis
Unemployment	Using the employment arrays, we summed the number of weeks a respondent was employed between ages 16 to 19 and ages 20 to 23. Respondents were considered unemployed if they reported actively searching for work during a within-job gap or between-jobs gap. We created six dichotomous variables based on the number of weeks a respondent spent unemployed during each age range: Spent 0 weeks unemployed; spent 1-8 weeks unemployed; spent 9-16 weeks unemployed; spent 17-24 weeks unemployed; spent 25-52 weeks unemployed; spent 53+ weeks unemployed.	Between ages 16 to 19 combined and 20 to 23 combined	6 months or more unemployed between ages 16-19 and 20-23
Industries	The NLSY97 classifies respondents' self-employed and employee-type jobs using respondents' verbatim descriptions and 4-digit industry 2002 Census codes. We translate these codes to NAICS codes and provide industry information at the following level of aggregation: Agriculture, forestry and fisheries (NAICS: 11); Mining (NAICS 21); Utilities (NAICS 22); Construction (NAICS 23); Manufacturing (NAICS 31-33); Wholesale trade (NAICS 42); (NAICS 44-45); Arts, entertainment, and recreation services (NAICS 71-72); Transportation and warehousing (NAICS 48-49); Information and communication (NAICS 51); Finance, insurance, and real estate (NAICS 52-53); Professional and related services (NAICS 54-56); Educational, health, and social services (NAICS 61-62); Other services (NAICS 81); Public administration (NAICS 92)	At ages 20, 25, and 30	Not used in the multivariate analysis

TABLE 5 CONTINUED

Coding of variables used in the descriptive statistics and multivariate models

	Coding of the variable	Ages used in the descriptive analysis	Ages used in the multivariate analysis
WORK PATTERNS AND EMPLOYER CHARACTERISTICS			
Occupations	The NLSY97 classifies respondents' self-employed and employee-type jobs using respondents' verbatim descriptions and 3-digit occupation 2002 Census codes. We provide occupation information for respondents' main jobs. We also created five major occupations groups which we use in the multivariate analysis: Management/professional (Census codes 0010-3540); Service (Census codes 3600-4650); Sales and office (Census codes 4700-5930); Construction, production, transportation, and other (Census codes 6000-9990); No occupation (not employed). For the multivariate analysis, the omitted category was Service.	At ages 20, 25, and 30	Age 25

Source: Child Trends analysis of NLSY97 data

Multinomial logistic risk models

Once the group-based trajectory modeling identified four groups, we conducted multinomial logistic risk models to examine the predictors of each trajectory, including measures of each individual's background characteristics (such as education, marital status, gender, race, and incarceration) to consider factors that may explain individuals' membership in the wage and benefits trajectory clusters. Using the disadvantaged sample, we ran multinomial regressions predicting membership to each group.

Table 5 describes all variables used in the multivariate analyses, including variables that appear in the final model. The predictors are grouped by the category in which they fall: Basic demographics; Education; Family characteristics; Wealth, income, and poverty; Training experiences; and Work patterns and employer characteristics. All variables were developed collaboratively by Child Trends and Brookings. Given that respondents' experiences may be qualitatively different depending on a respondents' age, variables were coded differently for the developmental period in which they occurred (i.e., ages 16 to 19, ages 20 to 24).

Due to attrition and the change of the NLSY97 bi-yearly data collection that happened in 2011, several predictors had significant levels of missing data. The team considered using multiple imputation methods, but, to the best of our knowledge, there is no research examining imputation techniques in contexts where the results of the group-based trajectory models are used as dependent variables. Previous research that conducted similar analyses with large panels has relied on listwise deletion when there is attrition or item-level missingness on the predictors.[8] We followed a similar approach but ran analysis with a different set of predictors to evaluate whether the results are sensitive to changes in the sample size. Also, to increase our sample size and avoid dropping observations, we recoded some of the predictors by combining several ages of data, to account for the missing observations. For instance, instead of coding a dummy for whether a respondent had limitations in work due to chronic illness at age 27, we coded a dummy for whether the respondent reported limitations at any time between ages 24 to 27. After combining variables we used listwise deletion for cases where there were still missing observations. Finally, all the analyses are weighted to account for attrition. Due to missing data in the independent variables, 3,923 of people who experienced adolescent disadvantage were dropped from the multivariate analysis in our preferred model. The results of a chi-square difference test indicate significant differences between those included and excluded from the multivariate analysis. Those included in the analysis are more likely to be female, white, and have higher levels of education (see Table 6). We estimated additional versions of the model with fewer variables and hence more observations to test whether our results are robust and found similar results in terms of direction and significance of the parameters (see below).

TABLE 6

Results of chi-square test of differences between the multivariate sample and analytic sample excluded from multivariate analysis

	Analytic sample included in multivariate analysis (N= 1,916)		Analytic sample excluded from multivariate analysis (N= 3,923)			
	N	%	N	%	Chi-square value	P-value
Gender						
Male	915	47.76%	2,063	52.59%	12.025	0.001
Female	1,001	52.24%	1,860	47.41%		
Race/ethnicity						
White, non-Hispanic	808	42.17%	1,460	37.30%	148.043	0.000
Black, non-Hispanic	593	30.95%	1,260	32.19%		
Latino or Hispanic	49	2.56%	120	3.07%		
Other, non-Hispanic	466	24.32%	1,074	27.44%		
Highest degree earned by age 27						
Drop out or GED	388	20.55%	958	36.63%	14.347	0.002
High school diploma	1,042	55.19%	1,244	47.57%		
Post-secondary degree	458	24.26%	413	15.79%		

Source: Child Trends analysis of NLSY97 data
Note 1: Ns and percentages in this Table are unweighted
Note 2: Missing data on a characteristic will result in Ns that do not add up to 1,916 or 3,923
Note 3: Percentages may not add to 100 due to rounding
Note 4: Those missing data include those missing data on any of our disadvantaged criteria, earnings, or benefits

Tables 7a, 7b, 7c, and 7d report average marginal effects from the results of four multinomial logistic risk models with increasing numbers of variables for each of the four groups that we identified. As the number of variables increases, the number of missing observations also increases. Model 1 uses a first set of sociodemographic characteristics (e.g., gender, race), as well as a few variables that are strong predictors (e.g., being in the military). Model 2 adds additional covariates, but excludes parental wealth, given the high levels of missingness in this variable. Model 3 adds parental wealth, and Model 4 includes interactions between parental wealth and race. As shown in the tables, most of the results did not change in terms of direction and significance as we added more covariates to the analysis. Model 3 is our preferred model because it includes most of the variables, excluding interactions between race and wealth that are not significant. We report the results of this model in our findings. Table 8 presents the results of Model 3 for each group, side by side.

TABLE 7A

Average marginal effects of multinomial models, Group 1

Dependent variable: Membership in Group 1

Variable name (reference group)	GROUP 1			
	Model 1	Model 2	Model 3	Model 4
Female	0.1135***	0.0554***	0.0568***	0.0564***
	(0.0148)	(0.0138)	(0.0167)	(0.0165)
Race/ethnicity (Black, non-Latino or Hispanic)				
White, non-Latino or Hispanic	-0.1049***	-0.0232*	-0.0187	-0.0329*
	(0.0151)	(0.0136)	(0.0162)	(0.0182)
Latino or Hispanic	-0.1004***	-0.0275*	-0.0310*	-0.0401**
	(0.0180)	(0.0144)	(0.0171)	(0.0194)
Other, non-Latino or Hispanic	-0.0575	0.0616*	0.0771*	0.0774
	(0.0404)	(0.0362)	(0.0442)	(0.0491)
Age at time of first interview	-0.0007	0.0005	0.0018	0.0017
	(0.0050)	(0.0040)	(0.0050)	(0.0050)
Family patterns				
First became parent at or before age 19	0.1737***	0.0425***	0.0407**	0.0408**
	(0.0183)	(0.0158)	(0.0188)	(0.0189)
First became a parent between ages 20-27	0.0561***	0.0142	0.0068	0.0070
	(0.0166)	(0.0135)	(0.0168)	(0.0168)
Married or cohabitating at 23	-0.0149	-0.0146	-0.0081	-0.0085
	(0.0179)	(0.0150)	(0.0180)	(0.0180)
Ever incarcerated by age 27	0.2440***	0.0680***	0.0730***	0.0733***
	(0.0228)	(0.0185)	(0.0225)	(0.0224)
Ever in the military	-0.3649***	-0.2386***	-0.2059***	-0.2053***
	(0.0671)	(0.0609)	(0.0636)	(0.0633)
Employment and training				
Weeks self-employed 20-23		0.0010***	0.0012***	0.0012***
		(0.0002)	(0.0002)	(0.0002)
Unemployed for at least 6 months, cumulative				
Ages 16-19		0.0207	0.0239	0.0232
		(0.0176)	(0.0206)	(0.0204)
Ages 20-23		0.0959***	0.0865***	0.0855***
		(0.0205)	(0.0234)	(0.0232)
Participated in a relationship-based career and technical education program		-0.0017	0.0009	0.0008
		(0.0126)	(0.0152)	(0.0152)

TABLE 7A CONTINUED

Average marginal effects of multinomial models, Group 1

Dependent variable: Membership in Group 1

Variable name (reference group)	GROUP 1			
	Model 1	Model 2	Model 3	Model 4
Participated in a training program				
Ages 20-23		-0.0049	-0.0037	-0.0051
		(0.0125)	(0.0152)	(0.0151)
Ages 24-27		-0.0513***	-0.0563***	-0.0565***
		(0.0138)	(0.0164)	(0.0163)
Occupation held at age 25 (Service)				
No occupation		0.1975***	0.2129***	0.2112***
		(0.0266)	(0.0318)	(0.0313)
Management/professional occupation		-0.0760***	-0.0611**	-0.0609**
		(0.0230)	(0.0269)	(0.0270)
Sales/office occupation		-0.0380**	-0.0352**	-0.0358**
		(0.0148)	(0.0179)	(0.0179)
Construction/production/transportation/other occupation		-0.0637***	-0.0652***	-0.0653***
		(0.0184)	(0.0213)	(0.0212)
Job is covered by a union contract at age 25		-0.0861***	-0.0786***	-0.0794***
		(0.0252)	(0.0291)	(0.0289)
Educational attainment at age 27 (No high school diploma)				
High school diploma		-0.0684***	-0.0728***	-0.0717***
		(0.0178)	(0.0213)	(0.0212)
GED		-0.0197	-0.0217	-0.0223
		(0.0206)	(0.0249)	(0.0248)
Associate degree		-0.1548***	-0.1480***	-0.1468***
		(0.0387)	(0.0474)	(0.0473)
Bachelor's degree		-0.1381***	-0.1254***	-0.1241***
		(0.0361)	(0.0422)	(0.0424)
Graduate degree		-0.1553**	-0.1433**	-0.1410**
		(0.0608)	(0.0696)	(0.0679)
Has a work-limiting health condition at age 27		0.1558***	0.1724***	0.1716***
		(0.0206)	(0.0238)	(0.0236)
Standardized ASVAB		-0.0314***	-0.0394***	-0.0386***
		(0.0087)	(0.0105)	(0.0102)
Parental wealth in 1997			-0.1010	-0.4594*
			(0.0646)	(0.2636)

TABLE 7A CONTINUED

Average marginal effects of multinomial models, Group 1

Dependent variable: Membership in Group 1

Variable name (reference group)	GROUP 1			
	Model 1	Model 2	Model 3	Model 4
Wealth and race/ethnicity interactions (Wealth x Black, non-Latino or Hispanic)				
Wealth x white, non-Latino or Hispanic				0.3928
				(0.2715)
Wealth x Latino or Hispanic				0.3164
				(0.2830)
Wealth x other, non-Latino or Hispanic				0.0444
				(0.3940)
Observations	3,674	2,348	1,916	1,916

Source: Child Trends analysis of NLSY97 data
Note 1: Standard errors in parentheses
Note 2: ***p < 0.01, **p <0.05, *p < 0.1
Note 3: Model 3 is preferred model

TABLE 7B

Average marginal effects of multinomial models, Group 2

Dependent variable: Membership in Group 2

Variable name (reference group)	GROUP 2			
	Model 1	Model 2	Model 3	Model 4
Female	0.0445**	0.1688***	0.1473***	0.1473***
	(0.0198)	(0.0322)	(0.0354)	(0.0354)
Race/ethnicity (Black, non-Latino or Hispanic)				
White, non-Latino or Hispanic	-0.0781***	-0.0816**	-0.0898**	-0.0776*
	(0.0217)	(0.0333)	(0.0375)	(0.0429)
Latino or Hispanic	-0.0744***	-0.0982***	-0.0996**	-0.0903*
	(0.0241)	(0.0359)	(0.0396)	(0.0475)
Other, non-Latino or Hispanic	-0.1405**	-0.1196	-0.0735	-0.1171
	(0.0551)	(0.0854)	(0.0976)	(0.1082)
Age at time of first interview	-0.0165**	-0.0202**	-0.0148	-0.0144
	(0.0066)	(0.0093)	(0.0104)	(0.0104)
Family patterns				
First became parent at or before age 19	0.0955***	0.0651	0.0749	0.0751
	(0.0278)	(0.0428)	(0.0470)	(0.0474)
First became a parent between ages 20-27	0.0398*	0.0053	0.0265	0.0262
	(0.0215)	(0.0303)	(0.0336)	(0.0337)

TABLE 7B CONTINUED

Average marginal effects of multinomial models, Group 2

Dependent variable: Membership in Group 2

Variable name (reference group)	GROUP 2			
	Model 1	Model 2	Model 3	Model 4
Family patterns				
Married or cohabitating at 23	-0.0183	-0.0370	-0.0598	-0.0591
	(0.0245)	(0.0347)	(0.0386)	(0.0386)
Ever incarcerated by age 27	0.2122***	0.3113***	0.3364***	0.3339***
	(0.0371)	(0.0578)	(0.0635)	(0.0634)
Ever in the military	0.0154	-0.0154	-0.0602	-0.0599
	(0.0546)	(0.0759)	(0.0894)	(0.0893)
Employment and training				
Weeks self-employed 20-23		0.0012*	0.0012	0.0013*
		(0.0007)	(0.0008)	(0.0008)
Unemployed for at least 6 months, cumulative				
Ages 16-19		0.1145**	0.1409**	0.1424**
		(0.0582)	(0.0620)	(0.0618)
Ages 20-23		0.2952***	0.2242***	0.2209***
		(0.0680)	(0.0715)	(0.0710)
Participated in a relationship-based career and technical education program		-0.0151	-0.0164	-0.0150
		(0.0297)	(0.0330)	(0.0331)
Participated in a training program				
Ages 20-23		-0.0132	-0.0006	0.0001
		(0.0298)	(0.0330)	(0.0332)
Ages 24-27		-0.0951***	-0.0830**	-0.0828**
		(0.0295)	(0.0329)	(0.0330)
Occupation held at age 25 (Service)				
No occupation		0.3138***	0.3802***	0.3815***
		(0.0869)	(0.1030)	(0.1029)
Management/professional occupation		-0.2051***	-0.1368***	-0.1383***
		(0.0465)	(0.0507)	(0.0509)
Sales/office occupation		-0.0981***	-0.0985**	-0.0974**
		(0.0373)	(0.0407)	(0.0409)
Construction/production/transportation/other occupation		-0.0822**	-0.0580	-0.0587
		(0.0413)	(0.0457)	(0.0458)
Job is covered by a union contract at age 25		-0.0959**	-0.0780	-0.0765

TABLE 7B CONTINUED

Average marginal effects of multinomial models, Group 2

Dependent variable: Membership in Group 2

Variable name (reference group)	GROUP 2			
	Model 1	Model 2	Model 3	Model 4
Educational attainment at age 27 (No high school diploma)				
High school diploma		-0.1486***	-0.1612***	-0.1631***
		(0.0536)	(0.0591)	(0.0590)
GED		-0.0120	0.0056	0.0056
		(0.0657)	(0.0734)	(0.0734)
Associate degree		-0.1226*	-0.1438*	-0.1456*
		(0.0677)	(0.0764)	(0.0762)
Bachelor's degree		-0.1943***	-0.2019***	-0.2023***
		(0.0696)	(0.0776)	(0.0775)
Graduate degree		-0.3911***	-0.4269***	-0.4377***
		(0.1338)	(0.1481)	(0.1367)
Has a work-limiting health condition at age 27		0.2042**	0.1669*	0.1667*
		(0.0796)	(0.0911)	(0.0910)
Standardized ASVAB		-0.0101	0.0062	0.0050
		(0.0176)	(0.0196)	(0.0196)
Parental wealth in 1997			0.0407	0.2867
			(0.1075)	(0.3645)
Wealth and race/ethnicity interactions (Wealth x Black, non-Latino or Hispanic)				
Wealth x white, non-Latino or Hispanic				-0.2907
				(0.3822)
Wealth x Latino or Hispanic				-0.2759
				(0.5400)
Wealth x other, non-Latino or Hispanic				0.2473
				(0.6236)
Observations	3,674	2,348	1,916	1,916

Source: Child Trends analysis of NLSY97 data
Note 1: Standard errors in parentheses
Note 2: ***p < 0.01, **p <0.05, *p < 0.1
Note 3: Model 3 is preferred model

TABLE 7C

Average marginal effects of multinomial models, Group 3

Dependent variable: Membership in Group 3

Variable name (reference group)	GROUP 3			
	Model 1	Model 2	Model 3	Model 4
Female	-0.0780***	-0.1509***	-0.1786***	-0.1772***
	(0.0198)	(0.0324)	(0.0363)	(0.0363)
Race/ethnicity (Black, non-Latino or Hispanic)				
White, non-Latino or Hispanic	0.1277***	0.0752**	0.0954**	0.0925**
	(0.0227)	(0.0339)	(0.0388)	(0.0445)
Latino or Hispanic	0.1486***	0.1103***	0.1205***	0.1167**
	(0.0247)	(0.0360)	(0.0405)	(0.0483)
Other, non-Latino or Hispanic	0.1373***	0.0099	-0.0224	0.0131
	(0.0532)	(0.0810)	(0.0947)	(0.1018)
Age at time of first interview	0.0132**	0.0162*	0.0122	0.0118
	(0.0065)	(0.0093)	(0.0106)	(0.0106)
Family patterns				
First became parent at or before age 19	-0.1969***	-0.0951**	-0.1073**	-0.1074**
	(0.0300)	(0.0447)	(0.0489)	(0.0493)
First became a parent between ages 20-27	-0.0696***	-0.0150	-0.0334	-0.0333
	(0.0214)	(0.0302)	(0.0340)	(0.0340)
Married or cohabitating at 23	0.0048	0.0318	0.0617	0.0612
	(0.0242)	(0.0344)	(0.0390)	(0.0390)
Ever incarcerated by age 27	-0.3170***	-0.2686***	-0.1851***	-0.1849***
	(0.0430)	(0.0635)	(0.0675)	(0.0675)
Ever in the military	0.3233***	0.2574***	0.2667***	0.2657***
	(0.0467)	(0.0754)	(0.0912)	(0.0912)
Employment and training				
Weeks self-employed 20-23		-0.0019***	-0.0023***	-0.0024***
		(0.0007)	(0.0008)	(0.0008)
Unemployed for at least 6 months, cumulative				
Ages 16-19		-0.1086*	-0.1573**	-0.1578**
		(0.0602)	(0.0655)	(0.0653)
Ages 20-23		-0.3046***	-0.2860***	-0.2824***
		(0.0764)	(0.0795)	(0.0792)
Participated in a relationship-based career and technical education program		0.0021	0.0123	0.0109
		(0.0293)	(0.0334)	(0.0334)

TABLE 7C CONTINUED

Average marginal effects of multinomial models, Group 3

Dependent variable: Membership in Group 3

Variable name (reference group)	GROUP 3			
	Model 1	Model 2	Model 3	Model 4
Participated in a training program				
Ages 20-23		-0.0047	-0.0043	-0.0041
		(0.0299)	(0.0341)	(0.0342)
Ages 24-27		0.1212***	0.1322***	0.1321***
		(0.0293)	(0.0333)	(0.0333)
Occupation held at age 25 (Service)				
No occupation		-0.4151***	-0.5676***	-0.5664***
		(0.0936)	(0.1143)	(0.1142)
Management/professional occupation		0.2066***	0.1752***	0.1753***
		(0.0464)	(0.0514)	(0.0515)
Sales/office occupation		0.1075***	0.1252***	0.1244***
		(0.0380)	(0.0418)	(0.0419)
Construction/production/transportation/other occupation		0.0890**	0.1043**	0.1042**
		(0.0421)	(0.0467)	(0.0467)
Job is covered by a union contract at age 25		0.1453***	0.1464***	0.1451***
		(0.0451)	(0.0509)	(0.0510)
Educational attainment at age 27 (No high school diploma)				
High school diploma		0.1845***	0.2264***	0.2268***
		(0.0570)	(0.0636)	(0.0634)
GED		-0.0000	0.0093	0.0096
		(0.0708)	(0.0798)	(0.0797)
Associate degree		0.2285***	0.2808***	0.2809***
		(0.0708)	(0.0798)	(0.0798)
Bachelor's degree		0.2681***	0.3128***	0.3114***
		(0.0718)	(0.0818)	(0.0816)
Graduate degree		0.4639***	0.5503***	0.5580***
		(0.1220)	(0.1377)	(0.1290)
Has a work-limiting health condition at age 27		-0.3212***	-0.3308***	-0.3295***
		(0.0840)	(0.0974)	(0.0972)
Standardized ASVAB		0.0309*	0.0286	0.0288
		(0.0172)	(0.0196)	(0.0197)
Parental wealth in 1997			0.0419	0.0902
			(0.1069)	(0.3793)

TABLE 7C CONTINUED

Average marginal effects of multinomial models, Group 3

Dependent variable: Membership in Group 3

Variable name (reference group)	GROUP 3			
	Model 1	Model 2	Model 3	Model 4
Wealth and race/ethnicity interactions (Wealth x Black, non-Latino or Hispanic)				
Wealth x white, non-Latino or Hispanic				-0.0383
				(0.3959)
Wealth x Latino or Hispanic				0.0110
				(0.5363)
Wealth x other, non-Latino or Hispanic				-0.2147
				(0.6110)
Observations	3,674	2,348	1,916	1,916

Source: Child Trends analysis of NLSY97 data
Note 1: Standard errors in parentheses
Note 2: ***p < 0.01, **p <0.05, *p < 0.1
Note 3: Model 3 is preferred model

TABLE 7D

Average marginal effects of multinomial models, Group 4

Dependent variable: Membership in Group 4

Variable name (reference group)	Group 4			
	Model 1	Model 2	Model 3	Model 4
Female	-0.0800***	-0.0734***	-0.0255***	-0.0265***
	(0.0103)	(0.0123)	(0.0047)	(0.0049)
Race/ethnicity (Black, non-Latino or Hispanic)				
White, non-Latino or Hispanic	0.0553***	0.0296**	0.0131**	0.0180***
	(0.0121)	(0.0129)	(0.0058)	(0.0069)
Latino or Hispanic	0.0262**	0.0154	0.0101	0.0137*
	(0.0133)	(0.0138)	(0.0061)	(0.0073)
Other, non-Latino or Hispanic	0.0607***	0.0480**	0.0188**	0.0266**
	(0.0234)	(0.0215)	(0.0095)	(0.0112)
Age at time of first interview	0.0040	0.0035	0.0008	0.0009
	(0.0030)	(0.0028)	(0.0010)	(0.0011)
Family patterns				
First became parent at or before age 19	-0.0723***	-0.0125	-0.0083	-0.0085
	(0.0166)	(0.0165)	(0.0070)	(0.0074)

TABLE 7D CONTINUED

Average marginal effects of multinomial models, Group 4

Dependent variable: Membership in Group 4

Variable name (reference group)	Group 4			
	Model 1	Model 2	Model 3	Model 4
Family patterns				
First became a parent between ages 20-27	-0.0262***	-0.0045	0.0001	0.0001
	(0.0095)	(0.0091)	(0.0034)	(0.0035)
Married or cohabitating at 23	0.0284***	0.0198*	0.0061	0.0063
	(0.0096)	(0.0102)	(0.0038)	(0.0039)
Ever incarcerated by age 27	-0.1392***	-0.1107***	-0.2243***	-0.2223***
	(0.0300)	(0.0372)	(0.0399)	(0.0396)
Ever in the military	0.0263*	-0.0034	-0.0005	-0.0005
	(0.0156)	(0.0176)	(0.0064)	(0.0067)
Employment and training				
Weeks self-employed 20-23		-0.0003	-0.0001	-0.0001
		(0.0003)	(0.0001)	(0.0001)
Unemployed for at least 6 months, cumulative				
Ages 16-19		-0.0266	-0.0076	-0.0078
		(0.0277)	(0.0109)	(0.0115)
Ages 20-23		-0.0866**	-0.0247*	-0.0240
		(0.0420)	(0.0147)	(0.0153)
Participated in a relationship-based career and technical education program		0.0147*	0.0032	0.0033
		(0.0084)	(0.0031)	(0.0032)
Participated in a training program				
Ages 20-23		0.0229**	0.0087**	0.0091***
		(0.0089)	(0.0034)	(0.0035)
Ages 24-27		0.0252***	0.0071**	0.0072**
		(0.0087)	(0.0033)	(0.0034)
Occupation held at age 25 (Service)				
No occupation		-0.0962**	-0.0255*	-0.0263*
		(0.0447)	(0.0153)	(0.0159)
Management/professional occupation		0.0745***	0.0228***	0.0238***
		(0.0167)	(0.0062)	(0.0065)
Sales/office occupation		0.0286*	0.0085	0.0087
		(0.0155)	(0.0060)	(0.0063)

TABLE 7D CONTINUED

Average marginal effects of multinomial models, Group 4

Dependent variable: Membership in Group 4

Variable name (reference group)	Group 4			
	Model 1	Model 2	Model 3	Model 4
Occupation held at age 25 (Service)				
Construction/production/transportation/other occupation		0.0570***	0.0188***	0.0197***
		(0.0152)	(0.0058)	(0.0060)
Job is covered by a union contract at age 25		0.0366***	0.0103**	0.0109**
		(0.0124)	(0.0046)	(0.0047)
Educational attainment at age 27 (No high school diploma)				
High school diploma		0.0325	0.0076	0.0079
		(0.0238)	(0.0089)	(0.0093)
GED		0.0318	0.0069	0.0071
		(0.0270)	(0.0103)	(0.0107)
Associate degree		0.0488*	0.0110	0.0116
		(0.0266)	(0.0097)	(0.0101)
Bachelor's degree		0.0643**	0.0145	0.0150
		(0.0269)	(0.0099)	(0.0103)
Graduate degree		0.0825**	0.0200	0.0207
		(0.0333)	(0.0122)	(0.0127)
Has a work-limiting health condition at age 27		-0.0388	-0.0086	-0.0088
		(0.0295)	(0.0105)	(0.0109)
Standardized ASVAB		0.0105**	0.0046**	0.0048**
		(0.0053)	(0.0019)	(0.0020)
Parental wealth in 1997			0.0184**	0.0825**
			(0.0084)	(0.0371)
Wealth and race/ethnicity interactions (Wealth x Black, non-Latino or Hispanic)				
Wealth x white, non-Latino or Hispanic				-0.0638*
				(0.0373)
Wealth x Latino or Hispanic				-0.0515
				(0.0457)
Wealth x other, non-Latino or Hispanic				-0.0770
				(0.0589)
Observations	3,674	2,348	1,916	1,916

Source: Child Trends analysis of NLSY97 data
Note 1: Standard errors in parentheses
Note 2: ***p < 0.01, **p <0.05, *p < 0.1
Note 3: Model 3 is preferred model

TABLE 8

Marginal effects of the preferred model (Model 3) for trajectory groups 1, 2, 3, and 4

Variable name (reference group)	Group 1	Group 2	Group 3	Group 4
Female	0.0568***	0.1473***	-0.1786***	-0.0265***
	(0.0167)	(0.0354)	(0.0363)	(0.0049)
Race/ethnicity (Black, non-Latino or Hispanic)				
White, non-Latino or Hispanic	-0.0187	-0.0898**	0.0954**	0.0180***
	(0.0162)	(0.0375)	(0.0388)	(0.0069)
Latino or Hispanic	-0.0310*	-0.0996**	0.1205***	0.0137*
	(0.0171)	(0.0396)	(0.0405)	(0.0073)
Other, non-Latino or Hispanic	0.0771*	-0.0735	-0.0224	0.0266**
	(0.0442)	(0.0976)	(0.0947)	(0.0112)
Age at time of first interview	0.0018	-0.0148	0.0122	0.0009
	(0.0050)	(0.0104)	(0.0106)	(0.0011)
Family patterns				
First became parent at or before age 19	0.0407**	0.0749	-0.1073**	-0.0085
	(0.0188)	(0.0470)	(0.0489)	(0.0074)
First became a parent between ages 20-27	0.0068	0.0265	-0.0334	0.0001
	(0.0168)	(0.0336)	(0.0340)	(0.0035)
Married or cohabitating at 23	-0.0081	-0.0598	0.0617	0.0063
	(0.0180)	(0.0386)	(0.0390)	(0.0039)
Ever incarcerated by age 27	0.0730***	0.3364***	-0.1851***	-0.2223***
	(0.0225)	(0.0635)	(0.0675)	(0.0396)
Ever in the military	-0.2059***	-0.0602	0.2667***	-0.0005
	(0.0636)	(0.0894)	(0.0912)	(0.0067)
Employment and training				
Weeks self-employed 20-23	0.0012***	0.0012	-0.0023***	-0.0001
	(0.0002)	(0.0008)	(0.0008)	(0.0001)
Unemployed for at least 6 months, cumulative				
Ages 16-19	0.0239	0.1409**	-0.1573**	-0.0078
	(0.0206)	(0.0620)	(0.0655)	(0.0115)
Ages 20-23	0.0865***	0.2242***	-0.2860***	-0.0240
	(0.0234)	(0.0715)	(0.0795)	(0.0153)
Participated in a relationship-based career and technical education program	0.0009	-0.0164	0.0123	0.0033
	(0.0152)	(0.0330)	(0.0334)	(0.0032)

TABLE 8

Marginal effects of the preferred model (Model 3) for trajectory groups 1, 2, 3, and 4

Variable name (reference group)	Group 1	Group 2	Group 3	Group 4
Participated in a training program				
Ages 20-23	-0.0037	-0.0006	-0.0043	0.0091***
	(0.0152)	(0.0330)	(0.0341)	(0.0035)
Ages 24-27	-0.0563***	-0.0830**	0.1322***	0.0072**
	(0.0164)	(0.0329)	(0.0333)	(0.0034)
Occupation held at age 25 (Service)				
No occupation	0.2129***	0.3802***	-0.5676***	-0.0263*
	(0.0318)	(0.1030)	(0.1143)	(0.0159)
Management/professional occupation	-0.0611**	-0.1368***	0.1752***	0.0238***
	(0.0269)	(0.0507)	(0.0514)	(0.0065)
Sales/office occupation	-0.0352**	-0.0985**	0.1252***	0.0087
	(0.0179)	(0.0407)	(0.0418)	(0.0063)
Construction/production/transportation/other occupation	-0.0652***	-0.0580	0.1043**	0.0197***
	(0.0213)	(0.0457)	(0.0467)	(0.0060)
Job is covered by a union contract at age 25	-0.0786***	-0.0780	0.1464***	0.0109**
	(0.0291)	(0.0523)	(0.0509)	(0.0047)
Educational attainment at age 27 (No high school diploma)				
High school diploma	-0.0728***	-0.1612***	0.2264***	0.0079
	(0.0213)	(0.0591)	(0.0636)	(0.0093)
GED	-0.0217	0.0056	0.0093	0.0071
	(0.0249)	(0.0734)	(0.0798)	(0.0107)
Associate degree	-0.1480***	-0.1438*	0.2808***	0.0116
	(0.0474)	(0.0764)	(0.0798)	(0.0101)
Bachelor's degree	-0.1254***	-0.2019***	0.3128***	0.0150
	(0.0422)	(0.0776)	(0.0818)	(0.0103)
Graduate degree	-0.1433**	-0.4269***	0.5503***	0.0207
	(0.0696)	(0.1481)	(0.1377)	(0.0127)
Has a work-limiting health condition at age 27	0.1724***	0.1669*	-0.3308***	-0.0088
	(0.0238)	(0.0911)	(0.0974)	(0.0109)
Standardized ASVAB	-0.0394***	0.0062	0.0286	0.0048**
	(0.0105)	(0.0196)	(0.0196)	(0.0020)
Parental wealth in 1997	-0.1010	0.0407	0.0419	0.0825**
	(0.0646)	(0.1075)	(0.1069)	(0.0371)

Source: Child Trends analysis of NLSY97 data
Note 1: Standard errors in parentheses
Note 2: ***p < 0.01, **p <0.05, *p < 0.1

Endnotes

1 Martha Ross and others, "Pathways to High-Quality Jobs for Young Adults" (Washington: Brookings and Bethesda, MD: Child Trends, 2018).

2 The NLSY97 also reports earnings data in a weekly basis; however, the team found some inconsistencies in these data when compared to the yearly data. After consulting with the NLSY97 team, the team was told that there were some coding errors in the NLSY97 that would need to be corrected within a timeframe that was out of the scope of the project.

3 Daniel Nagin. *Group-Based Modeling of Development.* Cambridge, Mass.: Harvard University Press, 2009.

4 Geoffrey J. McLachlan, Sharon X. Lee, and Suren I. Rathnayake, "Finite Mixture Models," *Annual Review of Statistics and its Application* 6 (1) (2018): 355—378.

5 FMM include repeated measure latent class analysis (RMLCA), group-based trajectory modeling (GBTM), and growth mixture models (GMMs). Our preferred method is GBTM, since it assumes that members of the same latent trajectory group share observed patterns among observed variables. GMM and other methods assume some random variation within groups, leading to more homogenous groups. However, previous studies have shown that the three methods show similar trajectories for economic well-being measures. See, for example: Andrew Halpern-Manners and others, "The Impact of Work and Family Life Histories on Economic Well-Being at Older Ages," Social Forces 93 (4) (2015): 1369—1396.

6 Nagin, *Group-Based Modeling of Development*.

7 Although the model with five groups had a better BIC, the model had a highly singular variance matrix, meaning that it was hard to distinguish some of the groups from each other. Visual inspection confirmed that two of the groups had very similar trajectories.

8 Andrew Halpern-Manners, "Measuring Students' School Context Exposures: A Trajectory-Based Approach," *Social Science Research* 58 (2016): 135—149; Katherine Weisshar and Tania Cabello-Hutt, "Labor Force Participation Over the Life Course: The Long-Term Effects of Employment Trajectories on Wages and the Gendered Payoff to Employment," *Demography* 57 (2020) 33—60.