Center for Demography and Ecology
University of Wisconsin Madison
4412 Social Science Building
1180 Observatory Drive
Madison, Wisconsin 53706-1393
Telephone (608) 262-1537

Fax (608) 262-8400

In response to your inquiry about adding contextual data to NSFH, this document outlines the process through which CASEID can be merged with data on geographic characteristics from another source. Please note that you must read and carefully follow the directions below. We do not have the resources to examine why a user's files will not merge.

The attached confidentiality forms must be signed by all persons and their staff who has access to NSFH files with added contextual data. We have assured our respondents that the information they provide will remain confidential and in no way associated with their name. We cannot emphasize enough how important it is for the respondents and for the field at large that all precautions be taken to preserve this confidentiality. It is your ethical obligation to follow these procedures meticulously. In signing this, you are affirming that neither you nor your staff will make any attempt to identify individuals. Further, you are agreeing that you will not share these files with anyone not covered by this agreement.

An overview of the study and descriptions of Waves 1, 2 and 3 can be found at http://www.ssc.wisc.edu/nsfh/home.htm. Before proceeding, it is important that you read all of the introductory files on our web page as well as the "User Support" page: http://www.ssc.wisc.edu/nsfh/support.htm. Links for both of these are located on our homepage.

The codebooks for the data files are **nsfh1gdt.doc, nsfh1gdt.doc,** and **nsfh3apl.doc,** links to both can be found on the user support page. The file containing the geographic locations of Wave 1 respondents and that for respondents, spouses, and new spouses in Wave 2, and that for respondents and spouses in Wave 3 will be provided to the user when the confidentiality agreement is complete. Additional geographic locations are represented in Waves 2 and 3 that were not in Wave 1 because some respondents moved, and because a former or current spouse may have been living in a different location than the respondent.

Performing a Merge of Respondent Identification Numbers with Contextual Data

Overview: You send us a file with the characteristics of the geographic areas from the *gdt.dat files that you wish to merge with the individual NSFH records. In return, we will send you a file which contains the CASEID for each respondent attached to the contextual variables you have sent us. The variables identifying geographic areas will have been removed. The purpose of this merge is to allow analysis including the characteristics of places of residence, and *not* geographical locations.

The Process

Please note: the process of locating, obtaining, and checking the variables to be added is your responsibility. While we want to be helpful about the process at this end, we cannot afford to provide consultation with respect to the variables to be added. Whatever variables you send will be returned exactly as you send them, with CASEID added.

1. Choose the level of geographical aggregation at which you wish to add contextual variables.

2. Find a source (City-County Data Book, for example) for the variables you wish to add, and create a file including those data at your desired level of aggregation.

3. Select the variables from the *gdt.dat file which are necessary for a match at that level (for example, GDTSFIPS and GDTCFIPS to add data for counties). Please tell us which .gdt variables you have used for this selection.

4. Fill out and sign the attached confidentiality forms. Note that these require that you specify the level of aggregation for which you desire us to add CASEID to your contextual variables. There are two different forms, one for the person in charge and one for all others who will have access. These forms require explicit commitments that no attempt whatsoever will be made to identify either respondents or their geo- graphic location.

5. Prepare your file(s) to be sent to us.

   a) Create a file that reproduces the .gdt records exactly and simply adds your variables to the end of this file (beginning in location 55). Hence, this must be an ASCII file written in character format. The file(s) you send us must be limited to only the geographical units represented in our gdt.dat files. Please note there can be only one record for each unique geographical unit you wish to match: i.e., all information for each area **must be written on a single record**.

If you have worked with the gdt file previously, you will note that the earlier versions of the state and zip-codes that were at the start of this file have been removed. The first variable on the file is GDTSTATE which remains in the same location as before, beginning with locations 12-13.

The file(s) you send us must be limited to only the geographical units represented in our gdt.dat files. To limit the possibility of disclosure, we reserve the right to recode data to missing if we believe the frequency counts for a geographical area are too small.

   b) Check the variables in your file for missing data. Remember that we will not attempt to check these for you or investigate what went wrong if there are errors.

   c) Send your file by email. Please include a list of variables with a format (layout file with column locations) and descriptive statistics of means, lowest and highest values for each variable.

   d) When creating your layout file please enumerate your variables (if they **do not contain decimal places**) with the name first and the column locations directly to the right of the variable name. For example:
      gdtsfips   19-20
      gdtcfips   21-23
      skip2 $24-58
   Please also include a file with a list of your variable names without column locations. However, **if your variables include decimal places** please format your layout file in the following manner (we still request that you send a separate file with just variable names and no column locations in addition to your layout file):
      @59 pctblk 8.2
      @67 pctwhite 8.2
      @75 pcthisp 8.2

   e) Be sure to use the FIPS geographical codes that apply to the 1990 Census. We've had several users who've used the 2000 Census codes, and when doing a merge at the county level, this causes problems (especially with Miami-Dade County in Florida).

6. We will then:

     a) Merge your data with the individual records in NSFH1, NSFH2, or NSFH3.
     b) Remove all identifiers for geographic areas.
     c) Return to you a file with CASEID added to the file that you sent to us. This file will have CASEID in locations 1-6 and include in locations 7-zzzz exactly the string of characters you sent to us in these locations in your original file.

Checklist of items that need to be returned with this application. Please make sure every item on this list is submitted with your application, or it will not be accepted.

1. This confidentiality agreement, signed by the applicant, an advisor who is a Professor if the applicant is a student, and a representative of the receiving institution. In the case of the University of Wisconsin, this representative would be from Research and Sponsored Programs:

https://www.rsp.wisc.edu/

Other institutions would need a signature from their analog.

2. A completed Institutional Review Board (IRB) Human Subjects approval.

3. An explanation of how you intend to protect this restricted data from public disclosure. Depending on what platform you plan to use to access this data, please use the form at the CPC Data Portal:

https://data.cpc.unc.edu/projects/2/view

It has a set of restricted data plans used by the National Longitudinal Study of Adolescent Health (ADDhealth). Fill out the form for whatever plan you use and make sure to send the completed form of your choice to NSFH with this application.

CONFIDENTIALITY AGREEMENT

Researchers requesting to have geographic characteristics merged with NSFH data files must sign this statement agreeing to the conditions of use, and return it to NSFHHELP.

Confidentiality agreement between the University of Wisconsin and _____

_____(Applicant)

Regarding the use of geographical data from Wave 1 (1987-1988), Wave 2 (1992-1994) and/or Wave 3 (2001-2003) of the National Survey of Families and Households, the National Survey of Families and Households (UW IRB Protocol 01-01-05) agrees to provide a data file containing geographical information omitted from the public release version of the file for reasons of confidentiality.

I, _____ agree that:
        (Applicant)

1. Permission to use the geographic characteristics data is granted to me at my current institution. Only I and my staff will have access to any NSFH file containing geographic characteristics. These data will not be duplicated or shared with other researchers, including researchers at my institution, without the submission and approval of a confidentiality agreement.

All project staff members will read and sign the staff confidentiality statement before they have access to the geographic characteristics data. These staff forms will be countersigned and retained by me, and given to the NSFH director on request. If I leave my current institution, I understand that I must reapply for permission for continued access to these data, and that this is also true for graduate students and other staff members who change their institutional affiliation.

2. My staff and I will take reasonable precautions to preclude persons not authorized to use the geographic information from having access to it. The data files with this information will be maintained in a secure manner, with access restricted to authorized persons.

3. No attempt will be made by my staff or me to identify either individual respondents or the geographic location of respondents for any purpose, including the appending of additional geographic data to the respondent records. I understand that any such attempt may void permission to use the geographic characteristics data, and that if I, or others, become aware of such an attempt, we are obligated to report it to the NSFH director.

4. Reports or publications deriving from these data will not identify specific geographic locations of NSFH respondents or groups of respondents.

4

CONFIDENTIALITY AGREEMENT page 2

In agreement with the above conditions, I request permission to merge geographic characteristic to individual NSFH records at the level(s) of aggregation of:

_____
(Aggregation level: For example, county or zip code)

_____   _____
Signature   Date

_____
Print Name

_____
Title   University/Institution

IF YOU ARE A GRADUATE STUENT REQUESTING NSFH WAVE 1, WAVE 2 AND/OR WAVE 3 GEOMERGE, YOU NEED TO ASK YOUR ADVISER TO CO-SIGN THIS FORM IN THE SPACE BELOW.

_____   _____
Signature   Date

_____
Print Name

_____
Title   University/Institution

Representative of the Receiving Institution: By signing this License, this institution agrees that access to these confidential data will be contract to authorized persons whose names appear on this License and the Supplemental Agreements with Research Staff, and that this institution is legally bound by the covenants and terms of this License.

_____   _____
Signature   Date

_____
Print Name

_____
Title   University/Institution

CONFIDENTIALITY AGREEMENT FOR AFFILIATED STAFF MEMBERS

Persons who will have access to NSFH data files that include geographic characteristics that have been merged with NSFH data files, must read and sign this form, and give it to the researcher for whom the merge was performed (herein referred to as the primary investigator).

I agree that:

1. I may only use these data as directed by the primary investigator, _____ , who has approved access to them. I will not share these data with other researchers or students, including those at my institution, except for those who have been approved by the above researcher and who have signed their own copy of this agreement.

I understand that if I leave my current institution, I must apply directly to the NSFH Director for permission for continued access to these data.

2. I will take reasonable precautions to preclude persons not authorized to use the geographic information from having access to it. The data files with this information will be maintained in a secure manner, with access restricted to authorized persons.

3. I will make no attempt to identify either individual respondents or the geographic location of respondents for any purpose, including the appending of additional geographic data to the respondent records. I understand that any such attempt may void permission to use the geographic characteristics data, and that if I, or others, become aware of such an attempt, we are obligated to report it to the NSFH director.

4. Reports or publications deriving from these data will not identify specific geographic locations of NSFH respondents or groups of respondents.

_____
  Project Staff Member's Signature   Please Print

_____   _____
              Title   Date

Approved by Primary Investigator

_____   _____
      Primary Investigator's Signature   Date

CONFIDENTIALITY AGREEMENT page 4

Office Use Only:

Application Approved

_____   _____
              NSFH Study Director   Date

Staff performing merge:

Date:

Data sent:

Notes: