

Research Opportunity: A New Cohort of the National Longitudinal Survey of Youth

Seth G. Sanders

Workshop on Structural Racism and Rigorous Models of Social
Inequity

National Academy of Sciences

MAY 16-17, 2022

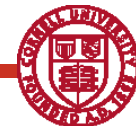
*The views expressed here are my own and may not
represent the views of BLS, NORC or CHRR.



Value of Population Representative Panel Data

Conclusions

1. Panel data in principle seems a powerful vehicle to study the effects of structural racism.
 - a. It has the ability to understand the interactions between different walks of life (e.g. criminal justice outcomes and late labor market success).
 - b. It has the ability to measure cumulative effects.
2. In practice these studies are less than ideal.
3. An effort to launch a new cohort of the NLSY seems like an opportunity to rectify some of the persistent issues.
4. Get involved!



Strengths of NLS for study of structural racism

1. Data collection generally starts when respondents are young:
 - Four cohorts started in 1967/68 young men and women (14-24), mature women, older men.
 - Cohort started in 1979, men and women 14-21.
 - Cohort started in 1997, men and women 12-16.
 - New cohort planned for 2026, age range TBA but likely similar to NLSY97.
2. This is helpful for life-course research and studies of cumulative disadvantage.



Strengths of NLS for study of structural racism

3. Annual or bi-annual data collection in multiple domains of life allowing studies of how racism in once set of institutions spills over to others.

- Education, Training & Achievement Scores (extensive)
- Employment (extensive)
- Household, Geography & Contextual Variables
- Family Background
- Dating, Marriage & Cohabitation; Sexual Activity; Pregnancy & Fertility; Children
- Income, Assets & Program Participation
- Health (historically weaker)
- Attitudes & Expectations
- Crime & Substance Use (rare in panel surveys)



Strengths of NLS for study of structural racism

4. Oversample of Racial and Ethnic Minorities makes important subgroup analysis possible. For Example NLSY97:
 - 2,335 Black individuals
 - 1,901 Hispanic individuals
 - 4,665 White individuals
5. Some intergenerational aspects can be studied:
 - NLSY79 has a child and young adult supplement
 - NLSY97 had a parents survey in year one.
6. Sampling design has many siblings and in 1979 many cousins allowing strong controls for family background.



Traditional weaknesses of NLS for study of structural racism: macro issues

1. To my knowledge, diversity in the planning process and design team has historically not been a specific priority.
2. This likely has affected content and design given the multipurpose nature of the survey and the inevitable time and resource constraints.
3. This can be changed by leadership and collective action; a willing BLS staff still needs to justify content and design decisions and hearing from the research community is key.
4. In my view, the current planning process is a big step forward to build a new cohort that better represents the research needs of a diverse user base.



Traditional weaknesses of NLS for study of structural racism: specific issues

1. NLS has done a better job of measuring outcomes (did you get a loan, were you arrested) than the process leading to those outcomes.
2. To some degree, content relevant to all subjects is prioritized; but some content may be highly relevant to studies of structural racism but perhaps less relevant generally (e.g. no question on eviction; no question on denial of loan).
3. Representative oversamples are helpful but are they large enough? Historically sample sizes of high SES minorities have been small.



Traditional weaknesses of NLS for study of structural racism and health: specific issues

1. Health data is generally weak relative to health surveys. Special emphasis on health every 10 years beginning at age 30 in the NLSY97 and age 40 in the NLSY97.
2. NLSY97 does a better job of collecting stressors than NLSY79 but no race-specific stressors are collected.
3. No biomarker assessments and limited self-reports on biomarkers; self-reports are often of the “told by a doctor” conflating health access and health conditions.
4. Lack of assessments includes typical assessments in health surveys, measured height, weight, blood pressure, pulse as well as genetic data (so no PRSs or telomeres etc.)
5. Even with this, health scientists are the fastest growing group of NLSY users.



Traditional weaknesses of NLS for study of structural racism in other areas: specific issues

1. NLSY is strong on the ability to link contextual data spatially (e.g. census tract level data)
2. NLSY is weaker on allowing the linkage of data in other dimensions (e.g. no school IDs, no firm IDs, no health care provider IDs, no law enforcement agency IDs.)
3. These are important institutions where structural racism may vary.
4. Intergenerational data is limited (without extensive supplementary surveys).



A New Cohort and Research Opportunities

1. Planning process well under way for a new cohort planned for 2026 data collection.
2. Planning contract let to NORC and CHRR to organize activities and engage a wide set of stakeholders:
 - a. Three content panels on Family Background and Cognition, K-12 Education and Health and the Environment with subject matter experts. A fourth panel targeted at Department of Defense needs and interest is planned.
 - b. Listening sessions on: childhood and family retrospective; mental health; physical health and the environment; innovations in international surveys; data needs of think tanks, non-profits and research organizations; and the nature of work.
 - c. Invitation to registered NLSY users to participate in user survey (well over 400 surveys completed).
 - d. An analysis of alternative data sources and underused variables.
3. Diversity in the planning process is a specific priority.



A New Cohort and Research Opportunities

1. Questions for discussions:

- a. Given the panel nature of the NLSY and its historic strength on measuring labor market outcomes and human capital, what complementary content might be collected to advance studies of structural racism?
- b. As health appears to be an increasing use of the NLSY, what health information would be most important to collect to advance studies of structural racism?
- c. Are there additional specific subgroups that are scientifically important enough to justify oversamples that are important to advance studies of structural racism?
- d. Are there alternative data sources that might be linked to survey data that would advance studies of structural racism?
- e. Are there new methods of data collection that would yield new data? What are they and how would you use them.



A New Cohort and Research Opportunities

Contact me and I will find the right people!

seth.sanders@cornell.edu

