National Student Clearinghouse

The National Student Clearinghouse (NSCH) (http://www.studentclearinghouse.org) is an organization that collects student degree and enrollment verification information from numerous colleges and universities in the United States. As of March, 2011, the NSCH obtains enrollment and completion data from 3,300 college and universities in the United States. The colleges that share data with the NSCH enroll 92% of all college students. Enrollment information from the enrolled post-secondary institutions is electronically relayed to the NSCH at multiple points through the year. Enrollment information is at least sent at the beginning and the end of each academic term. Not all students that are enrolled in a participating university or college are included in the NSCH database because some opt out of having their information remain private.

The UWBHS matched the senior survey respondents to the NSCH database based upon the student names and dates of birth. As name and date of birth is not a unique identifier there is the possibility that some of the records were incorrectly matched. However, in instances in which it seemed obvious that the information noted in the NSCH database did not match the respondent noted in the UWBHS database (e.g. completed college one year after the student attended college) the cases were not considered matches. The NSCH database contains the FICE codes for each of the schools attended by the students, which are easily matched to the rest of the Carnegie Codes.

For the purpose of this analysis we examined whether the respondents had enrolled in college from June 1st of the year that they graduated high school to June, 1st of the following year. We opted for this time frame as it roughly matches the time frame of the one year follow up, and it allows for a comparison. Obviously, the NSCH data, as it was requested in the fall of 2008, can be used to examine college enrollment and completion for all cohorts through 2008 (e.g. it is possible to examine a four year college graduation for all cohorts, save the 2005 cohort). Additionally, the NSCH has information on post-BA enrollment as well.

There was a maximum of 7 colleges matched for any student. We include the following variables for each match. You will see descriptions for the 1st college matched. The numeric value within the variable name changes for other colleges (for example, C2ENSTAT, C3ENSTAT).

C1ENSTAT From Nat Stu Clear:Status College 1

```
C1ENSTAT  From Nat Stu Clear:Status College 1

         Not matched                  2956
         A  Leave of absence             4
         F  Full-time                 2602
         H  Half-time                  714
         L  Less than half-time        327
         W  Withdrawn                  336
         Z  Information not provided  2713
         Total                        9652
```

C1MOBEG From Nat Stu Clear:Month Enrolled Beginning Coll 1
(Valid values of 1 through 12)

C1MOEND From Nat Stu Clear:Month Enrolled End Coll 1
(Valid values of 1 through 12)

C1NYEAR From Nat Stu Clear:Type of Coll 1 (2yr/4yr)

```
         Not Matched          2956
         2  2 year            3658
         4  4 year            3032
         L  Less than 2 year  6
         Total                9652
```

C1PUBLIC From Nat Stu Clear:College 1 is Public (Pub vs Priv)

```
  -999  Not matched in clearinghouse  2956
     0                                1325
     1                                5371
 Total                                9652
```

C1YRBEG From Nat Stu Clear:Year Enrolled Beginning Coll 1
(Valid values of 2000 through 2008)

C1YREND From Nat Stu Clear:Year Enrolled End Coll 1
(Valid values of 2000 through 2009)

C1COLSTg Clearinghouse State of college 1
(Code for state college is in)

The Clearinghouse provided information on the year that degrees were granted. These variables are:

YEARHS From Nat Stu Clear:Year of High School Gr
YEARCERT From Nat Stu Clear:Year of Certification
YEARAA From Nat Stu Clear:Year of AA
YEARBA From Nat Stu Clear:Year of BA
YEARMA From Nat Stu Clear:Year of Masters
YEARPHD From Nat Stu Clear:Year of PhD

The following variables are summary measures created by UWBHS staff.

The following variables indicate the number of years to degree. The SPSS syntax that can create the variables is included (YrToDeg YearBA ba6yr ba5yr ba4yr) . We collected the data in 2008, so not enough time had elapsed for some of the cohorts to calculate that a degree was not earned in 4, 5 or 6 years.

-- ba4yr is valid for the 2000, 2002 to 2004 cohorts. (2005 cohort is considered missing).
-- ba5yr is valid for the 2000, 2002 to 2003 cohorts. (2004 & 2005 is considered missing).
-- ba6yr is valid for the 2000 and 2002 cohorts. (2003, 2004 & 2005 cohorts are considered missing).

```
*** finished college -- from Clearinghouse data ***.
missing values YearBA (-999).
compute YrToDeg = YearBA  -  Year .
Variable label YrToDeg 'Years to Degree' .
if missing(YrToDeg ) YrToDeg = -999 .
value labels YrToDeg
   -999 'Match not found in Clearinghouse Data' .
missing values YrToDeg (-999) .
execute .

* Here a dummy variable is created (ba4yr) with a value of 1 for those who
received a BA in four years and zero for those who did not.
if year le 2004  ba4yr = 0 .
if year gt 2004  ba4yr = -1099 .
if YrToDeg le 4  ba4yr = 1 .
variable label   ba4yr 'fin coll in 4yrs-from yearba, year yeartodeg'.
```

```
* Here a dummy variable is created (ba5yr) with a value of 1 for those who
received a BA in five years and zero for those who did not.

if year le 2003  ba5yr = 0 .
if year gt 2003  ba5yr = -1099 .
if YrToDeg le 5  ba5yr = 1 .
variable label   ba5yr 'fin coll in 5yrs-from yearba, year yeartodeg'.


* Here a dummy variable is created (ba6yr) with a value of 1 for those who
received a BA in six years and zero for those who did not.

if year le 2002  ba6yr = 0 .
if year gt 2002  ba6yr = -1099 .
if YrToDeg le 6  ba6yr = 1 .
variable label   ba6yr 'fin coll in 6yrs-from yearba, year yeartodeg'.

execute .
missing values YrToDeg YearBA  () .

do repeat
 a = YrToDeg YearBA  ba6yr ba5yr ba4yr .
if YearBA = -999 a = -999 .
end repeat .

value labels ba4yr ba5yr ba6yr
  -999 'Match not found in Clearinghouse Data'
 -1099 'Missing because not enough years passed to test' .
```

The final two Clearinghouse variables created by UWBHS staff are:

CMAXSEQ From Nat Stu Clear:Max Number of Colleges attended
This is the number of colleges matched for each student – values range from 1 to 7.

CMXRECNM From Nat Stu Clear:Max number or records for person
This is the total number of attendance records found for a student. It represents the number of quarters or semesters that the students may have been enrolled. Valid values range from 1 to 35.

Applying Carnegie Codes to the College Attendance (F03a) and the NSCH Data

In responding to questions F03a students noted the name and location of the college(s) that they were attending/had attended since graduating from high school. In an effort to protect the respondent confidentiality the name, city, and state of college listed have been removed from the data file. However, the UWBHS team recoded the colleges listed in these items using the Carnegie Classifications. The Carnegie Classification of Institutions of Higher Education™ is the most widely recognized reference work of all colleges and universities in the United States. (The Carnegie Classification codes included in the UWBHS file come from the Carnegie Classifications Data File, as prepared May 30, 2006). For more information on the Carnegie variables, click http://depts.washington.edu/uwbhs/webdocs/Carnegie_memo.pdf .

The major advantage of using the Carnegie Classifications is that they provide a systematic and detailed classification of nearly all of the post-secondary institutions in the United States along 15 important dimensions that help to classify post-secondary institutions. The college that the student reported in the follow up survey has a value as calculated by Carnegie Classification for Institutions of Higher Education for each of 15 attributes that describe the school (accreditation type, level of institution, degree of urbanization, control, etc). In order to recode the respondents text response (name of the school) into these specific numerical values, the name of the school that the student reported on question F03a or the name of the school noted by the NSCH is matched to those found in the Carnegie Classifications data file (see http://depts.washington.edu/uwbhs/memos/carnALL05.xls) and then the values given by the Carnegie

Classifications data file are recoded into the new variable. Thus, for the school that a student listed in F03a or is listed by the NSCH, provided it is included in the Carnegie Classification, a host of variables that describe the specific school are available and included in the data file.

A small percentage of students were attending multiple schools at once. For example they would take one course at one school and the rest of their courses at another school. For the responses from the follow up, as we only coded one school, we would preference four year schools over two year schools. If both schools were of the same type (e.g. both two year or four year) we coded the school that was listed first.

For the data received from the NSCH, we examined all of the schools that the student had enrolled in (and attended for at least one day) during their first year after high school with a year defined as 6/1/year of senior survey to 6/1/one year post year of senior survey. The maximum number of schools that a student had enrolled in during this year was 3 schools (N=23). In the instances in which students had enrolled in more than one school in the year following high school (N=563) we gave preference to the school that they had most recently attended. Of the students that attended multiple schools in the NSCH data file, 44% attended the same level of institution (e.g. attended multiple two year schools), 42% transferred from a two year school to a four year school, and 14% transferred from a four year school to a two year school.

Matching NSCH data to UWBHS—using college attendance one year post-HS as an example. The cross-tabulation below illustrates the extent to which the student responses to the follow up and the results in the NSCH overlap. As you can see by examining the cross-tabulation, the two data sources provide complementary information. Specifically, for 82% of the cases (that have information from both sources) the data sources provide the same information in regards to college attendance and type of college attended (2 year, 4 year, none) during the year after high school. The dark blue boxes and circles indicate the 7,308 students (37 + 2,236 + 2,218 + 2,854 = 7,308) for whom the information from the two sources agree. Below is a discussion of the various groups for whom the data from the two sources is not in complete unison.

One of the reasons that the responses may vary across data sources is that the neither source is perfect—the follow up survey captured 92% of the original respondents, while the NSCH only covers 92% of the enrolled post-secondary population. Additionally, students, if they are not receiving financial aid, can request that their school records not be shared with the NSCH, so they would not show up in NSCH database.

maxf03 Max followup school attending, carn+NonCarn * CollType College Type from Clearinghouse Crosstabulation

			CollType College Type from Clearinghouse				
			-1 college in HS, none post	2	4	No coll	Total
maxf03 Max followup school attending, carn+NonCarn	-1.00 Not In follup	Count	8	150	70	542	770
		% of Total	.1%	1.6%	.7%	5.6%	8.0%
	.00 No Coll Rept	Count	37	197	37	2236	2507
		% of Total	.4%	2.0%	.4%	23.2%	26.0%
	2.00 2 year coll	Count	11	2218	39	503	2771
		% of Total	.1%	23.0%	.4%	5.2%	28.7%
	4.00 4 year coll	Count	41	109	2854	606	3610
		% of Total	.4%	1.1%	29.6%	6.3%	37.4%
Total		Count	97	2674	3000	3887	9658
		% of Total	1.0%	27.7%	31.1%	40.2%	100.0%

***Red Box*—One year follow-up non-respondents**: This group includes the 770 respondents that did not complete the UW-BHS one year follow up survey. The NSCH data notes that 29% of these respondents were enrolled in college one year after high school. Clearly the college attendance rate for this group is much lower than it is for the students that completed the one

year follow-up and the other respondents included in the NSCH sample. One benefit of using the NSCH measure is that it would include these students in the sample of valid response cases. As these cases were unable to be contacted for the follow-up there is no follow up data on college attendance, so they are considered missing in the follow-up measure.

***Orange Boxes*—Data sources agree on enrollment but differ on level:** It is possible that both data sources may be correct in that the student was enrolled in both types of schools. This would be particularly possible amongst students that were enrolled in multiple locations and students that transferred schools. It is quite possible that the date of the follow up indicated the students enrollment at a given point in time, while the school noted in the NSCH data come from a later (or earlier) point in time.

***Pink Box*—Student enrolled in college during HS but not in year post-HS:** These are students that reported attending college in the one year follow-up, however, the NSCH notes that the student attended college but they first enrolled while they were in high school. A review of the administrative records (for the students in district 1) indicates that virtually all of these respondents were in the Running Start program (Running Start is a program that allows high school kids to take courses at a local community college or university for high school and college credit—for more information click http://www.k12.wa.us/runningstart/default.aspx). Perhaps in answering the follow up survey these students misreported their college enrollment while they were in high school, as college enrollment in the year after high school. Possibly these students noted that they wanted to keep their enrollment information private which would prevent their college/university from sending the information to NSCH.

***Aqua Box*—Student enrolled in college in NSCH but not in the follow-up:** In the follow-up survey the student does not report attending a post-secondary institution, but the NSCH notes enrollment for these students. One explanation is that the follow up occurred before the student enrolled in college (e.g. we interviewed the student in January or February and they enrolled in March/spring term which the NSCH notes). Additionally, it is possible that some of these were proxy responses that were incorrect or errors in matching the NSCH to the UWBHS data (we matched on student name and date of birth, as the combination of name and DOB is not a truly unique ID it is possible that we incorrectly matched a few cases).

***Green Boxes*—Student Reports Attendance in follow-up but none noted in the NSCH:** These are cases in which the student notes enrollment in the one year follow-up but we are unable to find an enrollment record in the NSCH database. A couple of plausible explanations exist. Possibly these respondents are misreporting their college attendance in the follow up. Additionally, it is possible that these students noted that they wanted to keep their enrollment information private which would prevent their college/university from sending the information to NSCH. Alternate methods of enrollment verification for a subset of these students found that many of these students were enrolled.

In sum, the data looks to be pretty good. The advantage of clearinghouse is that we can look at 4 year completion, etc.

Citations

Day, Jennifer C. and Eric C. Newburger. 2002. “The Big Payoff: Educational Attainment and Synthetic Estimates of Work-Life Earnings.” *Current Population Reports* P23-210. Washington, D.C.: U.S. Census Bureau.

Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. “The Education and Early Occupational Attainment Process”. *American Sociological Review* 34:82-92.

Sewell, William H., Archibald O. Haller, and George W. Ohlendorf. 1970. “The Education and Early Occupational Attainment Process: Replication and Revision”. *American Sociological Review* 35:1014-1027.

Sewell, William and Robert Hauser. 1975. *Education, Occupation, and Earnings.* New York: Academic Press.