

Appendix B

Linkages to Administrative Data Sources

B.1 Develop Linkages with Administrative Data Sources

RTI will develop linkages with existing data sources to supplement the 2012 National Postsecondary Student Aid Study (NPSAS:12) interview data. RTI recognizes the great value added to the NPSAS:12 data file with the addition of data from specific administrative data sources. Certain data (for example, specific financial aid amounts and associated dates) can only be accurately obtained from sources other than the student or parent. Through our experience providing data for many National Center for Education Statistics (NCES) postsecondary studies, including previous NPSAS studies, Beginning Postsecondary Student (BPS), and Baccalaureate and Beyond (B&B), we have gained considerable knowledge performing file merges with many existing sources of valuable data, including Department of Education's (ED) Central Processing System (CPS) for Free Application for Federal Student Aid (FAFSA) data, the National Student Loan Data System (NSLDS), and ACT. For this study, we propose to perform file merges with the following datasets: CPS, NSLDS, American College Testing (ACT), SAT (the College Board), and National Student Clearinghouse.

The Family Educational Rights and Privacy Act (FERPA), (34 CFR Part 99) allows the disclosure of information without prior consent for the purposes of NPSAS:12 according to the following excerpts: 99.31 asks "Under what conditions is prior consent not required to disclose information?" and explains in 99.31 (a) "an educational agency or institution may disclose personally identifiable information from an education record of a student without the consent required by 99.30 if the disclosure meets one or more specific conditions. The NPSAS:12 collection falls under:

Sec. 99.31 (a)(3). The disclosure is, subject to the requirements of Sec. 99.35, to authorized representatives of--

- (i) The Comptroller General of the United States;
- (ii) The Attorney General of the United States;
- (iii) The Secretary; or
- (iv) State and local educational authorities.

NPSAS:12 is collecting data under the Secretary's authority. The personally identifiable information is collected from student record systems with adherence to the security protocol detailed in 99.35: "What conditions apply to disclosure of information for Federal or State program purposes?"

- (a)(1) Authorized representatives of the officials or agencies headed by officials listed in Sec. 99.31(a)(3) may have access to education records in connection with an audit or evaluation of Federal or State supported education programs, or for the enforcement of or compliance with Federal legal requirements that relate to those programs.
- (2) Authority for an agency or official listed in Sec. 99.31(a)(3) to conduct an audit, evaluation, or compliance or enforcement activity is not conferred by the Act or this part and must be established under other Federal, State, or local authority.
- (b) Information that is collected under paragraph (a) of this section must:
 - (1) Be protected in a manner that does not permit personal identification of individuals by anyone other than the officials or agencies headed by officials referred to in paragraph (a) of this section, except that those officials and agencies may make further disclosures of

personally identifiable information from education records on behalf of the educational agency or institution in accordance with the requirements of Sec. 99.33(b); and

(2) Be destroyed when no longer needed for the purposes listed in paragraph (a) of this section.

(c) Paragraph (b) of this section does not apply if:

(1) The parent or eligible student has given written consent for the disclosure under Sec. 99.30; or

(2) The collection of personally identifiable information is specifically authorized by Federal law.

As part of initial sampling activities, we will ask participating institutions to provide SSNs for all students on their enrollment list. Having an initial list of all student SSNs minimizes the time and burden on both the institutions and RTI. Institutions will only have to provide one enrollment list; if the SSNs were provided only for those students selected, the institutions would have to provide two separate enrollment lists. Immediately after the student sample is selected, the SSNs for non-selected students will be securely discarded.

Secure Data Transfers. NCES has set up a secure data transfer system, using their NCES member site with Secure Sockets Layer (SSL) technology, described above. We will use this electronic system for submitting data containing potentially identifying information (such as SSNs, names, and dates of birth of our sample members) along with their survey ID (not the same ID that is available on the restricted-use data). Before being transmitted, files will be encrypted using FIPS 140-2 validated encryption tools. We will receive data from the NCES system as well. The system requires that both parties to the transfer be registered users of the NCES Members Site and that their Members Site privileges be set to allow use of the secure data transfer service as described above. This process will be used for all file matching procedures described below, except in instances when the vendor already has a secure data transfer system in place.

B.2 File Merge with ED Central Processing System (CPS)

RTI proposes to perform file merges with the CPS data containing federal student aid application information. The merge with CPS can occur at any time for any number of cases, provided that the case has an apparently valid SSN associated with it. RTI sends a file to CPS and receives in return a large data file containing all students who applied for federal aid. We already have existing programs and procedures in place to prepare and submit files according to rigorous CPS standards. Similarly, we have already developed programs and procedures to receive and process data obtained from CPS.

RTI will electronically upload a file on the FAFSA secure web-site for matching which contains SSN and the first 2 letters of the sample member's last name (but no other information). Access to the site for the upload is restricted to authorized users who are registered and provide identification/authentication information (SSN, DOB, and personal identification number [PIN]) to the FAFSA data site. The file is retrieved by the Central Processing System or CPS (the FAFSA contractor data system) for linkage. The linked file, containing student aid applications for matched records, is then made available to us only through a secure connection (EdConnect) which requires username and password. All CPS files will be processed, edited, and documented for inclusion on the analytic data files.

The CPS data obtained for the NPSAS:12 data collection will cover academic years beginning with 2010–2011. All CPS files will be processed, edited, and documented for inclusion in the Electronic Codebook (ECB).

B.3 File Merge with National Student Loan Data System Disbursement

RTI will also conduct a file merge with the NSLDS to collect federal loan and Pell grant data for the 2010–2011 award year. The resulting file will contain cumulative amounts for each student’s entire postsecondary education enrollment. NCES has set up a secure data transfer system that uses their NCES member site and Secure Sockets Layer (SSL) technology. The system requires that both parties to the transfer be registered users of the NCES Members Site and that their Members Site privileges be set to allow use of the secure data transfer service. These privileges are set up and carefully controlled by the ED’s Institute of Education Sciences (IES) NCES Chief Technology Officer (CTO). This service has been designed by ED/NCES specifically for the secure transfer of electronic files containing personally identifying information (i.e., data protected under the Privacy Act or otherwise posing risk of disclosure) and can be used for NCES-to-Contractor; Contractor-to-Subcontractor; Subcontractor-to-Contractor; and Contractor-to-Other-Agency data transfers. The party uploading the information onto the secure server at NCES is responsible for deleting the file(s) after the successful transfer has been confirmed. Data transfers using this system will include notification to the ED/IES, the NCES CTO, and the NCES Deputy Commissioner as well as the ED/NCES project officer. The notification will include the names and affiliations of the parties in the data exchange/transfer and the nature and approximate size of the data to be transferred. We have programs to create the files for the merge and also programs to read the data we receive. All matching processes are initiated by RTI staff providing a file with one record per sample member to be merged.

B.4 File Merge with ACT

To obtain valuable admissions test data, RTI will perform a file merge with American College Testing (ACT) data. RTI has the same experience file matching for ACT data as described in the sections on CPS and NSLDS file merging. We successfully used this matching process for NPSAS since NPSAS:96.

Matching of students to ACT requires Social Security number (SSN), name and date of birth to assist the data vendor in performing confirmatory data quality checks. RTI will initiate this process by providing a file with one record per SSN. We will use the same procedures described above for the NSLDS linkage: creating a password-protected, encrypted file using FIPS 140-2 validated encryption tools; uploading the electronic file on the NCES server for pick-up by ACT. ACT will process the data on their database and provide the matched data on the NCES server for our secure download. This file merge will only be conducted during the full-scale study.

B.5 File Merge with the College Board

To obtain SAT data, RTI will perform a file merge with the College Board. Matching of students to SAT data requires Social Security number (SSN), name and date of birth to assist the data vendor in performing confirmatory data quality checks. RTI will initiate this process by providing a file with one record per SSN. We will use the same procedures described above for the NSLDS linkage: creating a password-protected, encrypted file using FIPS 140-2 validated

encryption tools; uploading the electronic file on the NCES server for pick-up by the College Board. The College Board will process the data on their database and provide the matched data on the NCES server for our secure download. This file merge will only be conducted during the full-scale study.

B.6 File Merge with the National Student Clearinghouse

RTI will use the National Student Clearinghouse to obtain the *Student Tracker* data on institutions attended, enrollment dates, and degree completions for the NPSAS:12 sample.

RTI will first set up an account with the Clearinghouse which will enable sending and receiving of files securely over encrypted FTPS connections. The file containing sensitive student identifiers (name, date of birth, and Social Security number) will be encrypted using FIPS 140-2 validated encryption tools then submitted to the Clearinghouse using their secure FTP site. All files received by the Clearinghouse will be securely stored using FIPS 140-2 validated AES encryption, the US federal encryption standard. Matched files, containing data on enrollment dates, institution names, and degrees completed, will be returned to RTI using the same secure FTP site. This file merge will only be conducted during the full-scale study.

B.7 Processing Administrative Data

With the exception of the early matching activities to improve FTB identification and one initial merge with CPS data prior to computer assisted data entry (CADE), we propose to send files for matching after CADE data collection to ensure the availability of the maximum number of verified Social Security numbers and to facilitate the batch mode processing that is suitable to many of these resources. We may need to match to a source (for example, CPS or NSLDS) more than once.

The data from all of these sources, as allowed by the vendor, will be delivered in the ECB and will be useful for creating derived variables. These derived variables will also be available on Powerstats and Quickstats. Both direct-pull and derived variables will be documented thoroughly for the ECB, Powerstats, and Quickstats.