

# **New Immigrant Survey Restricted Data Use Agreement**

## **Guide to Accessing the New Immigrant Restricted Data**

### **Description of the New Immigrant Survey**

The New Immigrant Survey (NIS) is a nationally representative, longitudinal study of immigrants and their children that seeks to provide new data to help answer many of the important questions about immigration and while shedding light on basic aspects of human development. Data from the NIS are useful for addressing scientific and policy questions about migratory behavior and the consequences of international migration.

### **Protection**

The NIS was designed to collect data to be used by the broadest possible group of researchers. At the same time, the study's investigators are committed to protecting the privacy of NIS respondents and the confidentiality of NIS data. When subjects were invited to participate in the survey and during the interviews themselves they were promised that all identifying information would be kept completely confidential and separated from their responses when the information was released. Since NIS respondents are all legal U.S. immigrants, they are not at risk of disclosures about their current legal status. However, during the course of the interview it is possible that certain respondents revealed a prior illegal status or information at variance with statements made in the application for permanent residence. The interview may also uncover illegal status among others in the sampled household or information relevant to determining eligibility for public services such as AFDC, Medicaid, or SSI programs, disclosure of which could jeopardize receipt. In order to release NIS data to the public and while avoiding disclosure of confidential information, NIS investigators used different procedures and treatments for variables in the Adult, Proxy Child, Spouse, and Child datasets, and for other variables that could potentially reveal the identity of respondents we did not release them on our public use data files. The public use data include only variables with low disclosure risk and low levels of sensitivity.

In order to minimize the likelihood of indirect identification of respondents while making the maximum amount of data available to the research community, the NIS created two additional Restricted Data Files, each of which requires a different level of access requirements and data security.

## Eligibility

Access to the NIS Round 1 and Round 2 Restricted Data (Version 1 and Version 2) is limited to researchers who need additional variables that are not in the Public Use Data and who agree to the terms and conditions contained in the Restricted Data Use License. Only faculty and research personnel at institutions which have an Institutional Review Board/Human Subjects Review Committee are eligible to receive access to the restricted data. The Institution's IRB must be registered with the U.S. Office for Human Research Protections (OHRP) or the National Institute of Health (NIH). NIS Restricted data should not be used, under any circumstances, for the purpose of archiving or distribution to others.

University students may gain access to the Contract Data for research but a faculty advisor must serve as the Principal Investigator and complete the application process for them. The faculty advisor must be a PI of a federally-funded grant or must work within a federally-funded research center in which the Center Director agrees to take responsibility for data protection. The faculty advisor and institution bear full responsibility for ensuring that all conditions of the license are met by the student. The student must also sign the Supplemental Agreement with Research Staff form.

There are two versions of the NIS Restricted Data<sup>1</sup> that differ in the amount and detail of confidential information they include. A summary description of these two versions and the Public Use Data is provided in the following table:

**Comparison of Procedural Requirements for Versions of NIS data**

<b>Procedural Requirements</b>	<b>Public Use Data</b>	<b>Restricted Data V.1</b>	<b>Restricted Data V.2</b>
On-line User Agreement	Yes	-	-
Abstract and Justification	-	Yes	Yes
Institutional Review Board Approval	-	Yes	Yes
Data Protection Plan	-	Yes	Yes
Processing Fee	-	Yes	Yes
Licensing Agreements	-	Yes	Yes
Supplemental Research Agreement	-	Yes	Yes
Curriculum Vitae	-	Yes	Yes
Human Participants Protection Education for Research Teams Certificate	-	Yes	Yes
Secure Data Enclave	-	-	Yes
Federally Funded Grant or Center	-	-	Yes

<sup>1</sup> In this document, references to NIS Restricted Data involves NIS Round 1 and NIS Round 2 data files unless otherwise specified.

To be given access to the Restricted Data Files, users must submit TWO copies of the following items to NIS:

1. An Extended Abstract describing the proposed project and what it seeks to accomplish, along with a one-paragraph justification for why access to the Restricted Data File is needed rather than Public Use Data.
2. A Restricted Data Protection Plan, detailing how files will be protected while they are being used, being stored on computer, and after findings are published.
3. Written assurance by the researcher that his/her institution has an Institutional Review Board (IRB) for Human Subjects which has a Multiple Project Assurance (MPA) or Federal-wide Assurance (FWA) from NIH. The MPA or FWA number must be submitted with the application.
4. Proof of IRB Approval. The applicant's institutional review board must approve both the final research plan (extended abstract) and the final data protection plan.
5. An application fee of \$50 for Version 1 and \$500 for Version 2 (payable by check, purchase order or money order to Princeton University). Please note that the fee is non-refundable under any circumstances after Restricted Data Files have been received.
6. A signed Restricted Data Agreement by the Principal Investigator.
7. A signed Restricted Data Agreement by a senior university official who binds the university/institution. This refers to an individual who has the authority to represent your organization in agreements of this sort, such as a Vice President, Dean, Provost, Center Director, or similar official.
8. A signed Supplemental Research agreement with Research Staff for each person who will have access to the data.
9. A curriculum vitae for each person who will be accessing the information.
10. A copy of the Human Participants Protection Education for Research Teams completion certificate from NIH for all research staff who will access the contract data. The online certification can be completed at <http://grants.nih.gov/grants/policy/hs/training.htm>  
Proof of equivalent training is also acceptable.

**Please note:** *If co-investigators are from different institutions, you will need separate Contract Data Use Licenses for each institution.*

Researchers who submit TWO copies of the above items and are Principal Investigators (PIs) of a federally-funded research project, or researchers working in a federally-funded research center may apply to receive Restricted Data Version 2. If researchers work in a federally-funded research center but are not PIs, the Director of the center must agree to

take responsibility for data protection and must co-sign the Restricted Data Use Agreement. Both of them, Data Investigator and Center Director are responsible for compliance with the agreement.

NIS Restricted Data Version 2 must be used in a secure data enclave. This secure data enclave will be a physical enclave set (physical location of the enclave).

The Secure Data Enclave must implement a complete set of physical and computer security measures. Researchers should propose to set up a physical enclave, with a dedicated computer (and printer, if needed) that is not connected to any type of network (LAN or otherwise) and that is kept in a locked room with limited access.

Preliminary applications containing a draft of extended abstract, description of data requested, and CV should be sent to NIS (c/o [nisopr@princeton.edu](mailto:nisopr@princeton.edu)) in order to get approval before beginning work on the full application.

As part of the Restricted Data Agreement, researchers will be required to:

1. Submit annual IRB updates;
2. Cite the restricted data on any written report or publication, as follows:

The New Immigrant Survey, a research project designed by G. Jasso, D. Massey, M. Rosenzweig, and J. Smith, and funded by NIH HD33843, NSF, USCIS, ASPE & Pew. Restricted Use Data, Version (1 or 2) August 2007. Retrieved {date}. Persons interested in obtaining NIS Restricted Use Data should see [http://nis.princeton.edu/data\\_restricted.html](http://nis.princeton.edu/data_restricted.html) for further information.

3. Submit to [nisopr@princeton.edu](mailto:nisopr@princeton.edu) electronic copies of any publications and/or presentations at professional meetings using NIS Restricted data;
4. Completion of a new Restricted Data Use Agreement when there are changes in the Restricted Data Investigator employment status or changes in Center Director, when agreements are co-signed by the Center Director;
5. Notify NIS when new staff are added and will have access to the restricted data. The Investigator should submit signed copies of the Supplemental Agreement with Research Staff and access to the data cannot be provided to these staff members until the Supplemental Agreements are signed by an NIS representative and returned to the Investigator.

The application process involves three different parties: the Restricted Data Investigator, the Institutional Review Board and contracting authority at the researcher's organization, and the NIS. The Restricted Data Use Agreement is a legal document between these three parties. NIS will have full discretion in deciding whether to approve an application for access to Restricted Data. NIS may request additional information from applications or request changes to the Data Protection Plan. If NIS decides all requirements are met, the approval will be granted and a representative from NIS will sign the Restricted Data Use Agreement and return a copy of the fully executed agreement to the Investigator with

instructions on how to download the data through the NIS secure web pages. The Restricted Data Use Agreement expires after three years, with the option of applying for an extension.

### **Delivery**

Data will be available through the NIS secure web pages in three different formats or platforms: SAS, Stata, or SPSS files. Researchers will be able to choose between these three formats for their files. Upon expiration of the Restricted Data Use License, access codes will expire as well and researchers should destroy any copies of the data that exist and should complete the NIS Certification of Compliance with contract for use of Restricted Data and send it to NIS.

For more information about the restricted-use datasets or the application process, [email \(nisopr@princeton.edu\)](mailto:nisopr@princeton.edu) or write to:

Monica Espinoza Higgins  
New Immigrant Survey  
Office of Population Research  
187 Wallace Hall  
Princeton University  
Wallace Hall  
Princeton, NJ 08544  
ATTN: New Immigrant Survey Restricted Data

## **Guidelines for the Contract for Obtaining Restricted Data Use from the New Immigrant Survey**

This contract for obtaining Restricted Data from NIS requires the signatures of everyone who will have access to the data as well as institutional representatives. Please read the entire package. Final application will be processed when two copies of all parts, fully completed, are received by NIS. After signed by all parties, this package constitutes a legally binding contract. The contract includes the following items:

Use the following checklist to ensure that your application package is complete.

- Restricted Data Use Application Cover Page;
- Restricted Data Use License Application;
- Completed Supplemental Agreement with Research Staff;
- Final Research Plan (approved by IRB)
- Restricted Data Protection Plan (approved by IRB);
- CVs of all research staff;
- Evidence of IRB's Certification with NIH or OHRP;
- Non-refundable application fee of \$50 for Version 1 and \$500 for Version 2 payable to "The Trustees of Princeton University";
- Copies of human subjects completion certificates for all research staff;

*Additional documents for Restricted Data Version 2:*

- Secure Data Enclave;
- Evidence of Federally funded Grant or Contract.

We must receive TWO complete sets of all the above documents with original signatures in each copy. Send complete package to:

Monica Espinoza Higgins  
New Immigrant Survey  
Office of Population Research  
Princeton University  
187 Wallace Hall  
Princeton, NJ 08544  
ATTN: New Immigrant Survey Restricted Data

**New Immigrant Survey  
Restricted Data Use Application Cover Page**

Date of Application:

Name of Investigator:

Title of Investigator:

Receiving Institution:

Department:

Street Address:

City/State/Zip Code:

Telephone Number:

Fax Number:

Email Address:

Title of Research Project:

---

**NIS Use Only**

Date access granted:	Expiration date:
Renewed:	Expiration date:
Renewed:	Expiration date:
Renewed:	Expiration date:

## **Agreement for Use of Restricted Data from the New Immigrant Survey**

Please note that you must submit two original, signed copies of this document; one will be countersigned and returned to you.

### **I. Definitions**

1. *The New Immigrant Survey* (NIS)- is a research project undertaken by four institutions: New York University, Princeton University, Yale University, and RAND supported by the Institutes of Health (NIH)/ National Institute of Child Health and Human Development (NICHD)/ National Institute on Aging (NIA)/Office of Behavioral and Social Science Research (OBSSR) under Grant No. HD33843, the National Science Foundation (NSF) under Grants No. SRS-9907421 and No. SES-0096867, the U.S. Immigration and Naturalization Service (now the U.S. Citizenship and Immigration Services). Additional support was provided by the Office of the Assistant Secretary for Planning and Evaluation (ASPE) and the Pew Charitable Trusts.
2. *Investigator* - The person primarily responsible for analysis and other use of restricted data obtained through this agreement. The Investigator must hold a faculty appointment or research position at the Receiving Institution and assumes all responsibility for compliance with all terms of this License by employee of the Receiving Institution. The investigator is the person who will serve as the primary point of contact for all communications involving this License.
3. *Receiving Institution* - The university or research institution employing the Investigator and at which the Investigator will conduct research using restricted data obtained through this agreement. The receiving Institution must have an Institutional Review Board/Human Subjects Review Committee registered with the United States Office for Human Research Protections or the National Institute of Health.
4. *Research Staff* - All individuals, excluding the Investigator, who will have access to restricted data obtained through this agreement. The Research Staff must be affiliated with the Receiving Institution.
5. *Representative of the Receiving Institution* – An individual authorized to enter into contractual agreement on behalf of the Receiving Institution, such as a Vice President, Dean, Provost, Center Director, or similar official.

**NOTE:** A Department Chair is not acceptable unless specific written delegation of authority exists.

6. *Sensitive/Restricted Data* - Includes any data from NIS that might compromise the anonymity or privacy of respondents to that study. Because of its sampling frame which is based on nationally representative samples of the electronic administrative records compiled for new immigrants by the U.S. government (via, formerly, the U.S. Immigration and Naturalization Service (INS)), its longitudinal nature, and the amount and type of information collected on immigrants and their families, NIS respondents are at higher risk of deductive disclosure. Therefore, all data collected from NIS are considered to be sensitive or restricted.
7. *Federally-funded* – Funding provided for research or institutional support through a grant or contract from an agency of the United States federal government. Such agencies include, but are not limited to, the National Institutes of Health and the National Science Foundation.

## **II. Limitations on Use and Disclosure of Restricted Data**

Data provided under this agreement shall be held by the Investigator, Research Staff, at the Receiving Institution in strictest confidence and can be disclosed only in compliance with the terms of this agreement.

In consideration of The New Immigrant Survey providing the Investigator access to the Restricted Data, the Receiving Parties agree as follows:

- A. That the data will be used solely for scientific and public policy statistical analyses, as described in the Research Plan submitted to and approved by NIS and attached to this agreement.
- B. Restricted Data will be safeguarded in accordance with the Restricted Data Protection Plan submitted to and approved by NIS and attached to this agreement.
- C. No persons other than those identified in this agreement, or in amendments subsequent to this agreement, as Investigator or Research Staff, be permitted access to the contents of restricted data files or any files derived from restricted data files.
- D. Under no circumstances will the Investigator use or disclose the Restricted Data for any purpose not stated in the Research Plan, including but not limited to any administrative or law enforcement purpose pursuant to a Certificate of Confidentiality.
- E. No attempt will be made to identify specific individuals, families, households, employers, or institutions; nor will any list of data at the individual or family

level be published or otherwise distributed. If the identity of any person, family, household, school, or institution should be discovered inadvertently, then (a) no use will be made of this knowledge; (b) the Principal Investigator(s) of NIS will be advised of the incident; (c) the information that would identify the person, family, household, school, or institution will be safeguarded or destroyed as requested by NIS; and (d) no one else will be informed of the discovered identity.

- F. Restricted Data will be used only to generate statistical summary information that does not permit the identification of any individual person, family, household, employer, or institution. To avoid inadvertent disclosure of persons, families, households, or institutions by using the following guidelines in the release of statistics derived from the dataset.
  - a. In no table should all cases in any row or column be found in a single cell.
  - b. In no case should the total for a row or column of a cross-tabulation be fewer than three.
  - c. In no case should a cell frequency of a cross-tabulation be fewer than three cases.
  - d. In no case should a quantity figure be based on fewer than three cases.
  - e. Data released should never permit disclosure when used in combination with other known data.
  - f. No geographic unit (state, county, metropolitan area, city, etc.) with fewer than 100 NIS respondents will be separately tabulated
  - g. No country of origin with fewer than 40 NIS respondents will be separately tabulated
- G. No attempt will be made to link Restricted Data with any other dataset without written authorization from the NIS.
- H. To supply NIS with two copies of each of the following:
  - a. completed Investigator Information form
  - b. Agreement for the Use of Restricted Data, each with original Institutional Signatures page

- c. Restricted Data Security Plan
  - d. Data File Order with specific files requested, and explanatory statements for Restricted Data Version 2
  - e. Supplemental Agreement with Research Staff for the Use of Restricted Data signed by each Research Staff person
  - f. A copy of the document, originated by the Investigator and signed by the Receiving Institution's Institutional Review Board (IRB), approving the research project and the secure use, storage, and handling of the Restricted Data outlined in the Restricted Data Protection Plan.
- I. If in the event the Investigator changes institutional affiliation during the period covered by this contract, the Investigator will take the following actions:
- 1. Inform NIS six weeks prior to the date of relocation.
  - 2. Execution of a new Restricted Data Use License, resubmission of a security plan for the new institution and obtain approval from NIS prior to moving any electronic or paper files from the originally approved site to the new location.
  - 3. Destroy all electronic and paper files at the originally approved site prior to the date of relocation.
  - 4. Until the new contract is executed at the new institution, the data cannot be installed or used at the new institution. The Investigator is responsible for the security of the data.
  - 5. Within three months of the effective date of the relocation, submit two copies of a new contract with appropriate supporting documentation.
- J. If in the event there are changes in the Research Staff, the Investigator will take the following actions:
- 1. When Research Staff leave the project, the Investigator will notify NIS that these individuals are no longer authorized to access the Restricted Data.
  - 2. When Research Staff join the project, they will submit the Supplemental Agreement with Research Staff. Such Supplemental Agreements must be submitted before the new Research Staff may have access to the Restricted Data.

- K. The investigator will provide NIS annual reports, which include:
1. a copy of the annual IRB approval for the research project;
  2. a list of public presentations at professional meetings using results based on these data;
  3. a list of papers accepted for publication using these data, with complete citations;
  4. a list of graduate students using the NIS data for dissertations or theses, the titles of these papers, and the dates of completion.
- L. To include in each written report or other publication based on analysis of restricted data from NIS, the following statement:

The New Immigrant Survey, a research project designed by G. Jasso, D. Massey, M. Rosenzweig, and J. Smith, and funded by NIH HD33843, NSF, USCIS, ASPE & Pew. Restricted Use Data, Version (1 or 2), 2007. Retrieved {date}. Persons interested in obtaining NIS Restricted Use Data should see [http://nis.princeton.edu/data\\_restricted.html](http://nis.princeton.edu/data_restricted.html) for further information.

### **III. Representations By Investigator**

The Investigator represents and warrants that:

- A. The Investigator has permanent, faculty appointments or faculty-equivalent research appointments at the Receiving Institution. “Permanent” in this agreement means a full time employment throughout the course of the proposed project.
- B. All research staff signing the Supplemental Agreement with Research Staff have a formal affiliation (i.e. employee, currently enrolled student, etc.) with the Receiving Institution and with the research project described in the Research Plan, and will have access to Restricted Data only under the supervision of the Investigator and subject to the terms of the Restricted Data Protection Plan.
- C. *Only for Restricted Data Version 2:* The Investigator has a current federally-funded research grant or contract that has been provided to NIS.

#### **IV. Representations By the Center Director**

The Center Director represents and warrants that:

- A. The Center Director is currently the Principal Investigator of a federally-funded grant or contract which provides institutional support to the research center and information on the center's federally-funded grant(s) and contract(s) has been supplied to NIS.
- B. The Center Director agrees to assume responsibility for ensuring compliance with the Restricted Data Agreement by the Investigator and all other individuals with access to the restricted data.
- C. The Center Director has reviewed and agreed with the Research Plan, the Restricted Data Agreement, and the Restricted Data Protection Plan.

#### **V. Representations By Receiving Institution**

The Receiving Institution represents and warrants that:

- A. The Receiving Institution has an IRB/Human Subjects Protection Committee in accordance with the DHHS regulations codified at Title 45 Part 46 of the Code of Federal Regulations. Proof of such certification and the Institution's Multiple Project Assurance (MPA) or Federalwide Assurance (FWA) number has been provided to NIS.
- B. The Receiving Institution's IRB/Human Subjects Protection Committee has reviewed and approved the Research Plan, the Restricted Data Agreement, and the Restricted Data Protection Plan in accordance with the DHHS regulations codified at Title 45 Part 46 of the Code of Federal Regulations, using the standards and procedures for live human subjects, and that certification of such approval has been provided to NIS.
- C. The Receiving Institution will treat allegations by NIS of violations of this agreement as it does allegations of violations of its policies on scientific integrity and misconduct. Formal written policies and procedures for resolving questions of scientific integrity and misconduct from the Receiving Institution have been provided to NIS.

#### **VI. Certificate of Confidentiality**

Research subjects who participated in NIS study are protected by a Certificate of Confidentiality issued by the Department of Health and Human Services in accordance with the provisions of section 301(d) of the Public Health Service Act

(42 U.S. C., 241(d)). The Receiving Institution is considered to be a contractor or cooperating agency of Princeton University under the terms of the Certificate of Confidentiality; as such, the Receiving Institution, the Investigator, and Research Staff are authorized to protect the privacy of the individuals who are the subjects of NIS by withholding their identifying characteristics from all persons not connected with the conduct of the study. Identifying characteristics are consider those data defined as restricted under the terms of this contract.

## **VII. Destruction of Data Upon Completion of Research Project**

The Investigator will ensure that all copies of Restricted Data, on whatever media, will be destroyed at the completion of the research project, or within 36 months from the date this License is accepted by NIS, or within 5 days of a written request from NIS.

## **VIII. Duration of License**

The Restricted Data Use Agreement expires after 36 months, with the option of applying for an extension. The License will go into effect upon approval of the License by NIS, and will remain in effect until the completion of the research project, or 36 months from the date the License is accepted by NIS, whichever comes first.

## **IX. Ownership of Data and Liability**

Ownership of the Restricted Data will be retained by NIS. NIS can revoke the permission to use the Restricted Data by the Investigator and Receiving Institution at any time, at their discretion. If permission is revoked, the Investigator and Research Staff will automatically lose permission to log into the restricted data web pages. The Investigator must destroy copies of the Restricted Data, and complete and send the Certification of Compliance form within 5 days of written request to do so.

The investigator will not make any claim to copyright ownership of the Restricted Data and accompanying documentation.

The Investigator and Receiving Institution jointly and severally shall indemnify Princeton University, their officers, agents, and employees against any liability, including costs and expenses, incurred as the result of the violation of copyrights, or right of privacy or publicity, arising out of the Institution's or Investigator's creation, delivery, publication, or use of data furnished under this license or the breach of any of the terms of this License. Princeton shall provide the Investigator and Receiving Institution of timely notice of any claim or suit, afford the Investigator and Receiving Institution an opportunity under applicable laws, rules, or regulations to participate in the defense thereof, and obtain the Investigator's

and Receiving Institution's consent to the settlement of any suit or claim other than as required by final decree of a court of competent jurisdiction.

#### **X. Violation of this License**

Ownership of the Restricted Data will be retained by NIS. NIS can revoke the permission to use the Restricted Data by the Investigator and Receiving Institution at any time, at their discretion. If permission is revoked, the Investigator must destroy copies of the Restricted Data, and complete and send the Certification of Compliance form within 5 days of written request to do so. Violations of this agreement will be reported to the National Institutes of Health. NIS personnel reserve the right to undertake unannounced site visits to verify continued compliance.

**Institutional Signatures (Please do not use black ink.)**

**Investigator**

**Representative of Your Institution**

\_\_\_\_\_  
Signature Date

\_\_\_\_\_  
Signature Date

\_\_\_\_\_  
Name typed or printed

\_\_\_\_\_  
Name typed or printed

\_\_\_\_\_  
Title

\_\_\_\_\_  
Title

\_\_\_\_\_  
Institution

\_\_\_\_\_  
Institution

\_\_\_\_\_  
Building address

\_\_\_\_\_  
Building address

\_\_\_\_\_  
Street address

\_\_\_\_\_  
Street address

\_\_\_\_\_  
City State Zip

\_\_\_\_\_  
City State Zip

**NIS Representative**

\_\_\_\_\_  
Signature Date

Douglas S. Massey, Ph.D.  
Principal Investigator  
Office of Population Research  
Wallace Hall  
Princeton, NJ 08544

**Data File Order for the Use of Restricted Data from the  
New Immigrant Survey Study**

The data will be available in the file formats specified below.

All data will be compressed using WinZip and password protected.

Instructions on how to download data from NIS secure web pages will be sent by email to the Investigator.

Contact person: \_\_\_\_\_ Contact email: \_\_\_\_\_

*Investigator*

_____	_____	_____
Name	Signature	Date
<hr/>		

Please check files that will be downloaded (descriptions are on the following pages):

- NIS 2003-1 Baseline survey
- NIS 2003-2 Follow-up survey
  
- Adult Sample respondents
- Proxy Child respondents (CPs and NAMCs)
- Spouse respondents
- Child Sample

**Restricted Data Version 1**

- Raw Country variables
- Raw State variables
- Raw Language variables
- Raw Relationship variables
- Raw Currency variables

- Education information
- NonCash Payment information
- Employment & Wage Rates
- Health conditions
- Raw Religion variables
- Occupation and Industry
- Raw Height and Weight Measurements

### **Restricted Data Version 2**

The following datasets are available by special request and additional requirements. In order to receive the datasets please follow carefully the guidelines for Restricted data Version 2.

- Geographic Identifiers
- INS variables (only available for NIS 2003-1)
- Migration Dates (only available for NIS 2003-2)

Please check the format or platform that will be used for data download:

- SAS Files
- SPSS Files
- STATA Files

### **Description of NIS files**

**Country Variables**—The survey has numerous questions about the country where the respondent and his or her relatives were born including country of citizenship, country where visa application was filed, country where respondent studied, worked, traveled and so on.

The public use version of the dataset includes recoded versions of these country variables, based on information obtained by the INS records CISCOBINS *Country of*

*birth of sampled person*, as well as information given by the respondent in question A9a, *Country of birth of sampled person for the Adult Sample*. After cleaning the *specify* variable and merging to the *original* variable<sup>2</sup>, the *total* variable was recoded based on the number of cases found in the variables *country of birth of sampled person*. A country name was included in the public use dataset only if more than 100 respondents were born in that specific country, the remaining countries were classified by regions. The restricted data file includes the *total* country variable unrecoded which merges the *original* variable and the cleaned *specify* variable. Approximately there are 897 variables related to country for the Adult Sample, and a similar number were recoded for Proxy Child, and Spouse.

**State Variables**—This file includes state names for the questions asked in the administration of the questionnaire that asked for state. A few examples include state where spouse/partner is enrolled in school; state where each of the respondents' children is living in at the time of the interview; states where homes are owned by respondents and states where the spouse/partner currently lives; states where respondents lived in their previous trips to the United States before becoming legal residents.

For the public-use datasets, these variables include the main six states (New Jersey, New York, Texas, Illinois, Florida, and California) and the rest of the states were grouped based on the Census Regions and Divisions.

These variables do not include current state of respondent at time of interview (adult sample and proxy child). Current state at time of interview is under the category Geographic Identifiers, Restricted Data Version 2.

**Currency Variables**—These variables are the raw *specify* variables for the currency variables in the Adult Sample, Proxy Child, and Spouse datasets. The raw *specify* variables are character variables and not based on a picklist, although most of the values are found in the picklist. In the public-use data files, these variables were cleaned and coded to the picklist for the public-use dataset.

**Relationship Variables**—These variables are the raw *specify* variables for the relationship variables in the Adult Sample, Proxy Child, Spouse datasets, and Child Interview. The raw *specify* variables are character variables and not based on a picklist, although most of the values are found in the picklist. There are approximately 486 questions for the Adult Sample, 486 questions for the Proxy Child, 243 questions for the spouse interviews, and 20 variables for the Child Sample.

In the public-use data files, these variables were cleaned and coded to the picklist for the public-use dataset.

---

<sup>2</sup> Many questions in the survey were compounded by two variables: an *original* variable which makes use of a picklist and a *specify* variable which is an open question if the answer in the original variable was 'Other'.

**Language Variables**—There are approximately 240 questions for the Adult Sample, 240 questions for the Proxy Child, 125 questions for the Spouse interviews, and 37 variables for the Child Sample about all the diverse languages and dialects immigrants and their families use to communicate. The raw *original* variables are based on a complete language picklist containing 642 languages. The raw *specify* variables are character variables and not based on a picklist, although most of the values are found in the picklist.

For the public-use datasets, the original and specify variables were merged and grouped, when there were not enough cases to count as one category. The raw variables released in the Restricted Data include the variables that merge the original and specify variables and are coded based on the complete language picklist.

**Education Variables**—Verbatim responses about the different levels of education attained/hope to obtain by the respondents from different countries are released in these variables. There is information about education of each respondent, husband/wife/partner, all biological children, adopted children, stepchildren, and parents.

For the public release, education variables were coded into categories based on respondent's answers and two education variables were created: degree and certificate obtained/hope to obtain.

It is important to mention that researchers should be very careful when using these raw variables for their research. The classification of levels of education and the concepts of levels of education vary among regions. If researchers want to analyze these raw variables, we suggest they look at other responses for each respondent, such as: field of degree, age, sex, occupation, kind of business or industry, kind of activities and duties at work, country of origin, etc in those cases with *unidentified* answers.

A sample of answers given to question A24 Highest degree, diploma or certificate that respondent received which has 5,930 observations for the Adult Sample: SOVIET COLLEGE DEGREE, SPECIALIST IN NURSING, SSC, SSC.PASS, SSCE LIKE GED HERE, SSE, SSHGD, STALISH BEAUTY, TACHNIC MECHANIC, TACHNICIAN-HOTEL MENAGER, PERITO CONTADOR<sup>3</sup>, ORDER OF ART, etc.

**Noncash Payment Information**—Raw variables with verbatim responses regarding noncash payments received by respondents when they first started their first job, just before they left their last job before coming to the United States, and when they first started their first job in the United States. For the public release, these variables were recoded based on the verbatim answers given by the respondents.

---

<sup>3</sup> Respondents were interviewed in their preferred language. Although this maximizes the response rate and data quality, it also presents challenges in preparing the data ready for release, since some answers were still in the language in which the interview was conducted.

**Employment & Wage Rates**—Employment variables are the raw *specify* variables for the current employment variables *CIa* in the Adult Sample, Proxy Child, and Spouse datasets. In the public-use data files, these variables were recoded in order to follow a similar structure as the original variable *CI*.

Wage rates are the raw original variables for the hourly wage rates (*C49\_X*) reported by the Adult respondent, Proxy Child, and Spouse. In the public-use data files, these variables were slightly modified where inconsistencies were found.

**Health Conditions**—Respondents were asked about the most serious health conditions that caused them to ever miss school for one month or more, to ever be confined to bed or home one month or more, or if sports or physical activities were ever restricted for three months or more. These files are also in the public-use datasets based on a picklist with 58 values. The raw data were coded to this picklist for the public-use dataset.

**Religion variables**—The survey provides information on the respondent's religious affiliations and the affiliations of the respondent's family members. These questions were asked in Section J – Social Variables. This section was not asked of Spouses.

The religion variables for the restricted release are the original responses provided by the respondent with verbatim replies cleaned of typos and misspellings. The religion variables for the Adult Sample and Proxy Child public release were recoded based on the verbatim information that is provided to the researcher in the restricted release.

**Occupation & Industry information**—For each Adult Sample, Proxy Child, and Spouse dataset, these files contain the verbatim original responses about occupation of respondents and kind of business or industry respondents worked on at their first job, last job before coming to the U.S., first job in U.S., current first and second job, occupation of parents, occupation of children, etc. There were approximately 58,000 separate entries of these variables, about 29,000 occupation responses and the same number of industry responses.

There is a public-use version of these variables. The entries were coded to a 2003 coding scheme. For Occupation, the Census 2003 codes are still based on the 2000 SOC (Standard Occupational Classification) codes and can be matched to Census 2000 codes. For Industry, the Census 2003 codes are based on 2002 NAICS (North American Industry Classification System) codes.

**Height and Weight Measurements**— These are the raw original variables for the height and weight self-measurements (*D129* to *D130C*) reported by the Adult respondent, Proxy Child, and Spouse. In the public-use data files, these variables were modified where inconsistencies were found.

**Geographic Identifiers**—The Restricted Data Version 2 contain geographic identifiers (State, City, and Zip Code of sampled respondent at time of interview).

The first variable was recoded as explained above for *state* variables and released in the public-use datasets.

These files also include the raw *city* variables for the corresponding *state* variables in the Adult Sample, Proxy Child, and Spouse datasets. In the public-use data files, these variables were not released.

**INS variables**—The Restricted Data Version 2 contain the original variables from the INS administrative record of the respondent. These variables are: year and month of birth; country of birth; country of chargeability; country of last residence; country of nationality; state, city, and zip code where green card was sent; class of admission; non-immigrant class, and visa category.

**Migration Dates**—These variables contain day/month information for questions about when main respondent became a U.S. citizen and when main respondent filed a petition to begin the process to bring someone to live permanently in the United States. Year information for these variables are released in the public data.

## Description of Parameters for Data Protection Plan

Researchers must provide a concise but detailed data protection plan as part of their application to receive NIS Restricted Data (Version 1 & Version 2). NIS will not provide Restricted Data if the plan is not written with sufficient specificity, or if NIS does not deem the data protections to be adequate.

The data Protection Plan applies to the original Restricted Data files received from NIS (regardless of its format), to any copies made by the research team, and to any new data derived solely or in part from the original Restricted Data files. The plan also should address how computer output derived from the data (for example, case listings), will be kept secure.

Users should follow description of parameters for Data Protection Plan for the Fragile Families Survey ([http://www.fragilefamilies.princeton.edu/documents/restricted\\_agreement.pdf](http://www.fragilefamilies.princeton.edu/documents/restricted_agreement.pdf)) (pages 12-14) which includes all the components that should be included in the Data Protection Plan, disclosure rules, and the types of protection expected.

Users can see detailed instructions for acceptable data protection plans for the National Longitudinal Study of Adolescent Health (<http://www.cpc.unc.edu/projects/addhealth/data/restricteduse/security>), Health and the Retirement Survey data ([http://hrsonline.isr.umich.edu/rda/rdapkg\\_prot.htm](http://hrsonline.isr.umich.edu/rda/rdapkg_prot.htm)), The Los Angeles Family and Neighborhood Survey (<http://www.lasurvey.rand.org/data/restricted/>) as references for developing an acceptable data protection plan.

**Supplemental Agreement with Research Staff  
for the Use of Restricted Data from the  
New Immigrant Survey**

- I. The undersigned Research Staff, in consideration of their use of sensitive data from the New Immigrant Survey, agree:
  - A. That they have read the associated Restricted Data Use Agreement Application from the New Immigrant Survey and the Restricted Data Protection Plan.
  - B. That they are “Research Staff” within the meaning of the agreement.
  - C. To comply fully with the terms of the agreement, including the Restricted Data Protection Plan.
  
- II. The undersigned Investigator agrees that the persons designated herein are Research Staff within the meaning of the associated Agreement for the Use of Restricted Data from the New Immigrant Survey.

*Research Staff:*

Name	Title	Signature
Telephone	Email Address	Date
Name	Title	Signature
Telephone	Email Address	Date
Name	Title	Signature
Telephone	Email Address	Date
Name	Title	Signature
Telephone	Email Address	Date

\_\_\_\_\_  
Name Title Signature

\_\_\_\_\_  
Telephone Email Address Date

\_\_\_\_\_  
Name Title Signature

\_\_\_\_\_  
Telephone Email Address Date

\_\_\_\_\_  
Name Title Signature

\_\_\_\_\_  
Telephone Email Address Date

*Investigator:*

\_\_\_\_\_  
Name Title Signature

\_\_\_\_\_  
Telephone Email Address Date

**New Immigrant Survey  
Certification of Compliance  
with Contract for use of Restricted Data**

This certifies that the requirements of my Contract for Use of Restricted Data from the New Immigrant Survey (NIS) has fully complied with, including but not limited to the following:

- A. NIS Restricted Data was used solely for the purposes specified in my approved Extended Abstract.
- B. No attempt was made to identify or contact any individual, family, or household participating in the NIS, nor any employer or benefit provider for such person.
- C. No persons, other than those identified in the Agreement as Investigator and Representative of our institution, or in a Supplemental Agreement as Research Staff, was permitted access to the contents of the NIS Restricted Data or any files derived from the NIS Restricted Data files, including intermediate, analysis, or backup versions of those files.
- D. The approved Restricted Data Protection Plan was fully complied with.
- E. All copies of NIS Restricted Data, including copies of originals, intermediate files, and analysis files, and all backup copies thereof, have been sent to the NIS or destroyed so that they cannot be “undeleted” or otherwise recovered.
- F. All exceptions or qualifications of any of the above certifications, if any, are noted in a separate document attached to this certification.

**Investigator:**

**Representative of Your Institution:**

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name typed or printed

\_\_\_\_\_  
Name typed or printed

\_\_\_\_\_  
Title

\_\_\_\_\_  
Title

\_\_\_\_\_  
Institution

\_\_\_\_\_  
Institution

\_\_\_\_\_  
Building address

\_\_\_\_\_  
Building address

\_\_\_\_\_  
Street address

\_\_\_\_\_  
Street address

\_\_\_\_\_  
City State Zip

\_\_\_\_\_  
City State Zip

