

APPENDIX G

Summary Notes from the 2019 SDR Sample Design Planning Activities

SECTION G.1.

**Summary Report from the SDR Sample Design Expert Panel
(May-September 2017)**

**Summary Report from
SDR Sample Design Expert Panel Report
(Meetings held May-September 2017; Report issued November 2017)**

A. SDR Survey Objectives

The Survey of Doctorate Recipients (SDR) has been used for cross-sectional, points-in-time estimation for U.S.-earned PhDs since its inception in 1973. Beginning in 2015, the SDR sample size was greatly increased to meet two overarching objectives: (1) to produce reliable estimates of employment outcomes by the fine field of degree taxonomy used in the Survey of Earned Doctorates (SED); and (2) to maintain the existing estimation capabilities associated with analytical domains defined by key demographic characteristics and currently used in National Center for Science and Engineering Statistics (NCSES) publications such as the Science and Engineering Indicators (SEI) report, Women, Minorities and People with Disabilities in Science and Engineering (WMPD) report, and detailed statistical tables (DST).

NCSES has become aware of a growing demand for longitudinal data, both for the generation of official statistics and for longitudinal analysis by the research community. While cross-sectional estimation remains the primary objective of the SDR, NCSES embarked on an effort to utilize the redesigned 2015 SDR sample, with its much larger sample size, for longitudinal estimation, including the development of longitudinal indicators by NCSES and longitudinal analyses by other data users. To guide this effort, the NCSES commissioned an SDR sample design expert panel to tackle the design challenges involved in balancing longitudinal and cross-sectional analytic objectives. The mandate of the expert panel was to assist NCSES in the design of a longitudinal component to the SDR. This longitudinal component is intended as a complement to the cross-sectional component, which is acknowledged as representing the primary focus of SDR.

Some of the main characteristics shared by longitudinal surveys relate to sampling. As an official government survey, a longitudinal survey is expected to represent a specific population, from which the sample is drawn according to a documented sampling design. In the longitudinal context, this population often corresponds to the finite population at the initial time point at which the sample is drawn, and then both the sample and this population remain fixed over time. In a growing population such as the population of PhD earners, this implies that as time moves on (away from the time of the sample draw), the longitudinal sample represents a decreasing fraction of the actual population. This raises questions about how to update a longitudinal design over time, which will be addressed later in this report.

Estimates produced from official longitudinal survey datasets are expected to follow the standard paradigms of design-based estimation. This means that they are fully weighted to account for the sampling design and unit nonresponse, and possibly imputed for item nonresponse and measurement errors as well. The goal of the weights and associated adjustments are to ensure that the estimates are statistically valid, with associated estimated measures of uncertainty, for the characteristics of the target population. Thus, for longitudinal characteristics, such as the temporal change across multiple survey cycles in variables of interest, the weighted estimates are for changes in the population as it was defined at the initial time point at which the sample was drawn.

After such a fully-weighted longitudinal dataset is created, it can be made available to data users, who can then perform their own longitudinal analyses. By having access to data that have been vetted and weighted, the validity and representativeness of these analyses can now be statistically evaluated. This

is currently not possible, because weights previously provided as part of the SDR cross-sectional datasets cannot be applied longitudinally.

In this report, we present a range of options for the design of a longitudinal component for the SDR. We first describe, in general, the approach and then how it could be applied to the SDR. The ultimate decision about which longitudinal design to implement will depend on the scientific and policy questions NCSES is most interested in answering.

B. Longitudinal Design Approaches

In this section, we will consider four design approaches: (1) fixed panel plus births; (2) rotating panel; (3) cohort panel; and (4) split panel. For each design approach, we will discuss the following aspects: possible design variations; the pros and cons of the approach; some issues and considerations of using this design for the SDR; and a summary of feedback from SDR stakeholders obtained at the August 2017 meeting of the NCSES Human Resources Expert Panel (HREP). We will also provide some examples of longitudinal studies that use these approaches.

1. Fixed Panel Plus Births Design

A fixed panel plus births design begins with a sample that represents the baseline target population. This sample is retained through all subsequent survey waves. Periodically, the sample is refreshed with a sample of units that have become eligible since the prior refreshment (“births”). This sample refreshment, which may occur every wave or on some other schedule, is done in order to make the sample representative of the population at the time of refreshment.

The main advantages of a fixed panel plus births design are: (1) after the baseline interview, retaining cases in the fixed panel portion of the sample is a cost-saving (compared to introducing new cases); (2) refreshing the sample with births allows for representation of a population that is changing over time; and (3) over time, rich longitudinal data are accumulated for sampled units. The main disadvantages of a fixed panel plus births design are: (1) the high burden imposed on sampled units; (2) the fact that the sample design is essentially fixed at inception; (3) with a population that is growing over time, sample deselections (“maintenance cuts”) may be necessary to limit the sample size from wave to wave; and (4) the potential for nonresponse bias as a result of attrition of units from the sample that differ systematically from those that remain in the sample.

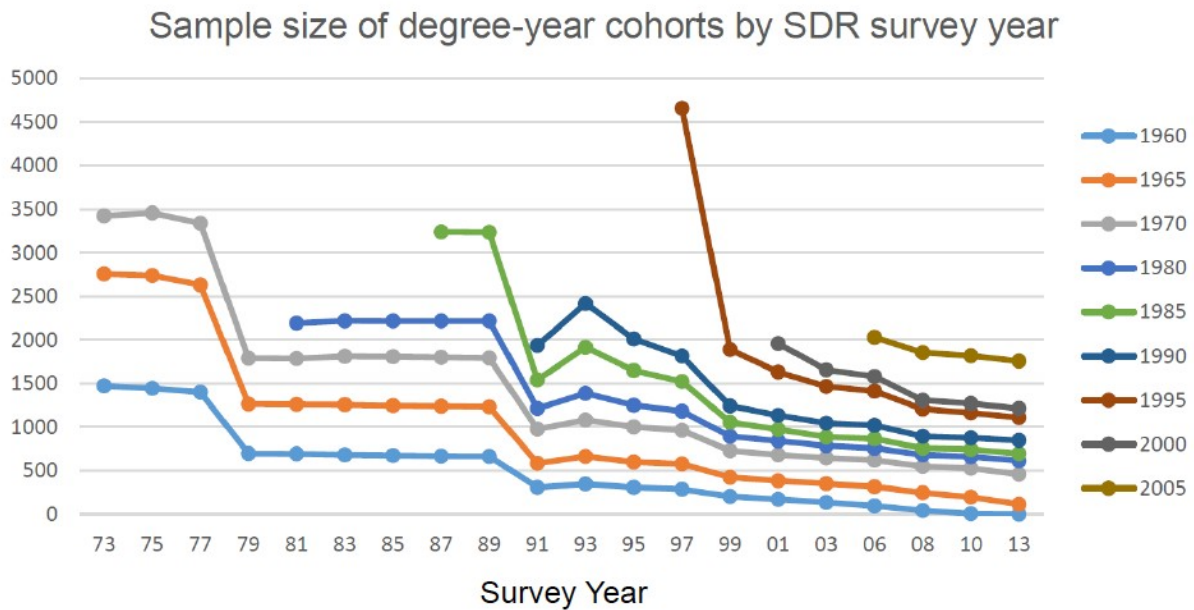
One study that uses a fixed panel plus births design is the National Health and Aging Trends Study (NHATS, <http://www.nhats.org/>), a study of older Medicare beneficiaries sponsored by the National Institute on Aging. NHATS initially sampled Medicare beneficiaries ages 65 and older as of September 30, 2010, and has followed up with sampled beneficiaries annually. The NHATS sample has been refreshed once to include births and to compensate for sample losses due to mortality and attrition. Since the time of sample refreshment, the NHATS sample represents Medicare beneficiaries ages 65 and older as of September 30, 2014.

Since its inception in 1973, the basic design of the SDR has been a fixed panel plus births design. Over that period, the target population has grown from about 300,000 in 1973 to 1.1 million in 2015, and the sample size has varied from about 30,000 to 120,000. However, unlike the target population, which has been nearly strictly increasing in size during that period, the sample sizes have oscillated over time depending on available resources. The fiscal inability of the sample to keep pace with the growth in the

population has resulted in the need for occasional sample maintenance cuts. As a result, the full longitudinal record for a given individual was susceptible to, and may have been attenuated by, these maintenance cuts. Exhibit 1 shows the sizes of degree-year cohorts by SDR survey year, illustrating that for a given cohort, the longitudinal sample sizes have been relatively limited.

The SDR stakeholders at the HREP meeting were generally satisfied with the longitudinal analytic capabilities of the fixed panel plus births design of the SDR through 2013. However, it is unclear how these stakeholders have been using this sample for longitudinal purposes. In particular, at least some of these analyses do not appear to have been based on a well-defined target population. Instead such analyses were conducted by simply pooling the available cases across years, without using weights.

Exhibit 1. Sample size of degree-year cohorts by SDR survey year



2. Rotating Panel Design

A rotating panel design is when a proportion of the initial survey sample is dropped from each subsequent survey wave (or round) and replaced with a fresh (representative) sample (of approximately equal size) of the population. The complete sample is thus replaced gradually over a period of months or years.

Rotating panel designs have been mostly adopted for purely cross-sectional surveys and surveys with limited time horizons, in which short-term changes are of interest. The Current Population Survey (CPS, <https://www.bls.gov/cps/>) and National Survey of College Graduates (NSCG, <https://www.nsf.gov/statistics/srvygrads/>) are among the surveys using a rotating panel design. The CPS sample design uses a 4-8-4 rotation scheme with 8 panels, ensuring that 75 percent of the sample is common from month to month, and 50 percent of the sample is common from year to year (U.S. Census Bureau 2006). The NSCG uses a rotating panel design with four panels, ensuring that 75 percent of the sample is common from round to round, and the sample is completely replaced after four rounds.

Compared to the fixed panel plus births design discussed above, a rotating panel design would be a less cost-effective option as it refreshes a substantial portion of the sample every wave. A main cause of survey nonresponse is the inability to reach sample members. Thus, locating members of the sample is key to successful data collection. A refreshment sample may require more locating efforts than a panel sample, for which contact information from the past survey cycle may still be valid.

On the other hand, a rotating sample design provides more flexibility for mid-course design changes than the fixed panel plus births design; however, an effective sample design must meet precision requirements for various key domains. The SDR is designed to provide survey estimates for career tracks of U.S.-earned doctorates. Over time, key domains may change to reflect significant drops/additions of specific groups (such as field of degrees (FODs) phasing out or new FODs emerging). Primary analysis domains may change (such as demographic groups or FODs). In such cases, a rotating panel design affords flexibility not offered by a fixed panel design, as it allows a portion of the sample to be refreshed in accordance with the updated design goals.

Assuming that 2017 is the base year for the newly redesigned SDR with four rotating panels, Exhibit 2 illustrates how a panel would be rotated out and a new panel added over future survey waves.

Exhibit 2: A Rotating Panel Design with Four Panels

	2017	2019	2021	2023	2025
<i>Panel 1</i>	+	Drop	--	--	--
<i>Panel 2</i>	+	+	Drop	--	--
<i>Panel 3</i>	+	+	+	Drop	--
<i>Panel 4</i>	+	+	+	+	Drop
<i>Panel 1</i>	--	+	+	+	+
<i>Panel 2</i>	--	--	+	+	+
<i>Panel 3</i>	--	--	--	+	+
<i>Panel 4</i>	--	--	--	--	+

The 2017 SDR sample is randomly partitioned into four panels with each panel representing the full population. Then, one fourth of the 2017 sample, namely panel 1, would be dropped from the sample as a newly refreshed sample is selected to take the place of panel 1 in 2019. Consequently, the 2017 SDR sample would get completely replaced with a new sample by 2025.

As indicated above, a rotating panel design is predominantly used for cross-sectional or short-term longitudinal surveys. Although a rotating panel design involves carrying a segment of the sample over a fixed period of survey cycles by design, it is also constrained by the life time of the fixed panel. It may not be suitable for collecting data on phenomena that change slowly or have long term effects. However, if key indicators for longitudinal panels of interest are based on data collected across a span of less than 10 years, a four-panel based rotating panel design may still be an appropriate solution. For example, if a new Ph.D. cohort's post-doc experience or a post-doc's academic career path to the tenure track is of interest, up to 8 years of data would be sufficient.

The rotating panel design, as compared to the fixed panel plus birth design, would add higher costs of selecting and fielding new panels, maintaining panels, and weighting individual and combined panels. In addition, with substantially high sampling fractions of the SDR samples over survey rounds, a significant overlap between SDR rotating panels would occur. To minimize sample overlap within the panel frame,

a negative sample coordination based on permanent random numbers could be considered (Ohlsson 1995).

On the other hand, the SDR’s sister survey, NSCG, currently uses a four panel rotating design and its methods for data processing could be adopted for the SDR with moderate efforts. The SDR population can be viewed as a subpopulation of the NSCG, with a substantially larger sample size than NSCG for this specific PhD population. The SDR could adopt methods developed by NSCG on sampling, weighting, and variance estimation for multi-frame based samples, as well as the use of a base questionnaire for newly added cases and follow-up questionnaires for subsequent rounds of the same cases.

The fixed duration of the panels also facilitates the implementation of appropriate estimation methods, including non-response adjustments, imputation and weighting. SDR stakeholders at the HREP meeting had concerns about a rotating panel design because the time period a person would be in the sample would be limited. Panelists had varying views on panel length, and suggested values ranging from 10 years to forever. A few panelists suggested varying the frequency of follow-ups – with fewer follow-ups during more settled career periods.

3. Cohort Panel Design

A cohort panel design follows a sample of a subset of the population (cohort) over time. The cohort can be defined based on any group of analytic interest. The first round of the study (base-year) can consist of only the cohort members, who are then followed in subsequent rounds. Alternatively, the base-year study can contain the larger population, and then only the cohort (or multiple cohorts) are followed. The cohort members are likely oversampled in the base year to have a sufficient sample size. As with other longitudinal designs, the length of the panel and the frequency of follow-up rounds can vary. A new cohort can be started after a cohort has completed the study, or cohorts can run concurrently, if there is sufficient funding. Exhibit 3 shows an example of what a cohort design could look like. Each new cohort could be similar or different subsets of the population. The length, frequency, and overlap of the cohorts in the example can vary.

Exhibit 3. Design example – cohort panel

	Year							
Sample	0	2	4	6	8	10	12	...
Base-year	B1			B2			B3	
Follow-ups		F11	F12		F21	F13, F22		

There are several advantages to a cohort panel design. The panel can be short or long and intervals between follow-ups can be lengthened if there are budget cuts. Also, different cohorts can be spun off from each base-year study and similar cohorts can be compared over time. In addition, intervals between follow-ups can be based on analytic interests, which may vary by cohort.

One disadvantage of a cohort panel design is that the frequent start of new cohorts may not be feasible with budget constraints. Also, the panel only follows a subset of the population. Moreover, oversampling cohort members in the base-year could affect overall design effects for cross-sectional estimates (when the longitudinal cohort is combined with the cross-sectional panel for cross-sectional estimation purposes) and the design might invite too much emphasis on the panel estimates and too little emphasis on the cross-sectional estimates.

The National Center for Education Statistics (NCES) conducts several cohort longitudinal panels. The National Postsecondary Student Aid Study (NPSAS, <https://nces.ed.gov/surveys/npsas/>) is the base-year study for two longitudinal studies, the Baccalaureate and Beyond Longitudinal Study (B&B, <https://nces.ed.gov/surveys/b&b/>) and the Beginning Postsecondary Students Longitudinal Study (BPS, <https://nces.ed.gov/surveys/bps/>). NPSAS is a cross-sectional study of how college students pay for their education, and it is typically conducted every four years with a sample size of approximately 120,000 students. Each NPSAS study alternates between spinning off a B&B and a BPS cohort.

B&B is a longitudinal study of baccalaureate recipients. After NPSAS, this cohort is typically followed 1, 3, and 10 years later. The B&B studies have ranged in sample size from 12,000 to 29,000 baccalaureate recipients. BPS is a longitudinal study of first-time beginning students (FTBs). After NPSAS, this FTBs cohort is typically followed 2 and 5 years later. The BPS studies have ranged in sample size from 10,000 to 37,000 FTBs. In addition, NCES has conducted the following five high school longitudinal studies going back to 1972:

- High School Longitudinal Study of 2009 (HSLs:09, <https://nces.ed.gov/surveys/hsls09/>)
- Education Longitudinal Study of 2002 (ELS:2002, <https://nces.ed.gov/surveys/els2002/>)
- National Education Longitudinal Study of 1988 (NELS:88, <https://nces.ed.gov/surveys/nels88/>)
- High School and Beyond (HS&B, <https://nces.ed.gov/surveys/hsb/>)
- National Longitudinal Study of the H.S. Class of 1972 (NLS-72, <https://nces.ed.gov/surveys/nls72/>)

Each of these studies starts with a cohort of 8th, 9th, or 10th grade students and follows the students over time at varying intervals. The sample sizes vary from 18,000 to 71,000 students.

The cohort panel design by itself will not work for the SDR, because it focuses on the longitudinal design rather than on the cross-sectional design. Moreover, SDR stakeholders at the HREP meeting agreed that the cohort panel design is not the way to go for the SDR. This option would limit longitudinal analyses to just those involving the subpopulation represented by the cohort (thus rendering longitudinal analyses of certain subpopulations infeasible) and may have gaps that are too large between the start dates of the new cohorts. The stakeholders prefer a panel that is representative of the population and not just a subset based on one point in time. In addition, they do not want to wait a long time to introduce new graduates into the panel. The B&B cohorts introduced in 1993, 2000, and 2008 happened to align with changes in economic conditions that affected employment of baccalaureate recipients.

On the other hand, the stakeholders noted that the cohort panel design does have benefits, such as each cohort member being at the same point in the lifespan, which allows for the same questions to be asked for all cohort members. Moreover, the cohort panel design could be incorporated into a split panel design (described below).

4. Split Panel Design

The split panel design can use a combination of the previously discussed designs. Under a split panel design, the overall sample is split so that multiple sub-samples are created (the “panels”), each of which is run independently. However, the sub-samples can target the same population, which makes this different from stratification. The structure of a split panel is very general and can accommodate many combinations of target samples and objectives, which makes it appropriate for large-scale, multi-purpose surveys over time. The main advantage of split-panel designs is the increased flexibility to handle multiple objectives with separate component samples, instead of having to design a single sample that tries to balance all of the objectives. The main disadvantage is an increase in complexity in overall survey operations, including sample selection and estimation.

A large-scale survey that uses a split-panel design is the National Resources Inventory (<https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/technical/nra/nri/>) of the Natural Resources Conservation Service (NRCS). This annual survey is composed of a “core panel” of land tracts that are revisited every year, and multiple rotating panels that are revisited with different frequencies, from “every 2 years” to “not scheduled to be revisited.” The revisit frequencies of the core and rotating panels are set to ensure precise estimation of land use changes over different time scales.

To make matters more specific to the SDR, we will focus the discussion of split panel designs here on a simpler two-panel sample with one panel a cross-sectional sample of the overall population, similar to the current SDR sample, and the other panel a separate longitudinal sample. The objective of the cross-sectional panel is to continue providing biannual estimates of the state of the overall SDR population, at a level of granularity similar to that of the current survey. Hence, the design would also be similar to that of the current survey, which would include the need to add new cases each survey cycle to account for the population growth over time (new graduates), and associated sample trimming to make room for these additions. To allow for the creation of the new longitudinal panel, the sample size for the cross-sectional panel could be modestly reduced relative to the current survey.

The new longitudinal panel should be designed around the longitudinal objectives of NCSES, which are still being developed. Possible objectives include determining the career tracks over time of recent graduates entering academia or those of individuals who left the United States (i.e., targeting specific subpopulations), or developing new measures of persistence or frequency of occurrences over time across the whole population (i.e. targeting new longitudinal estimates). In principal, both types of objectives can be handled within a single-panel survey (i.e., a fixed panel design), but with less control over sample size over time and decreased flexibility to customize the questionnaire and survey protocols to address the longitudinal estimates.

Regardless of the specific longitudinal objectives, the goal of the longitudinal panel is to design a sample that specifically targets these objectives and from which statistically valid longitudinal estimates can be created. This can be most readily achieved by designing this panel as a longitudinal cohort, as described in the previous section. Depending on the longitudinal objectives, more complicated split panel designs of this type can be considered. For example, multiple cohorts, either targeting the same population with different start and end dates or targeting different subpopulations of interest, are readily accommodated within the split panel design framework. This increased ability to target sub-samples towards specific estimation goals has to be balanced against the fact that the overall sample size is divided between the different panels. This increases the complexity of the overall design and the estimation methods, and results in decreases in precision for each of the estimation targets. A rotating panel design can also be incorporated for the longitudinal cohort.

Split panel designs provide great flexibility in creating what are, in effect, almost separate surveys targeting different aspects of a population over time. However, because they target the same population and their designs can be coordinated, it is often possible to construct combined estimates across panels, resulting in more precise estimates compared to completely separate surveys. In the example of a cross-sectional and a longitudinal panel mentioned above, cross-sectional estimates are improved by pooling data from both panels, using ideas of dual-frame estimation (accounting for the fact that, except in the first year of the longitudinal cohort panel, its population will not match that of the cross-sectional panel due to population changes over time). Hence, the effective sample size for cross-sectional estimates based on this approach will be almost identical to that of the current survey, if the two panels have similar designs. Lohr (2009) provides an overview of multi-frame estimation approaches that can be considered here.

For the longitudinal panel, it is not possible to combine the panel estimates using this dual-frame approach, because the cross-sectional panel will not collect the longitudinal variables in a systematic manner. However, longitudinal estimates can still be improved by calibrating them to select cross-sectional estimates, using ideas of multi-phase regression estimation. Legg and Fuller (2009) provide a useful explanation of how to apply this approach. In addition to improvements in precision for the longitudinal estimates, this calibration across panels also ensures concordance between key estimates across cross-sectional and longitudinal data releases.

Section B.5. Other Considerations for Longitudinal Design Approaches

In considering the four design options, to see if each could be adopted for the future SDR to provide key survey indicators for both cross-sectional and longitudinal purposes, the SDR sample design expert panel identified several key challenges and issues:

- **Adding longitudinal analysis goals while maintaining cross-sectional estimation targets as primary goals** – An ideal design for the SDR would be a compromise between the most optimal cross-sectional design and the most optimal longitudinal design. In order to develop a design that meets both cross-sectional and longitudinal analysis goals, it would be critical to have those goals clearly specified. However, longitudinal analysis goals are usually not clearly defined at the point of sampling. To address this challenge, NCSES is undertaking an effort to develop longitudinal indicators. The nature of such uncertainty for longitudinal analytic targets was a key challenge the sample design expert panel grappled with.
- **Growth in population size** – Like other demographic surveys, the size of the cross-sectional target population changes over time, usually increasing. This is the case for the SDR population, with a net increase in the target population of about 6% from the prior survey round (an addition of new doctorate recipients minus people over age 75 or deceased). This brings technical challenges to a long-term sustainable rotating panel design for the SDR. A couple of options are:
 - Increase the sample size while maintaining equal sampling rates over time
 - Decrease sampling rates while maintaining equal sample sizes over time

In the past, the SDR (based on a fixed panel plus births design) mostly employed a sample maintenance cut based on cross-sectional analytic targets. With longitudinal design aspects embedded, sample maintenance would be more complicated.

C. Longitudinal Design Options for the SDR

C.1 Split Panel Design with General Population Panel

As described in section B.4, a split panel design includes two or more sub-panels. The sub-panels may be designed to meet different objectives, including representing different subpopulations or supporting different analytic objectives. In the context of the SDR, the split panel design is conceptualized as two panels: one designed with the cross-sectional objectives of the study as its focus, and the other designed to support the longitudinal objectives of the study. (These two sub-panels are referred to as the “cross-sectional sample” and the “longitudinal sample” but, as discussed in section B.4, under certain conditions both sub-panels may contribute, in different ways, to both cross-sectional and longitudinal analyses.) To support cross-sectional analyses, the cross-sectional sample is designed to represent the target population at each wave. With the longitudinal sample, there are many options in terms of population representation, such as a split panel design with a general population panel in which the longitudinal sample represents the general target population (as opposed to any specific subpopulation) at baseline.

Section B.4 outlines the advantages and disadvantages of split panel designs in general. Here, we focus on the advantages and disadvantages of the split panel design compared to other designs. Compared to the split panel design with a specific subpopulation panel, the split panel design with a general population panel allows for greater flexibility in the analyses that may be done. If a particular subpopulation is of primary interest for longitudinal analyses, using a specific subpopulation panel will be most efficient. However, if several distinct subpopulations are of interest or longitudinal analytic objectives are not yet well-defined, using a general population panel will retain the ability to analyze different subpopulations, albeit with reduced efficiency, compared to a design that uses the full longitudinal sample size for a specific subpopulation panel (for that given subpopulation).

Compared to the fixed panel plus births design, a split panel design (whether a general population longitudinal panel or a specific subpopulation longitudinal panel is used) has advantages due to the explicit separation of the designs of the cross-sectional and longitudinal samples. If either the cross-sectional or longitudinal objectives change (e.g., as a result of budgetary increases or cut-backs), those changes may be applied to one sample without affecting the other. Another advantage of splitting the cross-sectional and longitudinal samples is that different survey procedures may be administered to each of the samples. For example, the longitudinal and cross-sectional samples may be interviewed at different frequencies or may be administered different questionnaires.

For the SDR, given the continuous growth in the population over time and the budgetary constraints and uncertainties, the flexibility of the split panel design is appealing for several reasons. First, with a split panel design, the longitudinal panel may be maintained unaffected by budgetary fluctuations, while the cross-sectional sample remains responsive to these effects. Second, the cross-sectional and longitudinal samples may be designed to support estimation at different levels based on the expected precision. For example, the design may support estimation at the fine field level for cross-sectional estimates, while longitudinal estimates may be supported only for broad fields. Third, each sample may oversample specific subpopulations of interest, and these subpopulations do not have to be the same (e.g., the cross-sectional sample might oversample smaller fine fields, while more recent graduates are oversampled in the longitudinal sample). Fourth, the ability to administer different questionnaires to the two samples may enhance the longitudinal capabilities of the survey, particularly as longitudinal analytic

objectives become better defined. However, even if the cross-sectional and longitudinal questionnaires were to diverge, keeping certain core questions in both will still enable both samples to contribute (at different levels) to both cross-sectional and longitudinal analyses.

The SDR stakeholders at the HREP meeting identified a variety of longitudinal analysis objectives that require different subpopulations. These include early career, mid-career, and late-career doctorates. With a split panel design, using a general population panel for the longitudinal sample for the SDR offers the flexibility of accommodating these varying analytic needs and permits NCSES to explore and develop official longitudinal indicators without the restriction of a specific subpopulation.

C.2 Split Panel Design with Subpopulation Panel

A split panel design with a subpopulation panel is similar to a split panel design with a general population panel, as described above. The cross-sectional portion of the sample would be a fixed panel with sample maintenance cuts over time. The longitudinal portion of the sample would follow a subsample containing one or more cohorts with no maintenance cuts over time. The baseline for the longitudinal portion could be either SDR 2015 or 2019. Potential cohorts could be based on career stage (such as early or late career doctorates); demographics (such as women or under-represented minorities); or outcomes (such as postdocs, academic sector, or tenure track). As with any longitudinal design, design parameters include length of panel, frequency of follow-ups, and sample size. Additionally, parameters include the number of cohorts to follow and how often to start a new cohort.

There are several advantages of a split panel design with a subpopulation panel. True longitudinal analysis can be done to look at both transitions and duration. Also, important subpopulations can be oversampled and analyzed with better power and similar cohorts can be compared over time (such as comparing the 2019 cohort vs. the 2029 cohort). In addition, the panel can be followed without large effects on cross-sectional estimates.

The split panel design with a subpopulation panel also has several disadvantages. For example, with this design it is not possible to follow all subpopulations, so a few subpopulations have to be chosen. Also, decisions have to be made now about what subpopulations may be important over next ten years, and the more cohorts that are followed, the smaller the sample size available for each cohort. Exhibit 4 shows examples of sample designs with one, two, or three cohorts.

Exhibit 4. Design examples – split panel with subpopulation panel

1 cohort						
Year	2019	2021	2023	2025	2027	2029
Cohort 1	B	F1	F2	F3	F4	F5

2 cohorts simultaneous						
Year	2019	2021	2023	2025	2027	2029
Cohort 1	B1	F11	F12	F13	F14	F15
Cohort 2	B2	F21	F22	F23	F24	F25

3 cohorts simultaneous						
Year	2019	2021	2023	2025	2027	2029

Cohort 1	B1	F11	F12	F13	F14	F15
Cohort 2	B2	F21	F22	F23	F24	F25
Cohort 3	B3	F31	F32	F33	F34	F35

2 cohorts staggered without overlap						
Year	2019	2021	2023	2025	2027	2029
Cohort 1	B1	F11	F12			
Cohort 2				B2	F21	F22

2 cohorts staggered with overlap						
Year	2019	2021	2023	2025	2027	2029
Cohort 1	B1	F11	F12	F13		
Cohort 2			B2	F21	F22	F23

3 cohorts staggered with overlap						
Year	2019	2021	2023	2025	2027	2029
Cohort 1	B1	F11	F12	F13		
Cohort 2		B2	F21	F22	F23	
Cohort 3			B3	F31	F32	F33

Longitudinal indicators for the split panel design with a subpopulation panel are similar to those for the other longitudinal designs (such as duration of employment-related outcomes), but for this design, the indicators are limited to the subpopulation (cohort) being followed rather than the full population.

A split panel design with a subpopulation panel could be beneficial for the SDR if a small number of subpopulations of interest can be identified and agreed upon. This design would allow for analysis with good precision of key domains of interest over time, such as early or late career doctorates. However, there were varying domains of interest among SDR stakeholders at the HREP meeting, so the stakeholders prefer that a split panel design follow a subset of the general population, as described in C.1 above, rather than a few subpopulations. This allows all types of doctorate recipients to be followed over time rather than needing to reach consensus now about which groups are of interest to follow.

C.3 Fixed Panel Plus Births Design

As described in B.1, a basic setting for this design is to have a fully representative baseline sample (fixed panel) at the onset of the survey that is supplemented by a sample to cover the ‘births’ in the population at each round. This design had been used for the SDR since its inception through 2013. A similar design was used for the SDR’s sister survey, NSCG from 1993 through 2008. In the past, even if the majority of the base sample were followed in subsequent surveys, both surveys were mainly used for cross-sectional purposes only.

For illustration, suppose that 2017 is the base survey round for the future SDR (see Exhibit 5). Then, the 2017 SDR sample is considered as the baseline sample and, at each round, a supplemental sample will be added to cover the new cohort of the population.

Exhibit 5. Fixed Panel plus Births Design with 2017 as the Base Year

	2017	2019	2021	2023	2025	2027
<i>Base sample</i>	√	√	√	√	√	√
<i>2019 new cohort</i>		√	√	√	√	√
<i>2021 new cohort</i>			√	√	√	√
<i>2023 new cohort</i>				√	√	√
<i>2025 new cohort</i>					√	√
<i>2027 new cohort</i>						√

As mentioned in section B.1, a key advantage of the fixed panel plus births design is to have survey methods and infrastructure stable over time, which encompasses the sample design, survey operation, and data processing. Thus, the fixed panel plus births design would be the most cost-effective option among design options considered by the panel.

In considering how to incorporate longitudinal design features into the future SDR design, a fixed panel plus births design can still be a viable option. As in the past, it will provide a representative sample for the cross-sectional population of the SDR and, at any given point in the survey cycle, longitudinal cohorts can be defined and followed in subsequent surveys. With well-planned sample maintenance in subsequent survey rounds, longitudinal panels can be kept as long as needed for longitudinal analyses.

However, because the longitudinal cases are embedded within the overall sample, the creation of longitudinal weights becomes substantially more complicated, with potentially different weights for each combination of years being considered. This can (and most likely should) be simplified by only producing such weights for pre-specified years that correspond to target NCSES longitudinal estimation objectives. In order to ensure that the sample sizes remain sufficient for these longitudinal objectives and that the necessary data are collected from those cases, they will need to be treated, in most respects, as a separate panel. Hence, when trying to incorporate longitudinal estimation capabilities within this design, it is likely to end up sharing many characteristics with the pure split-panel design.

As discussed above, the SDR population size monotonically grows over time, outpacing population size shrinkage. For repeated surveys like the SDR, with budget constraints, the sample size is usually fixed for a time period of several survey rounds. And even in the event of sample size changes, the sample tends to be reduced mainly due to budget cuts. In order to maintain an equal sample size over time, while decreasing the sampling rate, guiding principles for sample cut need to be prepared. In particular, incorporating cross-sectional and longitudinal analytics goals into sample size determinations would be a key to maintaining the sample size for such critical dual analytical purposes.

A fixed panel plus births design could be considered with the 2015 SDR sample as the base year. The 2017 SDR sample included all 2015 panel members, which established a solid foundation for building the longitudinal panel into the future. The overall sample size would be kept at 120,000, consisting of a panel sample and a new cohort sample at each cycle. The sample stratification and sample allocation

could be determined to meet both cross-sectional and longitudinal estimation goals. The 2019 SDR sample would then cover the full cross-sectional population as well as the maximum data pool for longitudinal cohorts of interest.

C.4 Special Designs Proposed by Data Users

Prior to the initiation of the work by this panel, data users had proposed two options for longitudinal designs for the SDR:

- **Option 1: Increase the sampling probability of previous respondents.** This implies redrawing the sample in 2019 and oversampling frame groups 1 and 3 (previous respondents in 2013 and in prior years) with higher selection probabilities.
- **Option 2: Replace people who are not respondents of the 2015 SDR with people who were respondents in past SDR cycles.** The replacement cases would be selected to match the sampled non-respondents based on basic demographics (e.g. gender, age, immigrant status).

These proposals arose to address the reduction in the number of SDR microdata records that contain survey data across multiple observation years starting in the 2015 SDR. Such records provide valuable data that were used to conduct analyses on the career tracks of members of the SDR population, and there was concern that such analyses would no longer be possible going forward. Design options 1 and 2 were considered but ultimately rejected by the expert panel. What follows is a discussion of why the expert panel did not recommend that options 1 or 2 be adopted.

Option 1 involves returning to the sampling design that was in place until 2013 (at least with respect to the inclusion of previously interviewed cases). While this is certainly feasible from a statistical perspective, it carries some significant drawbacks. First, it could result in a substantial reduction in the level of precision of cross-sectional estimates at the finest level of granularity, depending on the extent to which previously sampled cases are reintroduced. Second, while it will provide more records with data in prior years (albeit not continuous, given the gap between 2013 and 2019), it would substantially complicate the creation of an official longitudinal dataset as described in section A, by restricting the ability of NCSES to redesign the overall sample to satisfy both the cross-sectional and the new longitudinal requirements.

Option 2 is even more difficult to implement. A direct replacement of non-responders with prior responders who have matching demographics is a form of quota sampling. While this is often done in less rigorous surveys because it is cost- and time-effective, it is not considered a valid randomization-based method, resulting in the potential for biased estimates and the inability to quantify the sampling uncertainty.

A different way to implement option 2 that avoids quota sampling is by dropping non-responders as before, but replacing them through a formal randomization procedure to select replacement members from among previously eligible cases. This is conceptually possible, but complicated in practice, because valid estimation and inference would require that the overall random selection process, which is composed of two rounds of selection as well as a round of de-selection, be traced and quantified. This selection-deselection-selection process would need to be quantified not just for the newly added cases but for all cases in the current SDR sample. Given the number of changes that were applied over time, both to the sampling and to the de-selection procedures, it is likely that this option could not be feasibly implemented. An additional issue is that option 2 also specifies that the newly added cases be chosen

from among past respondents. This implies selecting cases based on their past behavior, not just based on externally observed characteristics such as their demographics. Such “selection endogeneity” is not considered valid from the standpoint of design-based theory, because of the potential for biased estimation.

One approach to implementing option 2 that avoids the methodological and practical issues mentioned above is to add cases with prior history to the SDR, but treat them as a self-representing subsample. Under this treatment, each case represents only itself in the population and is given a weight of 1. They can then be included in the SDR datasets with that weight, while the remaining observations are weighted as before, to represent the remainder of the population. Given the range of weights in the SDR, this implies that the contribution of these cases to the overall estimates for the doctoral population will be extremely small. Hence, it might be easier to treat them as a completely separate, non-representative sample and not add them as a self-representing subsample, for the purpose of producing population-level estimates. It should be noted that calling this sample “non-representative,” does not mean that they are inherently different from the rest of the doctoral population, but rather, that they cannot be incorporated into a representative design-based sample because of lack of formal and quantifiable random selection mechanism.

D. SDR design recommendations

We recommend that NCSES implement a split-panel design consisting of a cross-sectional panel and a single longitudinal panel. Because NCSES is still in the process of developing its longitudinal objectives, we recommend that the longitudinal panel be of modest size and duration, and continue to cover the full SDR population in its base year. This would provide flexibility for NCSES to continue working on defining longitudinal objectives without limiting their scope a priori, and also to start the development of longitudinal questionnaires and estimation methods. Starting a general longitudinal panel will enable NCSES to produce an initial set of official longitudinal estimates within a few years, which can then be expanded over time as appropriate.

Starting with the initial release and continuing in subsequent years, NCSES should also provide weighted longitudinal datasets for analytic purposes. With access to weights and accompanying design information to enable design-based inference, researchers will be able to perform longitudinal analyses that fully account for potential design informativeness if desired, something that is not currently possible.

However, these datasets will only become available several survey cycles after the start of the longitudinal panel, and then only contain a limited number of years of data. Therefore, to ensure that researchers continue to be able to conduct research on topics related to doctoral workforce trajectories with current data, we suggest investigating the feasibility of creating an additional non-representative analytic panel as described at the end of section C.4. That panel would only need to be maintained until the official longitudinal panel with sufficient years of data is ready for release, but would provide a valuable transition in the meantime.

E. Design Evaluation Criteria

Criteria will be developed to evaluate the split panel design recommended above. These criteria will include considerations such as evaluating the sample size and the ability to compute precise estimates over the duration of the panel. After important longitudinal outcomes are identified and the fine fields

are refined, data from past rounds of the SDR will be used to conduct power analyses to determine the sample size needed at the end of the ten-year period, to meet precision requirements for the outcomes and fine fields of interest. This sample size will then be inflated to account for expected attrition and response rates at each round of the study, to determine the starting sample size. Finally, the starting sample size will be evaluated to determine if it is reasonable, given the budget.

Estimates for splitting the overall panel sample size into various domains of interest will also be prepared, to see how well subsets of the general population can be analyzed. Given that the SDR stakeholders at the HREP meeting have interest in various subpopulations, a general population sample may not yield a sample size large enough for precise estimates for all of the subpopulations of interest. However, the final design will need to be able to produce precise estimates for a sufficient set of subpopulations.

To determine response rates, response patterns over time for the SDR will be examined. This will include an evaluation of the effect on bias and response rates of excluding sample members who do not respond to one or two rounds of the study.

After the sample size is determined, design effects will be estimated for both the cross-sectional and longitudinal portions of the sample using previous SDR data. These estimated design effects will be compared to the current SDR design effects, to determine the degree of similarity.

F. Future Research

After a design approach is selected, and as the longitudinal panel is implemented and progresses over time, we recommend that the panel be monitored so that adjustments can be made, when possible, and so that we can learn for future SDR longitudinal panels. We also recommend that the attrition and response rates be monitored for each round of the study to see if the assumptions used to set the sample size were accurate. Depending on the final decision made about following non-respondents, we recommend that a nonresponse bias analysis be conducted to evaluate the effect of non-respondents not being followed, or being followed but not having data for a given round.

In addition, as the SDR stakeholders use and analyze the longitudinal data, we recommend that NCSES consult with stakeholders to understand their analyses and to help determine what weights are needed for stakeholder analyses. Also, we recommend that NCSES ask the stakeholders to identify where SDR sample sizes are and are not sufficient for their analyses of specific outcomes and domains.

References

Finamore, J., Cohen, S., Hall, D., Walker, J., and Jang, D. (2011) "NSCG Estimation Issues when Using an ACS-based Sampling Frame," 2011 ASA Proceedings pp.4408 - 4422.

Legg, J.C. and W.A. Fuller (2009), "Two-phase sampling," in Handbook of Statistics, D. Pfeffermann and Rao, C.R. (Editors), Vol. 29A, 55-70, North-Holland.

Lohr, S.L. (2009), "Multi-frame surveys," in Handbook of Statistics, D. Pfeffermann and C.R. Rao (Editors), Vol. 29A, 71-88, North-Holland.

Ohlsson, E. (1995). Coordination of Samples Using Permanent Random Numbers. In *Business Survey Methods*, eds. B.G. Cox, D.A. Binder, D.N. Chinnappa, A. Christianson, M.J. Colledge, and P.S. Kott. New York: John Wiley & Sons, Inc., 153-169.

U.S. Census Bureau (2006). *Current Population Survey Design and Methodology*, Technical Paper 66.

SECTION G.2.

**Summary Report from the Human Resources Expert Panel
(August 28-29, 2017)**

Summary Report from the Human Resources Experts Panel
Determining an Effective, Efficient, and Sustainable Longitudinal Design
for the Survey of Doctorate Recipients
(Meetings held August 28-29, 2017)

The Human Resources Expert Panel (HREP) was held on August 28 and 29, 2017, in Arlington, VA. The panel was comprised of experts who are knowledgeable about the Survey of Doctorate Recipients (SDR) and other surveys of the nation's highly educated workforce. Panelists came from academia, professional organizations, and federal agencies. The purpose of this panel was to review and recommend improvements to the SDR, one of several National Center for Science and Engineering (NCSES) Human Resources Survey (HRS) surveys.

The SDR is a biennial survey conducted since 1973 that provides demographic and employment information about individuals with a research doctoral degree in a science, engineering, or health field from a U.S. academic institution. In 2015, NCSES significantly expanded the SDR's analytical objectives, resulting in a substantial sample expansion from 47,000 cases in 2013 to 120,000 cases in 2015. The main objective of this expansion was to produce reliable estimates of employment outcomes by a finer field of degree level than was possible in the past. Now, with the 2015 SDR cycle complete, NCSES is exploring sample design options for future SDR cycles that would support longitudinal goals alongside the survey's long-standing cross-sectional estimation goals.

This summary report identifies the goals of the HREP and the discussions at the meeting related to the goals.

Goal 1: Identify and discuss SDR cross-sectional estimation objectives that take advantage of the fine field of degree sample expansion

NCSES reviewed the design of the SDR and their estimation goals, which require cross-sectional data. Prior to the 2015 SDR, the numbers in many fine fields of degree were suppressed due to reliability or disclosure issues. Therefore, policy makers and other analysts were unable to explore employment outcomes for different subpopulations such as those defined by sex and/or race at these fields. The expanded 2015 sample size also allowed NCSES to measure these populations for many fine fields. The 2015 SDR sample also provided full coverage of sample members who lived abroad. The 2017 SDR used the same sample as the 2015 SDR with the addition of 7,000 members who became eligible since the 2015 sample draw.

NCSES uses SDR data to provide stakeholders a current and representative picture of the science, engineering, and health workforce. Stakeholders include federal agencies, professional organizations, and academic researchers, among others. Therefore, it is imperative that the survey sample represent the target population. Panel members said they use the data for a wide variety of uses that require biannually representative data. Generally, panel members examine trends over time. For one representative from a professional organization, the fine field expansion allows the examination of the subfields of their discipline. For a researcher who investigates racial and gender issues, the expansion reduces the number of items that are suppressed. Another researcher looks at emerging fields, which may appear sooner in the data with the expanded sample size. A federal government employee specified the need for investigation of fields by sector – some industry sectors employ a small population of certain types of respondents, so those numbers have been suppressed in the past. Another federal government employee seeks to understand the connection between education,

training, and job outcomes, which does not necessarily require a nationally representative sample, but requires resolution at the variables of interest.

Goal 2: Identify and discuss SDR longitudinal estimation objectives

Up until 2015, to maximize operational efficiencies, the same population was surveyed every cycle with the addition of a portion of the newly eligible population and the elimination of those who “age out” of the target population. NCSSES produced biennial cross-sectional estimates from this data, but also released microdata files for each cycle with the individual observations. Researchers have been able to link these microdata files to construct a file with repeated measurements for survey respondents. The researchers have used these repeated measurements for longitudinal analyses. Since the newly eligible population is much greater than the newly ineligible population, deselection was employed to maintain a sample size within the resource constraints. Since NCSSES did not have official longitudinal estimate objectives for the survey, the cuts were random and reduced the longitudinal capabilities of the data set each year.

Panel members’ longitudinal objectives focused on “time-to-event” and “time-in-state” variables and their relationship to outcomes. For example, one member looks at the changes in outcomes by the amount of time spent in a post doc position. The current questionnaire asks respondents about a specific point in time, so the data collection misses any changes that occurred between the two cycles. If the questionnaire was designed to collect items such as time spent in a specific state of interest (such as performing R&D), then a cross-sectional survey would suffice for this analysis. However, one panelist cautioned about respondent’s ability to recall past details. In addition, another pointed out that it is impossible to think of all the future research questions one may have, and so a longitudinal data collection would work better for this work.

Goal 3: Review the longitudinal design options NCSSES is considering and provide feedback on the advantages and disadvantages of each option as they relate to the uses of SDR data.

Four longitudinal designs were presented to the panel, each based on extensive work since May 2017 by the SDR Sample Design Panel: (1) a fixed panel plus births, (2) a rotating panel, (3) a cohort panel and (4) a split panel, which combines two of these different designs. Members felt the first option would meet their research needs, especially because it was the design of the SDR until 2015. One said that they would address any biases resulting from attrition with regression techniques. Panelists had concerns about a rotating panel design because the time period a person would be in the sample would be limited. Panelist had varying views on panel length, suggested values ranged from 10 years to forever. A few panelists suggested varying the frequency of follow-ups – with fewer follow-ups during more settled career periods. Panelists said that the cohort option would not meet their needs because of the long time between the starts of new cohorts, which would miss new research doctorates, and the potential lack of representativeness of the groups close to the end of the cohort period due to potential changes in the research doctorate population and/or panel attrition. Panelists said that the split panel provided flexibility, balancing the need for newly eligible members every cycle, the need for longitudinal data collection, and sample size constraints.

The panelists also reviewed the following potential subpopulations for a longitudinal panel: early career, late career, other demographics, and outcome-based panels. Groups of meeting attendees reviewed each potential subgroup and shared their resulting design choices with the larger group. Ultimately, there were challenges to each potential population and a clear preference was not revealed. However, the group felt that an outcome-defined group would be hard to define with the current survey

instrument and noted that most investigators are interested in measuring outcomes of a group, not pre-defining the outcome(s).

Conclusion

Discussions highlighted the trade-offs that must occur in determining an effective, efficient, and sustainable longitudinal design for the SDR. If the cross-sectional sample is effectively reduced, then more data suppression will need to occur. If old SDR cases are preferentially brought in at a high rate, the weighting complexity would increase. This would result in higher coefficients of variation, increasing the need for data suppression for reliability. If a deliberative longitudinal design is not implemented, then in time the effective longitudinal panel may not be representative of the population due to attrition and non-response biases. In addition, they may be cut through random deselection, reducing the power of the data analyses.

SECTION G.3.

**Summary Report from the Technical and Analytical
Support for the SDR Longitudinal Design
(April-December 2018)**

**Summary Report from the Technical and Analytical
Support for the SDR Longitudinal Design
(Meetings held April-December 2018)**

From April – December 2018, NCSES convened a panel of five experts in survey sample design to review plans for the National Center for Science and Engineering Statistics (NCSES) Survey of Doctorate Recipients (SDR). The goal was to review and vet several proposed changes to the SDR sample selection and the new longitudinal design.

NCSES staff gave presentations on the background of SDR, proposed changes to the survey design, and proposed changes to the sampling procedures to support that design. The SDR has been conducted since 1973 and is a panel study designed to provide biennial demographic and employment trend data about individuals with a research doctoral degree in a science, engineering, or health (SEH) field from a U.S. academic institution who work in the U.S. The frame and target population has changed over the years due to changing resources and analytical goals. For the 2015 SDR, for the first time NCSES went back to the current survey frame (the Doctorate Records File, which is updated with the Survey of Earned Doctorates) and redrew the entire sample, focused on obtaining reliable estimates for more than 200 fine degree fields. After two survey waves, NCSES could only produce a handful of tables with estimates at that level of detail.

For the 2019 SDR sample, NCSES proposed the following revisions to the sampling procedure.

- **Follow-up rule.** NCSES proposed to drop people who have not responded in the first two waves. For the 2019 sample, NCSES proposes to drop sample cases that did not respond to both the 2015 and 2017 wave. A weight adjustment will be incorporated to account for dropped cases.
- **Degree stratification.** NCSES proposes that for future sample selection (new cohorts and supplemental samples) to change the number of fields of degree in the stratification from over 200 to 77 levels.
- **Supplement.** With the follow-up rule proposed in the first bullet above implemented and a sample goal of 120,000 cases, NCSES proposes to supplement the current sample to reach the 120,000 sample cases.

The experts approved the revised sampling procedure, noting that they did not see any issues to the proposal nor did the experts propose any further improvements to the sampling approach.

A second goal was to gather feedback on the longitudinal design and sampling. NCSES will implement the 2017 SDR Expert Panel recommendation of a general-purpose longitudinal panel followed for 10 years. The panel will have the 2015 SDR as the baseline survey and thus be subsampled from the 2015 SDR respondents. The sampling approach NCSES proposed is as follows.

- **Strata:** employment sector x age group x underrepresented minority (URM) x sex (with some collapsing of the categories with very small populations)
- **Implicit stratification:** sex, residential location, race/ethnicity, disability indicator, career stage, U.S. citizenship at the degree time, minor field of study
- **Allocation:** (i) proportional allocation with a subsequent redistribution of a portion of the strata with large populations to small strata with small populations to meet the minimum precision requirement for each stratum, (ii) final cell allocation calculated to account for nonresponse.

- **Sample selection:** Probability proportional to size (PPS) systematic sampling where the size is equal to the SDR 2015 final weight.

The expert panel discussed a variety of facets of this approach. In the end they did not find any major issues with this approach either. NCSES plans on continuing the discussion past this meeting and into a larger stakeholder meeting to be held at a later date.