NORC at the University of Chicago

CHRR AT THE OHIO STATE UNIVERSITY

TASK 1.0.3 CONTENT AND MEASUREMENT OBJECTIVES

August 30, 2022, REVISED

Contract Deliverable

Presented by:
Roxanne Wallace, PMP

Document Change Log

Published Date	Document Version No.	Pages Affected	Description of Revision	Author
8/4/2022	v1	Pages with comments/tracked changes		Quentin Brummet, Rupa Datta, Seth Sanders
8/30/2022	v2	Pages with comments/tracked changes		Quentin Brummet, Rupa Datta, Seth Sanders

Authored by: Quentin Brummet, Rupa Datta, Seth Sanders, and Erik Scherpf

Executive Summary

This report synthesizes three separate activities conducted as part of a needs assessment in order to provide recommendations on content and measurement objectives for a new NLSY26: stakeholder outreach activities including a user survey and six listening sessions; a retrospective analysis of prior use of NLSY data, potential underused variables, and comparisons to other surveys; and a detailed evaluation of potential alternative data sources (ADS).

Taken as a whole, these activities identified a clear demand for a new cohort of the NLSY26 from academic researchers, policy-focused research organizations, and federal agencies other than BLS. Although other data exists for the over-time study of individuals' participation in labor markets, we confirmed a distinct gap in data availability. Other surveys are short in panel duration, narrow in topical coverage or have inadequate sample sizes of specific birth cohorts. There is a need for a rigorous nationally representative longitudinal survey covering broad domains of life, with adequate birth cohort sample sizes to study disparities and mechanisms of labor market success over the life cycle. While longitudinal linked administrative data has become an important source for analysis, this data tends to be narrow in domains and does not include data that can be collected only through interaction with respondents including cognitive ability, expectations, and criminal activity not recorded in administrative records. Because the NLSY79 and 97 filled similar gaps at their inceptions, the ability to compare analyses from those earlier cohorts would be uniquely valuable.

Below, we provide further details on what our exploratory activities suggest specifically about content considerations for a new NLSY cohort, the methodological decisions that will need to be made, the potential uses of alternative data, and how NLSY26 data should be disseminated. We note at the outset that our work focused on opportunities but very little on the constraints that will restrict what is possible. Chief among these is respondent attention and willingness to cooperate, which are essential to the long-term representativeness of the survey data. Additional constraints are respondent privacy concerns, and ability to negotiate ADS access. Budget, budget timing, and potentially escalating costs will of course be a substantial constraint as with any data collection today. To the extent that optimal choices are combinations of opportunities with constraints, the work in this report can only get the BLS partway until married with the relevant constraints.

Content Objectives

The NLSY program has historically served many audiences, and the results of our exploratory activities confirmed this will remain true for the NLSY26. As such, the survey design will need to make many difficult content tradeoffs and there will be no easy answers. Factors such as cross-cohort comparisons, relevance to labor market outcomes, ability to support cross-domain research, and whether the data is available elsewhere will all need to be considered; and each may need to be weighted differently depending on the context. Nonetheless, there were multiple specific content objectives that emerged during the needs assessment:

> Our explorations identified no domains from prior NLSYs that had receded in importance but did identify some that had emerged or increased in importance: mental health, use of technology, and

to a lesser extent, gender identity and sexual orientation, and interaction with human services (such as foster care, special education, etc.).
Although non-cognitive and cognitive skills have always been a priority for the NLSYs, these domains were identified for more frequent, updated, and systematic over-time inclusion.
Employment has gotten more complex and is likely to need more survey time than previous cohorts. Once the cohort enters the labor market in a significant way, there will be limited space available for non-employment topics. Research questions that require significant interview time concurrent with employment will be more challenging to incorporate in the design than research questions that can primarily be addressed through questions collected in the early years.
Outside of employment, meaningful measurement has grown more complex for a wide variety of phenomena in American society, including racial and ethnic classification, gender identity and sexual orientation, household structure, and educational and training experiences (due to on-line and hybrid offerings and growth of alternative certifications).
Our exploratory activities show that both cross-cohort comparisons and updated content are essential to a successful new cohort. A limited set of items will be invaluable to collect in ways that allow for comparability with prior NLSY cohorts, especially in key educational and employment outcomes. Measures from other domains that could serve as covariates for examining those outcomes would also be helpful. But many of these may be more valuable to update to reflect contemporary society especially if more complex contemporary practices can be mapped into historical ones. Important examples include marital/cohabitation status, gender identity, race/ethnicity, disability status, and cognitive ability, all of which have undergone substantial transformation since the start of the NLSY79. In these instances, retaining strict comparability to prior NLSY cohorts on these types of measures will not position the cohort for widest use.

Taken together, these above considerations imply that the distribution of administration time across domains will likely need to shift in the NLSY26 relative to the NLSY97, its immediate predecessor. To decide on how best to make this shift, BLS could use four approaches: rely on alternative data sources to expand the content that can be gathered, reduce coverage of all domains, reduce the depth/granularity of coverage in some domains, or reduce the frequency of coverage in some domains. The NLSY26 will likely need to use a mix of these approaches, and we discuss some of the specific possibilities below.

Methodological Considerations:

The needs assessment documented strong support for ensuring adequate sample sizes of Black and Hispanic youth to allow for within-subgroup analyses, as the prior NLSYs have done through oversampling. There are a wide variety of other oversamples that were mentioned across the activities, but additional oversamples will need to be considered carefully given that the relatively small size of the survey will only support a certain amount of oversamples and each additional oversample will come with additional budgetary and analytic cost. Importantly, the larger the group and the more precisely it can be identified from external data sources, the more feasible the oversample will be given cost constraints.

Parental interviews will be an important consideration for the NLSY26, and support two important themes that emerged during the course of our needs assessment: (1) the development of the respondent and the effects on later life outcomes such as long-term labor market success and (2) intergenerational studies. Given the importance of these themes and the additional value from having multiple measurements, multiple parental interviews would make a valuable contribution to an NLSY26.

Usage of Alternative Data Sources

Across many of our exploratory activities, it became clear that an NLSY26 should prioritize opportunities for researchers to link to other data sources, including individual-level administrative data from public programs, as well as private data sources such as credit bureaus, social media sites, or health insurance records. Although the NLSYs have always allowed for data linkages, it will be important for BLS to facilitate increased possibilities for linkages and reduced costs to researchers for conducting research with linked sources. Nonetheless, despite the substantial advances in data availability and linkage, the initial rounds of the NLSY26 likely should be designed to collect essential data via surveys until the feasibility of consent and linkages has been demonstrated. This is particularly true given that successfully using administrative data requires addressing significant privacy and administrative challenges, which would add to the risk of using administrative data to replace survey content in early rounds.

Dissemination

Throughout our assessment, making data access as easy and as democratic as possible was mentioned as a priority. Dissemination is no easy task in the face of protecting respondent confidentiality especially with the integration of alternative data sources into the survey. The NLSY26 will have the greatest impact if barriers to use are minimized. Modern information technologies may make it easier to disseminate and document data content in an easy-to-understand form so that analysis can be done by users without advanced analytic skills and extensive resources for analysis. Timely and easy to use data are very valuable and can greatly facilitate use of the data for high-profile, policy-informative work. This type of work, when published or especially if cited in policymaking, can also be valuable to show respondents and stakeholders the value of the cohort and encourage continued support of the cohort. Developing dissemination tools will also facilitate use in the classroom and will grow the next generation of NLSY users.

A second dissemination challenge is administrative rather than technical. Without a doubt, the NLSY26, as was true for previous NLSY surveys, will contain sensitive information which will need to be distributed under data use agreements. Developing a system of quickly assessing applications, granting access, and providing needed analysis files and reviewing final output for data disclosure avoidance will be essential to encouraging use of a new cohort. This system would ideally be developed so that the time and potentially monetary costs of data access are lower for accessing data items with lower disclosure risk. As examples, distributing geocoded data might be done through BLS with short turnaround time for project initiation and data that is covered by Title 13 will need to go through the additional hurdles imposed by the Census Bureau. Multiple channels of data distribution will likely serve the user community best. The BLS has multiple channels currently for NLSY data, but these can be improved and expanded to support broader and more creative use of an NLSY26.

1. General Statement of Purpose of an NLSY26

The National Longitudinal Survey of Youth 2026 (NLSY26) will be a nationally representative cohort sample of young men and women living in the United States at the time of the initial survey in 2026 and will collect extensive information on respondents' labor market behavior and educational experiences. This cohort will join a collection of previous NLS cohort studies fielded by the US Bureau of Labor Statistics (BLS) beginning in 1966-67, 1979 and 1997.

A major feature of the new cohort study, as has been true since the inception of the NLS program, is measuring the inter-relationships between various domains of life. The tradition of the NLS program has been to select domains that are related bidirectionally or have important driving factors in common with labor market behavior and educational experiences, and this tradition will be important for the NLSY26. For example, family formation and fertility timing are likely jointly determined with labor market behavior and educational choices: romantic partnerships change labor market choice from a purely individual one to a joint one; and children compete with work, schooling, and other types of training for an individual's time. Government programs are related as they serve as a safety net when families do not have adequate earned income or to facilitate human capital investments. Human productivity and earnings likely both effect and are affected by mental and physical health. Criminal behavior may be related directly to the scarcity of job opportunities or indirectly through shared factors such as an individual's set of cognitive and non-cognitive skills. While data on these domains and others is important to collect in the NLSY26, the amount and detail of data in each domain should be guided by the strength of the relationship with labor market behavior and educational experiences.

Information from the NLSY26 can meet the dual objectives of allowing researchers to compare the progress of this cohort with that of other NLS cohorts and to answer questions that are new but especially relevant to this cohort.

In this report we discuss the many design decisions needed to shape the NLSY26 and tradeoffs therein. After discussing principles governing our analysis, the unique role of the NLSY program, and central research questions for an NLSY26, we then discuss specifics of the NLSY26 design. This includes the content that the NLSY26 program could cover, the methodological details for the new cohorts, how the NLSY26 can be designed to best make use of alternative data sources, and how the data might be distributed in a way that lowers the cost of usage while maintaining respondent confidentiality. We end by discussing important high-level tradeoffs facing the NLSY26 program.

2. Considerations Regarding Delivery of High Scientific Value

To produce a data design and infrastructure that will have high scientific impact, it is important to establish a set of governing considerations. Inevitably, no one data product can address all the needs of the research community. Any survey must establish its focus; given financial, time and attention constraints of respondents, this inevitably requires sacrificing some lines of inquiry to strengthen others. In addition, modern survey data, especially from longitudinal surveys, is used both for description and causal inference. It is not possible to design a survey that allows for strong causal inference on all questions. Some designs can facilitate stronger causal inference on a wider set of topics, but this often comes at some cost as well. Costs can include increased data collection resource requirements; decreased resources for non-data collection activities such as data file linkages, construction, and dissemination; or perhaps omission of content areas where at least descriptive analysis might have otherwise been feasible. In addition, there are a host of other decisions that will affect what can be done best with the survey and what can be done with less precision or perhaps not at all.

The content and measurement objectives in this report are guided by a set of considerations in each of the following areas -- project and sample design, survey content, and the user experience. These considerations were originally drafted by BLS, with the last a suggestion from the NORC team. Throughout the course of the needs assessment, the considerations were discussed with the needs assessment Oversight Committee and organized into three categories: design, content, and user considerations.

The table below lists the 13 considerations that have informed our work identifying content and measurement objectives. The righthand column of the table discusses interactions between considerations and how they can inform the NLSY26 design.

Table 1: Guiding Considerations

Design Considerations	Implications for NLSY26 Design
The design of the NLSY should align with the mission of BLS.	When resources are allocated (such as data collection dollars or interview minutes), relevance to the BLS mission can be a basis for prioritization.
Data collection, processing, and distribution should protect the confidentiality of respondents.	This consideration is non-negotiable and must be embedded in the design. It may have more implications for how data is disseminated than for what is collected.
The design of the NLSY should promote working relationships with key users and organizations whose work is consistent with the BLS mission and strategic goals.	Together with considerations #5 on maintaining response rates, #7 on minimizing government costs, and #8 on current and prospective information needs, this consideration underscores that sustaining a cohort over many years is a challenging proposition. Factors that bring in additional support for the cohort and help to sustain it financially, technically, through user support, are worth embracing into the design of the study, and can improve its content through diverse perspectives as well as its long-term well-being. This consideration too, may be more relevant for data dissemination than for what data is collected. Timely data, easy to use, and including policy-relevant created variables such as poverty ratios and urbanicity were most often requested by non-government users. Successful dissemination of their policy-related work will assist in maintain the sample's engagement in the study.
Design should include oversamples that permit estimates of under-represented and high priority populations.	This consideration interacts with consideration #5 about statistically valid samples, consideration #7 about minimizing cost to the government, consideration #8 about the current and prospective needs of researchers and policymakers. The four together define whether the financial and analytic costs of oversamples are adequately balanced by the additional analytic capabilities and policy relevance of having a specific oversample.
The design should enable the maintenance of sample sizes and response rates that allow for statistically valid analysis over the life course of respondents.	This consideration as well as consideration #7 on minimizing costs to the government, and #13 on accurate and unbiased data represent the most critical constraints on the study outside of legal/confidentiality issues. Having an engaging interview that does not feel overly intrusive to respondents and which imposes only reasonable burden on respondents will be critical for maintaining response rates, managing costs, and ensuring that data can be accurate and unbiased. These three considerations balance against almost all of the others in imposing constraints on what can be done, therefore demanding that priorities be set and choices be made.
The design should support valuable cross-cohort analyses.	This consideration interacts with consideration #8 on meeting the current and prospective information needs, and #9 on reflecting recent changes. The ability to generate selected comparable estimates across the NLSY

	cohorts will be essential but should not constrain the updating of NLSY26 content to reflect updated research and practice as well as new phenomena that have arisen since 1997. Updates and new areas in the NLSY26 will ideally be constructed so that collapsed measures can be used for comparisons to prior cohorts.
Data collection, processing, and dissemination should be conducted in a way that minimizes government costs.	Costs to the government are sometimes immediate, for example, in the inclusion of selected oversamples. At other times, a longer time horizon may be more appropriate. Investments in in-person data collection in the first wave may facilitate response rates and sample representativeness in later rounds. Expenditures on analysis and data dissemination tools may encourage data use and findings dissemination in ways that help sustain respondent and user engagement of the NLSY26.
The design should meet the current and prospective needs of researchers and policy makers.	Although cross-cohort analyses as referenced in consideration #6 above will be of great value in an NLSY26, a balance is required between having data that can support comparisons to prior NLSY cohorts and having data that will inform current and future information needs. Implications can extend beyond the questionnaire content to a sampling approach that will support policy analysis and facilitating linkages to alternative data sources to maintain currency throughout the life of the cohort.
Content Considerations	
Content should reflect societal, cultural, economic, etc. changes occurring since the NLSY97 cohort was initiated.	This consideration interacts with consideration #6 on cross-cohort analyses, and #8 on meeting current and prospective information needs. The NLSY26 will need to ensure that it is updated and relevant to changes that have occurred since 1997 in order to provide the greatest value to data users. Reflecting recent changes can include entire topics -- such as the COVID-19 pandemic – as well as measurement issues pertaining to developments in the research literature or in societal practices such as how race and ethnicity or gender identity are classified. Although the primary concerns are about the analyses that are possible, reflecting recent changes also affects respondents' perception of the study and its relevance and their willingness to participate (for example, respecting how youths describe themselves and their experiences).
Content should be age-relevant based on the life course stage of respondents.	This consideration can help manage the evolution of the NLSY26 questionnaires over time. Although over-time labor market experiences are the primary focus, capturing detailed employment information at age 13 may not be. In addition, keeping in mind consideration #5 about maintaining response rates, some topics of great scientific interest may not be good choices for very young respondents because of parents' sensitivities (for example, sexual activity, substance use).
Content should not duplicate data for analysis that is available to users elsewhere.	Some overlap and harmonization of key measures with other data sources is desirable to facilitate extensions, replications, or comparisons. However, beyond these purposeful duplications, this consideration can assist the NLSY26 in achieving a distinctive niche and thereby maximize usership and optimize allocation of federal dollars.
User Considerations	
Data should be accessible to a wide variety of users / remove barriers to ensure equal access to NLS data.	This consideration primarily affects dissemination choices more than content choices. To the extent that linked alternative data sources may have more restricted access than questionnaire responses, this consideration may compel BLS to have minimal coverage of each domain in the questionnaire (where it would be widely accessible) even if more detailed coverage is available through linked data.
Data is accurate, reliable, and unbiased, and is presented in an accurate, clear, and unbiased manner.	This consideration can be informative when choosing the measurement approach for a given construct (for example, through direct measurement, self-report, or alternative data). In addition, prioritizing what data can be accurately collected can assist in making choices between content areas.

These considerations were applied to a set of exploratory activities. These included the following:

Stakeholder Outreach: We conducted a series of six listening sessions with experts that covered the following areas: employment, jobs, and the future of work; child and family retrospective; physical health, environment, and climate; mental health; innovations in international surveys; and think tanks, research organizations, and non-profits. In addition, we widely circulated a survey to the community of potential users of an NLSY26, resulting in 412 responses.

Retrospective Analysis: We conducted a bibliometric analysis to identify the research topics for which the NLSY data has been most often used, and the keyword or keyword phrases most often appearing in the abstracts of published articles that have used NLSY data. To supplement the bibliometric analysis, we also conducted an analysis of variables in the NLSY97 which are underused by comparison with the rest of the dataset. Finally, we compared the NLSY to several non-BLS longitudinal surveys in the U.S. and internationally, as well as to other BLS household surveys.

Assessment of Alternative Data Sources: We assessed 20 alternative data sources that could be integrated into the NLSY26 to improve the accuracy of data collected, reduce respondent burden, or expand the scope of the survey content. Each alternative data source was assessed according to relevance, accuracy, coherence, feasibility, and data access issues related to confidentiality and consent.

Throughout the course of these activities, an Oversight Committee of experts were consulted in order to refine our approach to these activities and provide input on how these activities inform the design of an NLSY26. At the end of each activity, separate reports have been prepared for BLS. This document synthesizes across these three activities to make recommendations and does not replace those reports.

3. The Unique Role of the NLSY

There is clear demand for a new NLSY26 cohort.

There is a need for a rigorous nationally representative longitudinal survey covering broad domains of life, with adequate birth cohort sample sizes to study disparities and mechanisms of labor market success over the life cycle.

American society has changed in significant ways since the previous NLSY97, and a new NLSY26 cohort will shed light on the different experiences and challenges facing this new cohort of youth.

The ability to compare across cohorts is another key strength of the NLSY program, and an NLSY26 will enable further cross-generation research.

Relative to many administrative data sources, a new NLSY cohort will provide a much wider breadth of domains, collect data that can only be gathered in a survey (e.g., cognitive ability, mental health, expectations), and can fill gaps in knowledge for key domains (e.g., measuring hours worked).

A successful NLSY26 will be widely used by researchers seeking to address the research questions at the core of a new cohort, which means that the cohort will need to distinguish itself from among the alternative data sources available to researchers. Over the several decades of the NLSY, the landscape of data sources available to researchers has changed dramatically and continues to do so.

When the last cohort of the NLSY begun in 1997, detailed data to understand the US labor market was limited. The most used data sources were cross-sectional surveys such as the Current Population Survey (CPS) or short panels such as the Survey of Income and Program Participation (SIPP). Both of these datasets are relatively large, but they sample workers across the lifespan, making the sample of young workers of modest size. While these surveys remain important sources of information, because of the short follow-up period they are not well suited to studies requiring lifecycle data such as how decisions and opportunities early in life affect an individual's dynamic path in the labor market and in other aspects of life. And because they sample individuals across the life cycle without significant collection of retrospective information about experiences prior to sample recruitment, they are not well suited to study the cumulative contribution of opportunities and choices (e.g., the total labor market experience of workers up to the age of observation).

Perhaps the only alternative panel dataset collecting extensive data on labor market outcomes on young workers was the Panel Study of Income Dynamics (PSID). Importantly, the PSID studies the children and all the dependents of the respondents of original households, and it has become one of the most important data sources for understanding intergenerational mobility in the U.S. But, as important and as influential as that dataset was and remains, it was (and remains) limited for the purpose of studying contemporary labor force issues. By its design, the sample cannot change as the American workforce changes except in ad hoc ways.[1] In addition, the descendants of the original sample span a very wide age range, making the sample of young individuals in a narrow age range small.

Of course, the NLSY79 was a pre-eminent data source about labor market experiences in 1997, but that sample was aging out of labor market entry into later life stages. For an NLSY26, the NLSY97 and NLSY79 are similarly important reference data sets, but they describe older cohorts whose experiences cannot be taken to represent a generation almost thirty if not fifty years younger. A new cohort allows for updated content that reflects changes to American society that have occurred since 1997 and were not captured in prior cohorts. In addition, content from the prior cohorts that is included in the NLSY26 will allow for a unique opportunity to examine how the experiences of this cohort differ from prior generations. Cross-cohort comparisons are a traditional strength of the NLSY program, and an NLSY26 will enable further research to shed light on how societal changes since 1997 have affected American youth.

In terms of this newer generation, our comparison of NLSY to other surveys suggested that the best longitudinal coverage may come from the Transition to Adulthood (TAS) supplement to the PSID. Nonetheless, due to the genealogical design of the PSID, the TAS will likely have a substantially smaller sample size and is unable to be as representative of the U.S. population as the NLSY26.

Large administrative datasets, such as the Longitudinal Employer-Household Dynamics (LEHD) Program, linked Federal Tax Records, and data from private data vendors such as Burning Glass, have

[1] For example, in 1997 and 2017 supplemental samples of Immigrants were added to the study.

captured the attention of researchers and have been increasingly important in our understanding of the labor market. These files have vast numbers of records. This large size not only allows analysts to define narrow birth cohort ranges to follow over time, but also conduct detailed analyses such as those that rely on variation within firms or neighborhoods. While use of these data sources is increasing among U.S. researchers and researchers worldwide, the limitations are widely recognized. These data files were almost always captured for a purpose other than research; the content is almost always limited to the data that was needed to administer the program for which the data was collected, so broader coverage is not available.[2] In addition, even within a topical area (e.g., employment, fertility), key pieces of information are not collected because they are not necessary to administer the program. For example, none of the three comprehensive administrative records collections addressing earnings in the U.S. – Social Security Earnings records, IRS tax records, or Unemployment Insurance records – collects the number of hours an individual works. Many other important factors, including detailed cognitive and non-cognitive skills and the multitude of factors leading to their development, affect labor market choices but are unlikely to ever be captured except by a designed data approach and direct contact with sampled individuals.

There has also been a revolution in both laboratory and field experiments studying a host of labor market issues. These include experiments that manipulate race or gender to study discrimination in hiring (audit studies); government-funded social experiments, such as Moving To Opportunity, that manipulate subjects' residence to study the effects of relocating poor families to higher socioeconomic status (SES) neighborhoods; and a host of smaller experiments including many that manipulate worker incentives to study workers' choices of effort. These studies have high internal validity where causal claims can confidently be made. The growth in these kinds of studies is partly a response to the "identification revolution," the movement beginning in the early 1990s, which emphasized design-based research with well understood variation in factors that are plausibly exogenous but that change opportunities or choices under study. Generally, however, these studies are narrow and opportunistic, and there is debate on their generalizability to the labor market as a whole.

The NLSY program has served and continues to serve an important role in this overall data ecosystem. It is characterized by collecting a large sample on a narrow set of age cohorts and following them over their lifecycle, covering a range of topics. The 1979 and 1997 cohorts, for example, collected extensive data on education and employment. Less detailed but still considerable was data collected on each respondent's household composition; history of marital status changes; sexual activity, pregnancy and fertility outcomes; income, assets and program participation; and attitudes and expectations in many domains. The 1997 cohort has become the best source of data on criminal involvement and substance use that can be linked to other aspects of life. While it was never designed as a health survey, it has some data on health outcomes, although this is limited relative to health studies like the National Health Interview Survey (NHIS) or the National Health and Nutrition Examination Survey (NHANES). Health scientists in recent years have become the fastest growing group of users of the NLSY data infrastructure. There is no other national representative survey that is of the size, scope, and planned length to study what leads to disparities in labor market outcomes over the life cycle for its generation.

Compared to other BLS household surveys, the NLSY cohorts have the smallest sample sizes and represent the narrowest portions of the population. At the same time, the NLSY surveys are unique in

[2] The increasing use of data linkages has been to address the shortcoming that most administrative data systems have limited data elements, typically only in one area (employment, health, etc.).

their topical breadth and the duration of their longitudinal coverage of individuals' lives. To the extent that there is content overlap between the NLSYs and other BLS surveys, the differences in samples and timeframes mean that the data are almost never redundant. Rather, content overlap in the NLSYs with other BLS household surveys offers opportunities to use the larger surveys to corroborate NLSY estimates, and to use the NLSY data to dig deeper into why we might be seeing the behaviors documented in the larger surveys.

While some kinds of data can only be collected by a well-designed instrument administered by a survey taker, there is much to learn from the choices the research community has made in adopting other data sources for their work. A design that increases the chance of quasi experimental variation that leads to variation in labor market outcomes reflects the identification revolution. Linkages to alternative data sources that would increase the useful constructs about schools, neighborhoods, and workplaces leading to better measurement in the new cohort seem wise. No doubt the added data detail would also help scholars using administrative records interpret the bigger but coarser datasets. Finally, the rise in interest in the health sciences as with the proven advanced instrumentation from health surveys may broaden the user base and provide new insights into the relationship between health and human productivity.

4. Central Research Topics for an NLSY26

The NLSY program has historically enabled a wide range of cross-disciplinary research and therefore supports a variety of research questions.
Areas such as labor market behavior and educational attainment are historic strengths of the NLSY program and will need to remain key measures in the NLSY26.
As societal trends and the frontiers of research areas continue to evolve, new research questions will continue to emerge and add to the diversity of topics studied by NLSY users.
Research areas such as intergenerational transmission, inequality, structural racism, and the role of place in economic outcomes will be important for an NLSY26.

Our bibliometric analysis suggested multiple areas in which prior NLSY cohorts have been heavily used:

> "Labor market" and "educational attainment," as expected, were among the most common two-word phrases appearing in the abstracts of articles in the bibliometric database. In addition, two-word phrases related to health (e.g., "mental health" and "depressive symptoms") also featured prominently in article abstracts. While the health research results may be somewhat skewed by the Children of the NLSY79 cohort, this signals opportunities for health research in a stand-alone youth cohort as well.
>
> A network analysis of keywords occurring in tandem revealed a large cluster of traditional NLSY topics including income, earnings, employment, marriage, and work. This cluster also included terms such as race, women, and inequality; likely reflecting the utility of the NLSY for studying racial and gender disparities in labor market outcomes. This analysis also showed a significant linkage between children and health, based on keyword co-occurrences.

In terms of topical trends, our analysis showed that the NLSY data has long been a choice for studying topics related to, for example, demography, epidemiology, and unemployment. In addition, the data have more recently begun to be used for studying issues like gender gaps, disparities, job skills, technology, and incarceration.

Our stakeholder outreach activities confirmed the great variety of research topics relevant to an NLSY26; some high-level research topics tended to come up many times. Particularly frequently mentioned topics included the following:

How schooling and training contribute to human capital development
Changing nature of work, particularly related to the effects of informal work arrangements, automation, and changing wage structures
Family formation
Decision-making about schooling and work choices
Intergenerational transmission (of skills, education, assets, among others)
Inequality, especially inequality in income, wealth, and wages
Racial disparities and structural racism
The role of place in economic and health outcomes
Impacts of the Covid-19 pandemic
Health and mental health

Some of these topics have been core to the NLSY program since its inception more than four decades ago, while others reflect more recent developments in society and social science research. Moreover, almost all these topics can stay within domain, cross domains, or link to labor market outcomes.

It is likely that that any design of the NLSY26 would address each of these topics to some extent. Keeping in mind consideration #1 above to align with the mission of the BLS, it would make sense to focus each of these in terms of their relationship with labor market experience. For example, the literature on family formation might include questions about the characteristics of cohabitations that lead to marriage or relationship dissolution, but these may not be central to labor markets. Questions of which cohabitations do or do not exhibit "marriage" wage premiums or penalties, would be more related to labor market experiences and so higher priority for inclusion in an NLSY26. Similarly, questions of the impact of the COVID-19 pandemic on residential housing choices are of great relevance to our society but may have limited value for understanding employment. In contrast, if widespread disruptions to schooling during the pandemic have compromised the skills and workforce readiness of a future generation of Americans, that would be of first order importance for an NLSY26 to assess.

When content from these various topics speaks to multiple domains, that will improve the cross-domain relevance of the NLSY. Nonetheless, this should not be the only consideration and the considerations noted above in Section 2 can help to prioritize content across domains. For example, some content in training and human capital development may be relatively domain specific but may not be able to be duplicated in other data sources and will speak to the mission of BLS.

5. Factors to Consider in Deciding the Content of an NLSY26

Available interviewing minutes will not meet demand in the NLSY26. Our explorations identified no domains from prior NLSYs that had receded in importance, In addition, skills and employment will be more complex to measure (and hence require more items) than in the NLSY97.

Survey designers will be faced with tough choices regarding content and can prioritize by considering the quality of individual measures, the availability of the data from other sources, as well as their ability to support longitudinal analysis, enable cross-cohort comparisons, be relevant to labor market activity, and provide for cross-domain research.

While cross-cohort comparisons will be extremely valuable, there are some domains that will need to be updated.

Examples of these domains include marital/cohabitation status, gender identity, race/ethnicity, disability status, and cognitive ability.

The distribution of administration time across NLSY domains will likely need to shift (for example, relative to the distribution in the NLSY97).

Options for accommodating more content in the same length questionnaire include relying on alternative data sources to expand the content that can be gathered, reducing coverage of all domains, reducing the depth/granularity of coverage in some domains, and reducing the frequency of coverage in some domains.

Our comparisons to other longitudinal surveys illustrated the broad variation in content coverage within a topic area that can be possible even among federally funded U.S. longitudinal surveys, depending on the goals and priorities of each survey. The content contained in each survey reflects the topical emphasis of the particular study, its other design features (such as school-based samples versus household samples), and other study-specific assets and priorities. Similarly, an NLSY26 design will need to reflect its own priorities and goals. Nonetheless, the information gathered in the needs assessment points to a few key considerations that the NLSY26 will need to consider when designing new and altered content for deciding the content of the NLSY26.

5.1 General Factors for Selecting Content

Level of Detail. A key objective for content in the NLSY (as described in design consideration #13) is to collect data of sufficiently high quality to support meaningful analyses. In addition, managing respondent burden in order to maintain high response rates was also noted as a key consideration (consideration #5). Aligning with these considerations, one specific area related to data quality that was raised during our stakeholder outreach activities was the granularity of data collection An NLSY26 will need to trade off the need for detailed information with the potential length that comes with extremely detailed data collections such as event histories. In the words of one participant, "it seems there are some questions at the end of complicated skip patterns may ultimately be rarely used. The most valuable detailed questions are those required to create composite measures such as net worth and those that are required to create true time series questions." For many topical areas, the NLSY97 contained very detailed data on the exact

timing of events and a large amount of data on factors surrounding those events. Examples of this detail include weekly or monthly event history arrays on school and employment as well as a large set of factors surrounding job transitions and schooling transitions including semester to semester. Collecting this data is time-intensive, and one potential method to reduce the length of the interview is to shorten these arrays. This can be accomplished by either reducing the amount of detail collected on the timing of events, or by reducing the level of detail collected on the events and transitions while still maintaining the same level of detail on the timing of events themselves. The availability of alternative data sources may change the granularity of data collection as well.

Possibility for Accurate, Reliable, and Unbiased Data. As is implicit in consideration #13, data sources vary in their ability to deliver accurate, reliable, and unbiased data. For example, despite enormous effort to record the timing of program participation, one participant specifically noted research demonstrating the large amount of measurement error when program participation is recorded from respondent reports, and how administrative information has provided a more accurate picture of program receipt and at a lower cost. Even when a topic is of high salience to the NLSY26 purpose, measures should be selected (whether in survey content, administrative data, or through other means) so that they can be used for statistically valid analyses over the life course as described in consideration #5.

Ability to Support Longitudinal Analyses. Another key consideration for an NLSY26 when deciding on content is the ability of the items to support longitudinal analyses. One distinguishing feature of the NLSY surveys is the long-time horizon over which data is collected, and many participants suggested prioritizing data elements that will help to shed light on the evolution of sample members across the life course.

This has two implications. First, BLS should prioritize items that can be asked multiple times. Considerations #5 on statistically valid analysis over the life course and #11 on not duplicating data available elsewhere may point toward prioritizing content that will primarily be used for longitudinal analyses. Cross-sectional measures, even closely related to labor markets, may not be as useful to include in the NLSY26, whether due to relatively small sample sizes, difficulty for users to find and use the data, or simply because items with longitudinal relevance tend to have greater comparative value for inclusion in an NLSY relative to other data sources. This was also a consistent theme in the stakeholder outreach, and multiple participants noted that, for example, questions related to respondents' physical and mental health being asked more consistently would have enabled more longitudinal analyses. Second, some content is inevitably going to be more valuable for longitudinal research but may not be immediately used. Especially in early rounds, some items may serve as a "baseline" that only becomes useful several rounds in the future.

Cross Cohort Comparisons. Design consideration #6 specifically calls out the importance of cross-cohort comparisons for the NLSY program, and indeed a consistent theme coming from our stakeholder outreach activities was that a unique value of the NLSY program is its ability to perform comparisons across cohorts. Nonetheless, maintaining content from previous cohorts will need to be considered alongside consideration #9 that the NLSY26 reflect important changes since 1997. This is corroborated by the results of our listening sessions and user survey, which both indicated a high demand for collecting data on new phenomena or using new measures which would make cross-cohort comparisons more difficult.

The tension between over-time comparability versus improved or more relevant measures may be particularly salient for issues relating to traditional NLSY topics such as the transition from school to the labor force (as mentioned previously, "labor market" and "educational attainment" were among the most common two-word phrases appearing in the abstracts of articles using prior cohort data). On the one hand, the prior NLSY cohorts have exceptionally rich and well-used data in these areas, so the value of comparability is substantial. On the other hand, many of those data types, such as weekly or monthly event history arrays or data on every school or job, are extremely time-intensive in the interview, making it difficult to retain comparability while also adding new measures. Adding to the tension is the importance of the NLSY26 accurately measuring how work and school are being experienced by this generation; comparability of a 'usual hours worked' measure may not be so valuable if gig work, remote work or other features of the modern labor market mean that many workers are not working stable numbers of hours from week to week as a usual hours measure assumes. Similarly, the continued popularity of online schooling at the secondary and higher education levels with variable intensity and duration may weaken the meaning of 'month-by-month enrollment status.'

A host of new work and educational arrangements have emerged and the listening sessions and user survey both indicated a high demand for these items. The content design of the NLSY26 will need to consider this value and navigate the trade-offs between including content that speaks to new and emerging research questions and maintaining the ability to perform cross-cohort comparisons.

In addition, there are a host of concepts that have evolved in important ways since the fielding of the NLSY97 and will necessitate updated measures. These include marital/cohabitation status, gender identity, race/ethnicity, and disability status. In addition, given the increased interest in cognitive skill measurements it is likely that updated measures of cognitive ability throughout the life course may be needed in a new cohort, perhaps in addition to an early fielding of the Armed Services Vocational Aptitude Battery, which would provide unparalleled opportunities for cross-cohort comparisons. Implementation of cognition measures in the NLSY26 will respond to considerations #6 on cross-cohort analyses, #9 on societal changes, #10 on age-relevance, as well as #5 on statistically valid analysis and #13 on data accuracy.

Considerations Due to Sample Size. The sample size of an NLSY26 may not be sufficient to measure rare events or understand the experiences of numerically small subgroups to create statistically valid analyses of these respondents over the life course (see design consideration #5). The size of the NLSY26 sample is likely to be small relative to large cross-sectional surveys such as the labor-focused CPS or the health-related Youth Risk Behavior Surveillance System (YRBSS) survey, or to short-duration panels such as the SIPP or the longitudinal surveys sponsored by the National Center for Education Statistics (NCES), where such questions may be more appropriately addressed.

Estimated prevalence of events can be helpful in articulating survey priorities. For example, the experiences of American Indian and Alaska Natives (AIAN) are an important domain for research for our society. Certainly, it will be important for the NLSY26 to record the race and ethnicity of all sample members, even those who are members of rare groups. Nonetheless, given that the percent of the U.S. population that is AIAN alone or in combination with another race is only roughly 3%, the NLSY26 may not likely be a useful data source for understanding the nuanced issues affecting these individuals and questions focused on this population are unlikely to be valuable.

Similar considerations of expected sample size apply to almost every domain. For example, which disabilities or health conditions are worth documenting, and which might merit further attention because they could be studied within the data beyond estimation of prevalence?

Involving the User Community. Multiple participants noted that involving the user community in design conversations could be an important consideration for the NLSY26. Mechanisms such as the technical review panel and other increased user engagement will provide multiple benefits to the NLSY26 program. Involvement of the user community could help the NLSY26 to achieve consideration #3, which is to promote working relationships with key users and organizations whose work is consistent with the BLS mission and strategic goals. It will also help the NLSY26 to stay on top of developments in the research literature in order to maintain value to the research community as respondents age. This value from this engagement will come at least as much from its contribution to overall study approach and evolving priorities as it does to individual contributed items.

Harmonizing to Other Surveys. Multiple participants in our stakeholder outreach suggested that the ability to have harmonized measures between surveys provides value. This could include other U.S. surveys – two primary examples brought up were including questions that align with major surveys (such as the employment questions in the Current Population Survey) or including measures that span across age groups (for example, constructing "bridges" of particular skill measures between the Early Childhood Longitudinal Study-Kindergarten and the NLSY26).

In addition, multiple participants suggested that harmonizing to international surveys could enable useful cross-country research. For example, gig employment questions in the Understanding Society – The UK Household Longitudinal Study and robust cognitive testing measures in surveys such as Growing Up Australia and the German National Education Panel Study (NEPS) may serve as useful models for the development of similar questions in a new NLSY survey.

Harmonization efforts may offer collaborative opportunities with other surveys and sponsoring agencies as in consideration #3, and also indicate how intentional duplication can sometimes be valuable for an NLSY26 (consideration #11).

Items that Support Respondent Engagement. One possibility noted in the listening sessions is that some items that respondents find interesting can encourage their ongoing participation, even if those items may not otherwise qualify as high priority based on analytic value. For example, the "Growing Up in Ireland" study got press coverage pick up from asking what 20-year-olds hope to achieve by age 30 (buying house, grad school, job, family). These types of items can be one strategy by which the NLSY26 can achieve design consideration #5 to enable the maintenance of sample sizes and response rates that allow for statistically valid analyses over the life course of respondents.

5.2 Factors by Topic Area

In the tables below, we describe some of these factors that are specific to individual topic areas but can help to inform the content of a new NLSY26 cohort. Across the activities of the needs assessment, there were few areas identified for reduced coverage relative to the NLSY97 and several areas identified for expanded coverage. Although we did not identify any topics that are no longer important, there were

consistent themes that weak coverage – such as measures collected with low frequency, measures with high measurement error, or outdated measures– would lead researchers to not use collected data. This implies that a challenge for the NLSY26 will be to choose topical areas for robust coverage, rather than spreading thin available questionnaire time in ways that may leave larger numbers of topical areas inadequately covered.

Table 2 Factors Regarding Youth Content by Topic Area

Topic	Factors Regarding Survey Content	Factors Regarding Non-Survey Primary Data Collection and Alternative Data	Potential Research Topics
Labor Market	Employment and job characteristics are traditional strengths of the NLSY. Continuing to collect detailed data such as week-by-week event history of jobs will likely be important in an NLSY26 but will consume a significant amount of survey time. The NLSY26 will need to consider forms of informal work as well as accounting for potential job training and skills development that do not come from formal employment.	Alternative collections such as event-triggered data collections may be useful for collecting information on job changes. Linking to data on businesses could open additional lines of research. Strengths of the NLSY26 relative to available administrative data will likely lie in its ability to a) measure job characteristics such as schedule flexibility that are not traditionally covered in administrative data; and b) to link to other measures not traditionally available in administrative data (e.g., skills, health, education, childhood experiences).	How prevalent and important is informal work in the careers of respondents over the life cycle? How does job tenure change over the life cycle, and how does this compare to previous cohorts? How do wages change when workers change jobs? How does this vary over the life cycle and depending on the life context of the respondent?
Skills	In addition to cognitive skills that have been historical strengths in prior NLSY cohorts (e.g., Armed Services Vocational Aptitude Battery (ASVAB)/Armed Forces Qualification Test (AFQT)), non-cognitive skill measures are increasingly relevant for researchers. Having multiple measures of skills to track the progression of skills over the life course would enable more longitudinal research.	Alternative data collections such as cell phone apps may be particularly well suited for measuring skills. Other comparison surveys have collected a variety of skills and competencies that could serve as useful examples, including surveys by NCES or the German NEPS. Skills are an area that is not well covered in alternative data sources and could be an important contribution for an NLSY26.	How much of the labor market returns to human capital is explained by increases in skills? How do cognitive and non-cognitive skills evolve over time, and how are they acquired? What are the factors associated with skill formation? What is the effect of parental investment in children's cognitive and non-cognitive skills?
Education	Education data can be collected in varying levels of detail. Examples of potential detail could include the following: School year * school attended combination	Many states have K-12 administrative data systems that could present linkage opportunities. For postsecondary education, direct data linkage opportunities exist from the	What are the returns to education and does this compare with past cohorts? How are schooling decisions related to other contextual factors

	College term Month-by-month event history of enrollment	National Student Clearinghouse and the National Student Loan Data System. Providing school identifiers allows for researchers to link to the Common Core of Data and the Integrated Postsecondary Data System. A new NLSY26 will need to consider the relative importance of nationwide survey data to education research given the increased availability of administrative data providing detailed progression of students through schools. Different levels of sample clustering or other sampling decisions could make collection and linkage to high school transcript data more or less difficult.	and life events, such as the birth of a child?
Behavior	Many traditional topics from past NLSY cohorts such as substance use will continue to be important for the new cohort. Some aspects of behavior such as social media use will be new to the new generation surveyed in the NLSY26.	Time use data or measures based on tracking of respondents' cell phones could provide useful measures of behavior. Many behavior measures are difficult to obtain in administrative records, so this is a domain where survey-based measures are particularly important.	What is the relationship between anti-social behaviors in adolescence and educational and labor market outcomes?
Criminal Justice	The NLSY97 is historically a strong source of information on engagement with the criminal justice system. A variety of measures related to criminal justice are important for enabling research into structural inequality (e.g., interactions with police officers).	The Criminal Justice Administrative Records System has compiled records from many states and serves as a potential linkage opportunity. Nationwide administrative data on exposure to the criminal justice system is still not widely available to researchers. Even administrative data that is available is limited in focusing on measures such as arrests, and survey data can provide a wider picture of interactions with the criminal justice system.	How are encounters with the criminal justice system related to education and labor market outcomes? What factors—such as education and family background—are associated with encounters with the criminal justice system?
Health	It will be important to separate measures of health at interview date from retrospective health	Biomarker data collections can provide a richer set of measures than would be	How do chronic health conditions in childhood and adolescence affect

	information – both have important uses and an NLSY26 will likely include a mix of both. Importance of mental and physical health – this is likely an area where an NLSY26 will contain more content than the prior cohorts. Frequent repeated measures of health are important for enabling longitudinal research. A robust literature has demonstrated the importance of social determinants of health, and a cross-domain survey such as the NLSY26 is well poised to speak to questions in this area.	obtained from only survey-based measures. Electronic medical records can provide detailed information, and administrative records on claims for Medicaid and other government programs also exist. One advantage of nationally representative survey data is its ability to represent the entire population. This can be an issue with health administrative records – it is well known that interaction with the health care system varies by socioeconomic status, and for programs like Medicaid eligibility criteria can vary substantially across states so survey-based measures provide important coverage across the U.S. population.	labor market outcomes as an adult? What is the association of industry/occupation with health outcomes?
Fertility	Participants in stakeholder outreach activities noted that fertility expectations were an important component of past NLSY cohorts.	Because of the lack of nationwide repositories for information such as birth certificates or marriage licenses, direct linkage to fertility will likely be limited. If the NLSY26 were to be available in a Census Federal Statistical Research Data Center (FSRDC) environment this could enable some direct linkage to household composition in decennial censuses.	What are the factors associated with premarital childbearing? How does premarital childbearing affect labor market outcomes and wealth accumulation? What is the relationship between teenage (or early adult) childbearing and high school (college) completion?
Assets/Debt	Participants noted that previous NLSY cohorts have high-quality assets and debt information and suggested that this continues in the future. Societal changes such as new credit instruments will mean that new measures of assets and debt will likely be important for the new generation surveyed by the NLSY26.	Student debt is of particular interest if feasible to gather, as these data sources are often different from those tracking other assets and debts. Assets and debt remain a domain in which administrative data is not widely available for researchers, and the NLSY26 will ideally be able to enable important contributions to research in this area.	What is the effect of assets and debt on co-habitation/marital decisions and outcomes? What is the effect of assets and debt on retirement decisions? What are the factors associated with wealth accumulation and financial well-being throughout individuals' lifecycle?
Environmental Exposures	Prior cohorts of the NLSY contained relatively little in the way of measures of environmental exposures but	Air sensors could potentially provide a more refined measure that would be	How are various environmental exposures—such as to lead—related to later

	given the increasing realization of the importance of these exposures in explaining labor market outcomes this may be an area for increased measurement in the NLSY26. Some survey measures could enhance the use of alternative data on environmental exposures – for example, knowing if the respondent works outside where they might be more exposed to poor air quality.	obtained from geographic aggregates. Potential other environmental measures may also be able to be collected (e.g., water and soil samples). A wealth of information is available on environmental exposures across geographic areas. Using this information with an NLSY26 requires detailed geographic location to be available. Relative to research using only administrative records, the comparative advantage of the NLSY26 in this domain would likely revolve around its ability to measure precise exposures for respondents and being able to trace those impacts over the life course.	life outcomes in education and the labor market? How do environmental exposures evolve over the life course, and how is this explained by changes in life context such as acquiring a new job?

Table 3: Factors Regarding Parent Content by Topic Area

Topic	Factors Regarding Survey Content	Factors Regarding Non-Survey Primary Data Collection and Alternative Data	Potential Research Topics
Family Structure	The NLSY97 collected detailed family structure retrospectives. But the complexity of family structure continues to evolve. A key question is which family members to document and potentially collect data on; non-custodial family members (fathers); other children of non-custodial family members; non-biologically related individuals in the respondent's home (stepchildren/stepparents) and father figures. A second key question for the NLSY26 will be what level of detail is best to collect. Given complexity of situations and data quality challenges, could collection of summary status at key dates serve as a satisfactory	Relatively few alternative data sources on family structure are available for direct linkage, though data in an FSRDC environment might be able to provide more intergenerational information from direct linkage. State-level information from child-protective services and foster care would be invaluable although myriad challenges exist in accessing and processing these data. More generally, because there are relatively few administrative sources on family structure this is a topic that may merit significant administration time in the questionnaire(s).	What is the effect of family structure on the returns to education? What are the educational, behavioral, and labor market outcomes of children raised in single-parent families?

	substitute for full event histories?		
Parental Labor Market Information	Labor market circumstances of parents will be important to collect in an NLSY26. Many of the same considerations related to labor market apply similarly to parent interviews as they do to youth interviews. Factors such as work flexibility for parents could be important for youth development. Schedule flexibility and autonomy has been asked in the NLSY97 for youths' own jobs and could be a valuable addition to an NLSY26 parent interview. Both (limited) retrospective and (more extensive) prospective information on parental employment could be relevant.	Consent for linkage to administrative information would be an important piece of a parent interview if that were planned. Given that consent could be obtained directly from parents, linkages to labor market information such as social security earnings may be more feasible early in the NLSY26 than linking to information for respondents.	What is the relationship between family income and youth labor market outcomes? How do expectations and experiences related to education differ between youth from different family income backgrounds?
Family Financial Situation	In addition to the considerations laid out above for the youth interview, an NLSY26 will need to contend with the best way to measure financial situations for households. It is important to remember that financial situations in households can change year to year, so multiple measures of household financial situations will tell a more complete story than single-year measures.	Linkage to Social Security Administration (SSA) data provides a potentially valuable opportunity to characterize parental earnings over time. Where family structure evolves through death or divorce, change of custody, or new family formation, care will need to be taken to ensure that linked data remain relevant to the life of the sampled youth (as opposed to describing a non-custodial parent whose resources are not available to the youth, for example).	What is the extent of intergenerational wealth transmission? What effect does parent resources play in the career trajectories of youth?
Health of Youth	The parent interview may be a good place to ask about the youth's health history given that parents are likely to be a better reporter on this than the youth themselves. As with other retrospectives, the parent survey will need to consider the amount of detail to collect (e.g., collecting event histories or focusing on big events/health at particular ages) and whether the parent	Electronic health records and claims data present some linkage opportunities but will miss certain segments of the population.	What is the impact of childhood and adolescent chronic health conditions on educational attainment and labor market productivity?

	is the best reporter in all cases.		
Health of Parents	Relative to health of the youth, health information for the parent may be able to be limited to summary measures, primarily those indicating limitations on financial resources.	Electronic health records and claims data present some linkage opportunities but will miss certain segments of the population.	What is the incidence of intergenerational persistence of health? How do parental health conditions when the respondent is a child affect later life labor market outcomes?

Note that for many content areas there is a question about who the appropriate respondent should be. While parents in traditional family structures may be more accurate than the respondent at reporting retrospective information on domains such as health of the child and family structure, it is an open question as to the best choices considering parent surveys may not achieve full coverage. Specifically, for subpopulations of interest like system-involved children, children may be able to provide more complete information about the family structures they have experienced.

6. Sampling and Data Collection Choices

The results of our activities suggest a structure similar to the NLSY97 for starting age, frequency, and mode.

Strong reasons exist for starting earlier in children's lives or starting closer to their labor market entry and when they can complete the ASVAB with validity. No alternative is clearly preferable to the 12-16 starting age posited by BLS initially.

In terms of mode, in-person contact with all youths' households is likely to be valuable for retention. Given changes in youths' attitudes toward technology and social norms toward in-person contacts, BLS may want to allow for some modules to be completed outside of in-person interviews, even in the first round of data collection.

The NLSY26 will need to prioritize oversamples for consideration.

Choice of oversamples will need to be guided by research priorities, but also informed by the availability of external data and the additional cost.

Given the importance of these themes and the additional value from having multiple measurements, multiple parental interviews would make a valuable contribution to an NLSY26.

These will enable research into intergenerational transmission and how the development of the respondent effects later life outcomes.

BLS will need to consider a number of key sampling and data collection choices when constructing the NLSY26. Below, we discuss a few of the key considerations that have been illuminated by the needs assessment along with potential implications for other aspects of the survey program.

6.1 Cohort Composition

The sample size of the NLSY26 will clearly be a key decision for BLS. Participants noted that analyses of small subgroups can be difficult with existing NLSY cohorts. Additional sample size will become cost

prohibitive after a certain point. This will create clear tradeoffs for the design of the NLSY26, as including questions that are specific to a particular subgroup will only be valuable to researchers if there is a sufficient sample size available.

In terms of age ranges, the NLSY97 age range of 12-16 was generally supported throughout the needs assessment. During stakeholder engagement, some participants suggested that moving the age range lower provides more ability to obtain information about early life determinants of labor market success. Nonetheless, with the focus of the NLSY on labor market outcomes, too early of an age range will mean that there is more time before the cohort reaches the labor market. Note that a wider age range will go hand in hand with having more siblings in the sample, and this was noted as a value of the NLSY program in the needs assessment. Also related to starting age is the possibility of fielding an ASVAB in conjunction with the NLSY26, since the ASVAB is documented as valid for a relatively narrow range of older adolescents and young adults. The value of the ASVAB and its role in distinguishing the NLSYs from other data sources suggests that ensuring that an ASVAB is feasible should be given higher priority than other research questions that might be facilitated through a different age range.

Finally, a third piece of the sample composition is the geographic distribution of the sample across states. Many participants in the stakeholder outreach activities noted the potential use of the NLSY to compare respondents across geographic areas. This is particularly true of research using "natural experiments" comparing changes over time. An NLSY26 that includes data from a diverse range of geographies could support this type of analysis but given the likely sample size of the survey it is highly unlikely that it will be able to create the robust samples within small geographies that might be necessary to support studies of local policy variation. Given the possibility of state-level data linkages, and the ever-present desire of researchers to use state-level policies in analyses, distribution of the sample across states may be more important to consider than would be typical for a nationally representative sample of this size (where states might not be considered as a factor in the sampling design).

6.2 Frequency of Data Collection

The needs assessment demonstrated the utility of frequent data collections to researchers. In the user survey, 69% of survey participants answered that they would like to see data collected every year. Nonetheless, more frequent data collection comes with clear costs; a model similar to that of prior NLSY cohorts, where survey frequency moved from annual to biennial as the cohort aged, might provide a useful balance.

During the stakeholder engagement task, there were some requests for more frequent data collection, especially those that could be used for alternative collection methodologies such as experience sampling. Off-cycle data collections might be able to provide some additional value to researchers and assist with panel engagement but would incur an additional cost. For example, the UK Understanding Society Survey is currently performing a brief survey each month asking if a series of life events have happened in the last month (job change, marital status, children, acquiring property, changing residence, health event, and more). With experience sampling in particular, our activities uncovered no examples of large longitudinal surveys using this methodology, and therefore this should be treated as a potential area for consideration as the NLSY26 develops and not a core component of the survey in early rounds. A possible off-cycle data collection that might be most valuable is the collection of mental health status, which many

researchers requested to have more frequently, and is easily collected via self-administered modes. The NLSY97 COVID Supplement is a relatively large-scale example of such a data collection (and did include a mental health measure), which could be triggered by respondent-specific events like childbirth, population-wide events like a pandemic, or in-between situations like natural disasters that might affect some but not all sample members.

6.3 Mode

Some sort of mixed mode approach is most likely to present the ideal solution for the NLSY26. For one, in-person contact with all households is likely to be valuable for retention as historically the highest response rates for longitudinal surveys have been achieved using in-person interviews of respondents and their parents. Nonetheless, given changes in youths' attitudes toward technology and social norms toward in-person contacts, BLS may want to allow for some modules to be completed outside of in-person interviews, even in the first round of data collection.

In part, this conclusion follows from the listening session related to international surveys, where participants expressed that mixed mode approaches will likely result in the highest data quality, referring not only to retention rate but also representativeness of the sample. As an example, the UK Understanding Society Survey was almost exclusively in person in early rounds and has transitioned to a multi-mode data collection where prior to the Covid-19 pandemic as many as 80 percent of respondents completed the web instrument. Nonetheless, participants still noted that mixed mode helps to improve representativeness as certain age groups are more likely to respond via different modes.

Note also that a mixed mode approach may enable more flexible measurement techniques. For example, many other surveys have developed cognitive assessments or other self-administered instruments that can be administered in-person or on-line, but not via telephone interviewing.

6.4 Oversampling

Past cohorts of the NLSY oversampled Black and Hispanic households, and the results of the needs assessment strongly support continuing with this approach. The scientific need to study these populations is no less today than it was in 1997 and that since the fraction of children who are Black is similar to when the last sample was drawn the need for this oversample is justified.

In addition, the needs assessment also identified groups that would provide analytic utility and are feasible to sample. While all of these groups provide scientific value, an NLSY26 will need to weigh the additional cost of oversampling some or all of these groups against other factors, such as investments in the ability to recruit and retain potential panel members. We discuss some of the details of these groups below:

> **Asian Youth.** In 2020 about 5.4% of children were of Asian descent, and represent an important subgroup in the U.S. While it may be possible to cost-effectively sample the Asian American population, a great deal more work is needed to determine feasibility. Although an oversample that would facilitate reporting on this group is likely to be policy relevant, it is unclear how much insight can be gathered analytically for such a diverse group within the likely size of oversample.

Youth Living in Poor Areas. A sample of youth living in poor areas would greatly facilitate our understanding of intergenerational mobility and the characteristic of areas that promote generational mobility. The spatial distribution of mobility has attracted enormous interest by scholars with a prime example being the Opportunity Atlas. This project, which relies on linked IRS data, documents which census tracts have been especially successful at promoting upward mobility for low-income children. But we know very little about why. Having a sufficiently large sample of children growing up in low-income areas would greatly facilitate our understanding of the dynamics of upward mobility in the U.S. and how this plays out over multiple domains of life.

Youth Residing in Rural Areas. Interest in rural areas has also grown in importance in the literature as well. For many years, rural areas have been declining in population and human capital has been flowing out of rural areas into U.S. cities. Place based policy to aid rural areas is a constant policy debate. In addition, the opioid crisis has taken an unspeakable toll on rural America, partly as one of many forms of deaths of despair and partly related to specific failings of the rural health care system to deal with opioid addition. An oversample of adolescents growing up in rural areas would help us understand who stays in rural areas and who leaves to large cities; and it would help us form policy to aid Americans in low density areas of the U.S.

For other groups, while there exist geographic data that could be used for constructing an oversample it is likely that the oversamples will be very costly due to the low prevalence of the population. These groups include AIAN youth, immigrant youth, and more finely defined racial/ethnic groups (e.g., oversamples of different Hispanic or Asian countries of origin, recency of immigration, or preferred language). A more detailed assessment of costs will be needed in order to determine whether the costs would be worth the additional gain. We also note that if the NLSY26 were to invest in one of these oversamples, it would make sense to pair this with additional questionnaire content to better understand the circumstances of these individuals. Doing so might put additional pressure on the survey from a questionnaire time perspective, and therefore these groups are likely only feasible to sample if they represent a special priority to BLS.

Finally, another set of groups were identified as potentially important oversamples, but it is not clear how an oversample would be drawn. These include youth with disabilities, youth in foster care and other institutionalized settings, and LGBTQA+ youth. It is clear that many researchers would derive value from these oversamples, but the details of how the oversample would be performed are not readily apparent based on the results of the needs assessment.

As examples of how other major longitudinal surveys have used oversamples, the table below documents the oversamples used by the PSID, NEPS, Longitudinal Study of Australian Youth (LSAY), and Understanding Society. Note that a wider variety of oversamples are feasible in longitudinal surveys that begin with extensive sample frames because they are school-based or located in countries that can sample from registries.

Table 4: Oversamples in Other Major Surveys

Survey	Oversamples
PSID	Low-Income Families
German NEPS	Teacher Training Students
LSAY	Indigenous Youth

Understanding Society	Immigrants and Ethnic Minorities

6.5 Survey Respondents

There was strong support during the needs assessment for continuing to interview both youth and their parents early in the NLSY26. In addition, many participants noted the value of interviewing parents multiple times to be able to better measure family circumstances and track changes over time. There is a clear trade-off here in terms of how much time is allocated to the parent versus youth interview, but there was some support during stakeholder outreach for the idea of allocating more time to interviewing the parents in the early years and then shifting more time to interviewing the sampled youth in later rounds.

In addition, multiple parental interviews would provide a significant amount of value to the NLSY26 and would accomplish three valuable goals. First, researchers are often interested in relating parental characteristics to their children's characteristics when their children reach the ages of their parents. But because parents are interviewed fewer years than the respondents, the precision of measurement on parents is likely to be lower than on children. Income serves as a prime example. Relating a measure of permanent income of parents to children has attracted a great deal of attention since Solon's (1992) pioneering work in the PSID. But that work clearly shows that the correlation across generations depends critically on the accuracy of measurement of permanent income; and this accuracy grows in the number of years each generation is measured. Second, parents may have the ability to report on a host of developmental milestones and issues of children. The set of domains of development are large and it may not be possible to collect all domains of interest on child development as reported by parents in a single survey. Finally, there is a great deal of interest in relating changes in parental circumstances to changes in child outcomes. For example, understanding how schooling outcomes change when parents lose jobs or transition from relationships is of great importance to economists, sociologists, and developmental psychologists. Relating these changes in child outcomes to changes in parental outcomes requires multiple contemporaneous measurement of both generations.

As shown by the comparison to other surveys, many other survey programs have additional respondents. For example, many surveys by the NCES include teachers or guidance counselors. In addition, the PSID CDS surveys primary caregivers either 2 or 3 times. These types of additional surveys provide potentially valuable information (e.g., information on student behavior in the classroom) but could come with significant additional cost. In addition, these NCES surveys are based on school samples, making these types of surveys operationally easier than in a household sample, such as was traditionally used in prior NLSY cohorts.

There were also some requests for interviews of other family members such as spouses as the youth ages, but these were not as common during the needs assessment.

6.6 Other Sampling and Data Collection Considerations

Event Triggered Data Collections. Many participants noted that event-triggered data collections could potentially provide improved measurement for specific topic areas. For example, a short (off-cycle) interview could be planned whenever a respondent receives a school diploma or degree, hits the one-year

anniversary of a new job, or gives birth. These types of data collections may be able to provide higher-quality data for key events by limiting recall issues, but it is unclear what the expected response rates from these collections would be. Therefore, they may serve as useful ancillary data collections but are risky from the standpoint of being a core component of the survey in early rounds.

Alternative Measurement Modes. Environmental sensors, biomarkers, and other alternative measurement modes could serve as a useful complement to traditional survey collection. Examples of the use of these modes in major longitudinal surveys include biomarker data collection in Add Health, a physical health assessment in Growing Up Australia, and biomarker collection in the UK Millennium Cohort Study. Participants also noted the potential of using mobile apps for data collection or using sensors such as FitBits to track respondent activity. For example, the UK Millennium Cohort Study used a smartphone app to maintain a time diary of respondent activities. Nonetheless, these alternative modes would need to be carefully considered to ensure that they provided sufficiently high-quality and novel data in order to justify the cost.

Respondent Engagement. Respondent engagement is a particularly important aspect of maintaining a longitudinal panel. Participants stressed the importance of creating a partnership with respondents that makes the respondents feel like they are engaged in the study and not a research subject. If possible, they suggest the NLSY26 could consider focus groups with prior or future study participants to better understand how recruitment materials, gaining cooperation approaches, and even study content may be perceived by individuals within the sample and could be improved.

7. Alternative Data Sources and Their Potential Value

The NLSY26 should pursue alternative data, but the initial rounds' questionnaires should collect essential data until feasibility of consent and linkages has been demonstrated.

An NLSY26 should prioritize opportunities for researchers to link to other data sources, including individual-level administrative data from public programs, as well as private data sources such as credit bureaus, social media sites, or health insurance records.

Engaging agencies to provide access to data and seeking parental consent for linkages in early rounds would likely be good investments.

Despite the substantial advances in data availability and linkage, the initial rounds of the NLSY26 are safest if designed without assuming linked data will be complete and adequate.

Whether from stakeholder input, review of other longitudinal surveys, or thinking about research questions and content areas, alternative data sources (ADS) consistently emerge as a priority to be included as an integral component in the design of an NLSY26. In particular, engaging agencies to provide access to data and seeking parental consent for linkages in early rounds will be important components to ensuring the long-term benefit of ADS to the NLSY26. Nonetheless, because there are relatively few existing examples where alternative data has been used as a core component of survey operations, initial rounds of the NLSY26 should be designed without assumptions regarding the availability of these data.

Below, we discuss further considerations on choosing alternative data sources, specific data sources that were investigated in our ADS assessment, and the potential implications of using these data for survey operations.

7.1 General Considerations for Choosing Alternative Data Sources

Federal versus State Data Sources. One of the key considerations for the feasibility of an alternative data source is whether the data are available from a single, national repository versus data that would require data sharing agreements with multiple data owners (e.g., states). Multiple data owners will not only increase the difficulty of obtaining initial access to the data but will also increase the risk of discontinuation later in the panel. Turnover in state agency staff can lead to changes in that state's willingness or ability to continue to provide the required data. This distinction also bears on data accuracy and coherence, as data drawn from multiple data systems would require more cleaning and harmonization.

BLS could potentially tap into state-level data collection efforts at the US Census or Coleridge Institute, but this may still bring additional complications. Since they come from different data systems, state-level data may require a good deal of data cleaning and harmonization to make them suitable for integration with the NLS. Data elements may need to be harmonized not only to make them consistent across states, but also potentially over time within a given state.

Given the typical size and geographic dispersion of the NLS cohorts, near national coverage of the target population would appear to be necessary for most data use cases. While for certain program data substantial coverage of the recipient population could be achieved with data from several of the largest states, these types of uses would likely need to be limited.

Prior Data Use. Evidence of prior data sharing and data linkage for research purposes helps us evaluate the feasibility of integrating that data source in the NLSY26. Data sources that have been part of data linkage efforts at Census or National Center for Health Statistics (NCHS) should be more feasible candidates for integration with the NLS. US Department of Housing and Urban Development (HUD), Social Security Administration (SSA), National Death Index (NDI), and Medicaid records are currently part of the NCHS data linkage program, and US Veteran Affairs (VA) is slated to become part of the NCHS data linkage program. HUD, SSA, Medicaid, and Criminal Justice Administrative Record System (CJARS) are available at the Census. Medicaid, VA, and SSA are also linked to the Health and Retirement Survey (HRS).

Intended Data Use. Even for alternative data sources for which there is national coverage of the target population, direct substitution for survey items is unlikely to be viable. An adaptive design for integrating alternative data sources might be one alternative. Under this design, rather than eliminating a survey item entirely (and replacing it with a value from an administrative record match), the survey instrument will only skip the question—and use the administrative record value—if a record match is present. If a record match is not present, the survey instrument will ask the respondent the survey item and the survey-reported value will be used. Despite the promise of this type of adaptive approach, we are aware of no major surveys that have used this approach and therefore it is unlikely that it will be a major component in the initial rounds of an NLSY26.

Short of direct substitution or adaptive collection, alternative data sources could also be used to edit and impute NLSY survey items, even when coverage in the ADS is less than complete. This could entail a more straightforward implementation of the survey editing process based on administrative data linkage post-data collection. For ADS with substantially less than national coverage, newer developments in imputation might make even these ADS potentially viable candidates for enhancing the accuracy of the NLSY26. For example, Rothbaum et al. (2021) used SNAP administrative records from only eight states—including some smaller states—to impute SNAP participation to the "rest of the country" in the Annual Social and Economic Supplement of the CPS. Mittag (2019) similarly employed a conditional distribution method to impute SNAP receipt in the American Community Survey using administrative records. While these serve as useful examples, there are relatively few examples of these types of approaches in major longitudinal surveys and the design of the questionnaire in initial rounds will likely be unaffected by this potential use of alternative data.

Lastly, an ADS could serve as an auxiliary source of data that would not otherwise be available in the NLSY26 or as a data source that could supplement data collected in the NLSY26 in some way. These uses are prevalent in other major surveys such as HRS and could serve to increase the value of the NLSY26 to end users. An example of the former might include information on respondent credit scores, an item that is not likely to be asked of respondents in the survey. The latter use case would be similar to the postsecondary transcript study in the NLSY97, which provided an administrative record analogue to a number of items in the NLSY97 but did not cover all NLSY97 respondents who attended a postsecondary institution. An example of an auxiliary data source for the NLSY26 might be CJARS. Though CJARS is not national in scope, linkage between an NLSY26 and CJARS would both expand the information on respondent encounters with the criminal justice system in the NLSY26 as well as provide administratively recorded alternatives to survey items in the NLSY26 for a subset of the sample.

7.2 Specific Data Sources

BLS prioritized the following data sources in the ADS assessment:

National School Clearinghouse (NSC)
National Student Loan Data System (NSLDS)
US Veteran Affairs (VA) data
Medicaid/CHIP data
Housing and Urban Development (HUD) data
Criminal Justice Administrative Record System (CJARS)
Income and Employment Data: National Directory of New Hires (NDNH), Social Security Administration (SSA) data, and Unemployment Insurance (UI) data
National Death Index (NDI)

National Student Clearinghouse. The NSC contains data elements that are a key part of the NLS. Although participation in the NSC is voluntary, the participation rate is very high, covering 97% of all students in public and private postsecondary institutions in the US. This participation has increased over time, and there does not appear to have been any appreciable attrition, as schools value the benefits of their (voluntary) participation in the NSC. In addition to strong coverage of the universe of students

enrolled in postsecondary education in the US, the data in the NSC is reported in a timely manner, typically 45 days after the start and finish of school semesters.

At minimum, the NSC has the potential to enhance measurement of enrollment. As noted in the 1.3 ADS report, if BLS can get approval for the program-level data elements, which are available "for special research projects," this would add valuable information on students' program of study.

National Student Loan Data System. The NSLDS provides a centralized, integrated view of federal student aid loans and grants that are tracked through their entire lifecycle from aid approval through disbursement and repayment. It provides timely and accurate data and is updated frequently, though there is some variation in update frequency across data elements. NSLDS is heavily audited to ensure it is properly tracking loans and their repayment through time, and therefore is generally considered the best source of student loan data.

The NSLDS has been used for research purposes, including linkage with survey data. It has been linked to the National Postsecondary Student Aid Study (NPSAS), where student-level data on Pell Grants and federal student loans from the NSLDS were matched to NPSAS sample members. Nonetheless, use of NSLDS in surveys outside of the Department of Education is to this date limited so potential use in the NLSY26 may need to overcome administrative barriers to sharing of data across federal agencies.

Department of Housing and Urban Development. HUD data is available at the federal level from a single data provider and has a strong track record of linkage with other national household surveys (both at Census and at NCHS). The data are updated frequently- ranging from weekly to quarterly, depending on the specific HUD data source. Based on the NCHS experience, HUD-linked data may need to be held in restricted-use files. Detailed PII from HUD program recipients, including SSN, are available for linking.

Because program receipt is underreported in household surveys, having administrative records available, even as a prompt to survey respondents, is likely to be of considerable value. There is substantial research interest in the effect of public program receipt in childhood and adolescence on later-life outcomes. Hence the ability to link HUD records, even historical records, to parents or guardians to cover any receipt during the respondents' childhood and adolescence would be of substantial research value and would ease the burden and improve the accuracy of retrospective data collection. As respondents age into adulthood, consent will likely need to be obtained to continue to link records.

Medicaid/CHIP. Although Medicaid/CHIP records are state-level data, they are collected in a centralized data system, where some standardization of variables occurs. Unlike with some other program data (e.g., SNAP/WIC/TANF), this means that BLS would only have to engage one data partner (i.e., Centers for Medicare & Medicaid Services) rather than up to 50. As with HUD data, Medicaid/CHIP data has been linked to NCHS cross-sectional surveys, and these linkages include historical Medicaid records.

Criminal Justice Administrative Records System. Linking CJARS to the NLS data would be a potentially important enhancement to a topical area in which under- and misreporting is likely significant. Records would not be available for juveniles, so linkage could only occur when a respondent ages into adulthood. Coverage is also an issue, given that 20 states are currently contributing data, and some of these states are only "partial" participants. There is some question regarding data access; namely, it is

unclear whether BLS could obtain the data from (or have the data linked) at the University of Michigan or whether the data would only be accessible through the Census Bureau.

Department of Veterans Affairs. US Veteran Affairs data, specifically the U.S.VETS database, covers the universe of veterans and contains a wide range of information on veterans that includes where they live, their finances, their education and employment, health insurance, and benefit utilization. Improving data accuracy appears to have been a priority for the National Center for Veterans Analysis and Statistics (NCVAS). The U.S.VETS data are regularly audited, and the VA Data Governance Council has adopted a transparent data quality framework, evaluating the U.S.VETS programs with respect to seven data quality dimensions: accuracy, completeness, consistency, traceability, uniqueness, validity, and timeliness. Timeliness may be somewhat of an issue, as the data are currently released with up to a two-year lag. VA data have been shared with the Census Bureau and are slated to be integrated into NCHS surveys.

Income and Employment Data. Information on employment and income is core to the NLS mission and was noted as a high-value topical area during stakeholder outreach. Members of the listening sessions in particular mentioned the benefit of capturing more accurate information on employment and income by linking to IRS records—an alternative data source for which negotiated data access was considered too difficult. Of the remaining alternative data sources that could enhance employment and income information in the NLS, SSA data would appear to be the most promising. Relative to UI wage records, SSA data have the advantage of being housed in a centralized data system, thus requiring engagement with only one data partner. Moreover, the coverage of the SSA data is national in scope, and data have been shared for the purpose of linking with other surveys. Linked SSA data might be particularly well suited to providing a more detailed employment and earnings history for parents and guardians of NLSY26 respondents.

The National Directory of New Hires (NDNH) has not been used extensively in survey data linkage but is an intriguing possibility that would provide a timely, accurate, and temporally granular source of employment and income data. Data deletion and deidentification requirements could pose potential hurdles to use, however. Specifically, the NDNH has required data partners to delete shared data after 2 years, and NDNH data used by US Department of Health and Human Services (HHS), US Department of Education (ED), HUD, and US Department of Agriculture (USDA) to conduct research or analyses of certain topics generally do not contain personal identifiers.

Finally, we noted that all income data sources share the drawback of not capturing income earned in the informal economy, a potentially important source of earnings for certain subgroups.

National Death Index. The NDI has been linked multiple times to many cross-sectional and longitudinal surveys. The NDI could be matched to both the NLSY26 respondent and the respondent's parents. The NDI records information on date of death, place of death, and cause of death all in a standardized and well understood way. The NDI received little discussion during the ADS analysis because of the high certainty of its value and use. The information needed and consent process for the NDI are well known and many studies have shown its feasibility.

7.3 Implications for Data Collection

Each ADS's implication for data collection will depend in large part on its envisioned data use: whether for direct or adaptive replacement of survey items, editing or imputation of survey items, or as an auxiliary source of data.

Respondent Burden. Both direct replacement and adaptive replacement design could reduce respondent burden during data collection, with the former entailing a greater reduction than the latter. Editing and imputation as well as using ADS as a linked auxiliary data source purely for research purposes would not bear on data collection directly, as these would occur post-data collection.

Survey Instrument Design. The implications of ADS for survey instrument (or item) design turn primarily on the degree of coherence of data elements in the ADS with survey items in the NLSY26. Whereas coherence typically refers to how well measures in the ADS align with corresponding measures in an extant survey, the NLSY26 offers an opportunity to potentially shape the survey instrument so that items in the survey conform more closely to those in ADS that have been identified as candidates for integration into the data collection process. There are limitations, of course, since a construct like quarterly earnings may appear in an ADS but would likely be extremely cognitively difficult for a respondent to report.

For direct replacement, the relevant data elements in the ADS would presumably inform the corresponding survey items in the NLSY26. Similarly, for adaptive replacement the instrument would need to be designed in such a way to build in the skip patterns for respondents with linked records. It may also be desirable to prompt respondents with the information gleaned from the ADS to confirm the accuracy of information in the linked records. For income data, additional questions would still be needed to probe for income from the informal sector or from non-UI covered employment. Hence, timeliness has greater salience for this data use. Even when respondent consent is granted for linkage to a given ADS, there will still be some failures to match or other reasons that ADS do not provide the desired data values. This will be a necessary cost of adaptive or replacement approaches. Efforts to minimize this type of loss will also reduce any reductions in respondent burden.

For ADS used to edit or impute NLSY26 survey items, the appropriate level of coherence with the relevant survey items would need to be established when the survey instrument is developed, though the degree of coherence for this data use may not need to be as high as with one of the direct replacement strategies.

Using an ADS solely as an auxiliary data source would likely have little direct effect on data collection, unless BLS intended some or all of data elements in the auxiliary data source to be directly comparable to the survey (as was the case for some measures in the NLSY97 school transcript study). When used solely as an auxiliary data, the data elements in the ADS would not need to cohere, and could simply supplement, the NLSY26.

Timing of Data Access and Linkage relative to Data Collection. Direct replacement—either full or adaptive—would require the highest degree of timeliness. Contemporary data from an ADS used for direct or adaptive replacement would need be accessed, cleaned, harmonized, and linked prior to data collection. For editing and imputation, the window for integration of the ADS with the NLSY26 would be longer than with direct substitution or adaptive design. For ADS used as an auxiliary data source, the timing of ADS access and linkage would not be as crucial. Linking retrospective data from an ADS mid-panel would still be of considerable utility to the research community.

8. Serving the User Community

Making data access as easy and as democratic as possible should be a priority for the NLSY26.
Ease of use is essential and requires well-structured data files and structure around how metadata are accessed and presented.
Multiple channels of data distribution will likely be best, including public use files, geocoded data through BLS with short turnaround time for project initiation, and data that is covered by Title 13 available in a Census FSRDC environment.

Dissemination plays an important role in the eventual use of survey data, and an NLSY26 will be faced with a changing dissemination landscape that brings both opportunities and challenges.

8.1 Ease of Use

Participants stressed that dissemination will need to be considered in the design phase of an NLSY26. One key example is whether to organize data by age or by round. Organizing data by age of the respondent as opposed to round of the survey brings tradeoffs across multiple aspects of the survey, which we discuss further in section 9 below. Allowing users to select the format for their data extraction will provide for greater ease of use. While examples of this are limited in longitudinal surveys, it would be similar in idea to a dissemination platform that allowed users to select whether they wanted housing unit-level variables or person-level variables; or to select different levels of nested geographies such as is available in the NCES Common Core of Statistics.

In addition, participants noted that the organization of data in loops can be difficult to navigate. As an example, one participant reported that trying to work with term level data for colleges is challenging due to the difficulties determining what timeframe the data correspond to and noted the monthly enrollment variable is easier to work with.

A consistent theme in stakeholder outreach was the need for an easy-to-use dissemination platform in the NLSY26. Issues mentioned included the ability of researchers to quickly determine whether or not relevant content is available, the efforts required to understand complex computer-assisted-interviewing skip patterns, ensuring that commonly used variables can be easily created by researchers, and flexibility in the platform to generate different structures of data files. Respondents noted that ease of access can be a driver of data use, and straightforward and well-documented data dissemination tools can greatly enhance the impact of a survey program. Stakeholders cited the IPUMS USA data files and NCES longitudinal surveys as good examples of data dissemination. Another way to facilitate users' ability to use the data is to make available illustrative tabulations and analyses.

8.2 Protecting Privacy

Protecting the privacy of respondents will be of utmost importance for the NLSY26 but brings with it inherent tradeoffs. Making more detailed data widely available enables ease of access by researchers for a diverse set of use cases but leads to further disclosure risk for respondents. Complicating this is the fact that the NLSY user community has a variety of needs, so there is no single method that will work best. When user survey respondents were asked directly about whether they would prefer access to administrative data linkages or public use data files, the results were relatively evenly split: 49% preferred administrative data linkages and 41% preferred accessible public use files (10% "don't know").

A tiered system is likely the best approach to trading off these conflicting desires, but many details will need to be worked out. It is likely that at least three separate access mechanisms would be ideal for serving the needs of different data users. The first tier would include public use files, which stakeholders mentioned as an important asset of prior NLSY cohorts. On the other end, a highly restricted environment such as a Census Bureau FSRDC would allow for novel analyses using linked survey and administrative data. Nonetheless, stakeholders mentioned the difficulties with accessing very restricted environments, and this combined with the consistent desire of stakeholders for geographic information suggests that an intermediate tier of access that provides detailed geographic information without detailed administrative data linkages might serve this user community. The current NLS access system is also three-tiered, with public use, geocode, and more highly restricted-use files such as the census tract file and the NLSY97 School Survey data, although the user community is strongly interested in more flexibility and greater opportunities for linkages. All of this will be complicated by the fact that protecting respondent privacy is likely much more challenging today than in the low-technology environment of the 20th century.

8.3 Timeliness

Multiple participants during stakeholder outreach expressed a desire to see more timely data releases but creating timely data releases comes with a cost. Quick tabulation files that offer subsets of high-interest variables in advance of full data releases may be one method for accommodating this conflict. In addition to improving the timely availability of data, these types of highly curated files might serve other objectives such as allowing data users a benchmark with which to check their results.

8.4 Paradata and Metadata

Other surveys' staff reported that they often deliver their data to a third party for storage and dissemination, and that third party is responsible for metadata standards. Many reported being in the early stages of moving to DDI Lifecycle but not having moved completely. On the whole, the results of the needs assessment indicate that this is an area still in development, but one for which a new survey such as the NLSY26 should pay attention to developing standards for the future.

Suggestions for further inclusion of paradata were relatively rare during stakeholder outreach. Aligning with this, participants from international survey programs generally noted that demand for paradata is low from their research communities, creating little pressure to clean and organize this data for public dissemination.

9. Key Tradeoffs

We now highlight a few particularly important tradeoffs that an NLSY26 will need to consider. We discuss high-level content tradeoffs first, then those associated with survey methodologies and incorporation of ADS. We close with tensions and decisions pertaining to data file dissemination.

9.1 High-level Content Tradeoffs

Overall Structure. Many content ideas are best achieved outside of a single periodic youth interview: supplements can include interviews with parents, teachers, or spouses/partners; off-cycle or event-triggered interviews or assessments; linkages to alternative data sources; or other non-survey data collection (such as biomarkers or air quality samples). Each of these will require resources. Especially in the initial years, the overall burden on a household and the impression that households form of participation in the NLSY may influence their participation for many years to come.

Should questionnaire content be anchored to rounds or to key ages? For some analyses, researchers will appropriately want to structure their data by age of the respondent rather than the round in which data were collected. For example, an analysis file might organize observations to measure family structure at a given age. At other times, it is the period of data collection that is salient, rather than the age of the respondent. This is especially true for analyses of the onset of policies that would need to be keyed to the calendar date.

The question of structuring analysis files can partly be about flexibility in data dissemination – allowing users to receive data organized by age or by round could significantly improve user experience. But the age versus round question also has implications for data collection: just as having contemporaneous data may be analytically valuable, one can also imagine content such as ASVAB scores would be more useful to collect at specific respondent ages rather than all in the same year. The existing NLSY cohorts have sometimes varied data collection timing in content-specific ways, either at fixed rounds or at fixed ages. Increasing consideration of which timing is more appropriate and extending age versus round data access flexibility where possible could both increase analytic value and user satisfaction.

Note also that the frequency of data collection may bear on the anchoring of content to rounds or ages. This decision may be less relevant for content that is measured more frequently, as analysts can more easily separate age and year effects. However, for content such as ASVAB that might be measured less frequently the question of whether to anchor to rounds or key ages will be more important.

How much should the NLSY26 prioritize comparisons to prior cohorts, and how much should it consider new measurement approaches or constructs? Although respondents to the user survey indicated that they would be willing to trade comparability to prior cohorts for new measurements and constructs, participants in the listening sessions frequently mentioned the value of being able to conduct cross-cohort comparisons. If all measures in the NLSY79 and NLSY97 could be replicated in an NLSY26, there would, of course, be no room for any new content. The design of an NLSY26 will need to identify a set of constructs for which cross-cohort comparability will be prioritized. Priority could be given to constructs that are distinctive within the NLSY data sets, such as job tenure, or usual hours of work. Another consideration might be how the construct has endured over the last several decades; being

able to construct comparisons of marital status between opposite sex individuals without considering non-marital cohabitation or same sex partnering may be less valuable. A critical practical consideration will likely need to be how much questionnaire administration time is required to achieve exact comparability with prior cohorts, and whether 'good enough' comparability may suffice. An example of a time-intensive construct might be 'jobs ever held,' which requires committing to enumerating every job the respondent has ever worked in every round of the survey.

How much should the NLSY26 consider content that only enables domain-specific research as opposed to enabling research that at least partially examines labor market outcomes? The bibliometric analysis made clear how much the NLSYs have contributed to vast areas of social science knowledge that are quite far removed from connections to the labor market. There is no question that there is demand in the research community for an NLSY26 that can provide data to answer basic science and policy questions for which there may be no other data sources. How much an NLSY26 should prioritize research questions that inform the mission of the BLS and the Department of Labor is a critical determination for the agency and its partners. In any event, tough choices will need to be made about survey content, especially within the constraints of how much time and attention tomorrow's youth may be willing to contribute. Even if the NLSY26 continues the NLSY tradition of being a broad resource across disciplines and topical areas, some scope definition will be necessary for the survey to have coherent content and a feasible data collection strategy.

To what extent should the NLSY26 overlap with other available surveys versus cover unique content? The exploratory activities did not reveal clear justification for intentional or minimal overlap with other surveys. On the one hand, it is scientifically valuable to be able to replicate key findings across independent data sources, so overlap is helpful. Overlap of items across distinct samples (for example, across age ranges or countries or in specialized data sets) also expands analytic possibilities. Methodologically, re-using items that have performed well on a previous survey minimizes risk in data collection. On the other hand, minimizing overlap can encourage data use by being the data set of choice for select research questions, as well as make optimal use of limited research resources across federal and private funding sources by maximizing the breadth of research questions that can be answered. This creates a policy choice for the designers of an NLSY26 in how to allocate limited financial resources and questionnaire length.

How much should an NLSY26 emphasize content that can only be covered in surveys at the expense of domains that may be well-covered in other data but of central importance to the NLSY (e.g., educational experiences, earnings)? The bibliometric analysis and much of the input from listening sessions documented the rise of found data sources, administrative data, firm- and school-level process data, and many other alternatives to survey data. The wide use of these resources raises the possibility that surveys could valuably use their direct contact with respondents to collect measures that are unlikely to appear in other types of data such as attitudes and expectations, cognitive skills, difficult-to-observe behaviors such as in-home production, sexual activity, and mental health. As in the above discussion of overlap with other surveys, this will boil down to a policy question for the designers of the NLSY26. Providing measures that cannot be found in other data sources would provide value but will need to be chosen to provide value given the rest of the data collected in the NLSY26, and not just because they cannot be found elsewhere.

Tradeoffs between Survey Methodology and Content

The period since the launch of the NLSY97 cohort has seen substantial expansion in data collection methodologies but also greater challenges to fielding survey data collections with representative samples. Methodology concerns will tradeoff with content through considerations of the optimal length of survey instruments, what is known in the survey methods literature about how best to ask certain types of questions, and which technological approaches should be incorporated into the survey. For example, the NLSY97 introduced use of audio computer-assisted self-administered interviewing into the NLSYs. This technology made it reasonable to expand questionnaire coverage on sensitive topics such as criminal activity, substance use, and sexual activity. Without that technology, content choices would likely have been quite different. Similarly, ideas emerged in the listening sessions and elsewhere about adding supplemental data collections. These additional opportunities will at a minimum require financial resources that could draw focus away from the primary periodic youth interview. In addition to survey methodological issues, choice of oversamples may drive content to the extent that measures that are specific to certain subgroups will only be relevant if there is a sufficient sample size for that subgroup.

9.3 Tradeoffs pertaining to incorporating Alternative Data Sources

The incorporation of ADS raises important tradeoffs for content collected on the survey. Many measures, especially ones that are cognitively challenging, may be more accurately measured in ADS (e.g., dates of job spells or social program receipt). But sample coverage will be incomplete because of lack of coverage in the ADS, absence of consent allowing the data linkage, or errors in the linkage process. Relying more on ADS may reduce the need for onerous probing of respondents but at the cost of data being absent for respondents without alternative data coverage. In theory, a combination of survey data with ADS may maximize accuracy and coverage, but these combinations often require more survey administration time and other survey resources than do survey-only or ADS-only approaches and may introduce new biases in the sample. For example, in theory it is possible to preload data from alternative data and ask fewer questions for respondents with fuller coverage. But if the lack of coverage is due to refusal to provide data linkage consent, the effect is putting the most burden on the likely least compliant respondents raising the real fear of introducing selective attrition into the sample. One compromise might be to collect coarser data on respondents that are not covered which precludes certain analysis on the entire sample (e.g., analysis that relies on the exact timing of events).

The stakeholder input is clear that ADS should be integrated into the design of an NLSY26 from the start, and the investigations into these data sources identify several promising approaches and data sources. At the same time, incorporation of ADS raises a couple of important tradeoffs related to dissemination. One has to do with user access to measures. While some ADS can be incorporated into the NLSY26 public use data files, many others will be available only in NLSY26 restricted use files, or even to researchers who are able to conduct the linkages themselves. There may be some measures that are deemed essential to have available in public use data files in some form, even if higher-quality versions may be available through restricted-use ADS solutions. These measures might include parental income, or respondent high school graduation status, among others.

Another tradeoff relates to the high disclosure risk associated with so many ADS and the resultant caution about interactions between ADS. Some surveys prohibit linkage to more than one ADS at a time (for

example, to conduct an analysis that combines earnings measures from one ADS, health measures from another ADS, and the questionnaire data as a third). Even when not in the same analysis, providing data at a certain level of geography or for certain types of higher education institutions could then constrain what can be released in later years. Policies will need to be developed to ensure that linkages to ADS do not excessively preclude future linkages or topical coverage, and vice versa, that early decisions about public release of information do not prohibit high priority ADS linkages later.

9.4 Dissemination Tradeoffs

In thinking about dissemination-related tradeoffs, we can think about user experience, questionnaire content and structure, protection of respondent privacy, and allocation of survey budget. Choices about dissemination often implicate one or more of these aspects of the survey.

Consider, for example, how easy the data should be to use. One approach to make easy to use data files is to have very simple questionnaires that do not collect event histories, do not loop through multiple constructs such as schools or terms, and do not allow for correction or confirmation. Such a questionnaire requires minimal documentation and processing, and is easy for users to work with, but all of this comes at the expense of the richness of the data, and possibly accuracy of measures. A different approach to make easy to use data files is to have complex questionnaires that facilitate respondent cognition and detailed analyses, but then also invest in variable creation and documentation to facilitate the user experience. The latter has been some of the approach of the NLSY97, although it requires substantial resources for data processing and data file documentation. The choices for faster data release are similar to those for easy to use files: have simple questionnaires and low data processing costs, or complex questionnaires and higher costs to offset the complexity for users.

While ease of use may be somewhat discretionary, protecting respondent privacy is both statutorily required and essential for the ongoing trust and cooperation of the sample. We discussed some tradeoffs related to ADS inclusion and respondent privacy in section 9.3 above. Here we reiterate that both the completeness of public use data files and the administrative processes associated with restricted use access affect user experience and therefore use of the dataset. Public use data files must be able to support informative analyses that can lead to publication of findings. If it is difficult for researchers to publish using only the public use data because of content gaps, the data files will not achieve the dissemination goals associated with public use data. In terms of restricted use processes, restrictions and limitations on actual analyses and linkage possibilities are important contributors to user experience, but so are practical factors such as fees charged, difficulty of data access logistics, and length and uncertainty of clearance processes.

Creating different file types and a tiered access mechanism may serve the needs of different users and may help to balance some of these considerations. Unlike with questionnaire length and respondent tolerance for survey length, there may not be as many hard and fast constraints for data dissemination. Instead, the primary constraint may be in how many resources to allocate to dissemination compared to other project activities such as data collection. For example, investing in information technologies that can deliver data to users in multiple ways, can aid data discovery, and can transparently document the content is now possible. But these technological investments will require resources and the NLSY26 will need to consider these expenditures in light of other program priorities.

References

Federal Interagency Forum on Child and Family Statistics. America's children: Key national indicators of well-being, 2021. U.S. Government Printing Office

Mittag, N. (2019). Correcting for misreporting of government benefits. *American Economic Journal: Economic Policy*, *11*(2), 142-64.

Rothbaum, J., Fox, L., & Shantz, K. (2021). Fixing Errors in a SNAP: Addressing SNAP Under-reporting to Evaluate Poverty.

Singer, Audrey (2009). The New Geography of United States Immigration. *Brookings Immigration Series*. No 3.

Solon, G. (1992). Intergenerational Income Mobility in the United States. *American Economic Review*, 82(3), 393-408.