Attachment E

2013 Survey of Doctorate Recipients: Sample Design and Implementation Report

2013 SURVEY OF DOCTORATE RECIPIENTS:

Sample Design and Implementation

PREPARED FOR:

Lynn Milan, SDR COTR

National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
(703) 292-5111

Lori Thurgood
SRI International
1100 Wilson Blvd., Suite 2800
Arlington, VA 22209
(703) 247-8528

PREPARED BY:

Brenda G. Cox, SRA

Karen Grigorian

Michael Yang

Mike Sinclair

NORC at the
University of Chicago
55 East Monroe Street
Chicago, IL 60603
(312) 759-4000

FEBRUARY 11, 2013

NORC
at the UNIVERSITY of CHICAGO

This report was prepared by NORC at the University of Chicago for SRI under Subcontract 52-000059, which is in turn conducted for the National Science Foundation under Prime Contract Number NSFDACS12C1299 (Task Order 8NSFSRS083228). The NORC internal Project Number is 7442.

This version of the report contains suppression of small cells for confidentiality reasons and may be publically released. Please contact NORC for further information regarding this document or for information about the suppressed version named *2013 SDR Sample Design and Implementation Report_11Feb2013_Final.docx*.

NORC Authors
Karen Grigorian
Michael Yang
Mike Sinclair

SRA, International Author
Brenda G. Cox

Table of Contents

1.   Overview of the 2013 SDR Sample Design

Since its inception in 1950, the National Science Foundation (NSF) has been charged to "Provide a central clearinghouse for the collection, interpretation and analysis of data on scientific and technical resources in the United States, and provide a source of information for policy formulation by other federal agencies" (NSF Web Site 2012). The Survey of Doctorate Recipients (SDR) has been an important means for the NSF to accomplish this objective. Conducted biennially since 1973, the SDR follows a sample of U.S. trained doctorates in science, engineering, and health (SEH) throughout their careers from shortly after degree award by a U.S. institution through age 75. The SDR is widely used by the U.S. Congress and Federal agencies, universities and professional societies, and other organizations, and individuals interested in knowing more about the nation's education, supply, and employment of doctorate recipients in SEH fields. Employers in universities, industry, and government sectors also use the SDR to understand and predict trends in employment opportunities and salaries for doctorates in SEH fields.

Until the 2003 survey cycle, the SDR restricted data collection to U.S. residents. The 2003 SDR included two methodological changes to determine whether data could be successfully collected from U.S. trained SEH doctorates that reside outside the U.S. The first change expanded the data collection for the traditional SDR by completing surveys with sampled cases discovered to be living outside the U.S. The second change attempted data collection regardless of country of residence for a sample of the non-U.S. citizens with degrees awarded in the 2001 and 2002 academic years. These non-U.S. citizens were ineligible for inclusion in the new cohort portion of the traditional SDR frame because they reported plans to emigrate after degree receipt (Grigorian and Hoffer, 2005). Collection of data from international residents proved to be feasible and a formal design was developed for the international survey and its longitudinal panel of U.S. trained SEH doctorates that reside outside the U.S. in the 2006 cycle and continued in the 2008 cycle. (Cox, Grigorian, and Yang, 2006). Both the main and international samples targeted U.S.-trained SEH doctorates younger than 76 years on the survey reference date, but the main SDR target population was restricted to those residing in the U.S., and the eligible international target population was restricted to those residing outside of the U.S. As a result of this sample

segregation, the potential analytic power of both samples was diminished in terms of coverage and sample sizes. Survey data collected from non-U.S. residing respondents from the main SDR sample and survey data collected from U.S. residing respondents from the international sample were not utilized for analysis.

To address this issue, the NSF decided to integrate the two surveys to create a unified survey of U.S. trained SEH doctorates that provides analysts with the capability of studying and comparing U.S. versus non-U.S. residents in the 2010 survey cycle. NSF decided to refer to the integrated data set as the Survey of Doctorate Recipients (SDR) and its two components as the National Survey of Doctorate Recipients (NSDR) and the International Survey of Doctorate Recipients (ISDR).[1]

The integrated sample design developed for the 2010 SDR has been maintained for the 2013 SDR. This report describes the 2013 SDR sample designs and implementation for both the NSDR and ISDR sample components. We begin by summarizing the main design changes from the 2010 SDR in *Section 2*. These changes were relatively minor and were generally restricted to the development of the sampling frame variables. We then discuss in detail the main parameters of the 2013 SDR design. *Section 3* describes the frame construction process for the different cohorts of the population. *Section 4* presents the 2013 SDR stratification scheme for two sample components. *Section 5* discusses the sample sizes of the 2013 NSDR and ISDR sample components. *Section 6* discusses the strategy and results of sample allocation across strata and substrata. *Section 7* reports the sample selection procedures. Finally, *Section 8* provides recommendations for future SDR sample design research.

---

[1] Analysts should note that study documentation for previous survey cycles (2003 to 2008) refers to the NSDR as the SDR and that the ISDR designation was not applied to this sample component as it was still in feasibility stage in the 2003 survey cycle.

2.   Design Changes from the 2010 SDR

Changes in a longitudinal study such as the SDR must be documented so that study planners can properly understand and use data from past survey cycles. In addition, analysts need such data to assess whether and to what extent differences in study design may impact on time-series analyses. The changes between the 2013 SDR and 2010 SDR sample designs are minimal as the 2013 design closely followed that of the 2010. The majority of the differences relate to differences in the construction of sampling frame variables.   This section documents the limited differences that do exist between the two designs.

Target Population Definition

- The reference date changed from 1 October 2010 used in the 2010 SDR to 1 February 2013 for the 2013 SDR.

Frame Construction

- The field of study taxonomy for the 2013 SDR new cohort included 15 new fields added by the Survey of Earned Doctorates (SED). Those new fields and how they map to the SESTAT fine field of study code are shown in Table 2.1.

- While the definition of the location status variable (LOCSTAT) did not change from 2010 to 2013 SDR, the method for creating it for the old cohort nonrespondent cases was improved in 2013 SDR. As part of the 2010 SDR post-processing procedures, the most current sample member location variable (RESPLO3) was created for located nonresponse cases using the NSF-approved algorithm.[2] This variable was used to assign LOCSTAT13 for the 2013 SDR old cohort cases that finalized as located nonrespondents in 2010 SDR. In the 2010 cycle, last known address was used to create LOCSTAT10 for the located nonresponse cases, but was not done using the RESPLO3 program code algorithm which is more precise.

- The region of birth variable used to sort ISDR cases was updated. A new region of birth variable (BIREGION) was developed and replaces the region of birth variable used in 2010 ISDR (ISDRCODE). The variable BIREGION more closely aligns with the other

---

[2] For details regarding the algorithm for creating RESPLO3 for the 2010 SDR nonresponse cases, see the memorandum entitled "2010 Survey of Doctorate Recipients Final Data Delivery" addressed to Lynn Milan (NSF) and Dave Edson (MPR) from Karen Grigorian and Lance Selfa (NORC) dated 11 January 2013 included with the 2010 SDR final data delivery files.

place of birth variables used in post-processing and standard publications and the code frame can be found in Appendix A.1.

Table 2.1   New SED Field of Study Codes mapped to SESTAT Codes

New SED Field of Study Code (PHDFIELD)		SESTAT Field of Study Code (NSDRMED)	
Code	Label	Code	Label
415	Robotics	D67	Computer/information sciences
509	Astronomy, Other	871	Astronomy and astrophysics
577	Medical Physics/Radiological Science	878	Physics, except biophysics
316	Structural Engineering	726	Civil engineering
168	Virology	637	Microbiological sciences and immunology
104	Computational Biology	642	OTHER biological sciences
155	Structural Biology	642	OTHER biological sciences
167	Environmental Toxicology	642	OTHER biological sciences
207	Oral Biology/Oral Pathology	786	Medicine (e.g., dentistry, optometry, osteopathic, podiatry, veterinary)
227	Gerontology	731	OTHER health/medical sciences
684	Gerontology	930	OTHER social sciences
806	Urban Education and Leadership	Not applicable; non-SEH field	
808	Educational Policy Analysis		
833	International Education		
912	Hospitality, Food Service and Tourism Management		

- The 2000 Census list of Hispanic surnames became available since conducting the 2010 SDR. As such, this updated Hispanic surname list was used to impute ethnicity for 2013 SDR new cohort cases in the frame when ethnicity was not reported. In 2010 SDR, the 1990 Census list of Hispanic surnames was used.

- While the 2013 SDR old cohort disability status frame variable was constructed as it was in the past, the 2010 SDR questionnaire included a new disability or functional limitation category contributing to an increase in the number of disabled old cohort frame cases. In SDR frame construction, the most recently reported data are used to construct the disability status variable. SDR respondents that classify themselves as having a moderate

or greater difficulty with any disability category in the survey are classified as disabled in the subsequent round's frame file.

Prior to the 2010 cycle, respondents could choose from four disability categories (i.e., difficulty with seeing, hearing, walking, or lifting). Starting with the 2010 SDR, a fifth disability category for reporting difficulty with concentrating, remembering or making decisions was added.

As a result of the added disability category in the 2010 SDR, the number of old cohort frame cases classified as disabled in the 2013 SDR frame file was noticeably greater. To assess the impact of the added category, the disability status was calculated for the 2013 SDR old cohort as it was defined for the 2010 old cohort frame cases using responses from just the four disability categories and compared to the disability status calculated using all five disability categories. This comparison showed that the fifth new cognitive disability category caused an increase in the number of disabled old cohort frame cases of 7.6 percent.

This difference was limited to the old cohort frame cases. The method for deriving disability status for the new cohort was unchanged from the prior cycle. For new cohort frame cases, disability status is derived from the SED variable HANDICAP. If at least one disability was indicated at HANDICAP, the new cohort frame case was coded as disabled. For the 2013 and 2010 SDR new cohort frame cases, the SED disability categories have consistently been blind/visually impaired, deaf/hard of hearing, physical/orthopedic disability, learning/cognitive disability, vocal/speech disability, and other self-specified disability.[3] New cohort frame cases not reporting disability status in SED are imputed to be non-disabled.

Sample Selection

- For the 2013 ISDR, the redefined region of birth variable (BIREGION) was used for sorting instead of the more aggregated birth region variable (ISDRCODE) that was used in the 2010 ISDR.

---

[3] Note that the SED definition of disability is more encompassing than the SDR definition. The SED definition includes a response option for vocal/speech disabilities in addition to an "other specified" response option. SDR has neither of these. Furthermore, SED does not differentiate the degree of disability difficulty; respondents simply report having a disability or not. In SDR, only individuals with a moderate or greater degree of disability are considered disabled.

3.   Frame Development

3.1   Sample Frame Construction

The sample frame construction for the NSDR and the ISDR components were done together reflecting the SDR's integrated sample design. While the target population definitions for these two sample components are different and sample selection is done separately, the frame construction requirements for the variables included in each frame are identical. Thus, this sample frame construction section discusses the frame construction process for the NSDR and ISDR together.

The target population of the 2013 SDR covered individuals who met the following requirements, regardless of residency location:

- Received a doctoral degree in an SEH field from a U.S. institution;
- Age 75 years or younger on February 1, 2013; and
- Living in a noninstitutionalized setting on February 1, 2013.

A sampling frame was constructed to represent the NSDR and ISDR target populations, henceforth referred to as the 2013 SDR frame. A *sampling frame* is a set of elements and a set of procedures for identifying and locating the population elements. The frame usually contains information for sample stratification and sample selection. The goals of frame construction are twofold: one is to represent all population elements in the frame so they all have some known non-zero probability of being selected into the sample; the other is to define auxiliary variables for sample stratification and survey operations. The old cohort frame was developed from the 2010 SDR sample and the new cohort frames were developed from the two most recent cohorts added to the Doctorate Records File (DRF). The DRF is a cumulative database of all U.S.-granted research doctorates constructed using data collected from the SED, an annual census of research doctorates awarded by U.S. academic institutions since 1920.

The 2013 SDR frame was constructed as two separate databases:

1. The old cohort frame was constructed from the 2010 SDR sample (n=45,697) including only those eligible for 2013 SDR (n=44,602) and

2. Approximately half of the new cohort frame was constructed from the 2010 SED records (n=48,034) including only those cases eligible for the 2013 SDR (n=35,242). The other half of the new cohort frame was constructed from the 2011 SED records (n=49,010) including only those cases eligible for 2013 SDR (n=36,664). Because the survey reference date was shifted from October 1 to February 1, 2013, the fully processed DRF was available for both new cohort years (2010 and 2011) at the time of frame-building. Unlike previous survey rounds, when new cohort frames were constructed separately from each SED year's database as it became available, the SDR team was able to build a single new cohort frame file covering both SED survey years.

3.1.1 Frame File Layout

While two separate files (i.e., the old and new cohort frame files) make up the total 2013 SDR sampling frame, the file layout for each frame file is the same and is shown in Table 3.1. The layout describes each variable and its code frame, where feasible. Variables with longer coding taxonomies, such the field of study variables, can be found in Appendix A. When this occurs, it is noted in "Values" column in Table 3.1.

Table 3.1   2013 SDR Sample Frame File Layout

Variable	Description	Format	Length	Values
Case Identifiers				
SU_ID	Survey ID	Char	8	Randomly assigned value
REFID	Reference ID	Char	9	Randomly assigned value
REFID_ORIG	Reference ID prior to 2010 SDR and integration[1]	Char	9	Randomly assigned value
DRF_ID	DRF ID	Char	7	Randomly assigned value
DRF_ID_ORIG	DRF ID initially assigned, and subsequently dropped when a duplicate DRF entry was detected after SDR sample selection.	Char	7	Randomly assigned value
Location Variables				
LOCSTAT13	Most current location indicator; for old cohort cases derived from SMLOC13 and LOCSTAT10, for new cohort cases, derived from PDUS13.	Char	1	1 = Located in U.S.; 2 = Located outside U.S.
SMLOC13	Most current sample member location; for located old cohort cases, derived from RESPLO3; for unlocated old cohort cases, set to 999; and for new cohort cases, derived from DRF variable PDLOC or reported address (if	Char	3	See 2010 SESTAT geocode code frame in Appendix A.1

Table 3.1   2013 SDR Sample Frame File Layout

Variable	Description	Format	Length	Values
	PDLOC is missing or unspecific).			
PDUS13	Post-graduation location derived from DRF variable PDUSFOR for new cohort.	Char	1	1 = Located in U.S. (includes missing in PDUSFOR)
				2 = Located outside U.S.
				9 = NA, old cohort
LOCSTAT10	Location status indicator for old cohort cases.	Char	1	1 = Located in U.S.
				2 = Located outside U.S.
				9 = NA, new cohort
Stratification Variables				
DROP13	Disposition for 2013 round sampling	Char	3	See DROP13 code frame in Table 3.17
STRATUM13	2013 Stratum assignment	Char	3	NSDR=001-150; ISDR=A6, C7-C9, D1-D12, F43-F70
NSDRSTRAT13	2013 NSDR Stratum assignment regardless of sample component membership	Char	3	NSDR=001-150
ISDRSTRAT13	2013 ISDR Stratum assignment regardless of sample component membership	Char	3	ISDR=A6, C7-C9, D1-D12, F43-F70
BASEWGT10	2010 SDR base weight	Num	8	Actual base weight for panel cases from 2010, 1.0-43.0
NSFGRP	NSF demographic group for NSDR	Char	1	1 = Hispanic, regardless of race, citizenship at birth, and disability status
				2 = NH black, regardless citizenship at birth and disability status
				3 = U.S. born, NH Asian regardless of disability status
				4 = NH American Indian, regardless of citizenship at birth and disability status
				5 = NH Pacific Islander, regardless of citizenship at birth and disability status
				6 = U.S. born, disabled, NH white
				7 = U.S. born, not disabled, NH white
				8 = Non-U.S. born, NH white, regardless of disability status
				9 = Non-U.S. born, NH Asian, regardless of disability status
ISDRGRP	ISDR demographic group	Char	1	1 = U.S. citizens at birth
				2 = Hispanic, non-U.S. citizen at birth
				3 = NH black, non-U.S. citizen at birth
				4 = NH Asian, non-U.S. citizen at birth
				5 = NH white, non-U.S. citizen at birth
				6 = NH other race, non-U.S. citizen at birth

Table 3.1   2013 SDR Sample Frame File Layout

Variable	Description	Format	Length	Values
PHDFIELD	Doctoral field of study from the current DRF	Char	3	See DRF field of study code frame in Appendix A.2
PHDFIELD_ORIG	Doctoral field of study from the DRF when initially sampled	Char	3	See DRF field of study code frame in Appendix A.2
PHDFIELD_SDR	Doctoral field of study from the DRF updated with degree changes reported in the SDR and approved by NSF	Char	3	See DRF field of study code frame in Appendix A.2
NSDRMED13	SESTAT field of study code; for old cohorts, this is derived from ND2MED for respondents, and NSDRMED10 for nonrespondents; for new cohorts, this is derived from PHDFIELD	Char	3	See NSDRMED code frame in Appendix A.2
SDRFLD15	15-level field of study used in sampling (formerly SDRFLD)	Char	2	See SDRFLD15 code frame in Appendix A.2
DSTFLD8	8-level field of study used in sampling (formerly DSTFLD)	Char	1	See DSTFLD8 code frame in Appendix A.2
MAJFLD7	7-level field of study used in sampling (formerly MAJFLD)	Char	1	See MAJFLD7 code frame in Appendix A.2
FOD3	3-level field of study used in sampling	Char	1	See FOD3 code frame in Appendix A.2
SEX13	Sex or gender indicator	Char	1	1 = Male; 2 = Female
HCAPIN13	Disability status indicator	Char	1	Y = Disabled; N = Not disabled
HISPANIC13	Hispanic ethnicity indicator	Char	1	1 = Hispanic; 2 = Not Hispanic
HISPCAT13	Hispanic group	Char	1	1 = Mexican; 2 = Puerto Rican; 3 = Cuban; 4 = Other Hispanic
ASIAN13	Asian race indicator	Char	1	1 = Asian; 2 = Not Asian
BLACK13	Black race indicator	Char	1	1 = Black; 2 = Not Black
NATIVE13	American Indian race indicator	Char	1	1 = American Indian; 2 = Not American Indian
PACIFIC13	Pacific Islander race indicator	Char	1	1 = Pacific Islander; 2 = Not Pacific Islander
WHITE13	White race indicator	Char	1	1 = White; 2 = Not White
RACE13	Race-only indicator, independent of ethnicity	Char	1	1 = Asian; 2 = Black; 3 = American Indian; 4 = Pacific Islander; 5 = White; 6 = Multi-race
RACETH13	Concatenated race/ethnicity value	Char	20	Concatenation of Ethnicity and Race in the form of ETH-RACE Ethnicity: HISP, NH Race: ASIAN, BLACK, NATIVE, PACIFIC, WHITE
BIRCIT13	Citizenship at birth indicator	Char	1	1 = U.S. citizen at birth ; 2 = Non-U.S. citizen at birth
Sort Variables				
PHDFY	Fiscal (academic) year of doctorate in the current DRF	Num	4	1958-2011, cases before 1958 have missing data
PHDFY_ORIG	Fiscal (academic) year of doctorate from the DRF when initially sampled	Num	4	1958-2011, cases before 1958 have missing data
SDRAYR	Fiscal (academic) year of doctorate with year changes reported in the SDR and approved by NSF	Num	4	1958-2011, cases before 1958 have missing data

Table 3.1   2013 SDR Sample Frame File Layout

Variable	Description	Format	Length	Values
BTHST13	Geocode for state/country of birth	Char	3	See 2010 SESTAT geocode code frame in Appendix A.1
BIREGION	Region of birth used for sorting of the new cohort; replaces ISDRCODE from 2010 SDR	Char	6	See Birth Region crosswalk in Appendix A.1
MOB_13	Month of birth known at start of 2013 round	Num	2	1-12, -3 = missing
DOB_13	Day of birth known at start of 2013 round	Num	2	1-31, -3 = missing
YOB_13	Year of birth known at start of 2013 round	Num	4	1934-1992, -3 = missing
AGE13	Age on the 2013 SDR reference date	Num	2	21-75
AGEYR13	Year of birth reported and imputed	Num	4	1934-1992
Data Source Variables				
INSDRMED13	SESTAT field of study code source flag	Char	2	
ISDRAYR	Fiscal year of doctorate source flag	Char	2	
ISEX13	Sex source flag	Char	2	
IHCAPIN13	Disability status source flag	Char	2	
IHISPANIC13	Hispanic ethnicity source flag	Char	2	
IHISPCAT13	Hispanic group source flag	Char	2	
IASIAN13	Asian race source flag	Char	2	
IBLACK13	Black race source flag	Char	2	
INATIVE13	American Indian race source flag	Char	2	See Source Flag code frame in Appendix A.3
IPACIFIC13	Pacific Islander race source flag	Char	2	
IWHITE13	White race source flag	Char	2	
ILOCSTAT13	Location status source flag	Char	2	
IPDUS13	Post-graduation location source flag	Char	2	
IBIRCIT13	Birth citizenship source flag	Char	2	
ICURCIT13	Current citizenship source flag	Char	2	
IBTHST13	Birth state/country source flag	Char	2	
IAGE13	Age source flag	Char	2	
Operational Variables				
SDRTYP13	2013 SDR sample component assignment	Char	4	NSDR or ISDR
SAMPTYPE13	Sample Type*	Char	2	01 = 2010 Refusal
				02 = 2010 Cooperative
				03 = 2010 NIR
				05 = New Cohort
				06 = New Cohort—SED SM Refusal
				07 = New Cohort—MIL/MIR/Other nonresponse
SURVEY10	Completed survey in 2010 round	Char	1	Y = Yes, completed survey; N = No, did not complete survey; L = new cohort

Table 3.1   2013 SDR Sample Frame File Layout

Variable	Description	Format	Length	Values
STRATUM10	Stratum assigned in 2010 round	Char	3	NSDR = 001-150; ISDR=A6, C7-C9, D1-D12, F43-F70; New cohort = XXX
SDRTYP10	2010 SDR sample component assignment	Char	4	NSDR, ISDR, or NEW (for new cohort)
ORIGCOMP	Sample component or frame to which a case was initially allocated.	Char	4	NSDR or ISDR
CURCIT13	Current citizenship indicator	Char	1	1 = Currently U.S. citizen; 2 = Not U.S. citizen currently
PDOCSTAT	Post-graduation status in the DRF.	Char	1	0 = Returning to, or continuing in, pre-doctoral employment
				1= Signed contract or made definite commitment
				2 = Negotiating with a specific organization, or more than one
				3 = Seeking appointment but have no specific prospects
				4 = Other full-time degree program
				5 = Do not plan to work or study
				6 = Other
				A = Has postdoctoral fellowship
				9 = Missing
PREVDOC	Flag to indicate if the SM has earned a U.S. research doctorate before the sampled degree according to the DRF denoted in DRF variables PHDCOUNT and PREVDRF	Char	1	1 = Sampled degree is first and only doctorate
				2 = Prior doctorate is non-SEH doctorate
				3 = Prior doctorate is SEH doctorate (ineligible)
PROFDEG	DRF variable that indicates if a professional degree is earned or in progress	Char	1	0 = MD from U.S. institution
				1 = DVM from U.S. institution
				2 = DDS, DMD from U.S. institution
				3 = Other medical from U.S. institution
				4 = All other doctorates from U.S. institution
				5 = MD from non-U.S. institution
				6 = DVM from non-U.S. institution
				7 = DDDS, DMD from non-U.S. institution
				8 = Other medical from non-U.S. institution
				9 = All other doctorates from non-U.S. institution
				M = No other degree reported
DRF_REF	Indicator of refusal to complete DRF	Char	1	Y = Explicitly refused to complete SED; N = Did not explicitly refuse
ETHN_REF_DRF	Indicator of refused ethnicity in DRF	Char	1	Y = Ethnicity refused in the DRF; N = Ethnicity reported in the DRF; M = SED nonrespondent

Table 3.1   2013 SDR Sample Frame File Layout

Variable	Description	Format	Length	Values
RACE_REF_DRF	Indicator of refused race in DRF	Char	1	Y = Race refused in the DRF; N = Race reported in the DRF; M = SED nonrespondent

DRF = Doctorate Records File; NH = Non-Hispanic.

[1] In 2010 SDR, REFID was reassigned for cases originally sampled for ISDR. Prior to 2010 SDR, ISDR REFIDs started with "30"; these cases are currently assigned REFIDs starting with "2I".

3.1.2 Missing Data Imputation Rules for Sampling Stratification and Sort Variables

While there are many variables in the sampling frame file, there are only a few sampling stratification variables which define the strata, and only five of these may have missing data. One sort variable is also imputed when there are missing data. The six sampling stratification and sort variables that might have missing data are as follows:

1. RACETH13, derived from ASIAN13, BLACK13, HISPANIC13, NATIVE13, PACIFIC13, and WHITE13
2. SEX13
3. LOCSTAT13
4. BIRCIT13
5. HCAPIN13
6. AGE13

The imputation rules and the amount of missing data for each of these sampling stratification variables in the 2013 SDR frame file are detailed below.

RACETH13. RACETH13 was constructed from the separate race/ethnicity variables ASIAN13, BLACK13, HISPANIC13, NATIVE13, PACIFIC13, and WHITE13 after they were fully imputed. RACETH13 is defined in the following hierarchical manner:

- If a case is Hispanic or Latino, assign the case to the Hispanic value regardless of race;
- If a case is not Hispanic (NH) and is black, assign the case to the NH black value regardless of other race selections;
- If a case is not Hispanic or black, and is Asian, assign the case to the NH Asian value regardless of other race selections;
- If a case is not Hispanic, black, or Asian, and is American Indian or Alaskan Native, assign the case to the NH American Indian value regardless of other race selections;

- If a case is not Hispanic, black, Asian, or American Indian, and is Native Hawaiian or other Pacific Islander, assign the case to the NH Pacific Islander value regardless of other race selections; and

- Otherwise, assign the case to NH white.

Race/ethnicity variables are reported in either the SED or the SDR. When multiple reports exist, the most current report was used. Despite attempts to obtain this information in the SED and SDR surveys, some amount of missing data existed. The rules used for defining the race and ethnicity variables in 2013 SDR frame are as follows:

1. Use reported data from the most current version of the SDR;
2. Use reported data from the SED;
3. When ethnicity is missing, use the U.S. Census Bureau Hispanic surname list and logically impute any matches as Hispanic ethnicity (if race is also missing and the surname is Hispanic, impute the race to white);[4]
4. When race is missing, and ethnicity is either missing or non-Hispanic, use the GENESYS Asian surname list[5], and logically impute any matches as NH Asian;
5. When ethnicity is still missing, but race is reported, use place of birth to logically impute ethnicity;
6. When race and ethnicity are both still missing, use place of birth to logically impute race and ethnicity;
7. Where hot deck imputation exists from a past survey cycle, use the hot deck imputed values; and
8. When race and ethnicity are both still missing and place of birth is missing, impute to NH white.

The crosswalk of birth places to race and ethnicity imputation assignments is located in Appendix A.1. The sources for race and ethnicity data in the 2013 SDR frame files are detailed in Tables 3.2 and 3.3. The distribution of the resulting race/ethnicity group assignments is shown in Table 3.4.

---

[4]The 2013 new cohort cases were updated using the Hispanic surname list based on the 2000 U.S. Census available as of 2011 located at http://www.census.gov/genealogy/www/data/2000surnames/index.html. The 2013 old cohort cases were updated using the Hispanic surname list based on the 1990 U.S. Census.
[5] Market Systems Group provides the GENESYS Sampling Systems suite of sampling tools, which includes this algorithm that matches surnames to an Asian surname list for a nominal fee (http://www.m-s-g.com/Web/genesys/index.aspx).

Table 3.2 Race Data Sources: 2013 SDR Frame

Race Data Source	Total Cases	2010 Panel	2010 SED	2011 SED
Self-reported	109,494	43,486	32,396	33,612
Surname imputation (Asian)	1,467	137	604	726
Birthplace imputation	2,175	782	738	655
Hotdeck imputation	51	51	0	0
Default imputation (white)	3,321	146	1,504	1,671
Overall	116,508	44,602	35,242	36,664

Table 3.3 Ethnicity Data Sources: 2013 SDR Frame

Ethnicity Data Source	Total Cases	2010 Panel	2010 SED	2011 SED
Self-reported	110,205	44,043	32,310	33,852
Surname imputation (Hispanic)	315	24	148	143
Birthplace imputation	1,560	299	761	500
Hotdeck imputation	51	51	0	0
Default imputation (non-Hispanic)	4,377	185	2,023	2,169
Overall	116,508	44,602	35,242	36,664

Table 3.4 Race/Ethnicity Assignment: 2013 SDR Frame

Race/ethnicity Group	Total Cases	2010 SDR	2010 SED	2011 SED
Hispanic	7,591	3,138	2,108	2,345
NH-American Indian	771	339	210	222
NH-Asian	33,487	10,525	11,131	11,831
NH-Black	5,778	2,636	1,545	1,597
NH-Pacific Islander	290	144	72	74
NH-White	68,591	27,820	20,176	20,595
Overall	116,508	44,602	35,242	36,664

SEX13. Sex is primarily obtained from the SED survey data, and is very complete. However, starting with the 2003 SDR, cases with missing sex information completing the survey in an online mode (i.e., telephone interview or web survey) have been asked to identify their sex. If sex information is not in the DRF or reported in the SDR, sex data are updated with results found through Internet searches that reveal the sample member's sex through pictures or other unambiguous documentation (e.g., a sample member is described with female pronouns and thanks her husband for support in her dissertation). Any remaining missing sex data cases are

imputed to be female by default, giving these cases with unknown sex a higher probability of selection.

The sources for the sex data in the 2013 SDR frame files are detailed in Table 3.5. The distribution of the resulting sex assignments is shown in Table 3.6.

Table 3.5   Sex Data Sources: 2013 SDR Frame

Sex Data Source	Total Cases	2010 SDR	2010 SED	2011 SED
Self-reported	116,437	44,562	35,231	36,644
Verified with Internet source	52	35	6	11
Default imputation (female)	19	5	5	9
Overall	116,508	44,602	35,242	36,664

Table 3.6   Sex Assignment: 2013 SDR Frame

Sex Assignment	Total Cases	2010 SDR	2010 SED	2011 SED
Male	70,107	28,945	20,125	21,037
Female	46,401	15,657	15,117	15,627
Overall	116,508	44,602	35,242	36,664

LOCSTAT13. The LOCSTAT13 variable indicates the last known residence location of the sample member prior to 2013 SDR sampling, either in or out of the U.S. For the located 2010 SDR panel cases, this information primarily comes from the survey for respondents and contacting data for nonrespondents. For panel cases not found in the 2010 cycle, the last known residence location is obtained from past SDR cycles or planned post-graduation location reported in the SED. For the new cohort frame, LOCSTAT13 is derived only from planned post-graduation location reported in the SED. Any cases with no residency data from the SDR and the SED are imputed to be in the U.S. by default. The 2010 SDR was the first cycle to use this variable.[6]

---

[6] For more details about the LOCSTAT variable development for the 2010 SDR and continued for the 2013 SDR, see the memoranda "2010 SDR Sample Frame Development Memo #3 – Sample Member Location Variable" sent to Daniel Foley and Steve Cohen, NSF, on April 23, 2010 from Karen Grigorian, NORC, and Brenda Cox, SRA, and "2013 SDR Frame Decisions – Frame File Layout" sent to Lynn Milan, NSF, on September 18, 2012 and finalized October 4, 2012 from Karen Grigorian and Lance Selfa, NORC and Brenda Cox SRA.

The sources for the location data in the 2013 SDR frame files are detailed in Table 3.7. The distribution of the resulting location assignments is shown in Table 3.8.

Table 3.7   Location Data Sources: 2013 SDR Frame

Location Data Source	Total Cases	2010 SDR	2010 SED	2011 SED
SDR	43,488	43,488	0	0
SED	67,820	898	32,867	34,055
Default imputation (in the U.S.)	5,200	216	2,375	2,609
Overall	116,508	44,602	35,242	36,664

Table 3.8   Location Assignment: 2013 SDR Frame

Location Assignment	Total Cases	2010 SDR	2010 SED	2011 SED
In the U.S.	103,087	39,132	31,300	32,655
Out of the U.S.	13,421	5,470	3,942	4,009
Overall	116,508	44,602	35,242	36,664

BIRCIT13. The BIRCIT13 variable indicates the sample member's citizenship at the time of birth, as either "U.S." or "non-U.S." Citizenship information is asked in each round of the SDR, and so for the majority of panel members, this information comes from the SDR survey. For nonrespondents to the SDR and new cohort sample members, this information is obtained from the SED. Cases that have never reported birth citizenship were imputed to be non-U.S. born.

The sources for birth citizenship data in the 2013 SDR frame files are detailed in Table 3.9. The distribution of the resulting birth citizenship assignments is shown in Table 3.10.

Table 3.9   Citizenship at Birth Sources: 2013 SDR Frame

Citizenship at Birth Data Source	Total Cases	2010 SDR	2010 SED	2011 SED
Self-reported in SDR	42,135	42,135	0	0
Self-reported in SED	69,760	2,026	33,298	34,436
Citizenship imputed from DRF with BIRTHPL and PDLOC	48	13	12	23
Default imputation (non-U.S. born)	4,565	428	1,932	2,205
Overall	116,508	44,602	35,242	36,664

Table 3.10   Citizenship at Birth Assignment: 2013 SDR Frame

Citizenship at Birth Assignment	Total Cases	2010 SDR	2010 SED	2011 SED
U.S. born	65,388	28,430	18,284	18,674
Not U.S. born	61,120	16,172	16,958	17,990
Overall	116,508	44,602	35,242	36,664

HCAPIN13. The HCAPIN13 variable indicates the sample member's most current disability status – either disabled or not disabled. Disability information is asked in each round of the SDR, and so for the majority of panel members, this information comes from the SDR survey. Any SDR survey respondent that reports having a moderate or greater disability of any type (e.g., seeing; hearing; walking; lifting; or concentrating, remembering, or making decisions) is considered disabled. For nonrespondents to the SDR and new cohort sample members, this disability information is obtained from the SED. If at least one disability was indicated in the SED disability variable HANDICAP, HCAPIN13 was coded as disabled. The SED disability categories are blind/visually impaired, deaf/hard of hearing, physical/orthopedic disability, learning/cognitive disability, vocal/speech disability, and other self-specified disability. Cases never reporting disability status are imputed to be non-disabled.

The sources for disability status in the 2013 SDR frame files are detailed in Table 3.11. The distribution of the resulting disability status assignments is shown in Table 3.12.

Table 3.11   Disability Status Source: 2013 SDR Frame

Disability Status Data Source	Total Cases	2010 SDR	2010 SED	2011 SED
Self-reported in SDR	42,126	42,126	0	0
Self-reported in SED	65,813	1,942	31,395	32,476
Default imputation (not disabled)	8,569	534	3,847	4,188
Overall	116,508	44,602	35,242	36,664

Table 3.12   Disability Status Assignment: 2013 SDR Frame

Disability Status Assignment	Total Cases	2010 SDR	2010 SED	2011 SED
Disabled	5,394	3,410	960	1,024
Not disabled	111,114	41,192	34,282	35,640
Overall	116,508	44,602	35,242	36,664

AGE13. The AGEYR13 variable indicates the sample member's year of birth and is used to create AGE13 and IAGE13. The primary sources of AGEYR13 are birth year data reported on the SED, supplemented with birth year information collected on the SDR. Any missing data on AGEYR13 are imputed from sample members' bachelor's degree year, if known, or from their doctorate award year, which is known for all sample members. The birth year imputation rules assume that sample members earned degrees at an age somewhat lower than average for the population; when based on bachelor's degree award year, sample members are assumed to be 18 when earning this degree, and when based on doctorate award year, sample members are assumed to be 21 when earning this degree. These younger age assumptions are intentional so to minimize any sample undercoverage caused by eliminating doctorates with missing birth year that may have earned a degree at a young age. During data collection, every effort is made to collect date of birth from sample members with an imputed birth date to confirm their eligibility for the sample. In the next survey cycle, newly obtained unimputed birth date data replace the imputed birth year estimate in frame construction.

The sources for age in the 2013 SDR frame files are detailed in Table 3.13. The distribution of the resulting age assignments is shown in Table 3.14.

Table 3.13   Age Source: 2013 SDR Frame

Age Data Source	Total Cases	2010 SDR	2010 SED	2011 SED
Self-reported in SDR	29,087	29,087	0	0
Self-reported in SED	81,994	15,014	32,875	34,105
BA Year Imputation	1,517	151	717	649
PhD Year Imputation	3,910	350	1,650	1,910
Overall	116,508	44,602	35,242	36,664

Table 3.14   Age Assignment: 2013 SDR Frame

Age Assignment	Total Cases	2010 SDR	2010 SED	2011 SED
Under 35	48,818	2,855	21,223	24,740
35-39	21,575	6,169	8,375	7,031
40-44	12,093	6,767	2,919	2,407
45-49	7,995	5,843	1,122	1,030
50-54	6,727	5,321	743	663
55-59	6,036	5,102	478	456
60-64	5,161	4,652	264	245
65-75	8,103	7,893	118	92
Overall	116,508	44,602	35,242	36,664

SUMMARY OF SAMPLING VARIABLES DATA SOURCES. Table 3.15 summarizes the data source type for the sampling stratification and sort variables subject to imputation. These results are shown by variable and by the three main sample frame components.

Table 3.15 Data Source for Sample Frame Variables Subject to Imputation: 2013 SDR Frame

Sample Frame Component	Sample Frame Variable	2013 SDR Sample Frame Cases		
		Reported Values in the Final Frame	Imputed from a Non-default Rule	Assigned Default Imputation
2010 SDR	Race (RACE13)	43,486	970	146
	Ethnicity (HISPANIC13)	44,043	374	185
	Sex (SEX13)	44,562	35	5
	Location (LOCSTAT13)	44,386	0	216
	Citizenship at birth (BIRCIT13)	44,161	13	428
	Disability status (HCAPIN13)	44,068	0	534
	Birth year (AGEYR13)	44,101	501	0
2010 SED	Race (RACE13)	32,396	1,342	1,504
	Ethnicity (HISPANIC13)	32,310	909	2,023
	Sex (SEX13)	35,231	6	5
	Location (LOCSTAT13)	32,867	0	2,375
	Citizenship at birth (BIRCIT13)	33,298	12	1,932
	Disability status (HCAPIN13)	31,395	0	3,847
	Birth year (AGEYR13)	32,875	2,367	0
2011 SED	Race (RACE13)	33,612	1,381	1,671
	Ethnicity (HISPANIC13)	33,852	643	2,169
	Sex (SEX13)	36,644	11	9
	Location (LOCSTAT13)	34,055	0	2,609
	Citizenship at birth (BIRCIT13)	34.436	23	2,205
	Disability status (HCAPIN13)	32.476	0	4,188
	Birth year (AGEYR13)	34,105	2,559	0
Overall	Race (RACE13)	109,494	3,693	3,321
	Ethnicity (HISPANIC13)	110,205	1,926	4,377
	Sex (SEX13)	116,437	52	19
	Location (LOCSTAT13)	111,308	0	5,200
	Citizenship at birth (BIRCIT13)	111,895	48	4,565
	Disability status (HCAPIN13)	107,939	0	8,569
	Birth year (AGEYR13)	111,081	5,427	0

3.2 Old Cohort Sample Frame Construction

The 2013 SDR old cohort population is composed of doctorates who received their SEH degree prior to July 2010. The old cohort frame is a secondary frame because it is derived from the

panel sample and each frame member carries a sampling weight to represent the old cohort population. The frame construction process for the 2013 old cohort was relatively simple. As noted in *Subsection 3.1*, SDR survey responses were used to update the sample frame variables, whenever possible, and to determine eligibility for either NSDR or ISDR frame inclusion. The following subsections provide the NSDR and ISDR old cohort frame definitions and show the final eligibility status of the 2010 SDR sample for the 2013 cycle.

3.2.1 NSDR Old Cohort Frame Definition

The 2013 NSDR old cohort frame was derived from the 2010 NSDR sample consisting of 40,000 cases. The 2013 NSDR old cohort frame included only cases that met the 2013 SDR target population requirements (e.g., received a doctoral degree in an SEH field from a U.S. institution, age 75 years or younger on the survey reference date of February 1, 2013, and living in a noninstitutionalized setting on the reference date) *and* were last located in the U.S. or one of its territories as defined by the LOCSTAT13 frame variable.

3.2.2 ISDR Old Cohort Frame Definition

The 2013 ISDR old cohort frame was derived from both the 2010 NSDR and ISDR samples. All 2010 ISDR cases that met the 2013 SDR target population requirements were included in the 2013 ISDR old cohort frame and selected for the 2013 ISDR sample with certainty. Additionally, all 2010 NSDR cases that met the 2013 SDR target population requirements and were last located outside of the U.S. or one of its territories (as defined by the LOCSTAT13 variable) were also included in the 2013 ISDR frame and selected for the sample with certainty.

3.2.3 2010 SDR Final Eligibility Status and Frame Assignment

Table 3.16 shows the 2013 SDR old cohort frame status for all 2010 SDR sampled cases. Ultimately, there are 44,602 cases included in the 2013 SDR old cohort frame – 38,424 included in the 2013 NSDR old cohort frame and 6,178 included in the 2013 ISDR old cohort sample.

Table 3.16   2013 SDR Old Cohort Frame Status by 2010 SDR Sample Type

2013 SDR Old Cohort Frame Status		2010 SDR Sample		
		Total	NSDR	ISDR
Eligible		**44,602**	**38,968**	**5,634**
00	NSDR Frame Eligible	38,424	38,424	0
00	ISDR Frame Eligible (selected with certainty)	6,178	544	5,634
Ineligible		**1,095**	**1,032**	**63**
01	Age ineligible	855	816	39
07	Age ineligible, according to BA year or PhD year	1	1	0
11	Non-SEH doctoral degree field per SDR	13	8	5
12	No doctorate degree earned per SDR	1	1	0
13	Duplicate case per SDR	0	0	0
13b	Double Doctorate; first SEH doctorate earned before SED 2010/2011	0	0	0
14	Frame ineligible, not otherwise defined	1	1	0
15	Deceased per SDR	196	179	17
16	Terminally ill per SDR	26	24	2
19	Institutionalized two consecutive SDR cycles	2	2	0
Overall		**45,697**	**40,000**	**5,697**

Further, Table 3.17 shows the 2013 SDR eligibility status of all 186,111 cases ever included in the SDR sample since its inception in 1973. Cases are classified as one of three types: (1) eligible, (2) permanently ineligible, or (3) eligible, but deselected (sampled out) in a previous SDR cycle. Note that permanently ineligible cases met the SDR eligibility criteria at one point in time, but due to changed circumstances became ineligible and are expected to never become eligible again (e.g., the case is over age 75 or known to be deceased).

It is important to note, that in addition to the 44,602 cases that were eligible for either the 2013 NSDR old cohort frame or 2013 ISDR old cohort sample, there were 2,824 cases classified as permanently ineligible which would have been age 75 or younger on the survey reference date. These 2,824 were not included in sampling, but retained for later use in response rate calculations and weighting adjustments of the age eligible SDR population.

Table 3.17   All Cases Ever Included in SDR by 2013 SDR Frame Status

2013 SDR Old Cohort Frame Status		Total Cases	Percent
Eligible for Old Cohort Frame Inclusion		**44,602**	**23.97%**
00	Eligible for 2013 NSDR Panel Sample Frame	38,424	20.65%
00	Eligible for 2013 ISDR Panel Sample Frame	6,178	3.32%
Age Ineligible for Old Cohort Frame Inclusion		**34,734**	**18.66%**
01	Age ineligible	29,037	15.60%
07	Age ineligible, according to BA year or PhD year	3,112	1.67%
Age Eligible, Otherwise Ineligible for Old Cohort Frame Inclusion		**2,824**	**1.52%**
11	Non-SEH doctoral degree field per SDR	78	0.04%
12	No doctorate degree earned per SDR	80	0.04%
13	Duplicate case per SDR	30	0.02%
13b	Double Doc; first SEH doctorate earned before SED 2010/2011	3	0.00%
14	Frame ineligible, not otherwise defined	26	0.01%
02	Permanently out of scope per SDR, not otherwise defined	110	0.06%
15	Deceased per SDR	1108	0.60%
16	Terminally ill per SDR	95	0.05%
19	Institutionalized two consecutive SDR cycles	6	0.00%
04	Non-US citizen, out of country 1993-1997 (dropped in 1999)	396	0.21%
05	Non-US citizen, out of country 1995-1997 (dropped in 1999)	71	0.04%
06	Non-US citizen, out of country (dropped in 1997)	391	0.21%
17	Non-US citizens, out of country (dropped in 2003)	128	0.07%
18	Non-US citizens, out of country (dropped in 2001)	297	0.16%
20	Other permanent ineligible in 1995, not otherwise defined	5	0.00%
Deselected Through Sampling		**103,951**	**55.85%**
21	Deselected in sampling 1973-1995 SDR	51,707	27.78%
22	Deselected in 1997 sampling	2,976	1.60%
23	Deselected in 1999 sampling	15,256	8.20%
24	Deselected in 2001 sampling	2,930	1.57%
26	Deselected in 2003 sampling	2,854	1.53%
28	Deselected in 2006 sampling	776	0.42%
29	Deselected in 2008 sampling	4,968	2.67%
30	Deselected in 2010 sampling	724	0.39%
25	Humanities sample dropped from SDR sample	21,760	11.69%
Overall		186,111	100.00%

3.2.4 Evaluation of Old Cohort Frame Strata Assignments

In a longitudinal survey sampling frame, it is desirable to have the variables used to stratify the sample remain consistent over time resulting in consistent strata assignments. Changes to stratification assignment should be justified. This is also true to the SDR.

All 2013 SDR old cohort frame cases which changed strata assignment from their 2010 strata assignment were evaluated to ensure that the change was accurate and correct. There were a total of 2,623 out of 44,602 old cohort frame eligible cases (5.9 percent) that changed strata assignment from 2010 to 2013. Some changes are expected as the SDR sample design updates stratification variables with the most current reported data and actively seeks to replace imputed data with reported data.

As is usually the case for the SDR, the primary reason for strata assignment changes in the 2013 frame are the differences in disability status coded from 2010 survey responses. Typically, an equivalent number of old cohort cases switch disability status to and from being disabled. However, in the 2013 SDR old cohort frame, a greater proportion of cases became disabled as a result of the change to the disability question in the 2010 survey (for more details see *Section 2*). The secondary reason for stratification assignment change resulted from a change in the sample member's location. Table 3.18 details the reasons why 2,623 2013 SDR old cohort frame cases changed from their 2010 SDR strata assignment.

Table 3.18 Reason for Strata Assignment Change from 2010 to 2013 SDR

Code	Reason for Strata Change	Overall	2010 and 2013 Frame Components		
			Both NSDR	Both ISDR	NSDR to ISDR
01	Only location changes to out of U.S., no other demographic changes	511	0	0	511
02	Became disabled	928	925	0	3
03	Became not disabled	580	577	0	3
04	Revised sex	9	7	0	0
05	Birth citizenship changed	115	96	14	5
06	Field of study changed	164	147	11	6
07a	Race/ethnicity changed from 2010 survey	66	34	18	14
07b	Race/ethnicity changed with 2001 reported data	205	191	12	2
09a	Birth citizenship and race/ethnicity changed	24	23	1	0
09b	Birth citizenship and field of study changed	5	5	0	0
10a	Disability status and race/ethnicity changed	10	10	0	0
10b	Disability status, race/ethnicity, and field of study changed	2	2	0	0
11	Race/ethnicity and field of study changed	4	4	0	0
	Total	2,623	2,021	56	544

3.3   New Cohort Sample Frame Construction

As noted previously in *Subsection 3.1*, the data source for constructing the 2010 SDR new cohort frame was the two most recent doctoral cohorts included in the DRF from the 2010 and 2011 SED rounds.

As with the old cohort frame, cases considered eligible for the 2013 SDR new cohort frame needed to first meet the 2013 SDR target population requirements of having received a doctoral degree in a SEH field from a U.S. institution, being 75 years or younger on the survey reference date of February 1, 2013, and living in a noninstitutionalized setting on the reference date. The variable LOCSTAT13 was used to assign the target population eligible cases into either the NSDR or the ISDR new cohort frames. Table 3.19 shows the 2013 SDR new cohort frame status for all 2010 and 2011 SED cases.

Table 3.19   2013 SDR New Cohort Frame Status by SED Cohort

	2013 SDR New Cohort Frame Status	SED Cohort		
		Total	2010	2011
Eligible		**71,906**	**35,242**	**36,664**
00	NSDR Frame Eligible	63,955	31,300	32,655
00	ISDR Frame Eligible	7,951	3,942	4,009
Ineligible		**25,138**	**12,792**	**12,346**
01	Age ineligible	10	5	5
03	Deceased, according to the DRF	12	6	6
11	Non-SEH doctoral degree field	25,056	12,753	12,303
13b	Double Doc; first SEH doctorate earned before SED 2010/2011	60	28	32
Overall		**97,044**	**48,034**	**49,010**

4. Sample Stratification

Sample stratification for the 2013 SDR sample design is identical to the approach used for the 2010 SDR. The NSDR portion of the frame was stratified into 150 strata and the ISDR portion was stratified into 44 strata. The NSDR and ISDR sampling frames are stratified and the sample allocated separately. Cases are assigned to the NSDR or the ISDR sampling frames based on the target population definitions that utilizes predicted residency location of in or out of the U.S. (as defined by the frame variable LOCSTAT13). For the detailed definition of LOCSTAT13, see page 15 of *Subsection 3.1.2* of this report.

4.1 NSDR Sample Stratification

The 2013 NSDR frame contained 38,424 panel and 63,955 new cohort members. The NSDR stratification scheme is presented in Appendix Table B.1 along with the distribution of the sampling frame by stratum. The NSDR stratification approach introduced in the 2003 cycle has been continually implemented through the 2013 cycle with one minor exception. The 2003 and 2006 NSDR cycles included missing race strata; these strata were eliminated for the 2008, 2010, and 2013 NSDR designs when logical imputation rules were used to impute missing race/ethnicity data during sampling frame development when this information was not previously reported in the SDR or SED (see page 13 in *Subsection 3.1.2* of this report for the detailed race/ethnicity imputation rules). Strata were defined based upon the cross of demographic group by gender by degree field.

Degree field was collapsed in varying ways depending upon the population size of doctorates in the demographic group, resulting in a total of 150 explicit strata. Within each stratum, the data records were sorted by citizenship, disability status, degree field, and year of degree receipt prior to sample selection. This created an implicit stratification of the sample within each explicit stratum to ensure the sample selected is balanced on these factors.

4.1.1 Demographic Group Recode

Demographic group is a composite variable based upon U.S. citizenship at birth, race/ethnicity, and disability status with collapsing as needed for small populations. After collapsing, the demographic group stratification variable was defined as follows:

1. Hispanics, regardless of race, citizenship at birth and disability status;
2. NH blacks, regardless of citizenship at birth and disability status;
3. U.S. citizen at birth, NH Asians (excluding Hawaiians and Pacific Islanders) regardless of disability status;
4. NH American Indians (including Alaskan natives), regardless of citizenship at birth and disability status;
5. NH Pacific Islanders (including native Hawaiians), regardless of citizenship at birth and disability status;
6. U.S. citizen at birth, disabled, NH whites;
7. U.S. citizen at birth, non-disabled NH whites;
8. Non-U.S. citizen at birth, NH whites regardless of disability status; and
9. Non-U.S. citizen at birth, NH Asians regardless of disability status.

These nine groups were defined in a hierarchical manner as the group definitions imply. For example, all Hispanics belong to the first demographic group regardless of other demographic characteristics. Similarly, all NH blacks belong to the second demographic group regardless of other characteristics.

4.1.2 Degree Field Recodes

As for the 2003 to 2010 NSDR, the 2013 NSDR used two degree field recodes for stratifying different demographic groups. The first recode is the 15-category SDR degree field variable (SDRFLD15) which was used to stratify the three largest demographic groups: (7) U.S. citizens at birth, nondisabled NH whites; (8) non-U.S. citizens at birth, NH whites; and (9) non-U.S. citizens at birth NH Asians. The second recode is the 7-category SESTAT major degree field variable (MAJFLD7) that was used to stratify the remaining demographic groups except for American Indians and Pacific Islanders which were not stratified by degree field. The mapping of both degree field recode variables to the detailed SED degree field code frame can be found in Appendix A.2 *2013 SDR Field of Study Coding Taxonomies Crosswalk*.

The SDR degree field recode has the following 15 categories:

1. Chemistry;
2. Physics/astronomy;
3. Earth/ocean/atmospheric sciences;
4. Mathematics;
5. Computer and information sciences;
6. Agricultural sciences;
7. Medical sciences;
8. National Institutes of Health (NIH) biological sciences;
9. Other biological sciences;
10. Psychology;
11. Economics;
12. Anthropology/archeology/sociology;
13. Other social sciences;
14. Electrical/electronics/communications engineering; and
15. Other engineering.

The SESTAT major degree field recode has these seven categories:

1. Information sciences/mathematics and statistics;
2. Biological and agricultural sciences;
3. Health sciences;
4. Physical and related sciences;
5. Social sciences;
6. Psychology; and
7. Engineering.

4.2   ISDR Sample Stratification

The 2013 ISDR frame contained 6,178 panel and 7,951 new cohort members. As noted above, the 2013 ISDR stratification approach was identical to the one used for the 2010 ISDR developed for the integrated sample design. The 2013 ISDR stratification scheme is presented in Appendix Table C.1 along with the distribution of the sampling frame by stratum.

The 2013 ISDR strata were defined by the cross of U.S. versus non-U.S. citizen at birth by race/ethnicity by gender by degree field. Race/ethnicity was defined as Hispanic, NH black, NH Asian, NH white, and NH other races, where other races combines American Indians and Pacific Islanders. U.S. citizens at birth were collapsed over race/ethnicity in stratification. Gender was defined as male and female. The 2013 ISDR collapsed over gender for non-U.S. citizens at birth that were of NH-black and NH-other races. The 7-category SESTAT major degree field recode (MAJFLD7) was used in stratification for non-U.S. citizens at birth that were NH Asians and NH whites. A three-digit degree recode, FOD3, was used to stratify U.S. citizens at birth and Hispanics and NH blacks that were non-U.S. citizens at birth. The 3-category degree field recode (FOD3) has these categories:

1. Computer and information sciences, mathematics, physical sciences, and engineering;
2. Biological and agricultural sciences, and health sciences; and
3. Psychology and social sciences.

Non-U.S. citizens at birth whose race was NH other races were collapsed over field of degree as well as gender.

The sort order for frame records within each ISDR stratum was defined based upon SESTAT major degree field and then by place of birth.

The place of birth sorting approach was introduced as a part of the 2006 ISDR redesign which redefined the ISDR strata used in sample selection after comparing race/ethnicity to country of origin (Cox, Grigorian, and Yang, 2006). To control for country of origin in sampling, the ISDR used a 12-level subcontinent of birth as a sort variable for stratum records in the 2006, 2008, and 2010 cycles. The subcontinent code frame was as follows and was used in this sort order: Oceania, Europe, Canada, Mexico and Central America, South America, Central Africa, South Africa, North Africa, Middle East, Southeast Asia, and Northern Asia. For the 2013 cycle, the subcontinent variable was replaced with the more detailed birth region variable shown in Appendix A.1. The revised region of birth variable offered more control in sorting. Note that the sort order for both the subcontinent and the birth region variables were chosen so that adjacent locations would tend to have similar ethnicity/race characteristics.

5. Sample Size

The sample size requirements for the 2013 SDR were consistent with those used in the 2010 SDR resulting in 47,078 sampled cases. The NSDR retained its usual sample size of 40,000 doctorates. As was introduced in the 2010 SDR design, the NSDR frame size was reduced by transferring potential non-U.S. residents to the ISDR frame using the *Last Location Rule*.[7] The ISDR panel frame size increased as a consequence of the addition of these transferred NSDR panel members and the ISDR continued its practice of sampling panel members with certainty. Thus, the transfer of panel cases from the NSDR to the ISDR frame resulted in a sample size increase for panel members in the ISDR sample component. However, the ISDR new cohort sample size was set to 900 as it was for the 2006, 2008, and 2010 survey cycles.

5.1 NSDR Sample Size

The NSDR sample size is defined in terms of sampled doctorates after permanent ineligibles have been removed from the frame prior to sample allocation.[8] Of the cases that become permanent ineligible, most are panel members that will be age 76 or older on the reference date for the next survey cycle. In addition, a small number of the remaining age eligible panel members are permanently ineligible for one or more of the following reasons:

- Deceased,
- Terminally ill/permanently incapacitated,
- No earned doctorate,
- Earned doctorate after the new cohort academic years (June 30, 2011 for 2013 SDR), or
- Earned doctorate in a non-SEH field.

Of the age-eligible permanent ineligible cases, most are deceased.

---

[7] The *Last Location Rule* categorizes cases, both U.S. citizens and non-U.S. citizens, as likely to be permanent non-U.S. residents when their residence was outside the U.S. in the previous survey cycle; panel frame cases previously included in the NSDR sample component transfer with certainty to the ISDR frame if they were last found outside of the U.S.

[8] A panel member is defined to be permanent ineligible when they are not now and will never again be a member of the SDR target population for a future survey cycle. Prior to 1999, permanent ineligibles (other than age ineligibles) were left in the frame and the desired sample size was expanded to account for their presence.

Beginning with the 1999 NSDR, the desired sample size for most survey cycles has been 40,000. An exception was made for the 2006 NSDR, which had a sample size of 42,955 expanded to accommodate three new SED cohorts. For the 2008 NSDR, NSF decided to return to the standard sample size of about 40,000 cases, which was also used for the 2010 and 2013 NSDR. Because new cohort population counts were available for both cohorts, the final sample size was exactly 40,000 (see *Section 7 Sample Selection* for details).

For the 2013 NSDR, we followed the 2010 procedures for defining the minimum desired sample size of completed interviews per stratum. For each stratum, the NSF has specified use of a minimum sample size that is the equivalent of 60 completed interviews, except for the two American Indian strata where a minimum sample size equivalent of 150 interviews has been set. This requirement recognizes that some strata are so small that after accounting for the finite population size effect on precision, much less than the specified amount of completed interviews may be needed to achieve the desired precision level. In comparison to an infinite population, a finite population of size N has its variance for a sample of size n reduced by a finite population correction factor (fpc) of $1 - n/N$. To meet this precision requirement, the minimum stratum sample size for most strata was set to the equivalent of the number of completed interviews when adjusted for the stratum's fpc. Under this approach the minimum sample size allocated to stratum h is set to

$$n_h = \frac{n_h' * N_h}{n_h' + N_h}$$

where N_h is the population size for stratum h, and n_h' is either 60 or 150 depending on the stratum receiving the minimum assignment. Note that n_h will be smaller than n_h', which reflects the reduction in variance due to the fpc. In other words, a sample of n_h cases in a stratum of limited size is equivalent to a sample of n_h' cases in an extremely large stratum. The effect of ignoring the fpc would be to overestimate the minimum required sample size for a stratum.

Appendix B.2 shows the estimated population size for each stratum, the desired respondent sample size with and without fpc adjustment, and when the stratum sample size was set to the

minimum respondent sample size with fpc adjustment. Note that the minimum respondent sample sizes were defined in terms of completed interviews with eligible doctorates. For the 2013 NSDR, yield rates (number of completed interviews with eligible doctorates divided by sample size) were estimated for the cross of demographic group by gender based upon the 2010 NSDR data collection experience for the equivalent frame population.[9] For strata with their sample sizes set to the minimum respondent sample size, the yield rates shown in Appendix B.2 were used to determine the total sample cases to be selected from the stratum.

This minimum stratum sample size requirement was introduced in the 2003 NSDR redesign (Cox 2003, Yang et al. 2004). That redesign also redefined the strata so that they conformed more closely to analysis domains used in reporting, particularly with respect to the collapsing of very small race/ethnicity groups over degree fields to achieve strata with populations of sufficient size for reporting.[10] Small race/ethnicity by sex domains such as Hispanics and NH-blacks have an additional domain sample size supplement that increases the sample size for the individual strata within the domain and for the overall domain (see Table 6.1 in the next section).

5.2   ISDR Sample Size

The actual 2013 ISDR sample component size is 7,078 cases. The ISDR new cohort sample size remained the same as in 2010 at 900 cases and the panel sample size increased from 4,797 in 2010 to 6,178 in 2013. The historical development of the 2013 ISDR panel sample can be described as follows:

- 600 cases selected for the 2003 ISDR from the 2001 and 2002 SED new cohorts that were non-U.S. citizens reporting plans to emigrate after graduation,

- 900 cases selected for the 2006 ISDR from the 2003, 2004 and 2005 SED new cohorts that were non-U.S. citizens reporting plans to emigrate after graduation,

- 156 non-U.S. citizen cases removed from the 2006 NSDR frame for being abroad for two consecutive rounds and transferred to the 2006 ISDR sample,

- 948 cases selected for the 2008 ISDR from the 2006 and 2007 SED new cohorts that were non-U.S. citizens reporting plans to emigrate after graduation,

---

[9] Yield rates and not response rates were used because we had to account for loss due to ineligibility and nonresponse.

[10] Generally, the strata represent populations of size 500 or more. A few strata were allowed to have smaller population sizes to prevent excessive collapsing over degree fields.

- 228 non-U.S. citizen cases removed from the 2008 NSDR frame for being abroad for two consecutive rounds and transferred to the 2008 ISDR sample,

- 900 cases selected for the 2010 ISDR from the 2008 and 2009 SED new cohorts that reported plans to emigrate (without regard to citizenship),

- 15 ISDR panel cases determined to be permanently ineligible in the 2003 to 2008 cycles removed from the 2010 eligible frame;

- 1,980 cases with most recent location outside the U.S. transferred from the 2010 NSDR frame to the 2010 ISDR frame,

- 63 ISDR panel cases determined to be permanently ineligible in the 2010 cycle removed from the 2013 eligible frame; and

- 544 cases with most recent location outside the U.S. transferred for the 2013 NSDR frame to the 2013 ISDR frame.

Once transferred into or sampled for the ISDR, panel cohorts have remained in the sample for future survey cycles. At present, the intention is to build up the longitudinal ISDR panel over several cycles and to establish a fixed sample size for this sample component when the characteristics of international residents are better understood.

6.   Sample Allocation

The 2013 SDR used essentially the same basic approach for sample allocation as the 2010 SDR. However, one change to a sampling stratification variable noted previously did have an impact on the 2013 SDR sample allocation. Specifically, the stratification variable measuring disability was modified to include cognitive disabilities in the 2013 sampling frame which expanded the number of frame members classified as disabled and increased the population sizes for the U.S. born, non-Hispanic white disabled strata.

6.1   Background on NSDR Sample Allocation Procedures

This section provides historical background on the development of the sample allocation procedures for the 2013 NSDR as they relate to the current sample design.

6.1.1 Introduction of the Maintenance Cut

Prior to 1995, the NSDR retained all eligible panel members in the sample with certainty and then selected a sample from the new cohort frame for each stratum to update the sample coverage for the current survey cycle. As a consequence the NSDR sample size increased steadily over time resulting in unacceptable increases in the total survey costs (Mitchell, Moonesinghe, and Cox, 1998).

In the 1995 survey cycle, the NSDR introduced the concept of a maintenance cut which required that the total sample size of selected new cohorts and panel members be fixed to a pre-specified number of attempted interviews in that survey cycle (Moonesinghe, 1998). Each subsequent survey cycle has implemented a maintenance cut, although the total specified sample size has varied over time. Since 1999, the total NSDR sample size has been fixed at 40,000 attempted interviews, with the exception of the 2006 cycle which had a sample size of 42,955 expanded to accommodate three new SED cohorts.

This maintenance cut only affects the total sample size being allocated and is not intended to be a uniform cut to the number of panel members selected from each stratum. Rather the total specified sample size is reallocated to each stratum's new cohorts and panel members following

the sample design in place for that survey cycle keeping, for the most part, a proportional allocation of the sample between new and panel cases based on their respective populations.

6.1.2 The 2013 NSDR and its Derivation from 2003 and 2010 NSDR Redesigns

The 2013 NSDR sample design is derived from the redesign implemented in the 2003 NSDR, together with the 2006, 2008, and 2010 modifications to the NSDR and sample selection procedures. The 2003 NSDR redesign redefined the strata to ensure adequate minimum population sizes for each stratum and to better respond to analysts data needs (Cox, 2003). About 75 percent of the sample is allocated with probability proportional to population size to maximize the precision in the survey estimates. The remainder of the sample is allocated disproportionally to ensure adequate estimation capability for small minority domains and to ensure that each stratum is allocated sufficient numbers of attempted interviews so that they can be expected to yield the equivalent of 60 completed interviews.

The 2006 NSDR modified the 2003 NSDR design to impute missing data for stratification variables like race/ethnicity but otherwise the design remained the same (Yang et al., 2006). The 2008 NSDR also used logical editing to impute missing data for all stratification variables including race and ethnicity.

In 2010, the NSDR and ISDR frames were integrated into one, although the samples for the two subpopulations are stratified and allocated separately (Cox et al., 2012b). The 2010 NSDR followed the same sample design and allocation procedures as the 2008 NSDR except that the 2008 old cohort NSDR sample members were moved to the 2010 ISDR frame when they were found to be living outside the U.S. The 2013 NSDR followed the 2010 NSDR sample design procedures exactly except for the redefinition of the disabled frame variable to include the cognitively impaired.

6.2   Allocation of the 2013 NSDR Sample to Panel Members and New Cohorts

The NSDR panel sample allocation procedure is an iterative process that first proportionally allocates the sample to each stratum, and then increases the initial sample sizes in certain strata to achieve the minimum samples sizes desired for the number of completed interviews and for the specified analytical domains as needed, which in turn requires the allocation for the

remaining strata to be decreased to maintain the overall sample size. Some recycling of these steps is required to make sure all of the sample targets are met. In addition, since the panel cases are selected using a probability-proportionate-to-size (PPS) selection procedure (see *Section 7*), once the sample is specified for each stratum, an iterative process is used to identify the certainty selections in each strata and then to select from the remaining cases the balance of the sample required. For the new cohort, the sample is allocated proportionally across the strata. Since there are no minimum sample sizes or domain target restrictions to apply, no further adjustment is required. The new cohort sample is also selected using systematic sequentially sorted sampling procedures rather than a PPS procedure so certainty identification is not required.

Appendix Table B.3 shows the total of 40,000 cases as they were finally allocated, including the 36,666 panel cohort sampled cases and the 1,632 and 1,702 new cohort sampled cases for the 2010 and 2011 academic years, respectively.

6.2.1 The NSDR Allocation Process

The NSDR sample consists of two cohorts: the panel cohort and the new cohort. The new cohort is further divided into two separate cohort groups, one for each new SED cohorts defined by the two academic years. Across the two cohorts, the total sample was allocated to the panel cohort and new cohort proportionately based on population size. The sample allocated to the new cohort was further subdivided by allocating it proportionately to the two new cohorts. Within each new cohort, the sample is allocated to the strata proportionately based on the population size per stratum. Within the old cohort, however, an iterative process was required to allocate the sample across the strata to ensure that the minimum sample size requirements are met for all selected domains and strata.

Specifically, the 2013 NSDR panel sample allocation consisted of five iterative steps:

1. Allocate the sample proportionally to each stratum;
2. Allocate extra sample to specific demographic groups by gender domains through supplemental domain allocation;
3. Allocate supplemental sample to the small strata if needed to achieve the minimum sample size requirement;
4. Adjust the allocation for the remaining strata that are not involved in steps 2 and 3 to maintain the overall sample size; and

5. Repeat steps 2 through 4 as needed to ensure the minimum sample size requirements are achieved for all domains and all strata.

While large strata received only the proportional allocation, the smallest strata could receive additional sample through the stratum supplemental allocation and the domain supplemental allocation. Both the stratum and domain supplemental allocations are designed to support subgroup analyses with sufficient sample size. The size of the domain supplemental allocations was the same in 2013 as had been since 2003. The final panel sample allocation was therefore a combination of a proportional allocation across all strata, a domain-specific supplement allocated proportionately across strata in that domain, and a stratum-specific supplement added to each stratum, if needed, to obtain the minimum stratum size.

Since the panel sample allocation is based on weighted population counts instead of the number of cases on the frame, some strata did not have enough cases to support the desired allocation. In that situation, the allocated sample size is the same as the number of cases available while the balance of the sample is allocated to the other panel cohort strata via the iterative steps described above. That is, as such changes took place, the iterative process was repeated as needed until all requirements are met.

For the new cohort sample allocation is a straight proportional allocation based on the number of cases per stratum.

The allocation process worked as follows: First, the domain supplemental samples totaling 4,550 sample cases overall were proportionally allocated to the strata associated with each designated small domain defined by gender and demographic group receiving a supplemental sample. The domain specific allocation was based upon the stratum's estimated total population size across all cohorts. This domain specific allocation was fixed and never changed under the subsequent sample size iterations. Second, the remaining sample (35,450) was allocated in an iterative process.

The iterative portion of the sample allocation process began with a proportional allocation of the remaining 35,450 sample cases based on the estimated population size of each stratum. The next step in the first iteration was to make additional stratum-level allocations as needed to ensure that each stratum had its minimum sample size allocation. For each stratum, the resultant total

sample size of proportional, domain-specific, and stratum-specific allocations was further allocated to the panel and new cohort substrata. When the stratum's panel cohort sample allocation exceeded the number of panel cohort frame members, the panel cohort allocation was reduced to the number of panel cohort frame members in that stratum.

To decide if the second iteration was needed, the total sample size allocated across all strata was compared to the desired sample of 40,000 cases. Because that total exceeded 40,000 cases (due to the stratum-level allocations made in the first iteration), a second allocation was needed. The second iteration began by redefining the number of sample cases to be proportionately allocated as 35,450 minus the total number of cases allocated across all the stratum-specific allocations of the first iteration. This reduced sample size for the proportional allocation was again proportionately allocated across all strata in this second iteration. As before, the next step was to make additional stratum-level allocations as needed to ensure that each stratum had their minimum size allocation. This step might lead to additional strata needing a stratum-level allocation as well as increasing the stratum-level allocations made in the first allocation. Again, the revised total stratum size allocation was further allocated to the old versus new substrata and the panel cohort substratum allocation was reduced when it exceeded the number of old cohort frame cases.

The iteration process continued following the pattern of the second iteration until the total sample allocated across all strata was 40,000 and all the minimum stratum-level sample size requirements were met. Ultimately, a total of 1,555 sample cases were allocated at the stratum-level to ensure that minimum stratum sample size requirements were met, leaving 34,154 cases to be proportionately allocated to strata after 4,550 cases had been allocated at the domain level. For further clarification of the iteration process, see Appendix D for detailed specifications and the final 2013 NDR allocation.

6.2.2 The 2013 NSDR Allocation Results

As noted earlier, the domain-specific allocation was fixed. The purpose of the domain allocation was to maintain the sampling rates for the small domains that were achieved in previous NSDR survey cycles. Analysts routinely combine design strata to form domains for separate estimation, which should be duly reflected in the sample design and allocation. Without the domain

allocation, we would have allocated far more sample to the U.S.-born, non-disabled, white strata than past surveys, and there would be insufficient old cohort cases in the frame to support such allocation. As reported in *Sample Design and Implementation for the 2003 Survey of Doctorate Recipients*, additional sample had been allocated to minority by gender subpopulations prior to the 2003 NSDR (Yang et al., 2004). Such purposeful oversampling was carried out to support NSDR analyses on these small domains. Similar domain allocation has been implemented in the 2006 to 2010 NSDR survey cycles.

Following this practice, the 2013 NSDR allocated 4,550 cases to ten demographic by gender domains, with the extra sample allocated proportionally to the strata composing each domain. This extra sample size was arbitrarily set to the sample sizes allocated in the 2003 NSDR, which in turn was set to yield approximately the same average sampling ratio of population size to sample size in each domain as was achieved in the 2001 NSDR, while avoiding allocation of old cohort sample sizes in excess of the available frame cases. Table 6.1 gives the size of the supplemental allocation to each of the domains that received such allocation.

Table 6.1   Domain Supplemental Allocation

Demographic Group	Sex	Supplemental Allocation
Hispanic	Male	750
	Female	750
NH black	Male	750
	Female	750
U.S. born, NH Asian	Male	500
	Female	500
U.S. born, non-disabled NH white	Female	250
Non-U.S. born, NH white	Male	50
	Female	50
Non-U.S. born, NH Asian	Female	200
Total Supplemental Allocation		**4,550**

Overall, a total of 34,154 cases were allocated through proportional allocation and the remaining 5,846 cases were allocated through stratum or domain level supplemental allocations. The final sample size allocated through the two supplemental allocations was smaller than the total supplemental allocation in the first iteration because a fraction of the supplemental allocation

was added back to the proportional allocation when there was a shortage of old cohort cases in the frame. For the same reason, it is not possible to divide the total supplemental allocation between stratum and domain level supplemental allocations.

The sample allocation took place in November 2012, when population counts were available for the 2010 and 2011 SED cohorts as well as the old cohorts. As a consequence, the 2013 NSDR sample size allocated was exactly 40,000 and the sample was allocated in one step. Prior to sample selection, allocations of less than 1 sample case to any 2010 or 2011 SED new cohort stratum with one or more frame members were rounded up to 1, still resulting in a final 2013 new cohort sample of 3,339 instead of the 3,334 originally allocated.

The overall impact of the revised 2010 NSDR frame building procedures used in 2013 frame building too was to reduce the frame size as panel cohort cases were transferred to the ISDR and new cohort cases were incorporated into the ISDR frame that would have been in the NSDR frame with the rules used in previous survey cycles. The impact was modest, given that the major transfer of emigrants had occurred in the 2010 frame building, which had the effect of reducing the need for stratum-specific allocated sample to 1,555 compared to the 1,591 used in the 2010 NSDR. The proportion of the sample being allocated proportionately decreased to 85 percent for the 2013 NSDR compared to 86 percent for the 2010 NSDR. Finally, the panel cohort stratum allocations were 101 percent of the panel cohort frame sizes, which was 102 percent for the 2010 NSDR. The 2013 NSDR frame building procedures remained the same, so we would have expected a modest decrease in panel cohort allocations in excess of available panel cohort sample cases between 2013 and 2010DR samples.

6.2.3 Trends over Time in the NSDR Sample Allocation

Each survey cycle the NSDR sample of 40,000 sampled cases has about 85 percent of the total sample allocated in proportion to current population sizes for each stratum. As a consequence, the sample allocation changes over survey cycles to reflect trends in the distribution of SEH doctorates by race/ethnicity, sex, and other stratification variables. This section discusses changes observed in the 2013 NSDR sample allocation as a consequence of the changing composition of the SEH population over time and changing definition for the disabled strata.

U.S. Citizen at Birth Males. At the inception of the NSDR, the vast majority of the nation's trained SEH doctorates were U.S. citizen at birth, white, and male. Since that time, there has been an ever increasing percentage of new cohorts which are non-U.S. citizen at birth, minority racial groups, and female. As a consequence, doctorates aging out of the NSDR population reduce the overall proportion of the total population of U.S. citizen at birth, white, males, while there is a somewhat reduced percentage of U.S. citizen at birth, white, male doctorates entering the NSDR population. The reduction in the relative population size of U.S. citizen at birth, white males led to a modest reduction in the number of old cohorts retained in the 2013 NSDR sample—96.2 percent of eligible old cohorts—in comparison to the 95.4 percent of all eligible old cohorts retained in the 2013 sample.

U.S. Citizen at Birth Asian Females. The overall population sizes for these strata in 2013 ranged from 12 to 30 percent when expressed as a percentage of the 2010 population sizes. These strata are growing at a higher rate than the strata for other domains which means that the new cohort cases needs to be assigned proportionately more of the stratum's sample and the subsampling rate for old cohorts increased slightly. The overall effect is stratum maintenance cuts that range from 8 to 13 percent which is about twice as large as the overall average maintenance cut of 5.4 percent across strata.

U. S. Citizen at Birth Disabled Whites. The disabled population presents a difficult problem for stratification as disabled status may change from one survey cycle to another. Disability is defined as reporting disability in the prior SDR cycle for the panel cases or in the SED for new cohorts. Various alternative definitions for disability have been studied, but this definition produces the best results. However, a not-insubstantial number of sample cases stratified as nondisabled later report being disabled in the survey *and vice versa*. The movement from nondisabled to disabled has the most negative consequences as these cases have large weights in comparison to sample cases selected from the disabled strata. This type of movement was observed in the 2013 NSDR frame in part due to the additional cognitive disability category added to the 2010 survey. Prior to the 2010 cycle, respondents could choose from four disability categories (i.e., difficulty with seeing, hearing, walking, or lifting). Starting with the 2010 SDR, a fifth disability category for reporting difficulty with concentrating, remembering or making decisions was added.   As a result of the added disability category in the 2010 SDR, the number

of old cohort frame cases classified as disabled in the 2013 SDR frame file was noticeably greater (also discussed in *Section 2*).

Specifically, 4.5 percent of cases initially stratified as non-disabled in the 2010 frame reported being disabled in the 2010 survey, while 38.8 percent of cases stratified as disabled in the 2010 frame reported being nondisabled in the 2010 survey. To assess the impact of the added category, the disability status was calculated for the 2013 SDR old cohort as it was defined for the 2010 old cohort frame cases using responses from just the four disability categories and compared to the disability status calculated using all five disability categories. This comparison showed that the fifth new cognitive disability category caused an increase in the number of disabled old cohort frame cases of 7.6 percent. However, disability status is only used to stratify U.S. born, white cases in the NSDR frame. Table 6.2 shows the impact of the cognitive disability category on the NSDR old cohort frame cases in the U.S. born white strata (strata 47 to 90 which include the disabled and non-disabled strata).

Table 6.2   2013 NSDR U.S. Born White Old Cohort Frame Cases by Disability Status Derived by the 4-Category and 5-Category Disability Definition

| Disabled Status | Old Cohort Disability Definition | | | |
| | Based on 4 categories | | Based on 5 categories | |
	Population Estimate	Percent	Population Estimate	Percent
Total	515,700	100.0%	515,700	100.0%
Not disabled	472,900	91.7%	470,000	91.1%
Disabled	42,800	8.3%	45,700	8.9%

| Disabled Status | Old Cohort Disability Definition | | | |
| | Based on 4 categories | | Based on 5 categories | |
	Case Count	Percent	Case Count	Percent
Total	22,032	100.0%	22,032	100.0%
Not disabled	20,153	91.5%	20,027	90.9%
Disabled	1,879	8.5%	2,005	9.1%

Finally, comparing the 2010 NSDR allocation results to the 2013 NSDR results, we see a 1.1 percent increase in the proportion of U.S. citizen at birth, white disabled; 2.1 percent of the 2010

NSDR allocated sampling frame was U.S. citizen at birth, white disabled, and 3.2 percent of the 2013 NSDR allocated sample frame was U.S. citizen at birth, white disabled.

Demographic Domains by Sex. Table 6.3 compares the percent of the population for each demographic by sex domain by panel and new cohort and overall for the 2013 and 2010 SDR population. The table also shows the relative increase or decrease in the population sizes. As noted, the biggest proportional change observed is a decrease of 2.3 percent in the population of U.S. citizens at birth, NH white, nondisabled males. Proportional growth can be seen in many of the non-white domains, particularly the non-U.S. citizen at birth Asian men and women when comparing the 2013 to 2010 SDR population distribution.

Table 6.3   Population Proportions by Demographic Domain: 2010 and 2013 NSDR

Demographic Group Defined by NSFGRP by Sex	2013 SDR Percent of Population			2010 SDR Percent of Population			Differences 2013 to 2010 SDR		
	Total	Old	New	Total	Old	New	Total	Old	New
Hispanic males, regardless of race, citizenship at birth, and disability status	2.2%	2.1%	3.4%	2.1%	2.0%	3.1%	0.1%	0.1%	0.3%
Hispanic females, regardless of race, citizenship at birth, and disability status	1.5%	1.4%	2.8%	1.3%	1.2%	2.5%	0.2%	0.2%	0.3%
NH black males, regardless citizenship at birth and disability status	1.7%	1.7%	1.9%	1.7%	1.7%	1.9%	0.0%	0.0%	0.0%
NH black females, regardless citizenship at birth and disability status	1.5%	1.4%	2.5%	1.4%	1.3%	2.3%	0.1%	0.1%	0.2%
U.S. citizen at birth, NH Asian males regardless of disability status	1.0%	1.0%	1.5%	0.9%	0.9%	1.3%	0.1%	0.1%	0.2%
U.S. citizen at birth, NH Asian females regardless of disability status	0.7%	0.7%	1.5%	0.6%	0.6%	1.3%	0.1%	0.1%	0.2%
NH American Indian males, regardless of citizenship at birth and disability status	0.4%	0.4%	0.3%	0.4%	0.4%	0.3%	0.0%	0.0%	0.0%
NH American Indian females, regardless of citizenship at birth and disability status	0.2%	0.2%	0.3%	0.2%	0.2%	0.3%	0.0%	0.0%	0.0%
NH Pacific Islander males, regardless of citizenship at birth and disability status	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%
NH Pacific Islander females, regardless of citizenship at birth and disability status	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%
U.S. citizen at birth disabled, NH white males	3.8%	4.0%	1.1%	3.4%	3.7%	0.2%	0.4%	0.3%	0.9%
U.S. citizen at birth disabled, NH white females	1.4%	1.4%	0.8%	1.3%	1.3%	0.3%	0.1%	0.1%	0.5%
U.S. citizen at birth, not disabled, NH white males	36.9%	38.1%	21.8%	39.2%	40.7%	22.1%	-2.3%	-2.6%	-0.3%
U.S. citizen at birth, not disabled, NH white females	18.4%	18.4%	18.7%	18.2%	18.2%	18.9%	0.2%	0.2%	-0.2%
Non-U.S. citizen at birth, NH white males, regardless of disability status	7.1%	6.9%	8.7%	7.1%	6.9%	9.0%	0.0%	0.0%	-0.3%
Non-U.S. citizen at birth, NH white females, regardless of disability status	2.9%	2.7%	5.6%	2.8%	2.6%	5.8%	0.1%	0.1%	-0.2%
Non-U.S. citizen at birth, NH Asian males, regardless of disability status	14.6%	14.3%	18.5%	14.1%	13.6%	20.0%	0.5%	0.7%	-1.5%
Non-U.S. citizen at birth, NH Asian females, regardless of disability status	5.6%	5.2%	10.5%	5.0%	4.5%	10.6%	0.6%	0.7%	-0.1%
Overall	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	0.0%	0.0%	0.0%

NH=non-Hispanic.

6.3   ISDR Sample Allocation

All 6,178 panel cohort ISDR cases were selected with certainty in 2013, following the practice of the previous survey cycles. As in the 2010 survey cycle, the SED 2010 and 2011 ISDR new cohort cases were also allocated in one pass. The 900 new cohort ISDR sample cases were allocated proportionally to strata based upon population sizes. As introduced in the 2010 ISDR,

the 2013 ISDR used 44 new cohort strata to replace the 10 strata used for new cohorts in the 2006 and 2008 survey cycles. Any unrounded stratum allocation less than 1 was forced to be 1 to make sure these strata got represented in the sample. The frame counts and actual allocation of the ISDR sample is shown in Appendix C.

6.4   NSDR and ISDR Probabilistic Rounding

The final sample allocations were rounded to integers before they were used in sample selection. Chromy's probability minimum replacement sampling algorithm was used to convert each stratum and substratum allocation to an integer while keeping the total sample size fixed to the desired sample totals for the NSDR panel and new cohorts and for the ISDR new cohorts (Chromy, 1979). Probabilistic rounding converts the sample size to integers without changing the ultimate unconditional selection probabilities. As a consequence, except for strata with insufficient panel cohort cases available for sampling, the ultimate unconditional probability of selection based on rounded sample allocations were the same for all panel and new cohort cases within each stratum.

<div style="border-left: 8px solid orange; padding-left: 1em;">

7. Sample Selection

</div>

The 2013 SDR sample selection procedures were unchanged from the prior rounds of the SDR.

7.1 NSDR Sample Selection

The 2013 NSDR sample selection was carried out separately for the panel cohort, the 2010 SED new cohort, and the 2011 SED new cohort. Prior to 2010 SDR, the NSDR sample was selected in two parts, with the Part 2 selection for the most recent cohort delayed until July when final counts were available. However, the reference date for the 2013 SDR was changed to 1 February 2013 allowing time for both new cohort years to be available for sample selection. Although the 2010 and 2011 SED new cohort frames could have been combined and just one new cohort sample selected, we chose to continue the practice of selecting independent samples for each new cohort year to maintain control over the stratum sample size selected from each new cohort year. Within each cohort, the sample was selected independently from each stratum based on the final allocation presented in Appendix B.3.

As for past survey cycles, the panel cohort sample was selected with probability proportional to size (PPS) where the measure of size was the 2010 SDR sampling weight (the inverse of the probability of selection). For each stratum, the sampling algorithm started by identifying and removing certainty cases through an iterative procedure. A panel cohort case was selected with certainty when its selection probability was equal to or greater than 1.0 based on its measure of size. These certainty cases were transferred to the sample and revised selection probabilities were calculated based upon the remaining frame cases. The measures of size of the remaining panel cohorts were then compared to the revised selection probability and additional certainty cases designated. Iteration terminated when all certainty selections had been identified and removed. Next, the noncertainty cases within each stratum were sorted by citizenship, disability status, 15-level SDR degree field[11], and year of doctoral degree award. Finally, the balance of

---

[11] Prior to 2003, the DRF field of degree variable (PHDFIELD) was used in sorting with no control imposed over year of degree receipt from 1991 to 2001. The intent had been to use SDR field of degree from 2003 on together with year of degree receipt, but the DRF field of degree continued to be used in 2003 and 2006 due to oversight. Use of the multi-level DRF field of degree in sorting left little potential for control over the year of

the panel cohort sample (i.e., the total stratum allocation minus the number of certainty cases) was selected from each stratum as a systematic PPS sample.

The 2010 SED new cohort sample was selected at the same time as the 2011 SED new cohort sample, using exactly the same systematic sampling procedures. Both new cohort samples were selected using the same sampling algorithm as used for selecting the old cohort sample. Every case in the new cohort frame was assigned 1 as the measure of size for the PPS selection. There were no certainty selections from the new cohorts, and the new cohort sample within each stratum was an equal probability systematic sample. Across strata, however, sampling probabilities vary.

Both the panel cohort and new cohort samples were selected systematically from the sorted list within each stratum, where the sorting variables operated as implicit stratification variables. The efficiency of a systematic sample can be increased if the units on the list are sorted by characteristics that are relevant to analysis. Sorting places similar cases next to each other on the list so that each stratum sample includes a mix of cases representative of their population with respect to the sorting variables. Because citizenship and disability status are of analytical interest but were not featured in the stratification of minority demographic groups, it made sense to use these as the first two sorting variables. Sorting by the 15-level SDR degree field variable provides discrimination over degree field for American Indians and Pacific Islanders that are not stratified by degree field and also greater control over the degree field distribution for minority groups that are only stratified by the 7-level SESTAT degree field recode. Because analysts frequently report for domains based upon age or years since degree award, the frame was also sorted by years since degree award to control the age distribution of the final sample.

7.2   ISDR Sample Selection

The ISDR panel cohort cases were selected with certainty. The 2010 and 2011 SED new cohort files were combined for selection purposes, using the final sample size stratum allocations presented in Appendix C.1. The new cohort sample was selected systematically from the sorted

degree receipt. The oversight was corrected beginning with the 2008 survey with the 15-level degree field variable used for sorting to reserve the potential for control over year of degree receipt within degree field as originally planned.

list within each stratum, where the sorting variables (SESTAT major degree field and continent/region) operated as implicit stratification variables.

8. Concluding Remarks

The 2013 SDR sample design closely followed the 2010 SDR design. The process of creating the 2010 SDR sample design which integrated the main or NSDR survey with the ISDR survey required many design changes from the 2008 SDR program, but was well worth the effort. Methodological research conducted using 2008 SDR data enabled the NSF to integrate the ISDR sample cases accrued over the 2003 to 2008 cycles with the NSDR cases to improve the coverage properties of the SDR. This in turn provided the ability for the SDR to produce estimates for all cases graduating in the 21st century whether they were residing in or out of the U.S. and likewise to report the estimates by this status. The integration research in 2008 also included the development of an integrated set of sampling strata that used the predicted location of the cases to create a more homogeneous segmentation. As a result, we expect improved survey precision of the estimates with this revised stratification approach. Furthermore, we aligned the strata with around the cases expected residency determined in the data collection operations of survey administration and locating. This research resulted in a new integrated survey weighting procedure for the combined NSDR and ISDR cases that adjusted for nonresponse using a logistic regression technique and incorporated a poststratification procedure to ensure the weighted estimates reproduced population totals from the combined NSDR and ISDR sampling frames. For a discussion of the integrated research and the creation of the predicted location see (Cox et al., 2012a).

No matter how carefully survey redesigns are researched and implemented, substantial design changes need to be evaluated after a cycle or two to allow for adjustments in the event deficiencies are recognized. For the 2003 NSDR redesign, the design strata were redefined to be more responsive to the domains commonly used by data analysts. This process identified the fact that the NSDR strata were often based upon imputed data for race/ethnicity. Steps were taken to obtain the missing data in the 2003 and 2006 survey cycles, but there was still more missing data for race/ethnicity than desirable for stratification. As a result, for the 2008 NSDR introduced a multistep imputation procedure to logically impute this missing data when it had yet to be collected from sample members. This imputation approach was found to be reasonably effective in predicting missing race/ethnicity (Selfa et al., 2012) and was adopted for use in the

2010 and 2013 cycles. With the 2013 cycle, the integration of the NSDR and ISDR sample has been completed as originally planned. However, discussed, we recommend additional steps be taken to revisit the study objectives to determine whether the current sample design best supports the SDR's estimation goals. We also recommend conducting research to explore additional steps to unify the components of the SDR samples, NSDR and ISDR, into a single sample design and allocation methodology. The 2013 SDR sampling procedures followed the methodology adopted in 2010 with the minor exceptions as noted in *Section 2* which leveraged research conducted on the 2008 cycle's selected sample and data collection results. In particular, the 2008 survey cycle was the first cycle to have sufficient ISDR interviews completed to facilitate the analysis of the two SDR components separately and together. In a related investigation, 2008 SDR integrated weights were developed to facilitate integrated analyses (Harter et al., 2012) based on a weighting class procedure and to bridge the changes to the traditional and integrated estimates. The weighting process was enhanced in 2010 using a logistic regression methodology which is expected to be applied to the 2013 sample (Sinclair and Batishev, 2012). This research as noted enables the ISDR and NSDR data sets to be used in combination to provide insight into key analytic issues for international residents and domains that are of special interest. We note that the 2010 NSDR and ISDR design strata were defined based upon input from the NSF analysts and the same stratification plan was adopted for 2013 as discussed in *Section 4*.

The 2013 design follows the 2010 design that adopted new procedures for the ISDR sample size and allocation. In 2010, to build up the ISDR sample size, eligible panel members from the previous survey cycle were taken with certainty into the ISDR sample. Most ISDR panel members were doctorates earning their degrees in the 21st century sampled as new cohorts. Other cases were transferred out of the NSDR frame for the 2010 survey cycle when they were identified as being international residents in the data collection for the previous survey cycle. Most of these transferred cases are doctorates earning their degree in the 20th century, although there are a small number of doctorates earning their degree in the 21st century transferred from the NSDR frame to the ISDR frame. The same approach was followed for the 2013 design. At a future date, the ISDR may need to establish a fixed total ISDR sample size and implement a maintenance cut in each survey cycle just as the NSDR has done since the 1995 survey cycle. The NSF has been considering this, but a specific ISDR sample size or specific survey round for

implementing these changes has not yet been established.[12] Also at that point, we recommend a review of the current sample allocation to ascertain whether the survey data results are fully meeting the NSF's analytic goals for the SDR.

The integrated SDR data set can be expected to provide valuable insights concerning migration of U.S. trained doctorates. International residency may be becoming more attractive for recent doctorates as well as for experienced doctorates. Some doctorates leave the U.S. permanently but others return. Still others move back and forth repeatedly across national boundaries. The integrated SDR data provides valuable guidance into the characteristics of doctorates who choose to be international residents on a temporary or permanent basis.

Here we recommended some next steps for future research and program improvements:

- Development of sample design and sample allocation statistical program (possibly coded in SAS or other portable software) that will enable the NSF and NORC to easily examine the impact of different design choices (using different stratification and/or sample allocation methodologies) on domain specific samples sizes and their corresponding precision levels. Results will generate suggested design changes to improve the precision levels for specific domains (to be specified based on a fresh review of the study objectives) and will evaluate the trade-offs associated with the effects of oversampling as warranted on aggregate estimates that cover multiple domains. In particular this approach would suggest an optimal sample size for the international students and how to best allocate the same between the panel and new cohort cases for determining the use of maintenance cuts.
- Evaluate the migration patterns and citizenship status of cases initially sampled for the ISDR to determine if ISDR sample members currently located in the U.S. appear to making a permanent residency change to the U.S. and should be transferred to the NSDR frame.

---

[12] See the memorandum entitled "2010 SDR Integration Memo 5 – Identification of International Residents Among 21st Century Doctorates" addressed to Dan Foley and Steve Cohen (NSF) from Brenda Cox (SRA) and Karen Grigorian dated 11 May 2010.

- To explore methods to evaluate additional unification of the ISDR and NSDR sample designs with the goal of creating a single SDR sampling methodology using stratification based on the predicted location to control the sample sizes for national and international cases.

- Generation of a reference document that describes the changes to the sample design, survey and sampling frame eligibility standards, weighting methodology and survey definitions during the last two decades of the SDR program that can be easily updated each survey cycle once started.

- Research into alternative methods for handling the longitudinal aspects of the eligibility status of the panel cases, use of the panel cases prior cycle eligibility status data, and how to best use this information in light of the changes to the survey eligibility standards between the sample members earning their degrees in the 20th and 21st centuries.

References

Chromy, J.R. (1979), "Sequential Sample Selection Methods," Proceedings of the American Statistical Association, Survey Research Methods Section, 401–406.

Cox, Brenda G. (2003). *The Survey of Doctorate Recipients: Redesigned for the 21st Century*. Report submitted to the National Science Foundation by RoperASW under subcontract from Mathematica Policy Research, Inc., Washington. DC

Cox, Brenda. G., Karen Grigorian, Fang Wang, Rebecca Wang, and Rachel Harter (2012a). *2010 Survey of Doctorate Recipients: Investigating an Integrated Design for the 21st Century*. Report submitted to the National Science Foundation by the National Opinion Research Center at the University of Chicago, Chicago, IL.

Cox, Brenda. G., Karen Grigorian, Fang Wang, and Rebecca Wang (2012b). *2010 Survey of Doctorate Recipients: Sample Design and Implementation.* Report submitted to the National Science Foundation by the National Opinion Research Center at the University of Chicago, Chicago, IL.

Cox, Brenda G., Karen Grigorian and Michael Yang (2006). *The 2006 International Survey of Doctorate Recipients (ISDR): Sample Design*. Report submitted to the National Science Foundation by Battelle under subcontract to the National Opinion Research Center at the University of Chicago, IL.

Grigorian, Karen and Tom Hoffer (2005). *Non-U.S. Citizen Undercoverage Feasibility Study Report*. Report submitted to the National Science Foundation by the National Opinion Research Center at the University of Chicago, Chicago, IL.

Harter, Rachel, Michael Sinclair, Karen Grigorian, Susan Hinkins, Brenda G. Cox, Rebecca Wang, Peter Kwok, Michael Yang, and Fang Wang (2012). *2008 Integrated Survey of Doctorate Recipients: Weighting and Variance Estimation Report*, Report submitted to the National Science Foundation by the National Opinion Research Center at the University of Chicago, Chicago, IL.

Mitchell, Susan, Ramal Moonesinghe, and Brenda Cox (1998). *Using the Survey of Doctorate Recipients in Time Series Analysis: 1989-1993*. Final report submitted to the National Science Foundation under a subcontract to the National Research Council. Washington, DC: Mathematica Policy Research, Inc.

Moonesinghe, Ramal (1998). *Sampling Design and Weighting Procedures for the 1995 Survey of Doctorate Recipients.* Final Report submitted to the National Science Foundation under a subcontract to the National Research Council. Washington, DC: Mathematica Policy Research, Inc.

National Science Foundation public website. (2012). NSF at a Glance. http://www.nsf.gov/about/glance.jsp.

Selfa, Lance, Jessica Knoerzer, Karen Grigorian, and Lynn Milan (2012). *Coping with Missing Data: Assessing Methods for Logically Assigning Race/Ethnicity*, Report presented at the American Association of Public Opinion Research 67th Annual Conference, May 2012, Orlando, FL.

Sinclair, Michael, and Julia Batishev. (2012). 2*010 Integrated Survey of Doctorate Recipients (NSDR/ISDR): Survey Weighting Methodology Using Logistic Modeling Procedures*, Report submitted to the National Science Foundation by NORC at the University of Chicago, IL, September 12, 2012.

Yang, Y. Michael, Brenda G. Cox, Karen Grigorian and Scott Sederstrom. (2006). *Sample Design and Implementation for the 2006 Survey of Doctorate Recipients*, Report submitted to the National Science Foundation by the National Opinion Research Center at the University of Chicago, Chicago, IL.

Yang, Y. Michael, Karen Grigorian, Scott Sederstrom, Rachel Harter, and Tom Hoffer. (2004). *Sample Design and Implementation for the 2003 Survey of Doctorate Recipients*, Report submitted to the National Science Foundation by the National Opinion Research Center at the University of Chicago, Chicago, IL.

Appendices

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
1.1.0	Central Africa	401	Angola	NH black
		408	Cameroon	NH black
		410	Central African Republic	NH black
		411	Chad	NH black
		413	Congo	NH black
		416	Equatorial Guinea	NH black
		419	Gabon	NH black
		443	Sao Tome and Principe	NH black
		459	Zaire	NH black
		462	Africa, not specified	NH black
		463	Central Africa, not specified	NH black
		465	Equatorial Africa, not specified	NH black
		466	French Equatorial Africa, not specified	NH black
1.2.0	Western Africa	403	Benin (formerly Dahomey)	NH black
		406	Burkina Faso	NH black
		409	Cape Verde	NH black
		420	Gambia	NH black
		421	Ghana	NH black
		423	Guinea	NH black
		424	Guinea-Bissau	NH black
		425	Ivory Coast	NH black
		429	Liberia	NH black
		433	Mali	NH black
		434	Mauritania	NH black
		439	Niger	NH black
		440	Nigeria	NH black
		444	Senegal	NH black
		447	Sierra Leone	NH black
		454	Togo	NH black
		467	French West Africa, not specified	NH black
		469	Western Africa, not specified	NH black
1.3.0	Eastern Africa	402	Bassas da India	NH black
		405	British Indian Ocean Territory	NH black
		407	Burundi	NH black
		412	Comoros	NH black
		414	Djibouti	NH black
		417	Ethiopia	NH black
		418	Europa Island	NH black
		422	Glorioso Islands	NH black

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		426	Juan de Nova Island	NH black
		427	Kenya	NH black
		431	Madagascar	NH black
		432	Malawi	NH black
		435	Mayotte	NH black
		437	Mozambique	NH black
		441	Reunion	NH black
		442	Rwanda	NH black
		445	Mauritius	NH black
		446	Seychelles	NH black
		448	Somalia	NH black
		453	Tanzania	NH black
		457	Uganda	NH black
		460	Zambia	NH black
		461	Zimbabwe	NH black
		464	Eastern Africa, not specified	NH black
		471	Eritrea	NH black
1.4.0	Southern Africa	404	Botswana	NH black
		428	Lesotho	NH black
		438	Namibia	NH black
		449	South Africa	NH white
		450	St. Helena	NH black
		452	Swaziland	NH black
		455	Tromelin Island	NH black
		470	Southern Africa, not specified	NH black
1.5.0	North Africa	400	Algeria	NH white
		415	Egypt	NH white
		430	Libya	NH white
		436	Morocco	NH white
		451	Sudan	NH black
		456	Tunisia	NH white
		458	Western Sahara	NH black
		468	North Africa, not specified	NH black
2.1.0	Middle East	201	Bahrain	NH white
		208	Cyprus	NH white
		213	Iraq	NH white
		214	Israel	NH white
		216	Jordan	NH white
		220	Kuwait	NH white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		222	Lebanon	NH white
		228	Oman	NH white
		232	Qatar	NH white
		233	Saudi Arabia	NH white
		237	Syria	NH white
		240	Turkey	NH white
		241	United Arab Emirates	NH white
		243	Yemen, Peoples Democratic Republic	NH white
		244	Yemen, Unified (1991 and after)	NH white
		245	Asia, not specified	NH white
		246	Asia Minor, not specified	NH white
		248	Gaza Strip	NH white
		250	Iraq-Saudi Arabia, Neutral Zone	NH white
		251	Mesopotamia, not specified	NH white
		252	Middle East, not specified	NH white
		253	Palestine, not specified	NH white
		254	Persian Gulf States, not specified	NH white
		256	West Bank	NH white
2.2.0	Southwest Asia	200	Afghanistan	NH Asian
		202	Bangladesh	NH Asian
		203	Bhutan	NH Asian
		210	India	NH Asian
		212	Iran	NH white
		225	Maldives	NH Asian
		227	Nepal	NH Asian
		229	Pakistan	NH Asian
		236	Sri Lanka	NH Asian
		257	Southwest Asia, not specified	NH Asian
2.3.0	Southeast Asia	204	Brunei	NH Asian
		205	Myanmar (formerly Burma)	NH Asian
		206	Cambodia	NH Asian
		211	Indonesia	NH Asian
		221	Laos	NH Asian
		224	Malaysia	NH Asian
		230	Paracel Islands	NH Asian
		231	Philippines	NH Asian
		234	Singapore	NH Asian
		235	Spratley Islands	NH Asian
		239	Thailand	NH Asian

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		242	Vietnam	NH Asian
		249	Indochina, not specified	NH Asian
		255	Southeast Asia, not specified	NH Asian
		258	Timor-Leste	NH Asian
2.4.0	East Asia	207	China	NH Asian
		209	Hong Kong	NH Asian
		215	Japan	NH Asian
		217	Korea, not specified	NH Asian
		218	South Korea	NH Asian
		219	North Korea	NH Asian
		223	Macao	NH Asian
		226	Mongolia	NH Asian
		238	Taiwan	NH Asian
		247	East Asia, not specified	NH Asian
3.1.0	Eastern Europe, including FSU	104	Bulgaria	NH white
		105	Czechoslovakia or Czech Republic	NH white
		117	Hungary	NH white
		128	Poland	NH white
		132	Romania	NH white
		147	Yugoslavia	NH white
		150	Eastern Europe, not specified	NH white
		155	Slovakia	NH white
		156	Serbia/Montenegro/Kosovo	NH white
		157	Slovenia	NH white
		158	Macedonia	NH white
		159	Bosnia-Hercegovina	NH white
		160	Croatia	NH white
		180	USSR	NH white
		181	Baltic states, not specified	NH white
		182	Estonia	NH white
		183	Latvia	NH white
		184	Lithuania	NH white
		185	Moldova	NH white
		186	Belarus (Byelarus)	NH white
		187	Russia	NH white
		188	Kazakhstan	NH white
		189	Armenia	NH white
		190	Azerbaijan	NH white
		191	Georgia	NH white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		192	Uzbekistan	NH white
		193	Ukraine	NH white
		194	Tajikistan	NH white
		195	Kyrgyzstan	NH white
		196	Turkmenistan	NH white
3.2.0	Central Europe	102	Austria	NH white
		110	Germany, not specified	NH white
		111	West Germany	NH white
		112	West Berlin	NH white
		113	East Berlin	NH white
		114	East Germany	NH white
		120	Italy	NH white
		122	Liechtenstein	NH white
		124	Malta	NH white
		146	Vatican City	NH white
		149	Central Europe, not specified	NH white
3.3.0	Western Europe	103	Belgium	NH white
		109	France	NH white
		123	Luxembourg	NH white
		125	Monaco	NH white
		126	Netherlands	NH white
		137	Switzerland	NH white
		148	Europe, not specified	NH white
		154	Western Europe, not specified	NH white
3.4.0	Northern Europe	106	Denmark	NH white
		107	Faroe Islands	NH white
		108	Finland	NH white
		118	Iceland	NH white
		119	Ireland	NH white
		121	Jan Mayen	NH white
		127	Norway	NH white
		135	Svalbard	NH white
		136	Sweden	NH white
		138	United Kingdom, not specified	NH white
		139	England	NH white
		140	Scotland	NH white
		141	Wales	NH white
		142	Northern Ireland	NH white
		143	Guernsey	NH white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		144	Jersey	NH white
		145	Isle of Man	NH white
		151	Lapland, not specified	NH white
		152	Northern Europe, not specified	NH white
3.5.0	Southern Europe	100	Albania	NH white
		101	Andorra	NH white
		115	Gibraltar	NH white
		116	Greece	NH white
		129	Portugal	NH white
		130	Azores Islands	NH white
		131	Madeira Islands	NH white
		133	San Marino	NH white
		134	Spain	Hispanic white
		153	Southern Europe, not specified	NH white
4.0.0	South America	375	Argentina	Hispanic white
		376	Bolivia	Hispanic white
		377	Brazil	NH white
		378	Chile	Hispanic white
		379	Colombia	Hispanic white
		380	Ecuador	Hispanic white
		381	Falkland Islands	NH white
		382	French Guiana	Hispanic white
		383	Guyana	NH black
		384	Paraguay	Hispanic white
		385	Peru	Hispanic white
		386	Suriname	NH black
		387	Uruguay	Hispanic white
		388	Venezuela	Hispanic white
		389	South America, not specified	Hispanic white
5.0.0	Caribbean	330	Anguilla	NH black
		331	Antigua and Barbuda	NH black
		332	Aruba	NH white
		333	Bahamas	NH black
		334	Barbados	NH black
		335	British Virgin Islands	NH black
		336	Cayman Islands	NH black
		337	Cuba	Hispanic white
		338	Dominica	NH black
		339	Dominican Republic	Hispanic white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		340	Grenada	NH black
		341	Guadeloupe	NH black
		342	Haiti	NH black
		343	Jamaica	NH black
		344	Martinique	NH black
		345	Montserrat	NH black
		346	Netherlands Antilles	NH black
		347	St. Barthelemy	NH black
		348	St. Kitts-Nevis	NH black
		349	St. Lucia	NH black
		350	St. Vincent and the Grenadines	NH black
		351	Trinidad and Tobago	NH black
		352	Turks and Caicos Islands	NH black
		353	Caribbean, not specified	NH black
		354	Antilles, not specified	NH black
		355	British West Indies, not specified	NH black
		356	Latin America, not specified	Hispanic white
		357	Leeward Islands, not specified	NH black
		358	West Indies, not specified	NH black
		359	Windward Islands, not specified	NH black
6.1.0	Central America, including Mexico	310	Belize	Hispanic white
		311	Costa Rica	Hispanic white
		312	El Salvador	Hispanic white
		313	Guatemala	Hispanic white
		314	Honduras	Hispanic white
		315	Mexico	Hispanic white
		316	Nicaragua	Hispanic white
		317	Panama	Hispanic white
		318	Central America, not specified	Hispanic white
6.2.01	USA - Pacific	002	Alaska	NH white
		006	California	NH white
		015	Hawaii	NH Asian
		041	Oregon	NH white
		053	Washington	NH white
		093	Pacific region, state suppressed	NH white
6.2.02	USA - Mountain	004	Arizona	NH white
		008	Colorado	NH white
		016	Idaho	NH white
		030	Montana	NH white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		032	Nevada	NH white
		035	New Mexico	NH white
		049	Utah	NH white
		056	Wyoming	NH white
		092	Mountain region, state suppressed	NH white
6.2.03	USA - West South Central	005	Arkansas	NH white
		022	Louisiana	NH white
		040	Oklahoma	NH white
		048	Texas	NH white
		091	West South Central region, state suppressed	NH white
6.2.04	USA - East South Central	001	Alabama	NH white
		021	Kentucky	NH white
		047	Tennessee	NH white
		090	East South Central region, state suppressed	NH white
		028	Mississippi	NH white
6.2.05	USA - South Atlantic	010	Delaware	NH white
		011	District of Columbia	NH white
		012	Florida	NH white
		013	Georgia	NH white
		024	Maryland	NH white
		037	North Carolina	NH white
		045	South Carolina	NH white
		051	Virginia	NH white
		054	West Virginia	NH white
		089	South Atlantic region, state suppressed	NH white
6.2.06	USA - West North Central	019	Iowa	NH white
		020	Kansas	NH white
		027	Minnesota	NH white
		029	Missouri	NH white
		031	Nebraska	NH white
		038	North Dakota	NH white
		046	South Dakota	NH white
		088	West North Central region, state suppressed	NH white
6.2.07	USA - East North Central	017	Illinois	NH white
		018	Indiana	NH white
		026	Michigan	NH white
		039	Ohio	NH white
		055	Wisconsin	NH white
		087	East North Central region, state suppressed	NH white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
6.2.08	USA - Middle Atlantic	034	New Jersey	NH white
		036	New York	NH white
		042	Pennsylvania	NH white
		086	Middle Atlantic region, state suppressed	NH white
6.2.09	USA - New England	009	Connecticut	NH white
		023	Maine	NH white
		025	Massachusetts	NH white
		033	New Hampshire	NH white
		044	Rhode Island	NH white
		050	Vermont	NH white
		085	New England region, state suppressed	NH white
6.2.10	USA - Territories	060	American Samoa	NH white
		066	Guam	NH white
		067	Johnston Atoll	NH white
		069	Northern Mariana Islands	NH white
		071	Midway Islands	NH white
		072	Puerto Rico	Hispanic white
		076	Navassa Island	NH white
		078	U.S. Virgin Islands	NH white
		079	Wake Island	NH white
		081	Baker Island	NH white
		082	Howland Island	NH white
		083	Jarvis Island	NH white
		084	Kingman Reef	NH white
		095	Palmyra Atoll	NH white
		096	U.S. State or Territory (Puerto Rico and Island Areas)	NH white
6.3.0	Northern North America	300	Bermuda	NH black
		301	Canada	NH white
		302	Greenland	NH native
		303	St. Pierre and Miquelon	NH black
		304	North America, not specified	NH white
7.0.0	Oceania	500	Ashmore and Cartier Islands	NH white
		501	Australia	NH white
		502	Christmas Island, Indian Ocean	NH white
		503	Clipperton Island	NH white
		504	Cocos Islands	NH white
		505	Cook Islands	NH white
		506	Coral Sea Islands	NH white
		507	Fiji	NH white

Appendix A.1 2013 SDR Birth Region Code Frame Mapped to SESTAT Geocodes and Race/Ethnicity Imputation based on Birthplace

SDR Birth Region (BIREGION)		SESTAT Location (Geocode)		Imputed Race/Ethnicity base on Birthplace
Code	Region Name	Geocode	SESTAT Location Name	
		508	French Polynesia	NH white
		509	Kiribati	NH white
		510	Marshall Islands	NH white
		511	Micronesia	NH white
		512	Nauru	NH white
		513	New Caledonia	NH white
		514	New Zealand	NH white
		515	Niue	NH white
		516	Norfolk Island	NH white
		517	Palau	NH white
		518	Papua New Guinea	NH white
		519	Pitcairn Islands	NH white
		520	Solomon Islands	NH white
		521	Tokelau	NH white
		522	Tonga	NH white
		523	Tuvalu	NH white
		524	Vanuatu	NH white
		525	Wallis and Futuna Islands	NH white
		526	Western Samoa	NH white
		527	Oceania, not specified	NH white
		528	Polynesia, not specified	NH white
		529	Melanesia, not specified	NH white
		550	Antarctica	NH white
		551	Bouvet Island	NH white
		552	French Southern and Antarctic Lands	NH white
		553	Heard and McDonald Islands	NH white
7.1.0	At sea/abroad	554	At Sea	NH white
		555	Abroad, not specified	NH white
8.2.0	Missing	999	Missing/Unknown	NH white

FSU = Former Soviet Union country; NH = non-Hispanic.

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
1	Computer and information sciences, mathematics, physical sciences, engineering	1	Computer and math sciences	1	Computer and information sciences	5	Computer/ information sciences	D67	Computer/information sciences	400	Computer Science
										410	Information Science/Systems
										415	Robotics
										419	Computer/Information Sciences, Other
				2	Mathematics and statistics	4	Mathematics	841	Applied mathematics	420	Applied Mathematics
								842	Mathematics, general	498	Mathematics/Statistics, General
								843	Operations research	363	Operations Research
										465	Operations Research
										930	Operations Research
								844	Statistics	450	Statistics
										690	Statistics
								845	OTHER mathematics	425	Algebra
										430	Analysis & Functional Analysis
										435	Geometry/Geometric Analysis
										440	Logic
										445	Number Theory
										455	Topology/Foundations
										460	Computing Theory & Practice
										499	Mathematics/Statistics, Other
		4	Physical sciences	5	Physical sciences	1	Chemistry	873	Chemistry, except biochemistry	520	Analytical Chemistry
										521	Agricultural/Food
										522	Inorganic Chemistry
										524	Nuclear Chemistry
										526	Organic Chemistry
										528	Medicinal/Pharmaceutical
										530	Physical Chemistry
										532	Polymer Chemistry
										534	Theoretical Chemistry

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										538	Chemistry, General
										539	Chemistry, Other
						2	Physics/ astronomy	871	Astronomy and astrophysics	500	Astronomy
										505	Astrophysics
										506	Astronomy/Astrophysics
										509	Astronomy, Other
								878	Physics, except biophysics	560	Acoustics
										561	Atomic/Molecular/Chemical Physics
										562	Electron Physics
										563	Electromagnetism
										564	Particle (Elementary) Physics
										565	Biophysics
										566	Fluids
										567	Mechanics
										568	Nuclear Physics
										569	Optics/Phototonics
										570	Plasma/Fusion Physics
										572	Polymer Physics
										573	Thermal Physics
										574	Condensed Matter/Low Temperature Physics
										575	Theoretical Physics
										576	Applied Physics
										577	Medical Physics/Radiological Science
										578	Physics, General
										579	Physics, Other
						3	Earth/ocean/ atmospheric sciences	872	Atmospheric sciences and meteorology	510	Atmospheric Chemistry & Climatology
										512	Atmospheric Physics & Dynamics
										514	Meteorology

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										518	Atmospheric Science/Meteorology, General
										519	Atmospheric Science/Meteorology, Other
									875	Geology	540	Geology
											548	Mineralogy & Petrology
											549	Mineralogy/Petrology/Geological Chemistry
											550	Stratigraphy & Sedimentation
											552	Geomorphology & Glacial Geology
											554	Applied geology
											555	Applied Geology/Geological Engineering
									876	Geological sciences, other	542	Geochemistry
											544	Geophysics & Seismology
											545	Geophysics, Solid Earth
											546	Paleontology
											547	Fuel Technology/Petroleum Engineering
											558	Geological & Earth Sciences, General
											559	Geological & Earth Sciences, Other
									877	Oceanography	590	Oceanography, Chemical & Physical
									D87	Earth sciences/other physical sciences	585	Hydrology & Water Resources
											595	Marine Sciences
											599	Ocean/Marine, Other
		7	Engineering	8	Engineering	14	Electrical/ electronic/ communications engineering	727	Computer and systems engineering	321	Computer Engineering
											372	Systems Engineering
									728	Electrical, electronics and communications engineering	318	Communications Engineering
											322	Electrical Engineering
											323	Electronics Engineering
											324	Electrical, Electronics & Communications Engineering
						15	Other engineering	721	Aerospace, aeronautical, astronautical engineering	300	Aerospace, Aeronautical & Astronautical

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
								722	Agricultural engineering	303	Agricultural Engineering
								724	Bioengineering and biomedical engineering	306	Bioengineering & Biomedical Engineering
								725	Chemical engineering	312	Chemical Engineering
								726	Civil engineering	315	Civil Engineering
										316	Structural Engineering
								729	Engineering sciences, mechanics and physics	327	Engineering Mechanics
										330	Engineering Physics
										333	Engineering Science
								730	Environmental engineering	336	Environmental Health Engineering
								731	Engineering, general	398	Engineering, General
								733	Industrial and manufacturing engineering	339	Industrial & Manufacturing Engineering
								734	Materials engineering, including ceramics and textiles	309	Ceramic Sciences Engineering
										342	Materials Science Engineering
										369	Polymer & Plastics Engineering
								735	Mechanical engineering	345	Mechanical Engineering
								736	Metallurgical engineering	348	Metallurgical Engineering
								737	Mining and minerals engineering	351	Mining & Mineral Engineering
								738	Naval architecture and marine engineering	354	Naval Architecture/Marine Engineering
								739	Nuclear engineering	357	Nuclear Engineering
								740	Petroleum engineering	366	Petroleum Engineering
								741	OTHER engineering	376	Engineering Management & Administration
								D74		360	Ocean Engineering
										375	Textile Engineering
										399	Engineering, Other
2	Biological and	2	Biological, agricultural,	3	Biological, agricultural,	6	Agricultural sciences	605	Animal sciences	005	Agricultural Animal Breeding
										007	Animal Husbandry

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
	agricultural sciences, health sciences		and environmental life sciences		and environmental life sciences					010	Animal Nutrition
										012	Dairy Science
										014	Animal Science, Poultry (or Avian)
										019	Animal Science, Other
									606	Food sciences and technology	040	Food Sciences
										042	Food Distribution
										043	Food Science
										044	Food Science & Technology, Other
									607	Plant sciences	020	Agronomy & Crop Science
										023	Agricultural & Horticultural Plant Breeding
										025	Agricultural & Horticultural Plant Breeding (2010 & 2011)
										025	Plant Breeding/Genetics (1920-2009)
										030	Plant Pathology/Phytopathology
										032	Plant Protection/Pest Management
										039	Plant Sciences, Other
										050	Horticulture Science
									608	OTHER agricultural sciences	045	Soil Sciences
										046	Soil Chemistry/Microbiology
										049	Soil Sciences, Other
										098	Agriculture, General
										099	Agricultural Science, Other
									680	Environmental science or studies	054	Fish and Wildlife Science
										055	Fishing & Fisheries Sciences/Management
										580	Environmental Science
										081	Environmental Science
									681	Forestry sciences	060	Wildlife
										065	Forestry Science
										066	Forest Sciences & Biology

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										068	Forest Engineering
										070	Forest/Resources Management
										072	Wood Science & Pulp/Paper Technology
										074	Natural Resources/Conservation
										079	Forestry & Related Science, Other
										080	Wildlife/Range Management
						8	NIH biological sciences	631	Biochemistry and biophysics	100	Biochemistry
								633	Botany	125	Plant Physiology
								636	Genetics, animal and plant	115	Plant Genetics
										170	Genetics/Genomics, Human & Animal
										171	Genetics
								637	Microbiological sciences and immunology	110	Bacteriology
										156	Microbiology/Bacteriology
										157	Microbiology
										168	Virology
								639	Pharmacology, human and animal	180	Pharmacology, Human & Animal
								640	Physiology and pathology, human and animal	158	Cancer Biology
										175	Pathology, Human & Animal
										185	Physiology, Human & Animal
										186	Animal/Plant Physiology
								642	OTHER biological sciences	130	Anatomy
										137	Evolutionary Biology
										160	Neurosciences
										166	Parasitology
						9	Other biological sciences	631	Biochemistry and biophysics	105	Biophysics
								632	Biology, general	198	Biology/Biomedical Sciences, General
								633	Botany	120	Plant Pathology/Phytopathology
										129	Botany/Plant Biology

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
								634	Cell and molecular biology	136	Cell/Cellular Biology & Histology
										154	Molecular Biology
								635	Ecology	139	Ecology
								638	Nutritional sciences	163	Nutrition Sciences
								641	Zoology, general	148	Entomology
										189	Zoology, Other
								642	OTHER biological sciences	102	Bioinformatics
										103	Biomedical Sciences
										104	Computational Biology
										107	Biotechnology
										133	Biometrics & Biostatistics
										140	Hydrobiology
										142	Developmental Biology/Embryology
										145	Endocrinology
										151	Immunology
										155	Structural Biology
										167	Environmental Toxicology
										169	Toxicology
										199	Biology/Biomedical Sciences, Other
3	Health	4	Health	7	Medical sciences	781	Audiology and speech pathology	200	Speech-Language Pathology & Audiology		
						782	Health services administration	212	Health Systems/Service Administration		
						786	Medicine (e.g., dentistry, optometry, osteopathic, podiatry, veterinary)	205	Dentistry		
								207	Oral Biology/Oral Pathology		
								225	Medical/Surgery		
								235	Optometry/Ophthalmology		
								250	Veterinary Sciences		
						787	Nursing (4 years or longer program)	230	Nursing Science		
						788	Pharmacy	240	Medicinal/Pharmaceutical Sciences		

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
								789	Physical therapy and other rehabilitation/ therapeutic services	245	Rehabilitation/Therapeutic Services
								790	Public health (including environmental health and epidemiology)	210	Environmental Health
										211	Environmental Toxicology
										215	Public Health
										219	Public Health/Epidemiology
										220	Epidemiology
								791	OTHER health/medical sciences	222	Kinesiology/Exercise Science
										224	Hospital Administration
										227	Gerontology
										298	Health Sciences, General
										299	Health Sciences, Other
3	Psychology, social sciences	5	Social sciences	6	Social sciences	11	Economics	601	Agricultural economics	000	Agricultural Economics
								923	Economics	666	Economics
										667	Economics
										668	Econometrics
							12	Anthropology/ archeology/ sociology	921	Anthropology and archaeology	650	Anthropology
										773	Archaeology
								922	Criminology	658	Criminology
								929	Sociology	686	Sociology
							13	Other social sciences	620	Area and ethnic studies	652	Area /Ethnic/Cultural/Gender Studies
										770	American/U.S. Studies
								771	Linguistics	676	Linguistics
										729	Linguistics
								902	Public policy studies	682	Public Policy Analysis
								924	Geography	670	Geography
								925	History of science	710	History, Science & Technology & Society
								927	International relations	674	International Relations/Affairs

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
								928	Political science and government	678	Political Science & Government
										679	Political Science/Public Administration
								930	OTHER social sciences	662	Demography/Population Studies
										684	Gerontology
										694	Urban Affairs/Studies
										698	Social Sciences, General
										699	Social Sciences, Other
		6	Psychology	7	Psychology	10	Psychology	704	Educational psychology	618	Educational Psychology
										822	Educational Psychology
								891	Clinical psychology	600	Clinical Psychology
								892	Counseling psychology	609	Counseling
								893	Experimental psychology	615	Experimental Psychology
								894	General psychology	648	Psychology, General
								895	Industrial/Organizational psychology	621	Industrial & Organizational
								896	Social psychology	639	Social Psychology
								897	OTHER psychology	603	Cognitive Psychology & Psycholinguistics
										606	Comparative Psychology
										612	Developmental & Child Psychology
										613	Human Development & Family Studies
										616	Experimental/Comparative Psychology/Physiology
										619	Human Engineering
										620	Family Psychology
										624	Personality Psychology
										627	Physiological/Psychobiology Psychology
										630	Psychometrics
										633	Psychometrics & Quantitative Psychology
										636	School Psychology
										649	Psychology, Other

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
					Not applicable; non-SEH field					002	Agricultural Business/Management
										657	Criminal Justice & Corrections
										695	Urban/City, Community & Regional Planning
										700	American History (U.S. & Canada)
										703	Asian History
										705	European History
										706	African History
										707	Latin American History
										708	Middle/Near East Studies
										718	History, General
										719	History, Other
										720	Classics
										723	Comparative Literature
										724	Folklore
										725	English and American Literature
										726	English Language and Literature
										732	American Literature (U.S. & Canada)
										733	English Literature (British & Commonwealth)
										734	English Language
										735	Creative Writing
										736	Speech & Rhetorical Studies
										738	Letters, General
										739	Letters, Other
										740	French
										743	German
										746	Italian
										749	Spanish
										752	Russian

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										755	Slavic (other than Russian)
										758	Chinese
										762	Japanese
										765	Hebrew
										768	Arabic
										769	Other Languages & Literature
										774	Art, Applied
										775	Art, Fine/Applied
										776	Art History/Criticism/Conservation
										780	Music
										785	Philosophy
										786	Music Theory & Composition
										787	Music Performance
										788	Musicology/Ethnomusicology
										789	Music, Other
										790	Religion/Religious Studies
										791	Religion and Theology
										792	Bible/Biblical Studies
										795	Drama/Theater Arts
										798	Humanities, General
										799	Humanities, Other
										800	Curriculum & Instruction
										805	Educational Administration & Supervision
										806	Urban Education and Leadership
										807	Educational Leadership
										808	Educational Policy Analysis
										810	Educational/Instructional Media Design
										814	Educational Measurement & Statistics

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										815	Educational Statistics/Research Methods
										820	Educational Assessment/Testing/Measurement
										825	School Psychology
										830	Social/Philosophical Foundations of Education
										833	International Education
										835	Special Education
										840	Counseling Education/Counseling & Guidance
										845	Higher Education/Evaluation & Research
										850	Pre-elementary/Early Childhood Teacher Education
										852	Elementary Teacher Education
										854	Jr. High Education
										856	Secondary Teacher Education
										858	Adult & Continuing Teacher Education
										860	Agricultural Education
										861	Art Education
										862	Business Education
										864	English Education
										866	Foreign Languages Education
										867	Physical Education, Health and Recreation
										868	Health Education
										870	Family & Consumer/Human Science
										872	Technical & Industrial Arts Education
										874	Mathematics Education
										876	Music Education
										878	Nursing Education
										880	Physical Education & Coaching
										882	Reading Education
										884	Science Education

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										885	Social Science Education
										886	Speech Education
										887	Technical Education
										888	Trade & Industrial Education
										889	Teacher Education & Professional Development
										898	Education, General
										899	Education, Other
										900	Accounting
										901	Finance
										905	Banking/Financial Support Services
										910	Business Administration & Management
										912	Hospitality, Food Service and Tourism Management
										915	Business/Managerial Economics
										916	International Business/Trade/Commerce
										917	Management Information Systems/Business Statistics
										920	Marketing Management & Research
										921	Human Resources Development
										925	Business Statistics
										935	Organizational Behavior
										938	Business Management/Administration, General
										939	Business Management/Administration, Other
										940	Communication Research
										945	Journalism
										947	Mass Communication/Media Studies
										950	Film, Radio, TV & Digital Communication
										957	Communication Theory
										958	Communication, General
										959	Communication, Other

Appendix A.2 2013 SDR Field of Study Coding Taxonomies Crosswalk

3-level FOS FOD3		7-level FOS MAJFLD7		8-level DST FOS DSTFLD8		15-level FOS SDRFLD15		SESTAT FOS NSDRMED13		DRF FOS PHDFIELD	
Code	Label	Code	Label	Code	Label	Code	Label	Code	Label	Code	Label
										960	Architecture/Environmental Design
										964	Family/Consumer Science/Human Science
										968	Law
										972	Library Science
										974	Parks/Sports/Rec./Leisure/Fitness
										976	Public Administration
										980	Social Work
										984	Theology/Religious Education
										988	Professional Fields, General
										989	Other Fields, NEC
										999	Unknown Field

DRF = Doctorate Records File; DST = detailed statistical table; FOS = field of study.

NOTE: PHDFIELD degrees shown in highlight were added to the Survey of Earned doctorates field of study taxonomy in the 2010 cycle.

Appendix A.3 2013 SDR Data Sources Used to Develop Sampling Frame Variables

Data Source Flag Code	Data Source Flag Code Description
10	DRF, reported data for variable
11	Citizenship imputed from DRF with BIRTHPL and PDLOC
12	DRF, data updated in DRF after used for SDR sample
13	Postdoc location imputed from DRF with PDEMPLOY
14	Sex assigned from name in DRF
20	Hispanic surname list
21	Asian surname list
22	Race reconciliation file (Not Imputed Data)
23	Race reconciliation file (Imputed Data)*
24	PI Category Not Available, Other Race Selected
26	Sex assigned during frame processing (Not Imputed Data)
27	Race/ethnicity imputed from birth place
28	Race/ethnicity imputed to default modal assignment (NH white)
29	SDR data delivery hot-deck imputation
30	Pre-1991 source data (Not Imputed Data)
31	2001 SDR sampling file (Not Imputed Data)
32	2003 chronic unlocatables (Not Imputed Data)
33	Master 2003 data base (Not Imputed Data)
34	Survey administration data
35	Permanent Ineligible database
40	SDR 1993
41	SDR 1995
42	SDR 1997
43	SDR 1999
44	SDR 2001
4X	SDR 2001, reported data used for Hispanic indicator
45	SDR 2003
46	SDR 2006
47	SDR 2008
48	SDR 2010
80*	Age imputed from PhD year, degree earned at 21 years
81*	Age imputed from BA year, degree earned at 18 years
90	ISEX: Missing data, imputed female
91	IHCAPIN: Missing data, imputed not handicapped
92	IBIRCIT: Missing data, imputed not native born
93	ICURCIT: Missing data, imputed current U.S. citizen
94	IPDUS: Missing data, imputed staying in U.S.

Appendix A.3 2013 SDR Data Sources Used to Develop Sampling Frame Variables

Data Source Flag Code	Data Source Flag Code Description
95	IHSPIN: Missing data, imputed not Hispanic
96	ILOCSTAT: Missing data, imputed to in U.S.
99	Missing data

* The birth year imputation rules assume that sample members earned degrees at an age somewhat lower than average for the population. This is intentional so that we minimize any sample undercoverage caused by eliminating doctorates with missing birth year's that may have earned a degree at a young age. During data collection, every effort is made to collect date of birth from sample members with an imputed birth date to confirm their eligibility for the sample, and in the next survey cycle the unimputed data replace the imputed birth year estimate in frame construction.

Appendix B.1 2013 NSDR Strata and Frame Counts

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Sample Cases	2010 New Cohort Population Size	2011 New Cohort Population Size	Total Frame Size
1	Hispanic	Male	Computer/math	101	76	99	276
2	Hispanic	Male	Biological and agri. sci.	338	251	251	840
3	Hispanic	Male	Health sci.	60	18	33	111
4	Hispanic	Male	Physical and related sci.	237	126	140	503
5	Hispanic	Male	Social sci.	191	141	150	482
6	Hispanic	Male	Psychology	144	64	79	287
7	Hispanic	Male	Engineering	266	223	240	729
8	Hispanic	Female	Computer/math	57	23	18	98
9	Hispanic	Female	Biological and agri. sci.	309	273	297	879
10	Hispanic	Female	Health sci.	82	55	61	198
11	Hispanic	Female	Physical and related sci.	99	78	79	256
12	Hispanic	Female	Social sci.	187	126	173	486
13	Hispanic	Female	Psychology	346	178	243	767
14	Hispanic	Female	Engineering	76	100	85	261
15	NH Black	Male	Computer/math	81	60	54	195
16	NH Black	Male	Biological and agri. sci.	275	139	155	569
17	NH Black	Male	Health sci.	77	30	45	152
18	NH Black	Male	Physical and related sci.	178	87	90	355
19	NH Black	Male	Social sci.	299	109	116	524
20	NH Black	Male	Psychology	128	38	56	222
21	NH Black	Male	Engineering	235	147	132	514
22	NH Black	Female	Computer/math	61	24	29	114
23	NH Black	Female	Biological and agri. sci.	253	213	210	676
24	NH Black	Female	Health sci.	153	129	138	420
25	NH Black	Female	Physical and related sci.	69	57	61	187
26	NH Black	Female	Social sci.	252	143	165	560
27	NH Black	Female	Psychology	294	187	183	664
28	NH Black	Female	Engineering	70	64	65	199
29	U.S. Born, NH Asian	Male	Computer/math	71	61	50	182
30	U.S. Born, NH Asian	Male	Biological and agri. sci.	255	137	146	538
31	U.S. Born, NH Asian	Male	Health sci.	56	21	22	99
32	U.S. Born, NH Asian	Male	Physical and related sci.	143	64	62	269
33	U.S. Born, NH Asian	Male	Social sci.	70	41	42	153
34	U.S. Born, NH Asian	Male	Psychology	64	26	28	118
35	U.S. Born, NH Asian	Male	Engineering	201	150	156	507
36	U.S. Born, NH Asian	Female	Computer/math	54	17	20	91
37	U.S. Born, NH Asian	Female	Biological and agri. sci.	270	197	180	647
38	U.S. Born, NH Asian	Female	Health sci.	62	37	34	133
39	U.S. Born, NH Asian	Female	Physical and related sci.	75	41	49	165
40	U.S. Born, NH Asian	Female	Social sci.	81	56	52	189
41	U.S. Born, NH Asian	Female	Psychology	145	84	87	316

Appendix B.1 2013 NSDR Strata and Frame Counts

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Sample Cases	2010 New Cohort Population Size	2011 New Cohort Population Size	Total Frame Size
42	U.S. Born, NH Asian	Female	Engineering	74	81	59	214
43	NH American Indian	Male	All Fields	170	89	98	357
44	NH American Indian	Female	All Fields	158	108	109	375
45	NH Pacific Islander	Male	All fields	63	30	41	134
46	NH Pacific Islander	Female	All fields	67	38	28	133
47	U.S. Born, Disabled NH White	Male	Computer/math	112	39	25	176
48	U.S. Born, Disabled NH White	Male	Biological and agri. sci.	366	83	94	543
49	U.S. Born, Disabled NH White	Male	Health sci.	53	13	19	85
50	U.S. Born, Disabled NH White	Male	Physical and related sci.	307	56	74	437
51	U.S. Born, Disabled NH White	Male	Social sci.	213	56	43	312
52	U.S. Born, Disabled NH White	Male	Psychology	186	39	39	264
53	U.S. Born, Disabled NH White	Male	Engineering	193	64	66	323
54	U.S. Born, Disabled NH White	Female	Computer/math	30	10	13	53
55	U.S. Born, Disabled NH White	Female	Biological and agri. sci.	122	76	89	287
56	U.S. Born, Disabled NH White	Female	Health sci.	66	42	27	135
57	U.S. Born, Disabled NH White	Female	Physical and related sci.	44	21	28	93
58	U.S. Born, Disabled NH White	Female	Social sci.	125	26	52	203
59	U.S. Born, Disabled NH White	Female	Psychology	159	69	60	288
60	U.S. Born, Disabled NH White	Female	Engineering	29	9	9	47
61	U.S. Born, Not Disabled NH White	Male	Chemistry	1,500	589	591	2,680
62	U.S. Born, Not Disabled NH White	Male	Physics/astronomy	1,048	514	606	2,168
63	U.S. Born, Not Disabled NH White	Male	Earth/ocean/atmos.	470	239	247	956
64	U.S. Born, Not Disabled NH White	Male	Math	696	428	441	1,565
65	U.S. Born, Not Disabled NH White	Male	Computer/info. sci.	277	372	378	1,027
66	U.S. Born, Not Disabled NH White	Male	Agricultural sci.	604	228	241	1,073
67	U.S. Born, Not Disabled NH White	Male	Medical sci.	326	233	258	817
68	U.S. Born, Not Disabled NH White	Male	NIH bio sci.	1,462	913	894	3,269
69	U.S. Born, Not Disabled NH White	Male	Other bio sci.	1,386	852	815	3,053
70	U.S. Born, Not Disabled NH White	Male	Psychology	1,675	662	639	2,976
71	U.S. Born, Not Disabled NH White	Male	Economics	511	199	206	916
72	U.S. Born, Not Disabled NH White	Male	Anthro/arch/sociology	449	257	286	992
73	U.S. Born, Not Disabled NH White	Male	Other social sci.	694	502	441	1,637
74	U.S. Born, Not Disabled NH White	Male	Electrical/electron/comm.	552	336	342	1,230
75	U.S. Born, Not Disabled NH White	Male	Other engineering	1,639	1,121	1,116	3,876
76	U.S. Born, Not Disabled NH White	Female	Chemistry	400	320	343	1,063
77	U.S. Born, Not Disabled NH White	Female	Physics/astronomy	109	138	133	380
78	U.S. Born, Not Disabled NH White	Female	Earth/ocean/atmos.	146	178	160	484
79	U.S. Born, Not Disabled NH White	Female	Math	145	154	142	441
80	U.S. Born, Not Disabled NH White	Female	Computer/info. sci.	68	79	74	221
81	U.S. Born, Not Disabled NH White	Female	Agricultural sci.	174	193	224	591

Appendix B.1 2013 NSDR Strata and Frame Counts

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Sample Cases	2010 New Cohort Population Size	2011 New Cohort Population Size	Total Frame Size
82	U.S. Born, Not Disabled NH White	Female	Medical sci.	655	757	739	2,151
83	U.S. Born, Not Disabled NH White	Female	NIH bio sci.	834	963	941	2,738
84	U.S. Born, Not Disabled NH White	Female	Other bio sci.	876	886	927	2,689
85	U.S. Born, Not Disabled NH White	Female	Psychology	2,010	1,445	1,568	5,023
86	U.S. Born, Not Disabled NH White	Female	Economics	132	98	64	294
87	U.S. Born, Not Disabled NH White	Female	Anthro/arch/sociology	460	415	421	1,296
88	U.S. Born, Not Disabled NH White	Female	Other social sci.	427	400	421	1,248
89	U.S. Born, Not Disabled NH White	Female	Electrical/electron/comm.	65	39	31	135
90	U.S. Born, Not Disabled NH White	Female	Other engineering	237	368	375	980
91	Non-U.S. born, NH White	Male	Chemistry	143	137	140	420
92	Non-U.S. born, NH White	Male	Physics/astronomy	188	206	240	634
93	Non-U.S. born, NH White	Male	Earth/ocean/atmos.	65	68	52	185
94	Non-U.S. born, NH White	Male	Math	143	172	157	472
95	Non-U.S. born, NH White	Male	Computer/info. sci.	96	197	226	519
96	Non-U.S. born, NH White	Male	Agricultural sci.	66	54	45	165
97	Non-U.S. born, NH White	Male	Medical sci.	68	88	69	225
98	Non-U.S. born, NH White	Male	NIH bio sci.	145	193	221	559
99	Non-U.S. born, NH White	Male	Other bio sci.	116	181	177	474
100	Non-U.S. born, NH White	Male	Psychology	92	122	127	341
101	Non-U.S. born, NH White	Male	Economics	94	119	124	337
102	Non-U.S. born, NH White	Male	Anthro/arch/sociology	58	48	47	153
103	Non-U.S. born, NH White	Male	Other social sci.	82	115	130	327
104	Non-U.S. born, NH White	Male	Electrical/electron/comm.	214	267	324	805
105	Non-U.S. born, NH White	Male	Other engineering	420	475	552	1,447
106	Non-U.S. born, NH White	Female	Chemistry	76	93	90	259
107	Non-U.S. born, NH White	Female	Physics/astronomy	77	61	45	183
108	Non-U.S. born, NH White	Female	Earth/ocean/atmos.	66	43	37	146
109	Non-U.S. born, NH White	Female	Math	81	62	69	212
110	Non-U.S. born, NH White	Female	Computer/info. sci.	74	68	56	198
111	Non-U.S. born, NH White	Female	Agricultural sci.	68	42	27	137
112	Non-U.S. born, NH White	Female	Medical sci.	72	140	137	349
113	Non-U.S. born, NH White	Female	NIH bio sci.	107	209	227	543
114	Non-U.S. born, NH White	Female	Other bio sci.	95	198	185	478
115	Non-U.S. born, NH White	Female	Psychology	138	281	333	752
116	Non-U.S. born, NH White	Female	Economics	75	57	72	204
117	Non-U.S. born, NH White	Female	Anthro/arch/sociology	76	53	81	210
118	Non-U.S. born, NH White	Female	Other social sci.	78	115	120	313
119	Non-U.S. born, NH White	Female	Electrical/electron/comm.	79	68	62	209
120	Non-U.S. born, NH White	Female	Other engineering	80	160	195	435
121	Non-U.S. born, NH Asian	Male	Chemistry	422	427	416	1,265
122	Non-U.S. born, NH Asian	Male	Physics/astronomy	328	395	445	1,168

Appendix B.1 2013 NSDR Strata and Frame Counts

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Sample Cases	2010 New Cohort Population Size	2011 New Cohort Population Size	Total Frame Size
123	Non-U.S. born, NH Asian	Male	Earth/ocean/atmos.	82	76	93	251
124	Non-U.S. born, NH Asian	Male	Math	238	300	328	866
125	Non-U.S. born, NH Asian	Male	Computer/info. sci.	227	439	459	1,125
126	Non-U.S. born, NH Asian	Male	Agricultural sci.	103	106	118	327
127	Non-U.S. born, NH Asian	Male	Medical sci.	97	152	150	399
128	Non-U.S. born, NH Asian	Male	NIH bio sci.	321	431	443	1,195
129	Non-U.S. born, NH Asian	Male	Other bio sci.	298	431	468	1,197
130	Non-U.S. born, NH Asian	Male	Psychology	70	62	56	188
131	Non-U.S. born, NH Asian	Male	Economics	110	103	117	330
132	Non-U.S. born, NH Asian	Male	Anthro/arch/sociology	61	23	38	122
133	Non-U.S. born, NH Asian	Male	Other social sci.	74	78	88	240
134	Non-U.S. born, NH Asian	Male	Electrical/electron/comm.	588	781	927	2,296
135	Non-U.S. born, NH Asian	Male	Other engineering	1,195	1,393	1,472	4,060
136	Non-U.S. born, NH Asian	Female	Chemistry	205	246	275	726
137	Non-U.S. born, NH Asian	Female	Physics/astronomy	84	113	117	314
138	Non-U.S. born, NH Asian	Female	Earth/ocean/atmos.	80	59	60	199
139	Non-U.S. born, NH Asian	Female	Math	99	208	195	502
140	Non-U.S. born, NH Asian	Female	Computer/info. sci.	89	130	156	375
141	Non-U.S. born, NH Asian	Female	Agricultural sci.	83	69	105	257
142	Non-U.S. born, NH Asian	Female	Medical sci.	105	202	208	515
143	Non-U.S. born, NH Asian	Female	NIH bio sci.	266	462	470	1,198
144	Non-U.S. born, NH Asian	Female	Other bio sci.	287	500	482	1,269
145	Non-U.S. born, NH Asian	Female	Psychology	98	187	170	455
146	Non-U.S. born, NH Asian	Female	Economics	83	109	134	326
147	Non-U.S. born, NH Asian	Female	Anthro/arch/sociology	78	63	66	207
148	Non-U.S. born, NH Asian	Female	Other social sci.	85	110	113	308
149	Non-U.S. born, NH Asian	Female	Electrical/electron/comm.	101	218	227	546
150	Non-U.S. born, NH Asian	Female	Other engineering	206	417	480	1,103
Total				**38,424**	**31,300**	**32,655**	**102,379**

NH = Non-Hispanic.

Appendix B.2 2013 NSDR Minimum Respondent Stratum Sample Sizes with and without Finite Population Correction Adjustment and Associated Yield Rates

Stratum	Demographic Group	Gender	Field of Degree	Minimum Sample Size Derivation				
				2013 Total Population	Minimum Respondent Sample Size Unadjusted for FPC	Minimum Respondent Sample Size with FPC Adjustment	Supplemental Sample Allocated to Meet Sample Size Minimum?	2010 NSDR Weighted Yield Rate Used Sample Size Allocation
1	Hispanic	Male	Computer/math	1,267	60	57		79.2%
2	Hispanic	Male	Biological and agri. sci.	4,274	60	59		79.2%
3	Hispanic	Male	Health sci.	365	60	52	Yes	79.2%
4	Hispanic	Male	Physical and related sci.	2,917	60	59		79.2%
5	Hispanic	Male	Social sci.	2,340	60	59		79.2%
6	Hispanic	Male	Psychology	1,820	60	58		79.2%
7	Hispanic	Male	Engineering	3,425	60	59		79.2%
8	Hispanic	Female	Computer/math	294	60	50	Yes	78.1%
9	Hispanic	Female	Biological and agri. sci.	3,300	60	59		78.1%
10	Hispanic	Female	Health sci.	841	60	56		78.1%
11	Hispanic	Female	Physical and related sci.	1,005	60	57		78.1%
12	Hispanic	Female	Social sci.	1,944	60	58		78.1%
13	Hispanic	Female	Psychology	3,738	60	59		78.1%
14	Hispanic	Female	Engineering	845	60	56		78.1%
15	NH Black	Male	Computer/math	967	60	56		70.2%
16	NH Black	Male	Biological and agri. sci.	3,178	60	59		70.2%
17	NH Black	Male	Health sci.	792	60	56	Yes	70.2%
18	NH Black	Male	Physical and related sci.	2,026	60	58		70.2%
19	NH Black	Male	Social sci.	3,294	60	59		70.2%
20	NH Black	Male	Psychology	1,752	60	58		70.2%
21	NH Black	Male	Engineering	2,669	60	59		70.2%
22	NH Black	Female	Computer/math	376	60	52	Yes	72.6%
23	NH Black	Female	Biological and agri. sci.	2,810	60	59		72.6%
24	NH Black	Female	Health sci.	1,704	60	58		72.6%
25	NH Black	Female	Physical and related sci.	742	60	56	Yes	72.6%
26	NH Black	Female	Social sci.	2,677	60	59		72.6%
27	NH Black	Female	Psychology	4,057	60	59		72.6%
28	NH Black	Female	Engineering	706	60	55	Yes	72.6%
29	U.S. Born, NH Asian	Male	Computer/math	796	60	56		80.2%
30	U.S. Born, NH Asian	Male	Biological and agri. sci.	2,731	60	59		80.2%
31	U.S. Born, NH Asian	Male	Health sci.	240	60	48	Yes	80.2%
32	U.S. Born, NH Asian	Male	Physical and related sci.	1,478	60	58		80.2%
33	U.S. Born, NH Asian	Male	Social sci.	741	60	56		80.2%
34	U.S. Born, NH Asian	Male	Psychology	600	60	55	Yes	80.2%
35	U.S. Born, NH Asian	Male	Engineering	2,211	60	58		80.2%

Appendix B.2 2013 NSDR Minimum Respondent Stratum Sample Sizes with and without Finite Population Correction Adjustment and Associated Yield Rates

Stratum	Demographic Group	Gender	Field of Degree	2013 Total Population	Minimum Sample Size Derivation			
					Minimum Respondent Sample Size Unadjusted for FPC	Minimum Respondent Sample Size with FPC Adjustment	Supplemental Sample Allocated to Meet Sample Size Minimum?	2010 NSDR Weighted Yield Rate Used Sample Size Allocation
36	U.S. Born, NH Asian	Female	Computer/math	211	60	47	Yes	80.1%
37	U.S. Born, NH Asian	Female	Biological and agri. sci.	2,403	60	59		80.1%
38	U.S. Born, NH Asian	Female	Health sci.	438	60	53	Yes	80.1%
39	U.S. Born, NH Asian	Female	Physical and related sci.	666	60	55		80.1%
40	U.S. Born, NH Asian	Female	Social sci.	702	60	55		80.1%
41	U.S. Born, NH Asian	Female	Psychology	1,257	60	57		80.1%
42	U.S. Born, NH Asian	Female	Engineering	714	60	55		80.1%
43	NH American Indian	Male	All Fields	3,501	150	144	Yes	76.4%
44	NH American Indian	Female	All Fields	2,083	150	140	Yes	84.0%
45	NH Pacific Islander	Male	All fields	703	60	55	Yes	84.8%
46	NH Pacific Islander	Female	All fields	546	60	54	Yes	80.8%
47	U.S. Born, Disabled NH White	Male	Computer/math	2,671	60	59		81.3%
48	U.S. Born, Disabled NH White	Male	Biological and agri. sci.	8,895	60	60		81.3%
49	U.S. Born, Disabled NH White	Male	Health sci.	1,027	60	57	Yes	81.3%
50	U.S. Born, Disabled NH White	Male	Physical and related sci.	7,376	60	60		81.3%
51	U.S. Born, Disabled NH White	Male	Social sci.	5,177	60	59		81.3%
52	U.S. Born, Disabled NH White	Male	Psychology	4,516	60	59		81.3%
53	U.S. Born, Disabled NH White	Male	Engineering	4,723	60	59		81.3%
54	U.S. Born, Disabled NH White	Female	Computer/math	446	60	53	Yes	81.6%
55	U.S. Born, Disabled NH White	Female	Biological and agri. sci.	2,976	60	59		81.6%
56	U.S. Born, Disabled NH White	Female	Health sci.	1,196	60	57	Yes	81.6%
57	U.S. Born, Disabled NH White	Female	Physical and related sci.	665	60	55	Yes	81.6%
58	U.S. Born, Disabled NH White	Female	Social sci.	2,975	60	59		81.6%
59	U.S. Born, Disabled NH White	Female	Psychology	3,853	60	59		81.6%
60	U.S. Born, Disabled NH White	Female	Engineering	431	60	53	Yes	81.6%
61	U.S. Born, Not Disabled NH White	Male	Chemistry	36,761	60	60		80.8%
62	U.S. Born, Not Disabled NH White	Male	Physics/astronomy	26,109	60	60		80.8%
63	U.S. Born, Not Disabled NH White	Male	Earth/ocean/atmos.	11,703	60	60		80.8%
64	U.S. Born, Not Disabled NH White	Male	Math	17,439	60	60		80.8%
65	U.S. Born, Not Disabled NH White	Male	Computer/info. sci.	7,362	60	60		80.8%
66	U.S. Born, Not Disabled NH White	Male	Agricultural sci.	14,871	60	60		80.8%
67	U.S. Born, Not Disabled NH White	Male	Medical sci.	8,139	60	60		80.8%
68	U.S. Born, Not Disabled NH White	Male	NIH bio sci.	36,724	60	60		80.8%
69	U.S. Born, Not Disabled NH White	Male	Other bio sci.	34,732	60	60		80.8%
70	U.S. Born, Not Disabled NH White	Male	Psychology	41,324	60	60		80.8%

Appendix B.2 2013 NSDR Minimum Respondent Stratum Sample Sizes with and without Finite Population Correction Adjustment and Associated Yield Rates

Stratum	Demographic Group	Gender	Field of Degree	2013 Total Population	Minimum Sample Size Derivation			
					Minimum Respondent Sample Size Unadjusted for FPC	Minimum Respondent Sample Size with FPC Adjustment	Supplemental Sample Allocated to Meet Sample Size Minimum?	2010 NSDR Weighted Yield Rate Used Sample Size Allocation
71	U.S. Born, Not Disabled NH White	Male	Economics	12,825	60	60		80.8%
72	U.S. Born, Not Disabled NH White	Male	Anthro/arch/sociology	11,140	60	60		80.8%
73	U.S. Born, Not Disabled NH White	Male	Other social sci.	17,700	60	60		80.8%
74	U.S. Born, Not Disabled NH White	Male	Electrical/electron/comm.	13,664	60	60		80.8%
75	U.S. Born, Not Disabled NH White	Male	Other engineering	40,930	60	60		80.8%
76	U.S. Born, Not Disabled NH White	Female	Chemistry	9,794	60	60		81.9%
77	U.S. Born, Not Disabled NH White	Female	Physics/astronomy	2,746	60	59		81.9%
78	U.S. Born, Not Disabled NH White	Female	Earth/ocean/atmos.	3,570	60	59		81.9%
79	U.S. Born, Not Disabled NH White	Female	Math	3,624	60	59		81.9%
80	U.S. Born, Not Disabled NH White	Female	Computer/info. sci.	1,631	60	58	Yes	81.9%
81	U.S. Born, Not Disabled NH White	Female	Agricultural sci.	4,310	60	59		81.9%
82	U.S. Born, Not Disabled NH White	Female	Medical sci.	16,366	60	60		81.9%
83	U.S. Born, Not Disabled NH White	Female	NIH bio sci.	21,048	60	60		81.9%
84	U.S. Born, Not Disabled NH White	Female	Other bio sci.	21,928	60	60		81.9%
85	U.S. Born, Not Disabled NH White	Female	Psychology	49,293	60	60		81.9%
86	U.S. Born, Not Disabled NH White	Female	Economics	3,185	60	59		81.9%
87	U.S. Born, Not Disabled NH White	Female	Anthro/arch/sociology	11,328	60	60		81.9%
88	U.S. Born, Not Disabled NH White	Female	Other social sci.	10,571	60	60		81.9%
89	U.S. Born, Not Disabled NH White	Female	Electrical/electron/comm.	963	60	56	Yes	81.9%
90	U.S. Born, Not Disabled NH White	Female	Other engineering	6,170	60	59		81.9%
91	Non-U.S. born, NH White	Male	Chemistry	3,607	60	59		66.6%
92	Non-U.S. born, NH White	Male	Physics/astronomy	4,814	60	59		66.6%
93	Non-U.S. born, NH White	Male	Earth/ocean/atmos.	1,317	60	57	Yes	66.6%
94	Non-U.S. born, NH White	Male	Math	3,666	60	59		66.6%
95	Non-U.S. born, NH White	Male	Computer/info. sci.	2,647	60	59		66.6%
96	Non-U.S. born, NH White	Male	Agricultural sci.	1,319	60	57	Yes	66.6%
97	Non-U.S. born, NH White	Male	Medical sci.	1,056	60	57	Yes	66.6%
98	Non-U.S. born, NH White	Male	NIH bio sci.	3,807	60	59		66.6%
99	Non-U.S. born, NH White	Male	Other bio sci.	3,031	60	59		66.6%
100	Non-U.S. born, NH White	Male	Psychology	2,396	60	59		66.6%
101	Non-U.S. born, NH White	Male	Economics	2,420	60	59		66.6%
102	Non-U.S. born, NH White	Male	Anthro/arch/sociology	875	60	56	Yes	66.6%
103	Non-U.S. born, NH White	Male	Other social sci.	2,141	60	58		66.6%
104	Non-U.S. born, NH White	Male	Electrical/electron/comm.	5,574	60	59		66.6%
105	Non-U.S. born, NH White	Male	Other engineering	10,804	60	60		66.6%
106	Non-U.S. born, NH White	Female	Chemistry	1,508	60	58	Yes	67.4%

Appendix B.2 2013 NSDR Minimum Respondent Stratum Sample Sizes with and without Finite Population Correction Adjustment and Associated Yield Rates

Stratum	Demographic Group	Gender	Field of Degree	2013 Total Population	Minimum Sample Size Derivation			
					Minimum Respondent Sample Size Unadjusted for FPC	Minimum Respondent Sample Size with FPC Adjustment	Supplemental Sample Allocated to Meet Sample Size Minimum?	2010 NSDR Weighted Yield Rate Used Sample Size Allocation
107	Non-U.S. born, NH White	Female	Physics/astronomy	812	60	56	Yes	67.4%
108	Non-U.S. born, NH White	Female	Earth/ocean/atmos.	437	60	53	Yes	67.4%
109	Non-U.S. born, NH White	Female	Math	1,138	60	57	Yes	67.4%
110	Non-U.S. born, NH White	Female	Computer/info. sci.	645	60	55	Yes	67.4%
111	Non-U.S. born, NH White	Female	Agricultural sci.	530	60	54	Yes	67.4%
112	Non-U.S. born, NH White	Female	Medical sci.	1,400	60	58	Yes	67.4%
113	Non-U.S. born, NH White	Female	NIH bio sci.	2,824	60	59		67.4%
114	Non-U.S. born, NH White	Female	Other bio sci.	2,479	60	59		67.4%
115	Non-U.S. born, NH White	Female	Psychology	3,731	60	59		67.4%
116	Non-U.S. born, NH White	Female	Economics	967	60	56	Yes	67.4%
117	Non-U.S. born, NH White	Female	Anthro/arch/sociology	948	60	56	Yes	67.4%
118	Non-U.S. born, NH White	Female	Other social sci.	1,593	60	58	Yes	67.4%
119	Non-U.S. born, NH White	Female	Electrical/electron/comm.	738	60	55	Yes	67.4%
120	Non-U.S. born, NH White	Female	Other engineering	1,693	60	58	Yes	67.4%
121	Non-U.S. born, NH Asian	Male	Chemistry	10,870	60	60		68.9%
122	Non-U.S. born, NH Asian	Male	Physics/astronomy	8,616	60	60		68.9%
123	Non-U.S. born, NH Asian	Male	Earth/ocean/atmos.	1,968	60	58	Yes	68.9%
124	Non-U.S. born, NH Asian	Male	Math	6,309	60	59		68.9%
125	Non-U.S. born, NH Asian	Male	Computer/info. sci.	6,324	60	59		68.9%
126	Non-U.S. born, NH Asian	Male	Agricultural sci.	2,657	60	59		68.9%
127	Non-U.S. born, NH Asian	Male	Medical sci.	2,552	60	59		68.9%
128	Non-U.S. born, NH Asian	Male	NIH bio sci.	8,498	60	60		68.9%
129	Non-U.S. born, NH Asian	Male	Other bio sci.	8,037	60	60		68.9%
130	Non-U.S. born, NH Asian	Male	Psychology	783	60	56	Yes	68.9%
131	Non-U.S. born, NH Asian	Male	Economics	2,836	60	59		68.9%
132	Non-U.S. born, NH Asian	Male	Anthro/arch/sociology	634	60	55	Yes	68.9%
133	Non-U.S. born, NH Asian	Male	Other social sci.	1,722	60	58	Yes	68.9%
134	Non-U.S. born, NH Asian	Male	Electrical/electron/comm.	15,735	60	60		68.9%
135	Non-U.S. born, NH Asian	Male	Other engineering	31,242	60	60		68.9%
136	Non-U.S. born, NH Asian	Female	Chemistry	4,812	60	59		68.1%
137	Non-U.S. born, NH Asian	Female	Physics/astronomy	1,661	60	58	Yes	68.1%
138	Non-U.S. born, NH Asian	Female	Earth/ocean/atmos.	679	60	55	Yes	68.1%
139	Non-U.S. born, NH Asian	Female	Math	2,482	60	59		68.1%
140	Non-U.S. born, NH Asian	Female	Computer/info. sci.	1,679	60	58	Yes	68.1%
141	Non-U.S. born, NH Asian	Female	Agricultural sci.	1,478	60	58	Yes	68.1%
142	Non-U.S. born, NH Asian	Female	Medical sci.	2,589	60	59		68.1%

Appendix B.2 2013 NSDR Minimum Respondent Stratum Sample Sizes with and without Finite Population Correction Adjustment and Associated Yield Rates

Stratum	Demographic Group	Gender	Field of Degree	Minimum Sample Size Derivation				
				2013 Total Population	Minimum Respondent Sample Size Unadjusted for FPC	Minimum Respondent Sample Size with FPC Adjustment	Supplemental Sample Allocated to Meet Sample Size Minimum?	2010 NSDR Weighted Yield Rate Used Sample Size Allocation
143	Non-U.S. born, NH Asian	Female	NIH bio sci.	6,535	60	59		68.1%
144	Non-U.S. born, NH Asian	Female	Other bio sci.	6,986	60	59		68.1%
145	Non-U.S. born, NH Asian	Female	Psychology	2,410	60	59		68.1%
146	Non-U.S. born, NH Asian	Female	Economics	1,524	60	58	Yes	68.1%
147	Non-U.S. born, NH Asian	Female	Anthro/arch/sociology	852	60	56	Yes	68.1%
148	Non-U.S. born, NH Asian	Female	Other social sci.	1,377	60	57	Yes	68.1%
149	Non-U.S. born, NH Asian	Female	Electrical/electron/comm.	2,559	60	59		68.1%
150	Non-U.S. born, NH Asian	Female	Other engineering	5,224	60	59		68.1%
Total				**845,574**				**77.4%**

NH = Non-Hispanic.

Appendix B.3 2013 NSDR Final Sample Allocation

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Cases	2010 SED New Cohort Cases	2011 SED New Cohort Cases	Total Allocation
1	Hispanic	Male	Computer/math	94	7	9	110
2	Hispanic	Male	Biological and agri. sci.	325	21	21	367
3	Hispanic	Male	Health sci.	56			65
4	Hispanic	Male	Physical and related sci.	229	11	12	252
5	Hispanic	Male	Social sci.	176	12	13	201
6	Hispanic	Male	Psychology	144	6	7	157
7	Hispanic	Male	Engineering	255	19	20	294
8	Hispanic	Female	Computer/math	55			64
9	Hispanic	Female	Biological and agri. sci.	281	28	31	340
10	Hispanic	Female	Health sci.	75	6	6	87
11	Hispanic	Female	Physical and related sci.	87	8	8	103
12	Hispanic	Female	Social sci.	170	13	18	201
13	Hispanic	Female	Psychology	342	18	25	385
14	Hispanic	Female	Engineering	68	10	9	87
15	NH Black	Male	Computer/math	78	6	5	89
16	NH Black	Male	Biological and agri. sci.	264	13	14	291
17	NH Black	Male	Health sci.	72			79
18	NH Black	Male	Physical and related sci.	169	8	9	186
19	NH Black	Male	Social sci.	280	10	11	301
20	NH Black	Male	Psychology	128			136
21	NH Black	Male	Engineering	219	14	12	245
22	NH Black	Female	Computer/math	61			70
23	NH Black	Female	Biological and agri. sci.	234	20	21	275
24	NH Black	Female	Health sci.	141	13	13	167
25	NH Black	Female	Physical and related sci.	64	6	6	76
26	NH Black	Female	Social sci.	232	14	16	262
27	NH Black	Female	Psychology	294	19	18	331
28	NH Black	Female	Engineering	62	7	7	76
29	U.S. Born, NH Asian	Male	Computer/math	67	5	5	77
30	U.S. Born, NH Asian	Male	Biological and agri. sci.	238	13	15	266
31	U.S. Born, NH Asian	Male	Health sci.	49	5	5	59
32	U.S. Born, NH Asian	Male	Physical and related sci.	132	6	6	144
33	U.S. Born, NH Asian	Male	Social sci.	64			72
34	U.S. Born, NH Asian	Male	Psychology	62			68
35	U.S. Born, NH Asian	Male	Engineering	185	15	16	216
36	U.S. Born, NH Asian	Female	Computer/math	48	5	5	58
37	U.S. Born, NH Asian	Female	Biological and agri. sci.	240	24	21	285
38	U.S. Born, NH Asian	Female	Health sci.	56	5	5	66
39	U.S. Born, NH Asian	Female	Physical and related sci.	68	5	6	79
40	U.S. Born, NH Asian	Female	Social sci.	70	7	6	83
41	U.S. Born, NH Asian	Female	Psychology	129	10	10	149

Appendix B.3 2013 NSDR Final Sample Allocation

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Cases	2010 SED New Cohort Cases	2011 SED New Cohort Cases	Total Allocation
42	U.S. Born, NH Asian	Female	Engineering	68	9	7	84
43	NH American Indian	Male	All Fields	170	5	5	180
44	NH American Indian	Female	All Fields	149	9	9	167
45	NH Pacific Islander	Male	All fields				65
46	NH Pacific Islander	Female	All fields				67
47	U.S. Born, Disabled NH White	Male	Computer/math				108
48	U.S. Born, Disabled NH White	Male	Biological and agri. sci.				360
49	U.S. Born, Disabled NH White	Male	Health sci.				55
50	U.S. Born, Disabled NH White	Male	Physical and related sci.				298
51	U.S. Born, Disabled NH White	Male	Social sci.				209
52	U.S. Born, Disabled NH White	Male	Psychology				182
53	U.S. Born, Disabled NH White	Male	Engineering				191
54	U.S. Born, Disabled NH White	Female	Computer/math				34
55	U.S. Born, Disabled NH White	Female	Biological and agri. sci.				119
56	U.S. Born, Disabled NH White	Female	Health sci.				71
57	U.S. Born, Disabled NH White	Female	Physical and related sci.				49
58	U.S. Born, Disabled NH White	Female	Social sci.				120
59	U.S. Born, Disabled NH White	Female	Psychology				156
60	U.S. Born, Disabled NH White	Female	Engineering				31
61	U.S. Born, Not Disabled NH White	Male	Chemistry	1,437	24	23	1,484
62	U.S. Born, Not Disabled NH White	Male	Physics/astronomy	1,010	21	25	1,056
63	U.S. Born, Not Disabled NH White	Male	Earth/ocean/atmos.	453	9	10	472
64	U.S. Born, Not Disabled NH White	Male	Math	669	17	18	704
65	U.S. Born, Not Disabled NH White	Male	Computer/info. sci.	267	15	16	298
66	U.S. Born, Not Disabled NH White	Male	Agricultural sci.	581	9	10	600
67	U.S. Born, Not Disabled NH White	Male	Medical sci.	309	10	10	329
68	U.S. Born, Not Disabled NH White	Male	NIH bio sci.	1,411	37	36	1,484
69	U.S. Born, Not Disabled NH White	Male	Other bio sci.	1,335	35	33	1,403
70	U.S. Born, Not Disabled NH White	Male	Psychology	1,617	27	25	1,669
71	U.S. Born, Not Disabled NH White	Male	Economics	502	8	8	518
72	U.S. Born, Not Disabled NH White	Male	Anthro/arch/sociology	428	10	11	449
73	U.S. Born, Not Disabled NH White	Male	Other social sci.	677	21	18	716
74	U.S. Born, Not Disabled NH White	Male	Electrical/electron/comm.	525	13	14	552
75	U.S. Born, Not Disabled NH White	Male	Other engineering	1,563	45	45	1,653
76	U.S. Born, Not Disabled NH White	Female	Chemistry	382	13	15	410
77	U.S. Born, Not Disabled NH White	Female	Physics/astronomy	104	6	5	115
78	U.S. Born, Not Disabled NH White	Female	Earth/ocean/atmos.	135	7	7	149
79	U.S. Born, Not Disabled NH White	Female	Math	140	7	6	153
80	U.S. Born, Not Disabled NH White	Female	Computer/info. sci.	64			70
81	U.S. Born, Not Disabled NH White	Female	Agricultural sci.	163	8	9	180

Appendix B.3 2013 NSDR Final Sample Allocation

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Cases	2010 SED New Cohort Cases	2011 SED New Cohort Cases	Total Allocation
82	U.S. Born, Not Disabled NH White	Female	Medical sci.	623	32	31	686
83	U.S. Born, Not Disabled NH White	Female	NIH bio sci.	802	41	40	883
84	U.S. Born, Not Disabled NH White	Female	Other bio sci.	842	37	39	918
85	U.S. Born, Not Disabled NH White	Female	Psychology	1,939	61	65	2,065
86	U.S. Born, Not Disabled NH White	Female	Economics	127			133
87	U.S. Born, Not Disabled NH White	Female	Anthro/arch/sociology	440	17	18	475
88	U.S. Born, Not Disabled NH White	Female	Other social sci.	408	17	18	443
89	U.S. Born, Not Disabled NH White	Female	Electrical/electron/comm.	64			69
90	U.S. Born, Not Disabled NH White	Female	Other engineering	227	15	16	258
91	Non-U.S. born, NH White	Male	Chemistry	138	6	5	149
92	Non-U.S. born, NH White	Male	Physics/astronomy	181	9	10	200
93	Non-U.S. born, NH White	Male	Earth/ocean/atmos.	65			73
94	Non-U.S. born, NH White	Male	Math	138	7	7	152
95	Non-U.S. born, NH White	Male	Computer/info. sci.	92	8	10	110
96	Non-U.S. born, NH White	Male	Agricultural sci.	66			72
97	Non-U.S. born, NH White	Male	Medical sci.	68	7	5	80
98	Non-U.S. born, NH White	Male	NIH bio sci.	140	8	9	157
99	Non-U.S. born, NH White	Male	Other bio sci.	111	8	8	127
100	Non-U.S. born, NH White	Male	Psychology	89	5	5	99
101	Non-U.S. born, NH White	Male	Economics	90	5	5	100
102	Non-U.S. born, NH White	Male	Anthro/arch/sociology	58			67
103	Non-U.S. born, NH White	Male	Other social sci.	78	5	6	89
104	Non-U.S. born, NH White	Male	Electrical/electron/comm.	207	11	13	231
105	Non-U.S. born, NH White	Male	Other engineering	405	19	23	447
106	Non-U.S. born, NH White	Female	Chemistry	75	5	5	85
107	Non-U.S. born, NH White	Female	Physics/astronomy	72	7	5	84
108	Non-U.S. born, NH White	Female	Earth/ocean/atmos.	64	7	6	77
109	Non-U.S. born, NH White	Female	Math	75	5	5	85
110	Non-U.S. born, NH White	Female	Computer/info. sci.	66	9	7	82
111	Non-U.S. born, NH White	Female	Agricultural sci.	68			78
112	Non-U.S. born, NH White	Female	Medical sci.	69	8	9	86
113	Non-U.S. born, NH White	Female	NIH bio sci.	102	9	9	120
114	Non-U.S. born, NH White	Female	Other bio sci.	89	8	8	105
115	Non-U.S. born, NH White	Female	Psychology	134	12	14	160
116	Non-U.S. born, NH White	Female	Economics	72	5	7	84
117	Non-U.S. born, NH White	Female	Anthro/arch/sociology	72	5	7	84
118	Non-U.S. born, NH White	Female	Other social sci.	73	7	6	86
119	Non-U.S. born, NH White	Female	Electrical/electron/comm.	67	8	7	82
120	Non-U.S. born, NH White	Female	Other engineering	68	8	10	86
121	Non-U.S. born, NH Asian	Male	Chemistry	405	17	17	439
122	Non-U.S. born, NH Asian	Male	Physics/astronomy	314	16	18	348

Appendix B.3 2013 NSDR Final Sample Allocation

Stratum	Demographic Group	Gender	Field of Degree	Old Cohort Cases	2010 SED New Cohort Cases	2011 SED New Cohort Cases	Total Allocation
123	Non-U.S. born, NH Asian	Male	Earth/ocean/atmos.	77			84
124	Non-U.S. born, NH Asian	Male	Math	230	12	13	255
125	Non-U.S. born, NH Asian	Male	Computer/info. sci.	219	17	19	255
126	Non-U.S. born, NH Asian	Male	Agricultural sci.	98			107
127	Non-U.S. born, NH Asian	Male	Medical sci.	91	6	6	103
128	Non-U.S. born, NH Asian	Male	NIH bio sci.	308	18	18	344
129	Non-U.S. born, NH Asian	Male	Other bio sci.	288	18	19	325
130	Non-U.S. born, NH Asian	Male	Psychology	69	6	5	80
131	Non-U.S. born, NH Asian	Male	Economics	106			115
132	Non-U.S. born, NH Asian	Male	Anthro/arch/sociology	61			69
133	Non-U.S. born, NH Asian	Male	Other social sci.	74			82
134	Non-U.S. born, NH Asian	Male	Electrical/electron/comm.	567	32	37	636
135	Non-U.S. born, NH Asian	Male	Other engineering	1,146	56	59	1,261
136	Non-U.S. born, NH Asian	Female	Chemistry	193	11	13	217
137	Non-U.S. born, NH Asian	Female	Physics/astronomy	73	6	6	85
138	Non-U.S. born, NH Asian	Female	Earth/ocean/atmos.	67	7	7	81
139	Non-U.S. born, NH Asian	Female	Math	94	9	9	112
140	Non-U.S. born, NH Asian	Female	Computer/info. sci.	70	7	8	85
141	Non-U.S. born, NH Asian	Female	Agricultural sci.	74			84
142	Non-U.S. born, NH Asian	Female	Medical sci.	98	10	10	118
143	Non-U.S. born, NH Asian	Female	NIH bio sci.	252	20	22	294
144	Non-U.S. born, NH Asian	Female	Other bio sci.	270	23	22	315
145	Non-U.S. born, NH Asian	Female	Psychology	92	8	8	108
146	Non-U.S. born, NH Asian	Female	Economics	71	6	8	85
147	Non-U.S. born, NH Asian	Female	Anthro/arch/sociology	70	6	6	82
148	Non-U.S. born, NH Asian	Female	Other social sci.	71	7	7	85
149	Non-U.S. born, NH Asian	Female	Electrical/electron/comm.	95	10	10	115
150	Non-U.S. born, NH Asian	Female	Other engineering	195	19	21	235
Total				**36,661**	**1,635**	**1,704**	**40,000**

NH = Non-Hispanic.

NOTE: Grayed out cells have been suppressed for confidentiality reasons.

Appendix C.1 2013 ISDR Strata with Frame Population Counts and Sample Cases

Stratum Number	Stratification Schema			Frame Population Size				Sampled Cases			
	Demographic Group	Sex	Field of Degree	Total Frame	20th Century Frame (estimate)	21st Century Frame (estimate)	New Cohort from 2010/2011 (actual)	Total Sample Size	20th Century Old Cohort Cases	21st Century Old Cohort Cases	New Cohort from 2010/2011 Cases
D1	U.S.Born; all race/ethnicities	Male	Computer & info. sciences/math/physical sciences/engineering	5,986	3,512	1,921	553	360	164	133	63
D2	U.S.Born; all race/ethnicities	Male	Biological, agricultural, or health sciences	3,306	2,056	1,059	191	185	96	67	22
D3	U.S.Born; all race/ethnicities	Male	Psychology or social sciences	4,516	3,330	987	199	240	154	64	22
D4	U.S.Born; all race/ethnicities	Female	Computer & info. sciences/math/physical sciences/engineering	1,098	524	410	164	93	29	45	19
D5	U.S.Born; all race/ethnicities	Female	Biological, agricultural, or health sciences	2,061	965	865	231	143	53	64	26
D6	U.S.Born; all race/ethnicities	Female	Psychology or social sciences	2,636	1,723	697	216	157	78	55	24
D7	Non-U.S. Born; Hispanic, any race	Male	Computer & info. sciences/math/physical sciences/engineering	1,804	363	1,203	238	202	24	151	27
D8	Non-U.S. Born; Hispanic, any race	Male	Biological, agricultural, or health sciences	1,260	332	807	120	140	20	106	14
D9	Non-U.S. Born; Hispanic, any race	Male	Psychology or social sciences	923	108	686	129	113	8	91	14
D10	Non-U.S. Born; Hispanic, any race	Female	Computer & info. sciences/math/physical sciences/engineering	336	8	277	51	49			6
D11	Non-U.S. Born; Hispanic, any race	Female	Biological, agricultural, or health sciences	655	82	500	73	85	7	70	8
D12	Non-U.S. Born; Hispanic, any race	Female	Psychology or social sciences	518	33	409	76	68			9
C7	Non-U.S. Born; NH-Black	All	Computer & info. sciences/math/physical sciences/engineering	424	185	188	51	44	13	26	5
C8	Non-U.S. Born; NH-Black	All	Biological, agricultural, or health sciences	729	230	424	75	79	14	56	9
C9	Non-U.S. Born; NH-Black	All	Psychology or social sciences	820	302	454	64	83	18	57	8
F43	Non-U.S. Born; NH-Asian	Male	Computer/information sciences or mathematics	2,809	818	1,612	380	237	35	159	43
F44	Non-U.S. Born; NH-Asian	Male	Biological/agricultural/environmental life sciences	3,834	1,854	1,743	237	264	82	156	26
F45	Non-U.S. Born; NH-Asian	Male	Health sciences	615	205	355	54	61	10	44	7
F46	Non-U.S. Born; NH-Asian	Male	Physical sciences	4,201	1,704	2,141	356	301	76	185	40
F47	Non-U.S. Born; NH-Asian	Male	Social sciences	4,251	1,464	2,350	436	394	71	274	49
F48	Non-U.S. Born; NH-Asian	Male	Psychology	259	51	174	35	31			
F49	Non-U.S. Born; NH-Asian	Male	Engineering	11,508	4,439	6,096	973	894	196	588	110

Appendix C.1 2013 ISDR Strata with Frame Population Counts and Sample Cases

Stratum Number	Stratification Schema			Frame Population Size				Sampled Cases			
	Demographic Group	Sex	Field of Degree	Total Frame	20th Century Frame (estimate)	21st Century Frame (estimate)	New Cohort from 2010/2011 (actual)	Total Sample Size	20th Century Old Cohort Cases	21st Century Old Cohort Cases	New Cohort from 2010/2011 Cases
F50	Non-U.S. Born; NH-Asian	Female	Computer/information sciences or mathematics	669	111	467	91	70	7	53	10
F51	Non-U.S. Born; NH-Asian	Female	Biological/agricultural/environmental life sciences	2,461	721	1,439	302	215	38	142	35
F52	Non-U.S. Born; NH-Asian	Female	Health sciences	823	88	624	111	94	5	77	12
F53	Non-U.S. Born; NH-Asian	Female	Physical sciences	1,219	337	736	146	118	20	81	17
F54	Non-U.S. Born; NH-Asian	Female	Social sciences	1,819	362	1,203	254	206	30	148	28
F55	Non-U.S. Born; NH-Asian	Female	Psychology	483	101	298	84	46	6	31	9
F56	Non-U.S. Born; NH-Asian	Female	Engineering	1,554	280	1,078	196	140	15	102	23
F57	Non-U.S. Born; NH-White	Male	Computer/information sciences or mathematics	2,296	764	1,319	213	207	36	147	24
F58	Non-U.S. Born; NH-White	Male	Biological/agricultural/environmental life sciences	2,303	1,101	1,064	138	190	53	122	15
F59	Non-U.S. Born; NH-White	Male	Health sciences	421	184	216	21	41	11		
F60	Non-U.S. Born; NH-White	Male	Physical sciences	2,799	1,079	1,528	192	230	52	156	22
F61	Non-U.S. Born; NH-White	Male	Social sciences	3,174	1,122	1,747	305	269	50	184	35
F62	Non-U.S. Born; NH-White	Male	Psychology	416	176	219	21	32	8		
F63	Non-U.S. Born; NH-White	Male	Engineering	4,815	2,185	2,335	295	373	99	241	33
F64	Non-U.S. Born; NH-White	Female	Computer/information sciences or mathematics	493	82	342	69	64	10	46	8
F65	Non-U.S. Born; NH-White	Female	Biological/agricultural/environmental life sciences	1,228	278	822	128	121	15	92	14
F66	Non-U.S. Born; NH-White	Female	Health sciences	327	12	273	42	40			5
F67	Non-U.S. Born; NH-White	Female	Physical sciences	692	125	490	77	85	10	66	9
F68	Non-U.S. Born; NH-White	Female	Social sciences	1,518	322	988	208	170	24	123	23
F69	Non-U.S. Born; NH-White	Female	Psychology	701	309	331	61	52	14	31	7
F70	Non-U.S. Born; NH-White	Female	Engineering	743	210	449	84	78	15	53	10
A6	Non-U.S. Born; NH-Other Races	All	All	132	24	97	11	14			
Overall				85,635	34,262	43,422	7,951	7,078	1,678	4,500	900

NOTES: Detailed cases counts for the sampled cases by cohort are suppressed for confidentiality reasons. Specific grayed out cells have been suppressed for confidentiality reasons

Appendix D. Detailed NSDR Allocation Algorithm and Final 2013 NDR allocation

The NSDR Allocation Algorithm

- Let h = 1 to H denote the NSDR strata where H = 150.

- Let N(h) denote the stratum h population size and N(+) the total population size across all strata.

- Let OLDN(h) denote the stratum h population size for old cohorts.

- Let NEWN(h) denote the stratum h population size for new cohorts.

- Let OLDCOUNT(h) denote the total old cohort cases in the stratum h frame.

- Let NEWCOUNT(H) denote the total new cohort cases in the stratum h frame.

- Let d = 1 to D denote the NSDR domains that receive a domain-level sample supplement.

- Let DOMSAM(hd) denote the fixed domain-level allocation made to each stratum in domain d and DOMSAM(+d) be the fixed sample size allocated across all strata in domain d and DOMSAM(++) denote the total domain-level sample size allocated across all strata and all domains.

- Let DOMN(hd) denote the population size of stratum h in domain d and DOMN(+d) denote the total population size of domain d.

- Let PROPSAM(+i) be the total sample size set to be allocated proportionately to all strata in Interation i where I = 1 to I and let PROPSAM(+i) be the proportional sample allocated to stratum h.

- Let STSPSAM(hi) be the stratum-level sample size supplement allocated to stratum h in Iteration i.

- Let DESSAM(hi) be the desired total sample to be allocated to stratum h in Iteration i.

- Let OLDDESSAM(hi) be the desired total sample to be allocated to the old cohort substratum of stratum h in Iteration i.

- Let NEWDESSAM(hi) be the desired total sample to be allocated to the new cohort substratum of stratum h in Iteration i.

- Let OLDACTSAM(hi) be the actual total sample that can be allocated to the old cohort substratum stratum h given the number of old and new cohort frame members in Iteration i.

- Let NEWACTSAM(hi) be the actual total sample that can be allocated to the new cohort substratum stratum h given the number of old and new cohort frame members in Iteration i.

- Let TOTACTSAM(hi) be the actual total sample allocated in Iteration i to old and new cohorts and TOTACTSAM(+i) be the total actual sample allocated in Iteration i across all strata.

- Let MINSAM(h) be the minimum sample size to be allocated to stratum h before the finite population adjustment.

- Let ADJMINSAM(h) be the minimum sample size to be allocated to stratum h after the finite population adjustment.

ITERATION 0

- Set the values for the fixed domain-level allocation as:

$$DOMSAM(hd) = DOMSAM(+d) * \frac{DOMN(hd)}{DOMN(+d)}$$

- Define the minimum number of attempted interviews to be allocated to each stratum as:

$$DJMINSAM(h) = \frac{MINSAM(h)*N(h)}{MINSAM(h)+N(h)}$$. (See *Section 5* for details.)

- Note that the starting sample size for Iteration i is determined at the end of Iteration i-1. The exception is for Iteration 1 where the starting value for the sample size to be allocated proportionately across strata is set to PROPSAM(1)=40,000-DOMSAM(++).

- Each iteration from i = 1 to I then follows the steps for "Iteration i" below.

ITERATION i

- Define the proportional sample to be allocated to stratum h in Iteration i as:

$$PROPSAM(hi) = PROPSAM(+i) * \frac{N(h)}{N(+)}.$$

- Define the stratum-level supplement to be allocated to stratum h as

$$STSPSAM(hi) = ADJMINSAM(h) - PROPSAM(hi) - \text{DOMSAM(hd)}.$$

- Calculate the desired sample to be allocated to stratum h in domain d as:

$$DESSAM(hi) = \text{PROPSAM(hi)} + \text{DOMSAM(hd)} + \text{STSPSAM(hi)}.$$

- Allocate the desired stratum h sample to the old and new cohort substrata:

$$OLDDESSAM(hi) = DESSAM(hi) * \frac{OLDN(h)}{N(h)}$$

$$NEWDESSAM(hi) = DESSAM(hi) * \frac{NEWN(h)}{N(h)}$$

- Determine the number of actual frame cases that can be allocated given the number of old cohort frame members:

    If OLDCOUNT(h)<OLDDESSAM(hi), set OLDACTSAM(hi)=OLDCOUNT(h)

    If OLDCOUNT(h)≥ OLDDESSAM(hi), set OLDACTSAM(hi)=OLDDESSAM(hi) .

- Determine the number of actual frame cases that can be allocated given the number of new cohort frame members:

    If NEWCOUNT(h)<NEWDESSAM(hi), set NEWACTSAM(hi)=NEWCOUNT(h)

    If NEWCOUNT(h)≥ NEWDESSAM(hi), set NEWACTSAM(hi)=NEWDESSAM(hi) .

- Calculate the total actual sample allocated to stratum h in Iteration i as:

$$TOTACTSAM(hi) = OLDACTSAM(hi) + NEWACTSAM(hi).$$

- Calculate the total sample allocated in Iteration i across all strata as:

$$TOTACTSAM(+i) = \sum_{h=1}^{150} TOTACTSAM(hi)$$

- If $TOTACTSAM(hi) = 40{,}000$ the allocation process is complete.

- Otherwise, determine the total sample size to be allocated proportionately in Iteration i+1 as:

$$PROPSAM(+, i+1) = PROPSAM(+i) + [OLDDESSAM(hi) - OLDACTSAM(hi)] + [NEWDESSAM(hi) - NEWACTSAM(hi)]$$

- And continue the iterations until the actual total allocated sample equals 40,000.

Appendix D. 2013 NSDR Allocation Plan Developed with Actual 2010 and Actual 2011 SED New Cohort Population Counts

Stratum	Demographic Group	Gender	Field of Degree	2013 NSDR Total Population (Actual for Both New Cohorts)	2013 NSDR Old Cohort Population	New Cohort Population (Actual for Both Cohorts)	Proportional Sample Size	Domain Added Sample	Stratum-specific Supplemental Sample	Desired Total Sample	Desired Old Cohort Allocation	Actual Old Cohort Allocation	Actual 2010 New Cohort Final Unround Allocation	Actual 2011 New Cohort Final Unround Allocation	Actual Combined New Cohort Allocation Unround	Total NSDR Unround Sampled Cases (Actual for All Cohorts)
1	Hispanic	Male	Computer/math	1,267	1,092	175	51	58	0	109	94	94	7	9	15	109
2	Hispanic	Male	Biological and agri. sci.	4,274	3,772	502	173	195	0	368	325	325	22	22	43	368
3	Hispanic	Male	Health sci.	365	314	51	15	17	34	65	56	56			9	65
4	Hispanic	Male	Physical and related sci.	2,917	2,651	266	118	133	0	251	228	228	11	12	23	251
5	Hispanic	Male	Social sci.	2,340	2,049	291	95	107	0	202	176	176	12	13	25	202
6	Hispanic	Male	Psychology	1,820	1,677	143	74	83	0	157	144	144	6	7	12	156
7	Hispanic	Male	Engineering	3,425	2,962	463	138	157	0	295	255	255	19	21	40	295
8	Hispanic	Female	Computer/math	294	253	41	12	18	34	64	55	55			9	64
9	Hispanic	Female	Biological and agri. sci.	3,300	2,730	570	133	207	0	340	281	281	28	31	59	340
10	Hispanic	Female	Health sci.	841	725	116	34	53	0	87	75	75	6	6	12	87
11	Hispanic	Female	Physical and related sci.	1,005	848	157	41	63	0	104	87	87	8	8	16	104
12	Hispanic	Female	Social sci.	1,944	1,645	299	79	122	0	200	170	170	13	18	31	200
13	Hispanic	Female	Psychology	3,738	3,317	421	151	234	0	385	342	342	18	25	43	385
14	Hispanic	Female	Engineering	845	660	185	34	53	0	87	68	68	10	9	19	87
15	NH Black	Male	Computer/math	967	853	114	39	49	0	88	78	78	5	5	10	88
16	NH Black	Male	Biological and agri. sci.	3,178	2,884	294	128	162	0	291	264	264	13	14	27	291
17	NH Black	Male	Health sci.	792	717	75	32	40	7	79	72	72			8	79
18	NH Black	Male	Physical and related sci.	2,026	1,849	177	82	104	0	185	169	169	8	8	16	185
19	NH Black	Male	Social sci.	3,294	3,069	225	133	168	0	301	281	281	10	11	21	301
20	NH Black	Male	Psychology	1,752	1,658	94	71	90	0	160	152	128			9	137
21	NH Black	Male	Engineering	2,669	2,390	279	108	136	0	244	219	219	13	12	26	244
22	NH Black	Female	Computer/math	376	323	53	15	22	35	71	61	61	5	5	10	71
23	NH Black	Female	Biological and agri. sci.	2,810	2,387	423	113	161	0	275	233	233	21	21	41	275
24	NH Black	Female	Health sci.	1,704	1,437	267	69	98	0	167	140	140	13	13	26	167
25	NH Black	Female	Physical and related sci.	742	624	118	30	43	4	76	64	64	6	6	12	76
26	NH Black	Female	Social sci.	2,677	2,369	308	108	154	0	262	232	232	14	16	30	262

Appendix D. 2013 NSDR Allocation Plan Developed with Actual 2010 and Actual 2011 SED New Cohort Population Counts

Stratum	Demographic Group	Gender	Field of Degree	2013 NSDR Total Population (Actual for Both New Cohorts)	2013 NSDR Old Cohort Population	New Cohort Population (Actual for Both Cohorts)	Proportional Sample Size	Domain Added Sample	Stratum-specific Supplemental Sample	Desired Total Sample	Desired Old Cohort Allocation	Actual Old Cohort Allocation	Actual 2010 New Cohort Final Unround Allocation	Actual 2011 New Cohort Final Unround Allocation	Actual Combined New Cohort Allocation Unround	Total NSDR Unround Sampled Cases (Actual for All Cohorts)
27	NH Black	Female	Psychology	4,057	3,687	370	164	233	0	397	360	294	18	18	36	330
28	NH Black	Female	Engineering	706	577	129	29	41	7	76	62	62	7	7	14	76
29	U.S. Born, NH Asian	Male	Computer/math	796	685	111	32	45	0	77	67	67	6	5	11	77
30	U.S. Born, NH Asian	Male	Biological and agri. sci.	2,731	2,448	283	110	155	0	266	238	238	13	14	28	266
31	U.S. Born, NH Asian	Male	Health sci.	240	197	43	10	14	37	60	49	49	5	5	11	60
32	U.S. Born, NH Asian	Male	Physical and related sci.	1,478	1,352	126	60	84	0	144	131	131	6	6	12	144
33	U.S. Born, NH Asian	Male	Social sci.	741	658	83	30	42	0	72	64	64			8	72
34	U.S. Born, NH Asian	Male	Psychology	600	546	54	24	34	10	68	62	62			6	68
35	U.S. Born, NH Asian	Male	Engineering	2,211	1,905	306	89	126	0	215	185	185	15	15	30	215
36	U.S. Born, NH Asian	Female	Computer/math	211	174	37	9	17	33	58	48	48	5	6	10	58
37	U.S. Born, NH Asian	Female	Biological and agri. sci.	2,403	2,026	377	97	188	0	285	240	240	23	21	45	285
38	U.S. Born, NH Asian	Female	Health sci.	438	367	71	18	34	14	66	55	55	6	5	11	66
39	U.S. Born, NH Asian	Female	Physical and related sci.	666	576	90	27	52	0	79	68	68	5	6	11	79
40	U.S. Born, NH Asian	Female	Social sci.	702	594	108	28	55	0	83	70	70	7	6	13	83
41	U.S. Born, NH Asian	Female	Psychology	1,257	1,086	171	51	98	0	149	129	129	10	10	20	149
42	U.S. Born, NH Asian	Female	Engineering	714	574	140	29	56	0	85	68	68	10	7	17	85
43	NH American Indian	Male	All Fields	3,501	3,314	187	141	0	47	188	178	170	5	5	10	180
44	NH American Indian	Female	All Fields	2,083	1,866	217	84	0	82	167	149	149	9	9	17	167
45	NH Pacific Islander	Male	All fields	703	632	71	28	0	37	65	59	59			7	65
46	NH Pacific Islander	Female	All fields	546	480	66	22	0	45	67	59	59			8	67
47	U.S. Born, Disabled NH White	Male	Computer/math	2,671	2,607	64	108	0	0	108	105					108
48	U.S. Born, Disabled NH White	Male	Biological and agri. sci.	8,895	8,718	177	359	0	0	359	352	352			7	359
49	U.S. Born, Disabled NH White	Male	Health sci.	1,027	995	32	41	0	28	70	68					55
50	U.S. Born, Disabled NH White	Male	Physical and related sci.	7,376	7,246	130	298	0	0	298	293	293			5	298
51	U.S. Born, Disabled NH White	Male	Social sci.	5,177	5,078	99	209	0	0	209	205					209

Appendix D. 2013 NSDR Allocation Plan Developed with Actual 2010 and Actual 2011 SED New Cohort Population Counts

Stratum	Demographic Group	Gender	Field of Degree	2013 NSDR Total Population (Actual for Both New Cohorts)	2013 NSDR Old Cohort Population	New Cohort Population (Actual for Both Cohorts)	Proportional Sample Size	Domain Added Sample	Stratum-specific Supplemental Sample	Desired Total Sample	Desired Old Cohort Allocation	Actual Old Cohort Allocation	Actual 2010 New Cohort Final Unround Allocation	Actual 2011 New Cohort Final Unround Allocation	Actual Combined New Cohort Allocation Unround	Total NSDR Unround Sampled Cases (Actual for All Cohorts)
52	U.S. Born, Disabled NH White	Male	Psychology	4,516	4,438	78	182	0	0	182	179					182
53	U.S. Born, Disabled NH White	Male	Engineering	4,723	4,593	130	191	0	0	191	186	186			5	191
54	U.S. Born, Disabled NH White	Female	Computer/math	446	423	23	18	0	47	65	61					33
55	U.S. Born, Disabled NH White	Female	Biological and agri. sci.	2,976	2,811	165	120	0	0	120	114	114			7	120
56	U.S. Born, Disabled NH White	Female	Health sci.	1,196	1,127	69	48	0	22	70	66					70
57	U.S. Born, Disabled NH White	Female	Physical and related sci.	665	616	49	27	0	41	67	62	44			5	49
58	U.S. Born, Disabled NH White	Female	Social sci.	2,975	2,897	78	120	0	0	120	117					120
59	U.S. Born, Disabled NH White	Female	Psychology	3,853	3,724	129	156	0	0	156	150	150			5	156
60	U.S. Born, Disabled NH White	Female	Engineering	431	413	18	17	0	47	65	62					32
61	U.S. Born, Not Disabled NH White	Male	Chemistry	36,761	35,581	1,180	1,485	0	0	1,485	1437	1437	24	24	48	1485
62	U.S. Born, Not Disabled NH White	Male	Physics/astronomy	26,109	24,989	1,120	1,055	0	0	1,055	1009	1009	21	24	45	1055
63	U.S. Born, Not Disabled NH White	Male	Earth/ocean/atmos.	11,703	11,217	486	473	0	0	473	453	453	10	10	20	473
64	U.S. Born, Not Disabled NH White	Male	Math	17,439	16,570	869	704	0	0	704	669	669	17	18	35	704
65	U.S. Born, Not Disabled NH White	Male	Computer/info. sci.	7,362	6,612	750	297	0	0	297	267	267	15	15	30	297
66	U.S. Born, Not Disabled NH White	Male	Agricultural sci.	14,871	14,402	469	601	0	0	601	582	582	9	10	19	601
67	U.S. Born, Not Disabled NH White	Male	Medical sci.	8,139	7,648	491	329	0	0	329	309	309	9	10	20	329
68	U.S. Born, Not Disabled NH White	Male	NIH bio sci.	36,724	34,917	1,807	1,483	0	0	1,483	1410	1410	37	36	73	1483
69	U.S. Born, Not Disabled NH White	Male	Other bio sci.	34,732	33,065	1,667	1,403	0	0	1,403	1336	1336	34	33	67	1403
70	U.S. Born, Not Disabled NH White	Male	Psychology	41,324	40,023	1,301	1,669	0	0	1,669	1617	1617	27	26	53	1669
71	U.S. Born, Not Disabled NH White	Male	Economics	12,825	12,420	405	518	0	0	518	502	502	8	8	16	518
72	U.S. Born, Not Disabled NH White	Male	Anthro/arch/sociology	11,140	10,597	543	450	0	0	450	428	428	10	12	22	450
73	U.S. Born, Not Disabled NH White	Male	Other social sci.	17,700	16,757	943	715	0	0	715	677	677	20	18	38	715
74	U.S. Born, Not Disabled NH White	Male	Electrical/electron/comm.	13,664	12,986	678	552	0	0	552	525	525	14	14	27	552
75	U.S. Born, Not Disabled NH White	Male	Other engineering	40,930	38,693	2,237	1,653	0	0	1,653	1563	1563	45	45	90	1653
76	U.S. Born, Not Disabled NH White	Female	Chemistry	9,794	9,131	663	396	15	0	410	383	383	13	14	28	410
77	U.S. Born, Not Disabled NH White	Female	Physics/astronomy	2,746	2,475	271	111	4	0	115	104	104	6	6	11	115

Appendix D. 2013 NSDR Allocation Plan Developed with Actual 2010 and Actual 2011 SED New Cohort Population Counts

Stratum	Demographic Group	Gender	Field of Degree	2013 NSDR Total Population (Actual for Both New Cohorts)	2013 NSDR Old Cohort Population	New Cohort Population (Actual for Both Cohorts)	Proportional Sample Size	Domain Added Sample	Stratum-specific Supplemental Sample	Desired Total Sample	Desired Old Cohort Allocation	Actual Old Cohort Allocation	Actual 2010 New Cohort Final Unround Allocation	Actual 2011 New Cohort Final Unround Allocation	Actual Combined New Cohort Allocation Unround	Total NSDR Unround Sampled Cases (Actual for All Cohorts)
78	U.S. Born, Not Disabled NH White	Female	Earth/ocean/atmos.	3,570	3,232	338	144	5	0	150	135	135	7	7	14	150
79	U.S. Born, Not Disabled NH White	Female	Math	3,624	3,328	296	146	5	0	152	139	139	6	6	12	152
80	U.S. Born, Not Disabled NH White	Female	Computer/info. sci.	1,631	1,478	153	66	2	2	71	64	64			7	71
81	U.S. Born, Not Disabled NH White	Female	Agricultural sci.	4,310	3,893	417	174	6	0	181	163	163	8	9	17	181
82	U.S. Born, Not Disabled NH White	Female	Medical sci.	16,366	14,870	1,496	661	25	0	686	623	623	32	31	63	686
83	U.S. Born, Not Disabled NH White	Female	NIH bio sci.	21,048	19,144	1,904	850	32	0	882	802	802	40	39	80	882
84	U.S. Born, Not Disabled NH White	Female	Other bio sci.	21,928	20,115	1,813	886	33	0	919	843	843	37	39	76	919
85	U.S. Born, Not Disabled NH White	Female	Psychology	49,293	46,280	3,013	1,991	74	0	2,065	1939	1939	61	66	126	2065
86	U.S. Born, Not Disabled NH White	Female	Economics	3,185	3,023	162	129	5	0	133	127	127			7	133
87	U.S. Born, Not Disabled NH White	Female	Anthro/arch/sociology	11,328	10,492	836	458	17	0	475	440	440	17	18	35	475
88	U.S. Born, Not Disabled NH White	Female	Other social sci.	10,571	9,750	821	427	16	0	443	408	408	17	18	34	443
89	U.S. Born, Not Disabled NH White	Female	Electrical/electron/comm.	963	893	70	39	1	29	69	64	64			5	69
90	U.S. Born, Not Disabled NH White	Female	Other engineering	6,170	5,427	743	249	9	0	258	227	227	15	16	31	258
91	Non-U.S. born, NH White	Male	Chemistry	3,607	3,330	277	146	4	0	149	138	138	6	6	11	149
92	Non-U.S. born, NH White	Male	Physics/astronomy	4,814	4,368	446	194	5	0	199	181	181	9	10	18	199
93	Non-U.S. born, NH White	Male	Earth/ocean/atmos.	1,317	1,197	120	53	1	32	86	78	65			8	73
94	Non-U.S. born, NH White	Male	Math	3,666	3,337	329	148	4	0	152	138	138	7	7	14	152
95	Non-U.S. born, NH White	Male	Computer/info. sci.	2,647	2,224	423	107	3	0	110	92	92	8	9	18	110
96	Non-U.S. born, NH White	Male	Agricultural sci.	1,319	1,220	99	53	1	32	86	80	66			6	72
97	Non-U.S. born, NH White	Male	Medical sci.	1,056	899	157	43	1	42	85	73	68	7	6	13	81
98	Non-U.S. born, NH White	Male	NIH bio sci.	3,807	3,393	414	154	4	0	158	140	140	8	9	17	158
99	Non-U.S. born, NH White	Male	Other bio sci.	3,031	2,673	358	122	3	0	125	111	111	7	7	15	125
100	Non-U.S. born, NH White	Male	Psychology	2,396	2,147	249	97	2	0	99	89	89	5	5	10	99
101	Non-U.S. born, NH White	Male	Economics	2,420	2,177	243	98	2	0	100	90	90	5	5	10	100
102	Non-U.S. born, NH White	Male	Anthro/arch/sociology	875	780	95	35	1	48	84	75	58	5	5	9	67
103	Non-U.S. born, NH White	Male	Other social sci.	2,141	1,896	245	86	2	0	89	78	78	5	5	10	89

Appendix D. 2013 NSDR Allocation Plan Developed with Actual 2010 and Actual 2011 SED New Cohort Population Counts

Stratum	Demographic Group	Gender	Field of Degree	2013 NSDR Total Population (Actual for Both New Cohorts)	2013 NSDR Old Cohort Population	New Cohort Population (Actual for Both Cohorts)	Proportional Sample Size	Domain Added Sample	Stratum-specific Supplemental Sample	Desired Total Sample	Desired Old Cohort Allocation	Actual Old Cohort Allocation	Actual 2010 New Cohort Final Unround Allocation	Actual 2011 New Cohort Final Unround Allocation	Actual Combined New Cohort Allocation Unround	Total NSDR Unround Sampled Cases (Actual for All Cohorts)
104	Non-U.S. born, NH White	Male	Electrical/electron/comm.	5,574	4,983	591	225	6	0	231	206	206	11	13	24	231
105	Non-U.S. born, NH White	Male	Other engineering	10,804	9,777	1,027	436	11	0	447	405	405	20	23	43	447
106	Non-U.S. born, NH White	Female	Chemistry	1,508	1,325	183	61	4	21	86	75	75	5	5	10	86
107	Non-U.S. born, NH White	Female	Physics/astronomy	812	706	106	33	2	48	83	72	72	6	5	11	83
108	Non-U.S. born, NH White	Female	Earth/ocean/atmos.	437	357	80	18	1	60	78	64	64	8	7	14	78
109	Non-U.S. born, NH White	Female	Math	1,138	1,007	131	46	3	36	85	75	75	5	5	10	85
110	Non-U.S. born, NH White	Female	Computer/info. sci.	645	521	124	26	2	54	81	66	66	9	7	16	81
111	Non-U.S. born, NH White	Female	Agricultural sci.	530	461	69	21	1	57	80	70	68			10	78
112	Non-U.S. born, NH White	Female	Medical sci.	1,400	1,123	277	57	3	26	85	68	68	9	8	17	85
113	Non-U.S. born, NH White	Female	NIH bio sci.	2,824	2,388	436	114	7	0	121	102	102	9	10	19	121
114	Non-U.S. born, NH White	Female	Other bio sci.	2,479	2,096	383	100	6	0	106	90	90	8	8	16	106
115	Non-U.S. born, NH White	Female	Psychology	3,731	3,117	614	151	9	0	159	133	133	12	14	26	159
116	Non-U.S. born, NH White	Female	Economics	967	838	129	39	2	42	84	73	73	5	6	11	84
117	Non-U.S. born, NH White	Female	Anthro/arch/sociology	948	814	134	38	2	43	84	72	72	5	7	12	84
118	Non-U.S. born, NH White	Female	Other social sci.	1,593	1,358	235	64	4	18	86	73	73	6	6	13	86
119	Non-U.S. born, NH White	Female	Electrical/electron/comm.	738	608	130	30	2	51	82	68	68	8	7	15	82
120	Non-U.S. born, NH White	Female	Other engineering	1,693	1,338	355	68	4	14	86	68	68	8	10	18	86
121	Non-U.S. born, NH Asian	Male	Chemistry	10,870	10,027	843	439	0	0	439	405	405	17	17	34	439
122	Non-U.S. born, NH Asian	Male	Physics/astronomy	8,616	7,776	840	348	0	0	348	314	314	16	18	34	348
123	Non-U.S. born, NH Asian	Male	Earth/ocean/atmos.	1,968	1,799	169	79	0	5	85	77	77			7	85
124	Non-U.S. born, NH Asian	Male	Math	6,309	5,681	628	255	0	0	255	229	229	12	13	25	255
125	Non-U.S. born, NH Asian	Male	Computer/info. sci.	6,324	5,426	898	255	0	0	255	219	219	18	19	36	255
126	Non-U.S. born, NH Asian	Male	Agricultural sci.	2,657	2,433	224	107	0	0	107	98	98			9	107
127	Non-U.S. born, NH Asian	Male	Medical sci.	2,552	2,250	302	103	0	0	103	91	91	6	6	12	103
128	Non-U.S. born, NH Asian	Male	NIH bio sci.	8,498	7,624	874	343	0	0	343	308	308	17	18	35	343
129	Non-U.S. born, NH Asian	Male	Other bio sci.	8,037	7,138	899	325	0	0	325	288	288	17	19	36	325

Appendix D. 2013 NSDR Allocation Plan Developed with Actual 2010 and Actual 2011 SED New Cohort Population Counts

Stratum	Demographic Group	Gender	Field of Degree	2013 NSDR Total Population (Actual for Both New Cohorts)	2013 NSDR Old Cohort Population	New Cohort Population (Actual for Both Cohorts)	Proportional Sample Size	Domain Added Sample	Stratum-specific Supplemental Sample	Desired Total Sample	Desired Old Cohort Allocation	Actual Old Cohort Allocation	Actual 2010 New Cohort Final Unround Allocation	Actual 2011 New Cohort Final Unround Allocation	Actual Combined New Cohort Allocation Unround	Total NSDR Unround Sampled Cases (Actual for All Cohorts)
130	Non-U.S. born, NH Asian	Male	Psychology	783	665	118	32	0	49	81	69	69	6	6	12	81
131	Non-U.S. born, NH Asian	Male	Economics	2,836	2,616	220	115	0	0	115	106	106			9	115
132	Non-U.S. born, NH Asian	Male	Anthro/arch/sociology	634	573	61	26	0	54	80	72	61			8	69
133	Non-U.S. born, NH Asian	Male	Other social sci.	1,722	1,556	166	70	0	15	84	76	74			8	82
134	Non-U.S. born, NH Asian	Male	Electrical/electron/comm.	15,735	14,027	1,708	636	0	0	636	567	567	32	37	69	636
135	Non-U.S. born, NH Asian	Male	Other engineering	31,242	28,377	2,865	1,262	0	0	1,262	1146	1146	56	59	116	1262
136	Non-U.S. born, NH Asian	Female	Chemistry	4,812	4,291	521	194	22	0	217	193	193	11	12	23	217
137	Non-U.S. born, NH Asian	Female	Physics/astronomy	1,661	1,431	230	67	8	10	85	73	73	6	6	12	85
138	Non-U.S. born, NH Asian	Female	Earth/ocean/atmos.	679	560	119	27	3	50	81	67	67	7	7	14	81
139	Non-U.S. born, NH Asian	Female	Math	2,482	2,079	403	100	12	0	112	94	94	9	9	18	112
140	Non-U.S. born, NH Asian	Female	Computer/info. sci.	1,679	1,393	286	68	8	9	85	71	71	7	8	14	85
141	Non-U.S. born, NH Asian	Female	Agricultural sci.	1,478	1,304	174	60	7	18	85	75	75			10	85
142	Non-U.S. born, NH Asian	Female	Medical sci.	2,589	2,179	410	105	12	0	117	98	98	9	9	18	117
143	Non-U.S. born, NH Asian	Female	NIH bio sci.	6,535	5,603	932	264	31	0	294	252	252	21	21	42	294
144	Non-U.S. born, NH Asian	Female	Other bio sci.	6,986	6,004	982	282	33	0	315	271	271	23	22	44	315
145	Non-U.S. born, NH Asian	Female	Psychology	2,410	2,053	357	97	11	0	109	93	93	8	8	16	109
146	Non-U.S. born, NH Asian	Female	Economics	1,524	1,281	243	62	7	16	85	71	71	6	7	14	85
147	Non-U.S. born, NH Asian	Female	Anthro/arch/sociology	852	723	129	34	4	44	82	70	70	6	6	12	82
148	Non-U.S. born, NH Asian	Female	Other social sci.	1,377	1,154	223	56	6	22	84	71	71	7	7	14	84
149	Non-U.S. born, NH Asian	Female	Electrical/electron/comm.	2,559	2,114	445	103	12	0	115	95	95	10	10	20	115
150	Non-U.S. born, NH Asian	Female	Other engineering	5,224	4,327	897	211	24	0	235	195	195	19	22	40	235
Total				845,574	781,619	63,955	34,154	4,550	1,555	40,259	36,925	36,666	1,632	1,702	3,334	40,000

NH = Non-Hispanic.

NOTE: Specific grayed out cells have been suppressed for confidentiality reasons.