

Appendix G

Linkages to Extant Data Sources

G.1 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. The CPS has recently migrated to an Internet website for file matching, and RTI has been successful sending and receiving files using the new application.

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:08 data collection will cover academic years beginning with 2006–2007. All CPS files will be processed, edited, and documented for inclusion in the Electronic Codebook (ECB).

G.2 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 2006–2007 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.

G.3 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 during NPSAS:96, NPSAS:2000, and NPSAS:04.

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 WinZip 10; 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.

G.4 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 WinZip 10; 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.

G.5 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:08 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 WinZip 10 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.

G.6 Processing Extant Data

With the exception of 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 the Data Analysis System (DAS). Both direct-pull and derived variables will be documented thoroughly for both the ECB and DAS.