Proposal

Ethan Weiland

2024-03-07

Statement of the Problem

There is widespread public concern regarding the finances of young American adults ("Millennials"). Commentators are especially worried about high levels of debt, with one going so far as to label this current crop of young adults "Generation Debt" (Kamenetz 2006). This public worry has been mirrored in the sociological literature (Kus 2015; Dwyer 2018; Wherry and Chakrabarti 2022). In light of the the new reality of widespread debt, scholars have studied its associations with fertility (Nau, Dwyer, & Hodson 2015), physical health (Lippert, Houle, & Walsemann 2022), home buying (Houle and Berger 2015), and cohabitation (Addo 2014). Houle (2014) empirically confirmed high levels of debt for Millennials in their mid-twenties using the National Longitudinal Survey of Youth - 1997. Specifically, he found that a significantly higher proportion of Millennials have a negative net worth compared to their predecessors (Late Boomers, Early Boomers), as well as significantly greater debt-to-asset and debt-to-income ratios. While insightful, sociological research has only studied Millennials in their twenties and has failed to track Millennials as they have aged into their thirties. This research updates the literature by studying debt trajectories of Millennials throughout their twenties *and* thirties. Specifically, this analysis will use the National Longitudinal Survey of Youth - 1997 and multilevel modeling to answer the following three research questions:

1. How much of the variance in debt is attributed to within-individual and between-individual differences respectively?
2. What is the association between parental income and debt?
3. Does the analysis differ when considering student debt as the dependent variable, rather than total debt?

Data

The NLSY97 - administered by the Bureau of Labor Statistics - is a nationally representative panel study of American youth between the ages of 12 and 16 in 1997. Today, respondents are between the ages of 39 and 43. The survey contains a wide variety of questions touching on practically every aspect of the respondents' lives, with a breadth similar to that of the General Social Survey. The NLSY97 was conducted annually in its earlier iterations but is now conducted biannually. The most recent round - Round 20 - was conducted in 2021 and released for researcher use in February 2024. The NLSY97 is made up of two independently selected, stratified multistage area probability samples. By combining these two samples, the NLSY97 ensures adequate representation of Black and Hispanic households in the survey. The complex sample design is accounted for using the `svydesign` command in the `survey` package in R.

Respondents are asked about their assets and debts every five years in what are called YAST modules. Respondents receive the YAST20 module when they reach the age of 20, the YAST25 module when they reach the age of 25, etc. To answer the first two research questions, respondents are included if they reported having debt in at least one YAST module. There are 7,523 respondents that reported debt in at least one YAST module from YAST20 up until YAST40. Total debt includes student debt, automobile debt, and other debt (debt owed to family/friends, credit card debt, store debt, etc.). Importantly, total debt does not

include housing debt as housing debt is generally seen as a positive wealth-building form of debt. This total debt variable is made available by the Bureau of Labor Statistics. To answer the third research question, respondents are included if they reported having student debt in at least one YAST module. There are 2,720 respondents reporting debt in at least one YAST module from YAST20 up until YAST40. To measure student debt, respondents were asked "You reported borrowing government-subsidized or other types of loans to attend a school or college. What is the total amount that you owe altogether on these educational loans?". Respondents answered with either the amount, a range (recoded to the midpoint), or choose from a showcard category (recoded to the midpoint of the category). Importantly, student debt does not include debt owed to family/friends as this is qualitatively different to the formal student debt owed to institutional creditors like banks. Table 1 and Table 2 in the Appendix reports the means, medians, standard deviations, minimums, and maximums for total debt and student debt respectively at each age measurement. Figure 1 and Figure 2 in the Appendix show the distribution of debt values among respondents reporting debt at least one time at each age. The points are jittered to help with visualization and the size of the points indicates their survey weight value.

The key independent variable in Research Question #2 is parental income. Parental income is operationalized as the gross household income in the first wave of the survey (1997), when the survey respondent was between the ages of 12 and 16. Parental net worth is the household assets minus the household debts in the first wave of the survey. Parental education is equal to the highest grade completed by either of the respondent's residential parents on an interval scale. For example, if parental education is equal to 16, this indicates that the highest level of education achieved by either of the respondent's residential parents is two years of undergraduate study. Sex is a factor variable with "Male" as the baseline category. Race is factor variable with four-levels: Black, Hispanic, Multiracial (Non-Hispanic), and Non-Black / Non-Hispanic (baseline category). Many of the financial variables were top coded, oftentimes at 2%, by the Bureau of Labor Statistics to ensure confidentiality. All financial measures (parental income, parental net worth, total debt, student debt) were adjusted for inflation. The appropriate conversion rates to adjust for inflation were obtained from the Bureau of Labor Statistics CPI Inflation Calculator (https://www.bls.gov/data/inflation_calculator.htm). Table 1 in the Appendix reports the descriptive statistics of the variables used to answer the first two research questions (where the dependent variable is total debt). Table 2 in the Appendix reports the descriptive statistics of the variables used to answer the third research question (where the dependent variable is student debt).

Analysis Plan

These data follow a fixed occasion design - rather than a variable occasion design - because there are 5 fixed measurement occasions (Snijders and Bosker 2012). In fixed occasion designs, the level-two units are individuals and the level-one units are measurement occasions. A hierarchical linear model will be implemented. There are three main ways to implement a hierarchical linear model in this context: the compound symmetry model (random intercept model), the random slope model, and the fully multivariate model. These models only differ in their random component, with covariance matrices ranging from most restrictive (compound symmetry) to least restrictive (fully multivariate).

The first research question is "How much of the variance in debt is attributed to within-individual and between-individual differences respectively?". To answer this research question, I will implement the compound symmetry model with total debt as the dependent variable, individuals as the level-two units, measurement occasions as the level-one units, and including the control variables described in the above section. Then, I will calculate the intraclass correlation coefficient. A compound symmetry model is appropriate because the variance of total debt is homoskedastic under this specification. In the random slope and fully multivariate models, the variance and covariance of total debt is dependent on the values of the explanatory variables, which makes the intraclass correlation coefficient difficult (if not impossible) to interpret. The second research question is "What is the association between parental income and debt?". To answer this research question, I will run a random slope model with total debt as the dependent variable, individuals as the level-two units, measurement occasions as the level-one units, a random slope for time, a main effect for parental income, an interaction effect between parental income and time, and the control variables described

in the above section. I will then test the main effect of parental income and the interaction between parental income and measurement occasion for significance. Finally, I will explore polynomial trend analysis in the random slope model and compare both the compound symmetry model and random slope model to the fully multivariate model to test model fit.

The third research question is "Does the analysis differ when considering student debt as the dependent variable, rather than total debt?". To answer this research question, I will rerun the above steps with a different dependent variable (student debt) and a different sample specification (those that report student debt at at least one measurement occasion rather than those that report any debt at at least one measurement occasion).

GitHub Link

https://github.com/ethanphilipweiland/millennial-debt

References

1. Addo, Fenaba R. 2014. "Debt, Cohabitation, and Marriage in Young Adulthood." Demography 51(5):1677–1701. doi: 10.1007/s13524-014-0333-6.
2. Dwyer, Rachel E. 2018. "Credit, Debt, and Inequality." Annual Review of Sociology 44:237–61. doi: 10.1146/annurev- soc- 060116- 053420.
3. Houle, Jason N. 2014. "A Generation Indebted: Young Adult Debt across Three Cohorts." Social Problems 61(3):448–65. doi: 10.1525/sp.2014.12110.
4. Houle, Jason N., and Lawrence Berger. 2015. "Is Student Loan Debt Discouraging Homeownership among Young Adults?" Social Service Review 89(4):589–621. doi: 10.1086/684587.
5. Kamenetz, Anya. 2006. Generation Debt: How Our Future Was Sold Out for Student Loans, Bad Jobs, No Benefits, and Tax Cuts for Rich Geezers–And How to Fight Back. New York, New York: Riverhead Books.
6. Kus, Basak. 2015. "Sociology of Debt: States, Credit Markets, and Indebted Citizens." Sociology Compass 9(3):212–23. doi: 10.1111/soc4.12247.
7. Lippert, Adam M., Jason N. Houle, and Katrina M. Walsemann. 2022. "Student Debt and Cardiovascular Disease Risk Among U.S. Adults in Early Mid-Life." American Journal of Preventive Medicine 63(2):151–59. doi: 10.1016/j.amepre.2022.02.002.
8. Nau, Michael, Rachel E. Dwyer, and Randy Hodson. 2015. "Can't Afford a Baby? Debt and Young Americans." Research in Social Stratification and Mobility 42:114–22. doi: 10.1016/j.rssm.2015.05.003.
9. Snijders, Tom A. B., and Roel J. Bosker. [1999] 2012. Multilevel Analysis: An Introduction to Basic and Advanced Modeling. 2nd ed. Thousand Oaks, California: SAGE.
10. Wherry, Frederick F., and Parijat Chakrabarti. 2022. "Accounting for Credit." Annual Review of Sociology 48:131–47. doi: 10.1146/annurev-soc-030320-114444.

Appendix

Figure 1. Bubble Plot of Total Debt Across Age

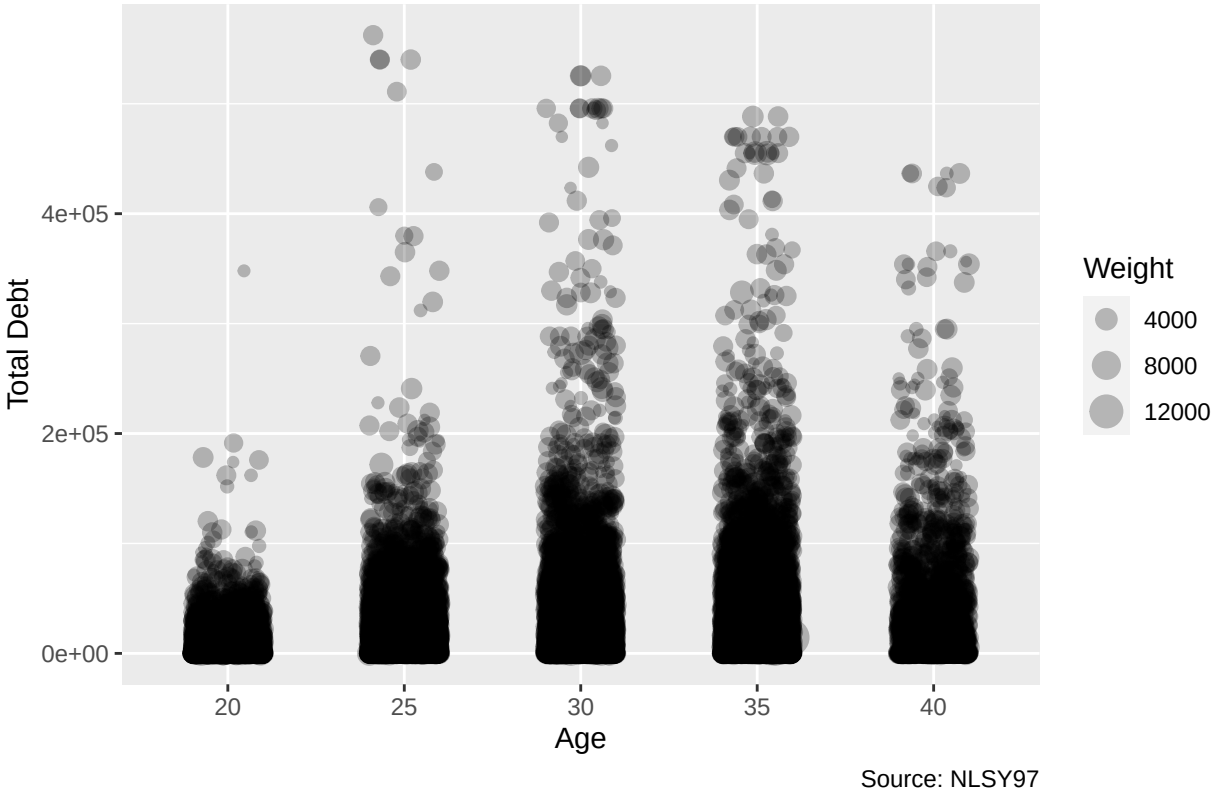

Table 1: Descriptive Statistics for Respondents Reporting Total Debt (N=37615)

	Mean	Median	SD	Min	Max
Total Debt at Age 20	6,781.54	340	14,002.61	0	348,000
Total Debt at Age 25	20,790.90	9,720	33,268.37	0	562,400
Total Debt at Age 30	30,661.70	13,400	50,337.31	0	525,400
Total Debt at Age 35	32,762.70	13,970	52,892.59	0	488,400
Total Debt at Age 40	34,523.43	13,688	55,955.56	0	436,600
Female	0.51			0	1
Parental Income	91,256.03	75,660	79,992.51	-93,314	478,159.60
Parental Net Worth	174,850.20	73,235	260,309.00	-1,814,387.00	1,164,000
Parental Education (yrs of schooling)	13.22	13	3.03	1	20
Black	0.23			0	1
Hispanic	0.21			0	1
Multiracial (Non-Hispanic)	0.01			0	1

Source: NLSY97. Among respondents reporting debt in at least one YAST module. Unadjusted for survey design.

Figure 2. Bubble Plot of Student Debt Across Age

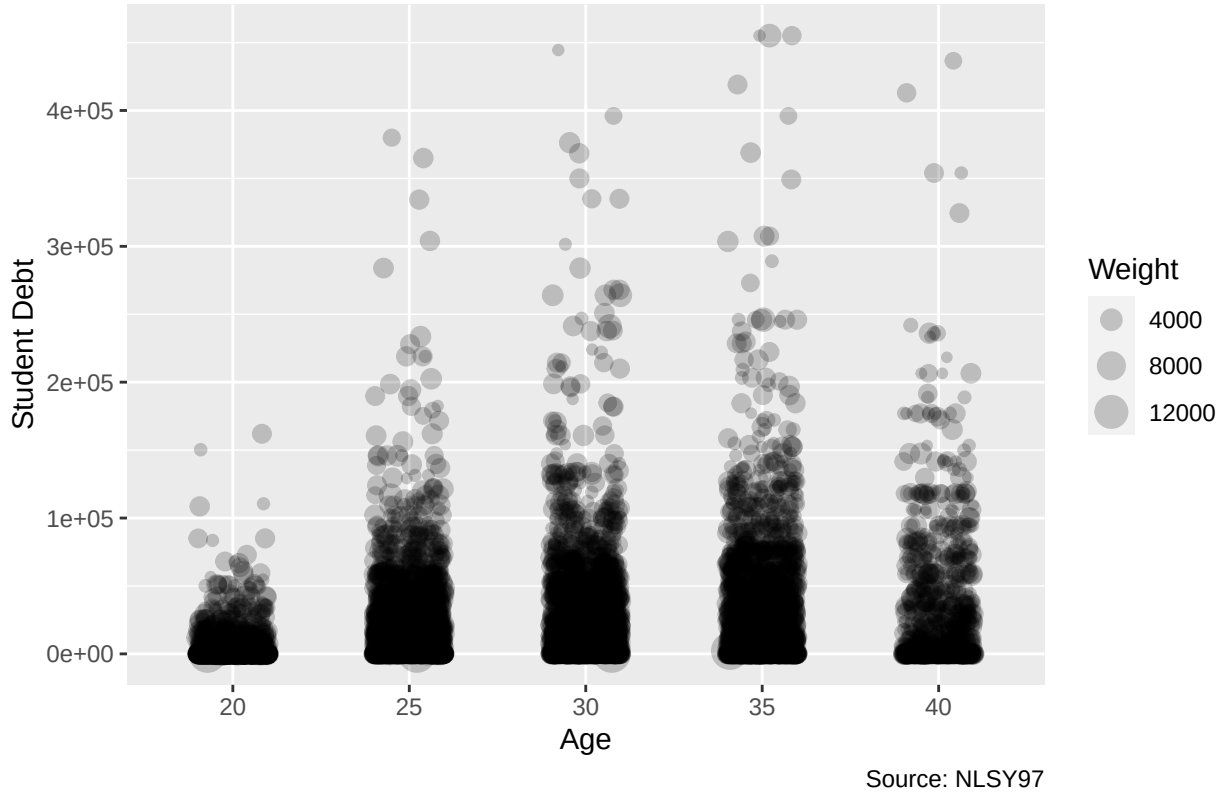

Table 2: Descriptive Statistics for Respondents Reporting Student Debt (N=13600)

	Mean	Median	SD	Min	Max
Student Debt at Age 20	2,683.84	0	9,774.04	0	162,000
Student Debt at Age 25	21,542.83	11,680	31,836.67	0	380,000
Student Debt at Age 30	29,547.42	15,840	41,649.57	0	444,500
Student Debt at Age 35	30,518.07	12,700	45,878.81	0	455,100
Student Debt at Age 40	33,533.04	9,440	52,614.27	0	436,600
Female	0.59			0	1
Parental Income	95,508.41	83,420	71,446.03	-15,617	478,159.60
Parental Net Worth	179,215.60	92,732	241,144.90	-607,705	1,164,000
Parental Education (yrs of schooling)	13.91	14	2.79	2	20
Black	0.29			0	1
Hispanic	0.16			0	1
Multiracial (Non-Hispanic)	0.01			0	1

Source: NLSY97. Among respondents reporting student debt in at least one YAST module. Unadjusted for survey design.