Proposed Derivations for Data Collection and Analysis

WICHE Multistate Data Exchange

January 9, 2012 ~ Boise, Idaho
Defining Derived Variables

• Two types of project derived variables:
  – Those that will be defined by WICHE and collected directly (as specified) from the state
  – Those that will be defined from existing data once WICHE gets the data set (we are likely to make additions based on data set)

• State derived variables:
  – Variables that a state or institution defines that it may be helpful to WICHE to collect
Some Things to Keep in Mind

• Research Questions
  – What are the patterns of postsecondary enrollment and employment of high school graduates from each participating state?
  – What are the patterns of postsecondary enrollment and employment of students and former students in public postsecondary institutions in participating states?
  – By more fully accounting for individual mobility across state lines, to what extent does sharing data among states supplement existing state data resources available for conducting evaluations leading to policy and program improvements?

• Data are likely to inform additional derived variables related to a students status (all can be derived as long as term credits are collected by institution)
Data Files to be Collected

1. Cohort files
2. Enrollment files
3. Award files
4. UI Wage Records
# Data Files Contents: Cohort Files

**Variables**
- Social Security Number
- State/sector student ID
- First name
- Middle Name
- Last Name
- Suffix
- Birth Date
- Sex
- Race/ethnicity
- ACT/CEEB code of high school
- Date of high school diploma
- Institutional IPEDS UnitID
- Institutional OPE ID
- Date of first postsecondary enrollment

**Comments**
- Two cohorts: high school graduates from 2004-05; first-time postsecondary students in 2005-06
- All instances of each variable to be included for matching
- State resolves inconsistencies for analysis
- Selection criteria for inclusion in cohort overlaps with enrollment and awards files
- State determines whether secondary or postsecondary data are used for selection into cohort
- Elements in green and red are there strictly to assist with matching records – green for Cohort A; red for Cohort B
Data Files Contents: Enrollment File

Variables
Social Security Number
State/sector student ID
Institutional IPEDS Unitid
Academic Term Start Date
Academic Term End Date
CIP Code for First Program or First Major (6-digit)
Academic calendar type
Instructional activity hours type (credit, contact)
Instructional activity hours attempted
Instructional activity hours completed
Total cumulative credits earned at start of term
Student level (Undergraduate, graduate)
Pell recipient
Degree-seeking status
Full-time status

Comments
• All data to be supplied by term (fall, spring, summer). Term is based on whatever academic calendar applies.
• Selection of students enrolled in non-standard terms to be made according to specific dates, i.e., include students enrolled on MMDDYYYY
Data Files Contents: Awards File

Variables
Social Security Number
State/sector ID Number
High school (CEEB code)
High school graduation date
High school diploma type (Regular or G.E.D.)
Institutional IPEDS UnitID
Postsecondary academic award level (certificate, associate’s, bachelor’s, etc.)
Postsecondary academic award date
CIP for postsecondary academic award (6-digit)

Comments
• This file may actually be two files or more files, depending on whether the data are merged across agencies/data sources within the state
• States submit data for all awards conferred on each student; analysis will rely on highest award earned and on most recent award earned
Enrollment Period

• Need a standard way to collect and examine enrollment over time
• Two dichotomous variables
• Collect by academic term
  – If student is enrolled as of Oct. 1 → enrolled in fall
  – If student is enrolled as of Feb. 15 → enrolled in spring
  – If student is enrolled as of July 1 → enrolled in summer
• Analyze by academic year (trailing summer)
  – If student enrolled in any of fall, spring, summer → he/she is enrolled during that academic year
• Enrolled is true if instructional activity hours attempted > 0
Continuous Enrollment

• Derived variable
• Equal to 1 if student is enrolled in all academic years under examination
Enrollment Status Postsecondary

• CEDS Element: Full-time/part-time status
• Definition: The status given to a student in relation to the course load that he or she is carrying
• Dichotomous using Pell guidelines (i.e., full-time is true if credits attempted in a term is \( \geq 12 \) for credit hour programs
Enrollment Status Postsecondary Analysis Derivations

- **Timing**
  - Define full-time during first-term (cohort approach)
  - Average credit hours attempted in a term for all semesters enrolled (average)
  - Divide into: exclusively full-time, mixed full- and part-time, exclusively part-time (federal longitudinal sample survey)
  - Modal enrollment status from among fed. long. categories - **PROPOSED**

- **Enrollment across institutions**
  - Determine students’ full-time status based on their credit hours attempted in a single institution
  - Determine students’ full-time status based on their credit hours attempted summed across institutions - **PROPOSED**

- **Institutional Differences**
  - Collect institution specific flag for full-time status where available

- **Other**
  - Examine students who are more than full-time (greater than 16 Attempted credit hours in a term)
Degree-Seeking Status

• CEDS Element: Degree or Certificate-Seeking Student
• Definition: Individual is enrolled in courses for credit and recognized by the institution as seeking a degree, certificate, or other formal award. High school students also enrolled in postsecondary courses for credit are not considered degree/certificate-seeking
• Collect the element as captured in state data
• “Behavioral” definition (for analysis and comparison) - PROPOSED
  – Student beginning at a two-year college attempts at least 12 credits within their first two academic years OR attempts at least 6 credits in their first term at a single institution OR receives a degree or certificate as a degree-seeking student
  – Students beginning at a four-year college are assumed to be degree-seeking unless they take fewer than 6 credits in their first academic year AND are not found in the data in subsequent years
• Additional considerations:
  – Distinguishing remedial and ESL students
  – Is there a clear course-based proxy, i.e., any math course?
Pell Recipient

• CEDS Elements: Title IV Participant and Recipient status (Yes, No) and Financial Aid Award Type (List includes Pell)

• Additional Clarification:
  – Use Financial Aid Award Type to determine if a student receives a Pell award during each term. If they do, Pell Recipient = Yes

• Collect by term. For analysis, if student ever received Pell in the first two academic years, his/her Pell recipient status is true - PROPOSED
Institution Type

- **PROPOSED** – use Carnegie Basic 2010 classifications as follows:
  - Associate’s colleges
  - Baccalaureate colleges
  - Master’s colleges and universities
  - Doctorate-granting universities

- Students concurrently enrolled in multiple institutions are “assigned” to the one where they are attempting the majority of credits; if equal, assignment is based on where they’d previously been but, if no prior record exists, they will be randomly assigned to one of the institutions they attended that term
Student Mobility/Pathways Through Postsecondary Education

- **PROPOSED** – Students “assigned” to enrollment pathways as follows.
  - Single transfer (courses taken a la carte at other institutions ignored for this assignment)
    - Traditional transfer
    - Reverse transfer
    - Lateral transfer
  - Empirically examine pathways of students who attend more than two institutions, as well as a la carte course taking.
- Examine these pathways for cross-state activity
Limitations on Employment Analysis

• Without *Hours worked* or *Wage rate* information, unable to derive employment level/intensity, e.g., PT or FT
  – As such, not able to attribute changes in wages over time to wage rate—increases in Gross wages could be from increase in hours worked

• Not possible to measure educational attainment effect on employment/earnings with no control group. Limited descriptive/correlation analysis for Awardees v. Non-awardees
Present in the State

• From UI wage records, if individual is present in the state file (and gross wages are > 0), state presence is true.

• State presence does not mean an individual resides in that state, but he/she is nevertheless contributing to that state’s economy.
Employment

- Employed at any given point:
  \[ Y = \text{Gross wages} \] for the UI quarter not blank/missing, greater than $0
  \[ N = \text{Gross wages} \] blank, missing or $0

- Over time:
  Continuous = \( \text{Gross wages} \) in 20 or more of 25 UI quarters
  Intermittent = \( \text{Gross wages} \) in 5-19 of the 25 UI quarters

- Employment status after award (within 2 UI quarters after date of award)*
  \[ Y = \text{Gross wages} \] reported in at least 1 of 2 UI quarters after date of award
  \[ N = \text{No Gross wages} \] reported within 2 UI quarters after date of award
  - Note: if award dates are lagged, will be imperfect
Employment (cont.)

• Employed upon enrollment?
  Wages reported in UI quarter that corresponds to initial date of enrollment? (Y/N)

• History of employment while enrolled
  No. and % of terms enrolled and at the same time Gross wages reported in corresponding/overlapping UI quarter
  – History of employment, and wage level as a proxy for intensity of employment, might then be compared to educational progress

Primary analyses for Employment status: a) Numbers, b) Proportions and c) selected other Distribution analyses
Wage Levels

• **Mean reported quarterly wages**
  – Analyze by a) All employed students, b) Awarded v. Not awarded, c) NAICS and/or CIP groupings
  – And other enrollment, award or employment category or pattern ??

• **Wage changes within 2 UI quarters of points of interest**
  – a) After award
  – b) After last date of enrollment

• **Wage changes and rate of wage changes over time**
  – For a subset of students who were last enrolled/awarded midway through study period?
  – Elapsed-time-to-employment after Award/Last date of enrollment (particularly, students not employed most or all of enrollment)

• **Annual mean wages**

• **Distribution of income (very exploratory in nature)**
  – Is there a threshold level or levels to examine? (i.e., 200% of poverty level, state median per capita income)
  – Can we reasonably break gross wages out by quintile? Or a normal distribution?
Location of Employment

• State of enrollment compared to State of employment? At any given point for individuals where Employment Status=Y, is Enrolled_State = UI_State? (Y/N)

• State of employment at end of enrollment, compared to initial state of residence
  Is UI_State = State of residence at start of cohort
  – Analyze by student's status at last point of enrollment, e.g.,"Award" or "Not Award"

• Migration patterns
  Distribution of students by Sending state (initial cohort) and Receiving state (State where wages were reported on or after last date of enrollment)
Field of Employment

• **Industry of employment**
  – NAICS code clustered by sectors of interest (e.g., State’s EcDev foci, STEM, highest wage sectors, etc.)

• **Compared to field of study**
  – Crosswalk and comparison of NAICS and CIP codes, e.g., Perkins IV Pathways Crosswalks, CEW Clusters
    • For CIP code contiguous to last date of enrollment, or most frequently indicated, if more than one across all enrolled periods

• May be analyzed at two points:
  – a) Ever
  – b) At end of enrollment, or upon award

• To a limited extent, may be able to do some comparison of wage levels by NAICS