Administrative Data Matching

Use of Administrative Records in
NCES Secondary and Postsecondary

Sample Surveys

Kristin Dudley

Overview of the Presentation

- NCES studies requiring administrative record matching
- Sources of administrative data
- Purposes and uses of the data
- Performing a match
- Match and response rates

NCES Studies Requiring Administrative Record Matching

Study	Cohort	Survey Years
Education Longitudinal Study (ELS:02)	10th grade students as of the spring of 2002	2002, 2004, 2006, 2012
Beginning Postsecondary Students Longitudinal Study (BPS:04)	Students who first began postsecondary education in 2003-04; identified in NPSAS:04	2004, 2006, 2009
2007-08 National Postsecondary Student Aid Study (NPSAS:08)	Undergraduate, graduate, and professional students enrolled in Title IV institutions in 2007-08	2008
Baccalaureate and Beyond Longitudinal Study (B&B:08)	Students who completed baccalaureate degrees in 2008-09; identified in NPSAS:08	2008, 2009, 2012
High School Longitudinal Study (HSLS:09)	9th grade students as of the fall of 2009	2009, 2012

RTI INTERNATIONAL

Sources of Administrative Data: Publicly Available Data

Source	Description of data source	Data elements
Integrated Postsecondary Education Data System (IPEDS)	A system of NCES surveys that collects aggregated data from all postsecondary institutions that participate in federal financial aid programs.	Institutional characteristics; price; enrollment; degrees and certificates; financial aid; retention and graduation rates; human resources and finances
Common Core of Data (CCD)	NCES survey that annually collects data on all elementary and secondary public schools in the nation.	School characteristics; enrollment; staff and student descriptive/demographic information; fiscal data
Private School Universe Survey (PSS)	NCES survey that collects data on all elementary and secondary private schools every two years.	School characteristics; total enrollment; number of high school graduates; enrollment by gender; number of teachers employed; program emphasis

RTI INTERNATIONAL

Sources of Administrative Data: Education Financing and Enrollment

Source	Description of data source	Data elements
Central Processing System (CPS)	Free Application for Federal Student Aid (FAFSA)	Personally-identifying info; college code; demographics; high school information; grade level; degree type; parent/student income; dependency status; EFC
National Student Loan Data System (NSLDS)	Federal loan and Pell grant history files	PII; school codes; loan/grant amount; loan/grant status; disbursement and deferment dates; enrollment period
National Student Clearinghouse (NSC)	Institutions voluntarily submitting academic information to the NSC	College code; term-level enrollment by school; major; class level; degrees earned

RTI INTERNATIONAL

Sources of Administrative Data: Institution Records

Source	Description of data source	Data elements
Student Records from Institutions	Collected directly from postsecondary institutions as a part of NPSAS.	Student contact info and demographics; enrollment; major; financial aid; budget
High School Transcripts	Collected directly from high schools as a part of ELS:2002	Courses taken; grades; credits in Carnegie units; GPA; test scores
Postsecondary Transcripts	Collected directly from postsecondary institutions attended by both BPS:04 and B&B:08 cohort members as part of the 2009 Postsecondary Education Transcript Study (PETS:09).	Courses taken; grades; credits earned; GPA; major; degrees

RTI INTERNATIONAL

Sources of Administrative Data: Standardized Tests

Source	Description of data source	Data elements
SAT	SAT survey and score data – College Board	SAT scores; honors and number of years by subject; education experience/plans; activity experience; demographics
ACT	ACT survey and score data	Demographics; high school information; ACT scores/subscores; educational plans
AP	AP exam score data – College Board	AP exam scores by subject
GED	GED survey data – American Council on Education	Demographics; test passed and date; reason for taking test; preparation for test

Why Match to Administrative Databases?

- Obtain data from the best, most accurate source

- Screen or identify specific types of students
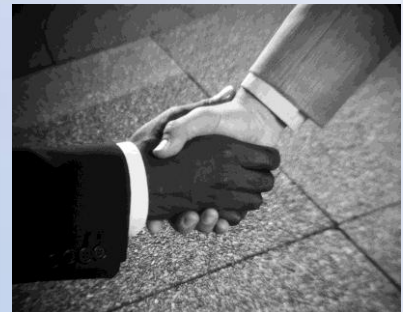
- Skip interview items in order to decrease burden

Why Match to Administrative Databases?

- Provide a reliable base for data comparison

- Collect data on interview nonrespondents

- Build on existing data and fill in missing data points

- Use in conjunction with similar data from other sources to create derived or composite variables

Performing a Match: Building Relationships and Processes

- NCES initiated relationships with key staff for each federal data source; RTI has maintained those relationships, as well as those with private sources, over time.

- Often needs are similar across studies and existing file formats and processes can be used.

- Maintain continuity with institution staff and state systems for student record and transcript collections.

Performing a Match: Systems and Processes Used

- Publicly-available data: data are downloaded and securely stored

- Student records from institutions: data are requested and securely uploaded or entered directly into a secure web application

- Transcripts: data are requested, securely uploaded or faxed, and entered into an application

- FAFSA, NSLDS, and NSC: utilize established and secure systems for automatic processing

- Standardized test data: secure server at NCES is used for data transfer

Performing a Match: Processing of Results

- Ensure that the number of records returned is as expected.

- If personally identifying information is returned from the matched database, verify that a correct match was made.

- Edit and reformat data to better suit the needs of the study and derived variable development.

Performing a Match: Data Use Agreements and Release of Data

- Signed data use agreements are often required by the agency or vendor.

- NCES requires that all non-government staff members with access to sample member data sign affidavits of nondisclosure.

- Some organizations will not allow their data to be released (even for restricted use) in its raw form. Data must be combined with similar data from other sources to create composite variables.

Match Results: CPS, NSLDS Loan and Pell Files, NSC

Database	NPSAS:08 Percentage of eligible students matched	B&B:08/09 Percentage of sample members matched	BPS:04/09 Percentage of sample members matched	ELS:2002 Percentage of sample members matched
CPS	68.6 (2007-08)	23.8 (2008-09)	19.9 (2008-09)	40.0 (2004-05)
NSLDS Loan	60.1	74.6	59.6	29.9
NSLDS Pell	47.4	51.6	47.9	22.9
NSC	75.4	94.4	58.0	n/a

Match Results: Tests

Database	Study	Percentage of sample members matched
SAT	NPSAS:08	35.3
ACT	NPSAS:08	21.5
AP	ELS:2002	14.2
GED	ELS:2002	3.5

Response Rates: Institution Records

Data	Study	Percentage of sample members with associated data
High school transcripts	ELS:2002	91.4
Postsecondary transcripts	B&B:08/09	94.2
Postsecondary transcripts	BPS:04/09	91.5
Student records	NPSAS:08	95.3

Summary

- NCES studies do take advantage of additional sources of data that can enhance interview data.

- These data can be used for a variety of different purposes to add value, increase accuracy, and decrease burden.

- Arranging for and implementing administrative data matching vary widely.

Questions?

More Information

Kristin Dudley

RTI International

marvill@rti.org

Jim Rogers

RTI International

jerogers@rti.og