

NEGLECTING THE “L” IN A LONGITUDINAL INTEGRATED DATA SYSTEM¹ CAN BE A COSTLY MISTAKE²

Submitted to:

U.S. Department of Labor
Employment and Training Administration
Office of Policy Development and Research
Division of Strategic Planning and Performance
200 Constitution Avenue NW
Washington DC 20210

Submitted by:

Grace Fendlay
Director, Discretionary Grants
Division of Workforce Development and Adult Learning
Maryland Department of Labor, Licensing and Regulation
1100 North Eutaw Street
Baltimore, MD 21201

January 2013

¹ Throughout this report, the term longitudinal *integrated data system* is used instead of *P-20W SLDS* that refers to early childhood (pre-school, or P) through postgraduate education (20) and workforce (W) coverage of a State Longitudinal Data System (SLDS). While acknowledging the U.S. Department of Labor’s funding of our research, through the Workforce Data Quality Initiative (WDQI) that supports aggressive development of the “W” component of *P-20W SLDS* initiatives, the new information presented here has wider applicability than the P-20W SLDS initiatives *per se*.

² The authors of this report are Ting Zhang, PhD, Research Assistant Professor, and David Stevens, PhD, Research Professor, The Jacob France Institute, University of Baltimore. John Janak, Sang Truong and Jing Li participated in the data processing conducted for this research. The Institute is a sub-award recipient of Workforce Data Quality Initiative (WDQI) funds received by DLLR from the U.S. Department of Labor, Employment and Training Administration. The authors accept full and sole responsibility for the content of this report. Agreement or disagreement with the views expressed here should not be attributed to any other person or organization.

EXECUTIVE SUMMARY

The title of this report—*Neglecting the “L” in a Longitudinal Integrated Data System Can be a Costly Mistake*—is intended to send a clear message to the funders, designers and managers of P-20W SLDS initiatives that immediate short-term coverage capabilities only permit access to tip-of-the-iceberg return on investment rewards. Hidden from view in the early years of these initiatives is the long-term value that can be achieved *if* attention is paid to the importance of sustained “L” coverage.

Valid public concern about administrative record confidentiality and permissible use is recognized by a prudent integrated data system manager as a warning that inattention to content can result in irreversible untoward consequences. One expression of public concern has been definition of a maximum time span that integrated data system content can be maintained. A related concern has been expressed through restriction of data use to an original defined purpose. Our research design responds to both the time limit and original purpose restrictions by exploring how opportunities to achieve public and private benefits may be lost without an offsetting gain in public comfort about the integrity of integrated data system security.

The content of this report builds on a foundation of multiple continuing collaborations and interagency partnerships that began in 1990 with authorized access to confidential administrative records for all 1984 fall semester public community college first-time enrollees in Maryland, initially matched with the State’s UI wage records for the years 1985-1990.

Linkage of former community college student administrative records with a state’s UI wage records, using a Social Security number identifier, is not new. Pioneering initiatives began in the 1970s. Our contribution here is new because we received authorized access to additional confidential administrative data sources that expand the scope of actionable insights that a robust integrated data system offers.

Summarizing what we have learned to date, the combination of Maryland UI wage record coverage, out-of-state UI wage record coverage, federal civilian employee coverage, University System of Maryland data files, social safety-net program coverage, and workforce program coverage, resulted in our ability to still find relevant information about at least 75% of the 1984 community college enrollees in 2011, after more than 25 years have elapsed. We hasten to add that at no time has the identity of an individual former student been disclosed as our research has proceeded. All retained records are encrypted and findings are carefully reviewed for potential disclosure risk prior to release.

The interim results we report here alert newcomers to the presentation of employment and earnings information that concurrent education and employment statuses can have important impacts on what is observed. Similarly, sequential education engagements, such as community college followed by a four-year university, that are not properly documented and taken into account can result in inaccurate conclusions and actions. This mistake can impact how return on investment and other *outcome* attributions are made and acted upon.

Our conclusions focus on making a strong case for maintenance of strategically constructed integrated data systems over extended time spans; spans that are long enough to observe downstream events and statuses that constitute teachable lessons pointing the way to actions that can be expected to have positive individual and social impacts.

High on our list of long time span coverage and robust data source content benefits that are derivable from a longitudinal integrated data system is improvement of rate of return and outcome estimates associated with education and training investments. We show that 1984 Maryland public community college enrollees appear in multiple confidential administrative databases over the next 27 years. In addition to those that continued their education in the University System of Maryland, some of these engagements have been with social safety-net programs; some appearances indicate movement out-of-state; and others document participation in *workforce* programs, undefined further up to this point.

Strong interest is expressed in two topics that we will study next, using the updated longitudinal integrated data system:

1. Public, business and educator concern is expressed that public postsecondary education entities in the U.S. are not preparing enough science, technology, engineering and mathematics (STEM) graduates to satisfy current and projected demand for these competencies.
2. Widespread concern is expressed that students unfortunate enough to have graduated since 2006 will suffer lasting untoward employment and earnings effects traceable to their immediate struggle to establish an appropriate entry-level affiliation.

Neither of the two topics described above can be adequately studied in the absence of a robust longitudinal integrated data system capability covering a time span of at least 5 years; and preferably more years as time passes.

TABLE OF CONTENTS

| | |
|--|----|
| INTRODUCTION | 1 |
| WHEN AND HOW THE ADMINISTRATIVE DATA SOURCES WERE ASSEMBLED | 2 |
| SELECTED ADMINISTRATIVE DATA SOURCE DETAILS | 6 |
| Maryland public community college first-time enrollees fall 1984 | 7 |
| University System of Maryland (USM) | 7 |
| State UI wage records | 7 |
| Maryland employment, training and labor exchange transactions | 8 |
| Other employment | 9 |
| Social safety net programs | 10 |
| Summary of the <i>Figure 2</i> data sources | 10 |
| INTEGRATED DATA SYSTEM INSIGHTS | 11 |
| CONCLUSIONS | 22 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1: A Robust Integrated Database | 3 |
| Figure 2: Components of the Integrated Data System Used in this Report | 5 |
| Figure 3: Percentage of Participants by Source | 12 |
| Figure 4: Number of Participants by Selected Source | 14 |
| Figure 5: UI and SNAP Participants | 16 |
| Figure 6: Median Earnings | 18 |
| Figure 7: The 1984 Community College Enrollees' Appearance in Other Sources | 19 |
| Figure 8: Appearance in Other Sources for the University System of Maryland Enrollees who were the 1984 Community College Enrollees | 20 |
| Figure 9: Appearance in Other Sources for the University System of Maryland Degree Holders who were the 1984 Community College Enrollees | 21 |

INTRODUCTION

Valid public concern about administrative record confidentiality and permissible use is recognized by a prudent integrated data system manager as a warning that inattention to content can result in irreversible untoward consequences. A wise manager exercises strategic caution when deciding what data sources and data fields within each source are going to be integrated for an authorized purpose.

One expression of public concern has been definition of a maximum time span that integrated data system content can be maintained. A related concern has been expressed through restriction of data use to an original defined purpose.

Our research design responds to both the time limit and original purpose restrictions by exploring how opportunities to achieve public and private benefits may be lost without an offsetting gain in public comfort about the integrity of integrated data system security.

Two criteria motivated the research design that led to this interim³ report:

- Demonstrate the importance of extended integrated data system time coverage to find individual statuses and transition movements among education enrollment, engagement in employment and related training and labor exchange activities, and episodes of social safety-net program participation.
- Show by example how revealed concurrent status and transition flow combinations can be translated into timely actions that offer a favorable prospect for public and private cost savings and quality of life improvements for those served.

Two unrelated meanings of the three-letter acronym SSN—Social Security number and social safety net—are valuable elements when an integrated data system is intended for use to identify opportunities for strategic policy or administrative actions. The importance of a Social Security number identifier for some linkage applications is well known, but unevenly promoted as a necessity.⁴

³ This is an *interim* report because several steps remain to arrive at the end-point that was defined at the outset of the research project. These remaining steps are described in the last section of this interim report.

⁴ See Zhang, T. and D. Stevens (2012), *Integrated data system person identification accuracy requirements and methods*, Baltimore, MD: University of Baltimore, The Jacob France Institute, 19 pp.; <http://www.jacob-france-institute.org/documents/MD-WDQI-Person-Identification-Report.pdf>.

Acknowledged less often is the relevance of social safety net program administrative data for some integrated data system uses. The federal SLDS and WDQI investments to date have not specifically targeted inclusion of these data⁵.

Figure 1 on the next page sets the stage for the remainder of this report.⁶ Three features stand out. First, *Figure 1* conveys a clear sense of the number of relevant administrative databases that exist in each state; and more are relevant but do not appear here. Second, *Figure 1* highlights the many transition flows that are possible among the education, workforce and social safety-net statuses that are shown; flows in multiple origin-destination directions and sequence orders. And third, the content of *Figure 1* contains an unstated time dimension, potentially over many years, as we demonstrate in the following pages.

The next section offers a brief overview of the organizational and individual collaborations over more than two decades that enabled us to undertake the research reported on here. This is followed by a more in-depth treatment of the individual data files that have been integrated so far. We then turn to presentation of selected combinations of educational attainment, engagement in employment and related training and labor exchange activities, and participation in social safety net programs. A final section pulls together what we consider to be decision-relevant findings to date and how we intend to fill some remaining data gaps to extend the policy and program management 'reach' of our research.

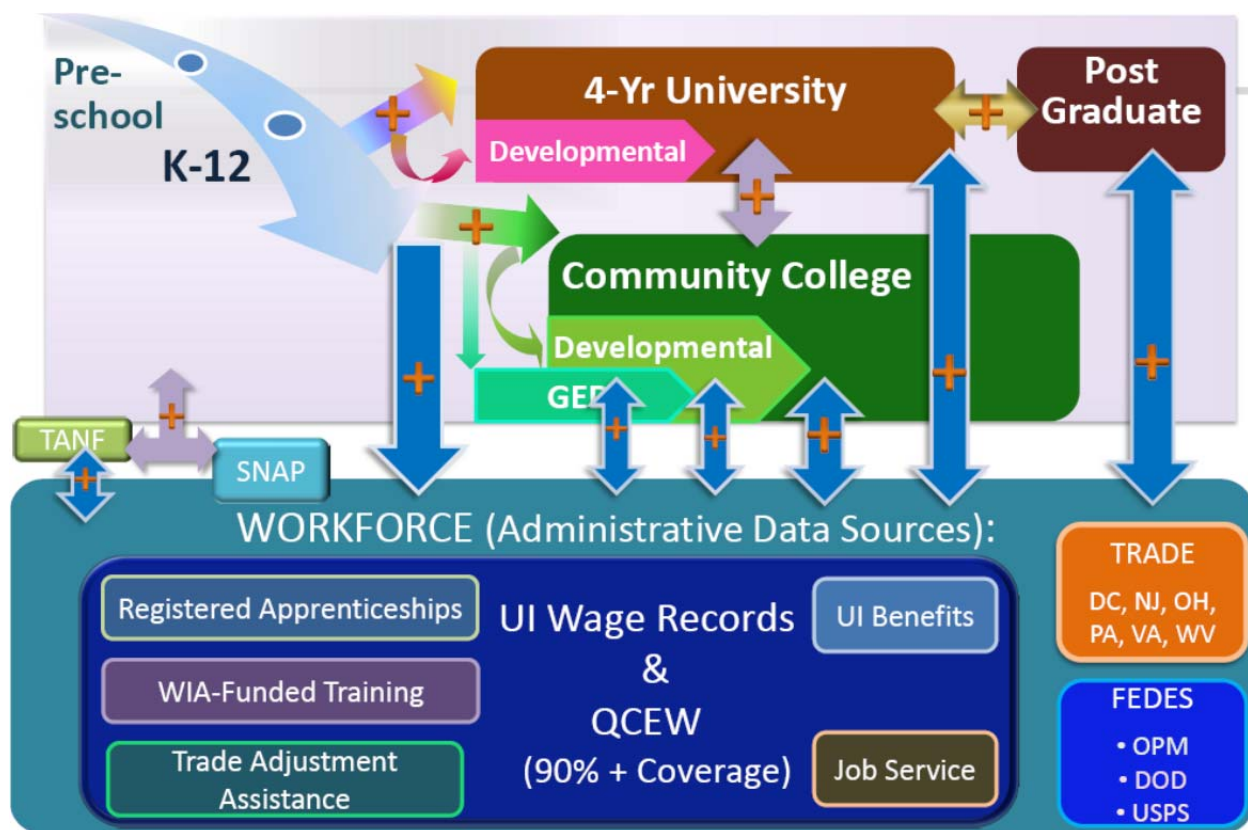
WHEN AND HOW THE ADMINISTRATIVE DATA SOURCES WERE ASSEMBLED

The first of a series of collaborations over more than two decades that enabled our research to proceed occurred in 1990. This was one year before the independent Maryland State Board for Community Colleges (1969-1991) was absorbed into the Maryland Higher Education Commission (1988-).

⁵ See: <http://nces.ed.gov/programs/slids/> and <http://www.doleta.gov/performance/workforcedatagrants09.cfm>.

⁶ This *Figure*, originally designed in 2011 by Ting Zhang, and then updated in June 2012, has received wide circulation in national forums.

FIGURE 1: A Robust Integrated Database



“+” indicates possibility of concurrent and/or sequential life-cycle participation in these activities.

Source: The Jacob France Institute, University of Baltimore (June 2012).

The State Board’s Director of Planning and Research in 1990 entered into an agreement with one of us to conduct a research study using two administrative data files:

- The Board maintained a secure confidential administrative database for the 28,876 first-time 1984 fall semester enrollees in Maryland’s public community colleges.
- Also in 1990, one of us, in partnership with the Maryland Department of Economic and Employment Development (1987-1995)⁷, maintained a secure confidential database of quarterly Maryland Unemployment Insurance (UI) wage

⁷ In 1995 the Maryland Department of Economic and Employment Development (DEED) was split into two state agencies—the Maryland Department of Business and Economic Development (DBED) and the Maryland Department of Labor, Licensing and Regulation (DLLR). Both continue today.

records, beginning with the April-June 1985 file. This database continues to be updated as each new quarter's information becomes available.⁸

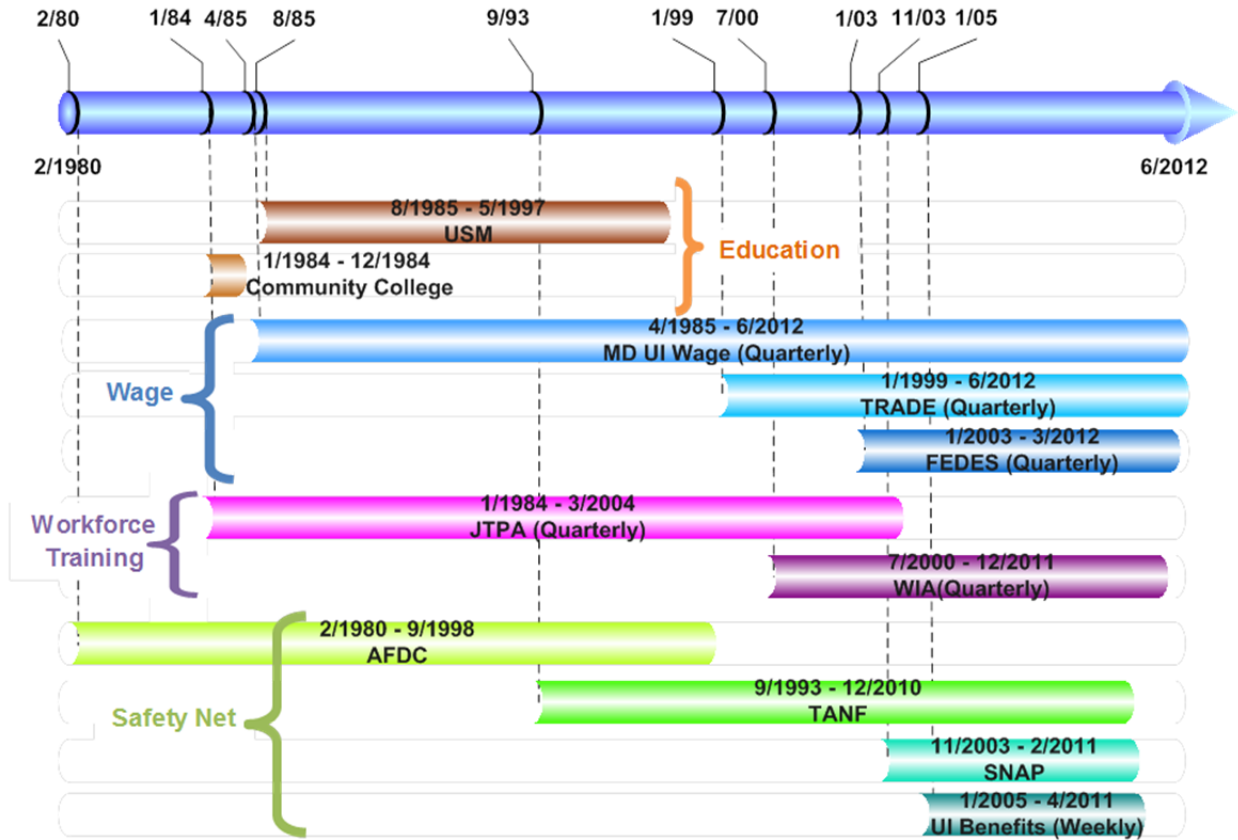
Authorization to conduct research using this file was initially sought from 5 of the State's 17 community colleges. Subsequently, in 1994, authorization to conduct new and continuing research was sought and received from all 17 of the community colleges. Since 2005, The Jacob France Institute has maintained a *Memorandum of Understanding* with the Maryland Association of Community Colleges that includes all of the State's public community colleges as members.

Linkage of former community college student administrative records with a state's UI wage records, using a Social Security number identifier, is not new. Pioneering initiatives began in the 1970s. The number of states engaged in administrative data linkage of this type then grew at an uneven and slow pace until 2006. That is when federal SLDS funds began to accelerate state commitments to create a basic P-20W SLDS capacity.

Our contribution is new because we received authorized access to additional confidential administrative data sources that expand the scope of actionable insights that a robust integrated data system offers. *Figure 2* on the next page shows the time span covered for each of the 11 administrative data sources available to us for processing and analysis in this report up to this point.

⁸ The Jacob France Institute at the University of Baltimore is defined in an interagency MOU as a subordinate agent of the Maryland Department of Labor, Licensing and Regulation, for the purpose of maintaining the Maryland Unemployment Insurance Wage Records database for DLLR authorized performance accountability, research and evaluation uses. The two party MOU is renewed annually with a new authorized scope of work and list of Institute staff members that are allowed to receive and process unit-record DLLR data. All files are encrypted and maintained in an off-line secure Institute facility. Individually identifiable information is not disclosed in any reports that are released to the public.

FIGURE 2: Components of the Integrated Data System Used in this Report



Additional data sources will be added in 2013 and some time coverage gaps will be partially or fully eliminated for selected available data sources. Details appear in the final section of this report.

Eleven administrative data sources⁹ appear in *Figure 2*:

Maryland public postsecondary education

- Maryland public community college first-time enrollees fall semester 1984.
- University System of Maryland Enrollment Information System (EIS) and Degree Information System (DIS) extracts, 1985-1997.

⁹ The University System of Maryland (USM) source actually contains two databases—Enrollment Information System (EIS) and Degree Information System (DIS) extracts. Maryland Job Service data are maintained, but this source and time coverage is not included in *Figure 2*.

Maryland employment, training and labor exchange transactions (DLLR)

- Maryland UI wage records, 1985 2nd quarter-2012 2nd quarter.
- Job Training Partnership Act (JTPA) transactions, January 1984-March 2004.
- Workforce Investment Act (WIA) transactions, July 2000-December 2011.

Other employment

- The Regional Area Data Exchange (TRADE) UI wage record exchange among DC, MD, NJ, OH, PA, VA and WV, January 1999-June 2012.¹⁰
- Federal Employment Data Exchange System (FEDES), federal government civilian personnel and U.S. Postal Service employee record extracts, January 2003-March 2012.¹¹

Social safety net programs

- Aid for Families with Dependent Children (AFDC), February 1980-September 1996.
- Temporary Assistance for Needy Families (TANF)¹², October 1996-December 2010.
- Supplemental Nutrition Assistance Program (SNAP), November 2003-February 2011.
- Unemployment Compensation, or Unemployment Insurance (UI) benefit data, January 2005-April 2011.

SELECTED ADMINISTRATIVE DATA SOURCE DETAILS

In the first paragraphs of this report we described a basic goal of the report as advancement of the business case for long-term integrated data system integrity and

¹⁰ A particular state's participation has not necessarily been consistent for each quarterly cycle of exchange throughout this time span.

¹¹ Active duty military personnel data extracts are also received through the FEDES, but only for defined subpopulations that satisfy the restricted use stipulations in the MOU between the Maryland Department of Labor, Licensing and Regulation and the Department of Defense.

¹² Maryland's official terminology is *Temporary Cash Assistance* (TCA) administered by the Family Investment Program of the Family Investment Administration, Maryland Department of Human Resources.

strategic attention to content parsimony. Unstated until now, parsimony should be defined with respect to each intended application of an integrated data system capability. What data fields, with what content quality, are needed to achieve a specific objective?

We turn in this section to brief descriptions of the data sources and content fields that were drawn upon to compile and present the findings that follow in the remainder of the report. Links to more information about some of these sources appear in footnotes.

Maryland public community college first-time enrollees fall 1984

The Maryland State Board for Community Colleges administrative records contain data fields that were not needed, or used, to create the integrated data system that enabled us to complete this interim report. The basic fact of fall 1984 enrollment is all that was drawn from this much more robust database.

University System of Maryland (USM)

The separate Enrollment Information System (EIS) and Degree Information System (DIS) administrative files¹³ each contain many data fields. However, again, to satisfy our immediate research need, only a 1984 community college cohort member's appearance in either of these files, or both, in a reference year between 1985 and 1997 was extracted for inclusion in our integrated data system.

Updating of USM EIS and DIS information through 2012, and consideration of possible extraction of other data fields, remains for future negotiation with USM officials. The underlying principle going into such negotiations is always *need to know*—honoring the parsimony guide, making a strong case for why additional information is needed.

State UI wage records¹⁴

UI wage record information is collected and maintained by each state to administer the state's unemployment insurance program. This is the sole original purpose for collecting the information. The rules governing other access rights and privileges appear in the Federal Code of Regulations (CFR)¹⁵ and relevant state-specific

¹³ <http://data.mhec.state.md.us/MACInfo/MACManuals/EIS%202012%20Manual.pdf>;
<http://data.mhec.state.md.us/MACInfo/MACManuals/DIS%20for%20independents.pdf>.

¹⁴ See: Stevens, David W. (2007), *Employment that is not covered by state unemployment insurance laws*, Technical Paper No. TP-2007-04, Suitland, MD: Census Bureau, LEHD Program, 37 pp., available at: <http://lehd.did.census.gov/library/techpapers/tp-2007-04.pdf>; and U.S. Department of Labor (April 2012), *Unemployment Compensation Federal-State Partnership*, Washington, DC: Office of Unemployment Insurance, Division of Legislation, 16 pp., available at: <http://workforsecurity.doleta.gov/unemploy/pdf/partnership.pdf>.

¹⁵ The relevant Code of Federal Regulations reference is Title 20, PART 603—FEDERAL-STATE UNEMPLOYMENT COMPENSATION (UC) PROGRAM; CONFIDENTIALITY AND DISCLOSURE OF STATE UC INFORMATION, available at: <http://www.ecfr.gov/cgi-bin/textidx?c=ecfr&SID=4712ca010b58c1639cf81d0ca454ff9&rgn=div5&view=text&node=20:3.0.2.1.3&idno=20>.

legislation, laws and regulations that may apply. Today's applicable rules are subject to change by legislative, legal or administrative action. The treatment of UI wage record access in the context of integrated data system design and practice requires constant due diligence to become and remain informed about current laws, regulations and administrative practices.

The content of state UI wage records is controlled by each state, although the content is very similar among the states. A full nine-digit Social Security number identifier is included in all cases. Beyond the uniform Social Security number data field, some state UI wage record content differences are found.¹⁶

Each UI wage record that has been submitted by or on behalf of a reporting entity includes a reference year/quarter designation, the reporting entity's state-specific unemployment insurance tax account number, an employee's Social Security number, and the dollar amount that was paid by the reporting entity to this employee during the reference year/quarter.

For our immediate research use, design of the limited purpose longitudinal integrated data system required only that we link fall 1984 Maryland public community college first-time enrollees' Social Security numbers with Maryland UI wage records beginning with the earliest available reference year/quarter of April-June 1985 and continuing through the April-June 2012 reference cycle. For each year/quarter person-specific match found, the sum¹⁷ of reported earnings was added to the longitudinal integrated data system being assembled.

Maryland employment, training and labor exchange transactions

The Job Training Partnership Act (JTPA) and Workforce Investment Act (WIA) administrative records include a combined 27 years of continuous coverage between 1984 and 2011. The content of each is extensive and changed over this time span.

For our immediate research purpose, we have only included an annual *participated* data field indicating whether a 1984 community college enrollee cohort member appears in a particular calendar year's JTPA or WIA administrative data files that are maintained by The Jacob France Institute in continued partnership with the Maryland Department of Labor, Licensing and Regulation.

¹⁶ None of these interstate differences impact our research design, but some might be of interest in the design of future integrated data system content. States differ in what combination of full or partial surname, first name and middle initial appears in a UI wage record. This can be important for probabilistic matching and identity validation purposes. A few states include reported hours or weeks worked during a reference year/quarter. And some states include a North American Industry Classification System (NAICS) code for each reporting entity.

¹⁷ Necessary when more than one employing entities have reported earnings for a reference employee.

DLLR's Workforce Exchange IT capability has expanded in recent years, and enhancements continue. *Figure 2* does not show the Employment Service component of this dynamic IT capability. However, we are interested in when and how often 1984 community college enrollee cohort members appear as registrants with Maryland's public labor exchange over subsequent decades.

Other employment

The Maryland Department of Labor, Licensing and Regulation and The Jacob France Institute continue as partners to collect additional employment information; that is, beyond the coverage of the Maryland UI wage records database. One of these additional data sources is regional and the other is international in scope.

Since 1999, DLLR and its counterpart state agencies in the District of Columbia, New Jersey, Ohio, Pennsylvania, Virginia and West Virginia¹⁸ have participated in a quarterly cycle of UI wage record exchange for authorized purposes defined in an interagency MOU that has been renewed during the ensuing years.

The interstate initiative is known as The Regional Area Data Exchange (TRADE).¹⁹ The only administrative data that is delivered through the secure TRADE portal is the sum of reference year/quarter reported earnings found in one or more of the seven participating state UI wage record files based on an individual Social Security number.

DLLR and the Institute are also partners in managing the national U.S. Department of Labor-funded Federal Employment Data Exchange System (FEDES).²⁰ DLLR has negotiated *Memoranda of Understanding* with the federal government Office of Personnel Management (OPM), the Department of Defense (DOD), and the U.S. Postal Service (USPS). States that have negotiated their own *Memorandum of Understanding* with DLLR then participate in quarterly cycles to receive pre-defined administrative data fields from OPM, DOD and the USPS based on individual Social Security numbers that satisfy the authorized use stipulations of the OPM, DOD and USPS *Memoranda of Understanding*.

¹⁸ As we noted earlier, not every state has participated continuously in each year/quarter cycle of data exchange. Since each year/quarter cycle includes the most recent two years of a participating state's UI wage record coverage, a gap in participation only matters in principle if the non-participation continues for more than seven quarters, which would mean that the oldest non-participation quarter's information could not be recovered in the next new cycle of participation.

¹⁹ There is continuing discussion among the participating states about whether the recent pilot program WRIS.2, which broadens the scope of permissible uses of UI wage record information exchanged on a voluntary basis among participating states, will negate a continued need for the TRADE capability. Information about WRIS.2 is available at: <http://www.doleta.gov/performance/wris2.cfm>. Currently, Maryland, New Jersey and Pennsylvania are the only three of the seven TRADE participant states that are included in the continuing pilot phase of WRIS.2.

²⁰ Detailed information about the FEDES is available on-line at: <http://www.ubalt.edu/jfi/fedes/>.

Once again, for our immediate research purpose we included a yes-no data field code in construction of our integrated data system for each year/quarter of available FEDES coverage—January-March 2003 through January-March 2012.

Social safety net programs²¹

Four social safety net programs appear in *Figure 2*. Program administrative data extracts are received by The Jacob France Institute based on long-standing and periodically renewed *Memoranda of Understanding* and Interagency Agreements. These extracts have been designed, and modified upon mutual agreement of the parties involved, on a need to know basis.

For our immediate research purpose we again included a yes-no data field code in construction of our integrated data system for the time-based unit that is applicable for each of the data sources—monthly for the AFDC, TANF and SNAP program data; and weekly for the Unemployment Insurance benefit data.

Summary of the *Figure 2* data sources

The title of this report is *Neglecting the “L” in a Longitudinal Integrated Data System can be a Costly Mistake*. This is intended to send a clear message to the funders, designers and managers of P-20W SLDS initiatives that immediate short-term coverage capabilities only permit access to tip-of-the-iceberg return on investment rewards. Hidden from view in the early years of these initiatives is the long-term value that can be achieved *if* attention is paid to the importance of sustained “L” coverage.

One often hears the lament “if I had only known then what I know now”, which is a painful reminder that mistakes happen. We cannot change the past, but we can learn lessons from what has transpired, so our theme in the next section is: We will show by example how to apply what we know now to limit negative impacts of some future decisions.

Our two stated research goals are to:

- Demonstrate the importance of extended integrated data system time coverage to find individual statuses and transition movements.
- Show by example how revealed concurrent status and transition flow combinations can be translated into time administrative actions that offer a

²¹ The network of partners affiliated with the *Actionable Intelligence for Social Policy* (AISP) initiative headquartered at the University of Pennsylvania exemplifies how a strong business case built on sound benefit-cost principles can be promoted for sustained investment in robust integrated data systems that include social safety net content. See: <http://www.ispc.upenn.edu/documents/RecordQuality.pdf>; and <http://bethblinburydesign.com/intelligence/>.

favorable prospect for public and private cost savings and quality of life improvements for those served.

Fortunately, both of these goals are achievable here. Looking once more at *Figure 2*, it should now be clear that our longitudinal integrated data system has been constructed in stages over more than two decades.

As we begin to present and describe selected status combinations in the next section attention should be focused on *actionable* insights that emerge; insights that offer hope for positive future cost savings and quality of life improvements. What lessons can be taken away from our 1985-2012 perspective on education, workforce and social safety net engagements of 1984 fall semester first time enrollees in Maryland's public community colleges?

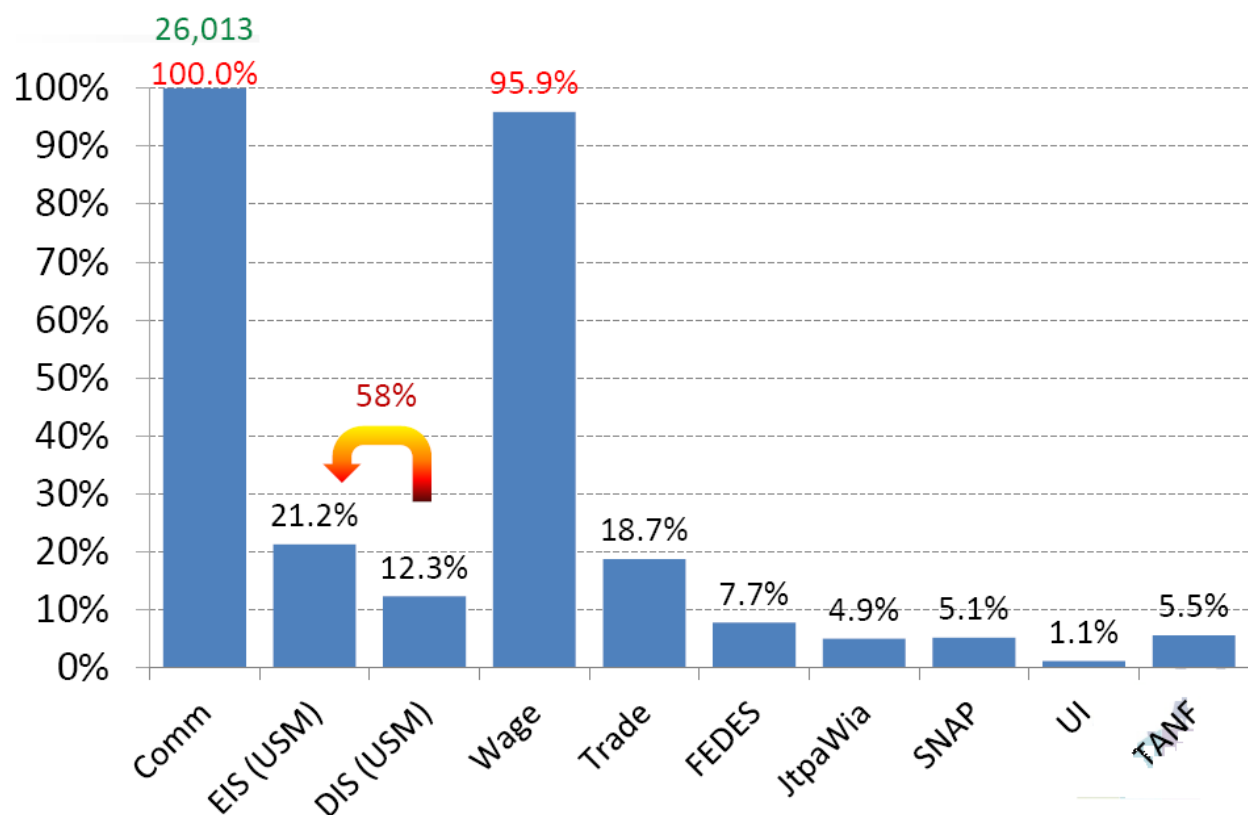
INTEGRATED DATA SYSTEM INSIGHTS

Figure 3 on the next page delivers an overview of the percentage of the 1984 Maryland community college first time enrollees that subsequently appeared in each of nine other administrative databases at some time between 1985 and mid-2012. Please note that the nine other administrative databases shown here split the USM data files into EIS and DIS, but merged AFDC with TANF data files and merged JTPA and WIA into one data file. Therefore, instead of ten other databases shown in *Figure 2*, there are nine other datasets shown in *Figure 3* and later figures. We begin with this figure because it reflects the time coverage gaps shown in *Figure 2*.

The left-to-right order of the *Figure 3* x-axis is intentional—starting with *education* statuses on the left, and then progressing sequentially through *employment* and *JTPA/WIA* participation, to safety net program engagements on the right segment of the x-axis. *Figure 3*, alone, does not shed light on the frequency of concurrent statuses, and it does not show the frequency of particular from-to transition combinations.

Figure 3 does show that over a decade long time span, 1985-1994, immediately following the community college enrollees first recorded entry into Maryland public postsecondary education, 21% had enrolled in a University System of Maryland member institution, and 58% of those that enrolled had been awarded a degree.

FIGURE 3: Percentage of Participants by Source



Over the much longer April 1985-June 2012 time span of Maryland UI wage record availability, 96% of the 1984 community college enrollees had been reported as an employee of at least one employing entity that is required to report covered employee earnings to DLLR.

Next, as we move on to the out-of-state (TRADE) employment data, and the federal civilian employee data (FEDES), irreversible gaps in time coverage of our integrated data system come into play in an accurate interpretation of what *Figure 3* shows. While almost one out of five (19%) of the 1984 community college enrollees has shown up in the UI wage records maintained by the District of Columbia, New Jersey, Ohio, Pennsylvania, Virginia and West Virginia, this coverage did not begin until 1999. We do not know if the incidence of this out-of-state employment status would have been higher if 1985-1998 data had been acquired and retained. We will explain in the final section of this report how a gap of this type limits an integrated data system's capacity to support some studies of policy and program management importance.

The time span coverage issue addressed in the previous paragraph for out-of-state employment is encountered again with regard to the federal civilian employee coverage topic. *Figure 3* shows only 8% of the 1984 community college enrollees as appearing in the maintained FEDES database, but the time span of currently available coverage in our integrated data system is January 2003 through March 2012. This enables us, in principle, to separate federal civilian government employment within Maryland from employment elsewhere in the world.

A governor, state legislators, educators, business leaders and taxpayers often ask: Are the students we educate in our state remaining here to work, and pay taxes?²²

The combined time coverage of JTPA/WIA in *Figure 3*, currently January 1984 through December 21011, shows that only 5% of the 1984 community college enrollees subsequently appeared as a registrant in either of these federally funded workforce programs. We consciously use the word *workforce* program, rather than *training* program, because more in-depth analysis of each program's database is needed to separate out the type(s) of engagement for each registrant. This additional step is necessary to answer some relevant policy and program management questions. We will return to this topic in the final section of the report.

We finally arrive at the three social safety net participation rates shown in *Figure 3*—Supplemental Nutrition Assistance Program (SNAP) benefits, Unemployment Insurance (UI) benefits, and Aid to Families with Dependent Children (AFDC)/Temporary Assistance for Needy Families (TANF) benefits. Time span coverage is very different here—1984 through 2010 for the sequential AFDC/TANF programs; November 2003 through February 2011 for the SNAP benefits; and January 2005 through April 2011 for UI benefits.²³

Thoughtful study of *Figure 2* and *Figure 3* generates more questions than answers, but this is what we intended. Provoking early thought is expected to focus attention and heighten interest in what follows in the remaining pages of this section.

Figure 4 on the next page adds important insights that advance us toward a point where actionable lessons learned come into view. Here, the x-axis is chronological time coverage from 1985 on the left to 2012 on the right. The vertical y-axis in this case shows the number of 1984 community college enrollees that appear in at least one of the administrative databases contained in our integrated data system. Maryland employment (UI wage record); out-of-state employment status (TRADE UI wage record);

²² Rate of return estimates that are based on projected lifetime earnings are often rejected because there is no confirmed factual basis for the *lifetime* trajectory and duration of in-state earnings. An integrated data system that offers long-term coverage can overcome this deficiency. Confidence in accurate rate of return estimates can translate into better policy and program management decisions.

²³ We expect to update each of these safety net data files in 2013 (and beyond).

and federal civilian employee status (FEDES) have been color coded to focus attention on several interpretation challenges.

Looking first at the red bars in *Figure 4* we see that after 27 years half of the 1984 community college enrollees are still being found in the Maryland UI wage record database. Using the year-of-birth information to examine the 1984 community college cohort, almost all (99.8%) enrollees were aged at least 18 in 1984, which means almost all of those that show up in the Maryland UI wage record database in 2012 would be at least 45 years old.

FIGURE 4: Number of Participants by Selected Source

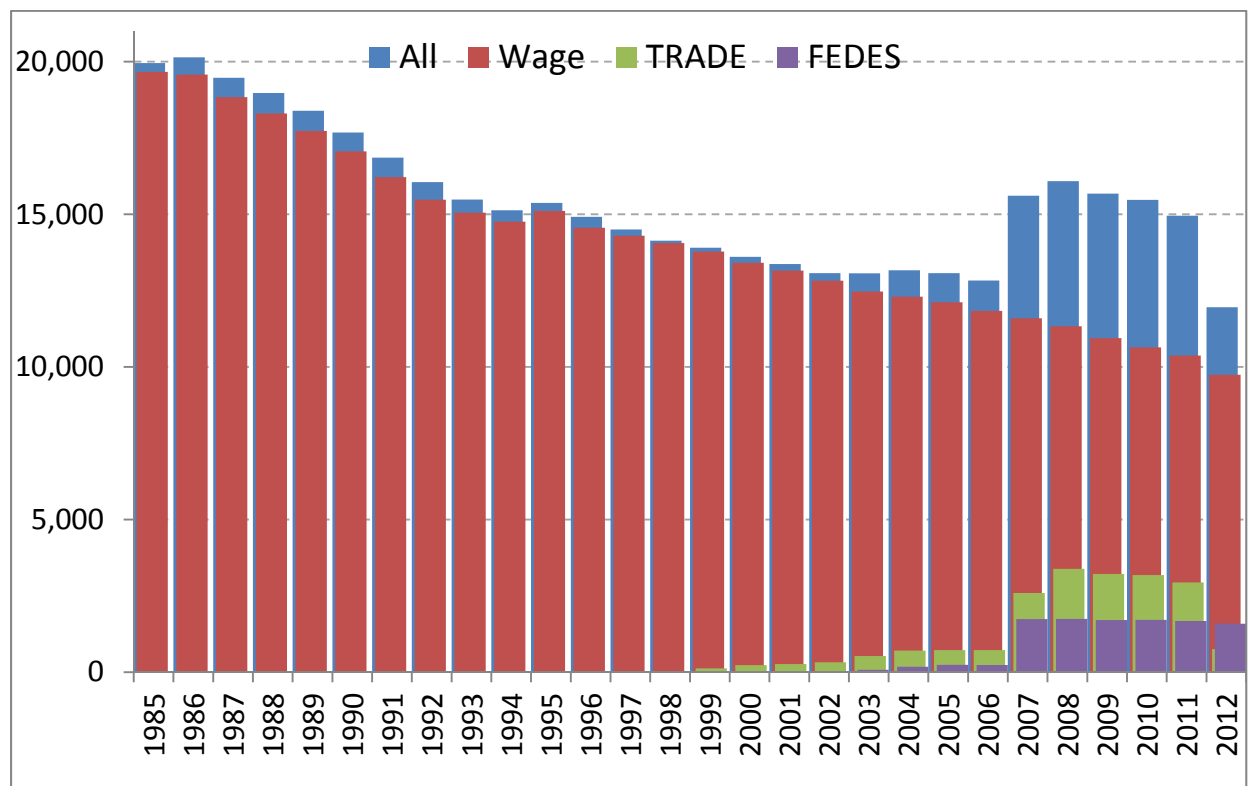


Figure 4 shows some community college cohort members' annual appearances in our longitudinal integrated data system (color coded blue) in addition to the Maryland UI wage record database (color coded red). Referring back to *Figure 2* we see that in the time span 1985 through 1997 our longitudinal integrated data system included appearances in the University System of Maryland Enrollment Information System (EIS) and Degree Information System (DIS), as well as in the JTPA database and the

AFDC/TANF database. So, during this 13 year time span any combination of these statuses during a calendar year is included in the count color coded blue.

Beginning in 1999, *Figure 4* shows community college cohort member appearances in the out-of-state UI wage record coverage; the TRADE database (color coded green). The count of participation in recorded out-of-state employment is shown to have increased in 2007 and remained high through the last year, 2011, of out-of-state employment data currently included in our longitudinal integrated data system. Because we have warned that the seven-state exchange of UI wage record information may not have included all of the states in every year/quarter cycle, care has to be exercised in speculation about the cause(s) of this increase and persistence. One plausible contributing factor is the December 2007-June 2009 recession.²⁴

Finally, *Figure 4* shows the appearance of community college cohort members in the database of federal civilian employees (FEDES), beginning in 2003. Here, too, we find a large increase in this status beginning in 2007 and continuing through the most recent month, March 2012, of FEDES data currently included in our longitudinal integrated data system.

Summarizing what we have learned from *Figure 4*, the combination of Maryland UI wage record coverage, out-of-state UI wage record coverage, federal civilian employee coverage, and the other administrative data sources, results in our ability to find at least 75% of the 1984 community college enrollees in 2011, after more than 25 years have elapsed.

What temporarily lies hidden beneath the surface of what is shown in *Figure 4*? We address this question next.

We speculated above that the December 2007-June 2009 recession is probably one contributing factor explaining the sudden increase and persistence of 1984 community college enrollee cohort members in our longitudinal integrated data system.

²⁴ The dating of this recession of 19 months was made by a committee of National Bureau of Economic Research (NBER) members. "The NBER does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales." See: <http://www.nber.org/cycles.html>.

FIGURE 5: UI and SNAP Participants

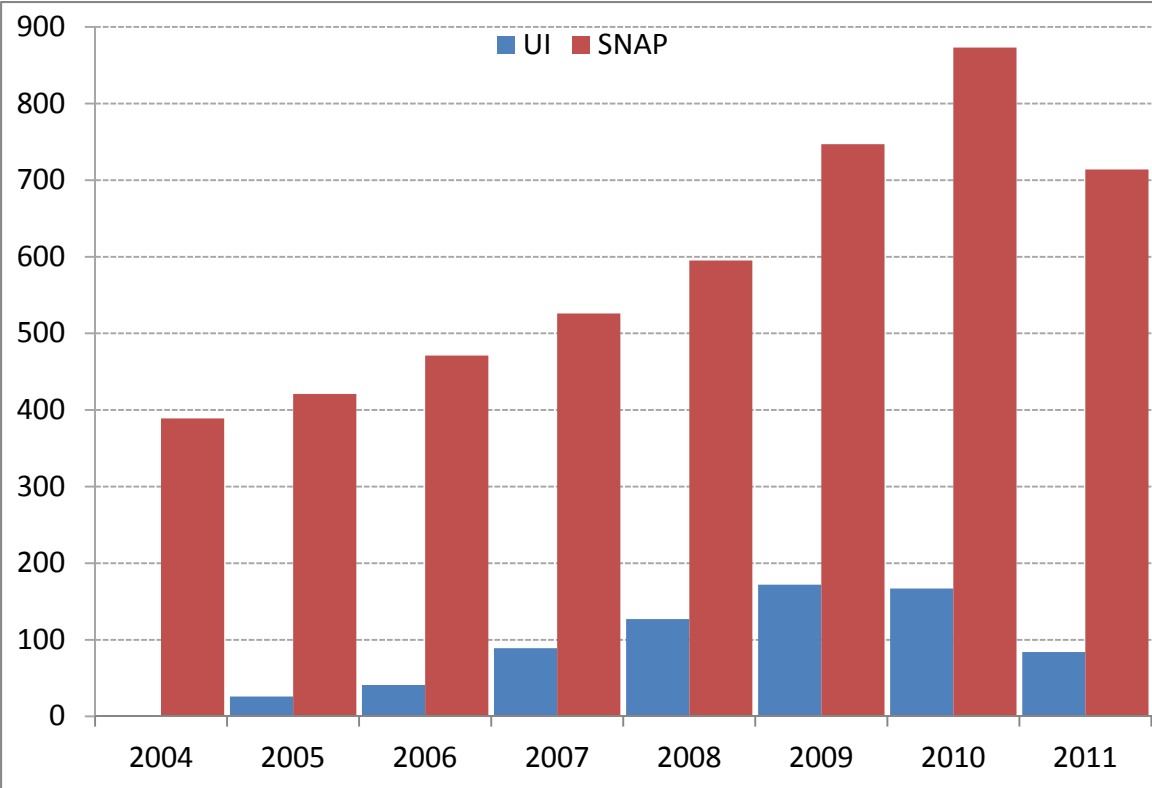


Figure 5 reveals two additional contributing factors to the increase and persistence of 1984 community college enrollee cohort members in our longitudinal integrated data system—receipt of Unemployment Insurance benefits and/or Supplemental Nutrition Assistance benefits. It is important to pay attention to the time span covered by each of these administrative data sources. Shown in our longitudinal integrated data system, Maryland UI benefit coverage begins in January 2005 and ends in April 2011; Maryland SNAP coverage begins in November 2003 and ends in February 2011. Please note that our UI benefit coverage data used for this report only covers UI benefit recipients who also received SNAP benefits, though not necessarily at the same time. In both cases, only partial coverage of 2011 is represented, so the vertical UI benefit and SNAP benefit participant counts are not comparable to the full-year coverage of the earlier years.

We expect to update both of these administrative databases during 2013, but neither can be filled-in for earlier years than those already received.²⁵

Together, *Figures 3, 4 and 5* have begun to sharpen and narrow our attention to analytical next steps that can be expected to result in actionable recommendations. We will address these opportunities in the final section of this report. Before doing so, however, we delve further into a few of the combinations of administrative data that compose our longitudinal integrated data system.

Figure 6 introduces our only coverage in this report of the earnings data field from the Maryland UI wage record database. Here, the x-axis is chronological time from 1985 through June 2012 (the light shading of the 2012 vertical bar signals a half year coverage).

The vertical y-axis of *Figure 6* measures median annual reported earnings in 2012 dollars.²⁶ Currently, only Maryland UI wage record earnings are included because the TRADE and FEDES sources of earnings data only cover recent years, not the full time span that is of interest when the reference population is the 1984 community college cohort members. Only those cohort members that had some amount of reported earnings in at least one of the four quarters in a reference year are included in the median calculation; those with no reported earnings in a particular calendar year are not included in that year's calculation of a median dollar amount. Those with positive reported earnings include full-time and part-time employment, year-round and seasonal employment, and concurrent jobs statuses.

The vertical red bars in *Figure 6* show what one would hope and expect to see for a cohort of former public community college enrollees—a steady increase in median reported inflation adjusted (real) annual earnings over 27 years. The vertical blue bars show the median reported annual earnings for the cohort members that appear in the University System of Maryland Enrollment Information System and/or Degree

²⁵ This is a timely opportunity to alert newcomers to the availability and use of *administrative* data that record destruction, or purging, is a common practice when the data are no longer needed to satisfy the administrative reason for collecting and maintaining the data. Most interagency *Memoranda of Understanding* stipulate that data are to be destroyed or returned to the owning agency when the authorized purpose stated in the MOU has been satisfied. A fundamental goal of our research and this interim report is to heighten awareness in the P-20W SLDS community that extended longitudinal integrity of these integrated data systems is necessary to document the interplay of education, workforce and social safety net program participation. This documentation, in turn, is necessary to communicate a strong business case for policy and administrative actions that can improve the quality of participants' lives.

²⁶ The CPI inflation calculator is used to convert current dollars of each corresponding year into 2012 dollars. The CPI inflation calculator, retrievable at the World Wide Web http://www.bls.gov/data/inflation_calculator.htm, uses the average Consumer Price Index for a given calendar year. This data represents changes in prices of all goods and services purchased for consumption by urban households.

Information System in one or more reference years currently covered by these databases, 1985-1997.²⁷

FIGURE 6: MEDIAN EARNINGS

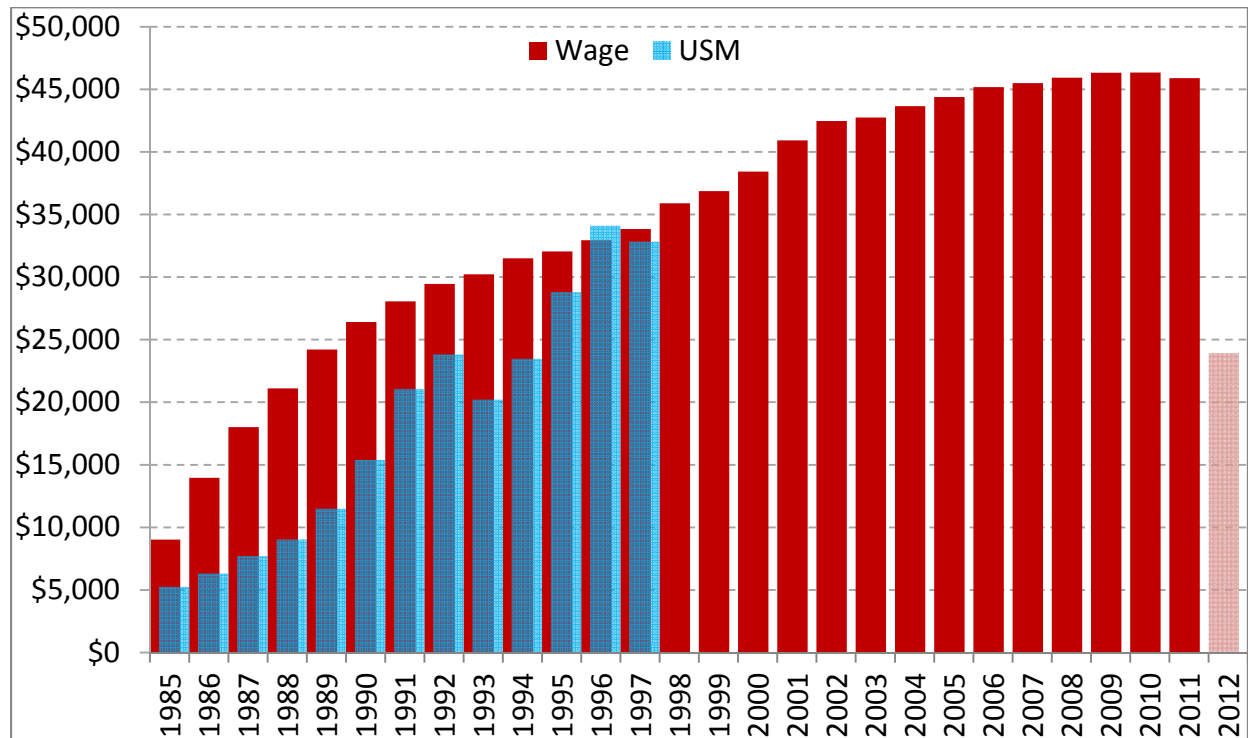


Figure 6 is intended to alert newcomers to the presentation of employment and earnings information that concurrent education and employment statuses can have important impacts on what is observed. Similarly, sequential education engagements, such as community college followed by a four-year university, that are not properly documented and taken into account can result in inaccurate conclusions and actions. This mistake can impact how return on investment and other *outcome* attributions are made and acted upon. This figure also prompts the need for more years of postsecondary education data to refine our understanding of how the economic recession impacted particular segments of the 1984 community college enrollee cohort.

Next, we turn attention to the education impact on job service and social safety-net program participation.

²⁷ Individuals can move in and out of the year/quarter cycles of the Maryland UI wage record database. This means that the employees included in a particular reference year in Figure 6, and Figure 4, can differ from those included in previous or subsequent years.

Figure 7 shows the percentages of the 1984 community college enrollees who appeared in one or two other data sources. Both the horizontal x- and vertical y-axes reflect the data sources. The size of the bubble reflects the percentage level of the 1984 community college enrollees who appeared in the corresponding data sources labeled by the x- and y-axes. For example, 4.9% of the 1984 community college enrollees reported positive wages in Maryland and appeared as a registrant in the JTPA-WIA job service programs; 5.4% of the 1984 community college enrollees reported Maryland wages and appeared as a TANF registrant. The proportions in the labeled square area reflect the percentages of the 1984 community college enrollees who appeared as a registrant in the job service or social safety-net programs (including JTPA-WIA, UI, SNAP, and TANF).

FIGURE 7: The 1984 Community College Enrollees' Appearance in Other Sources

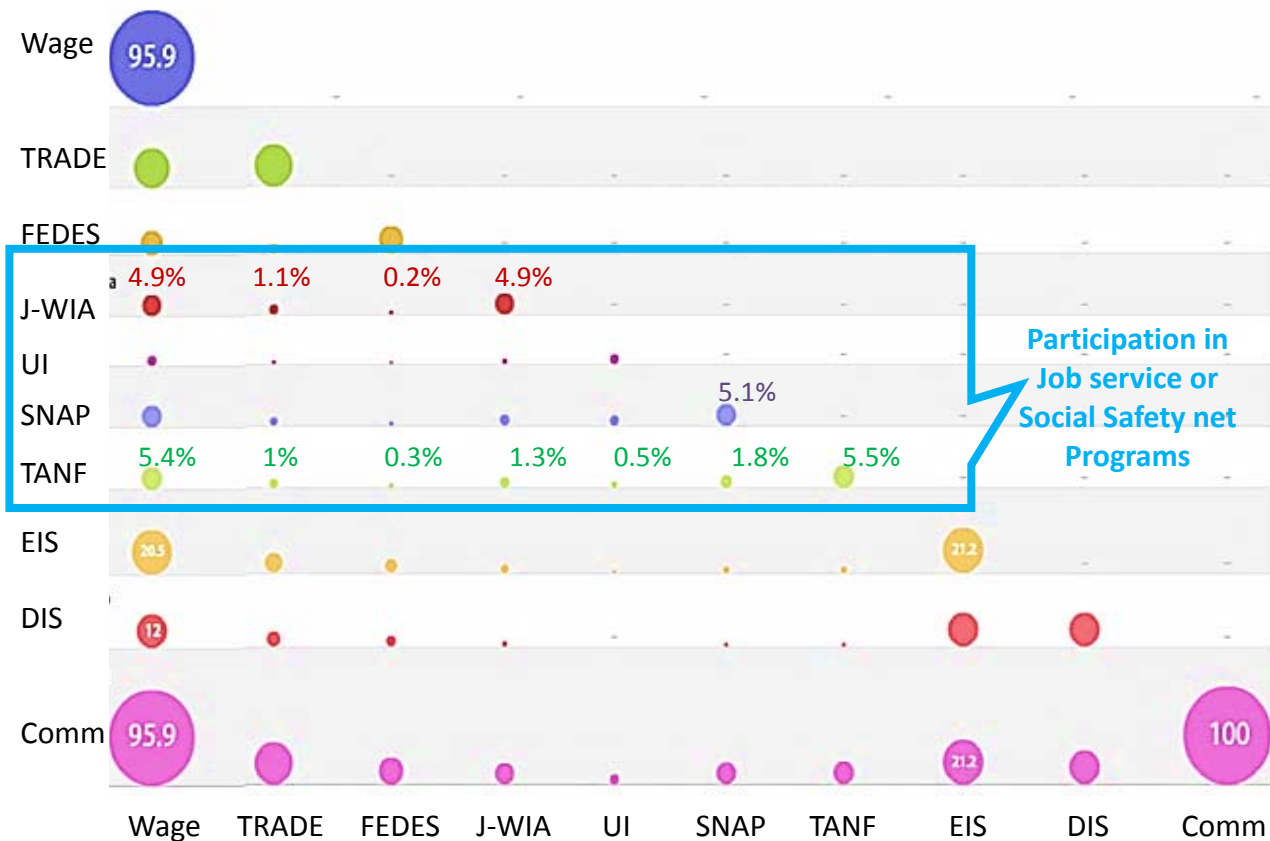


Figure 8 and Figure 9 follow the same layout as Figure 7, but with a subset of the 1984 community college cohort: Figure 8 shows the appearance of the University System of Maryland enrollees (i.e. EIS members) in one or two other data sources and Figure 9 focuses only on the University System of Maryland degree holders (i.e., DIS members).

Figures 7 through 9 compare the proportions that reflect job service and social safety-net program participation among the 1984 community college enrollees (see Figure 7), the subset cohort University System of Maryland enrollees (see Figure 8), and the subset cohort University System of Maryland degree holders (see Figure 9). The job service and social safety-net program registration rate shrank from the 1984 community college enrollees to the subset cohort University System of Maryland enrollees and shrank further to the University System of Maryland degree holders.

FIGURE 8: Appearance in Other Sources for the University System of Maryland Enrollees who were the 1984 Community College Enrollees

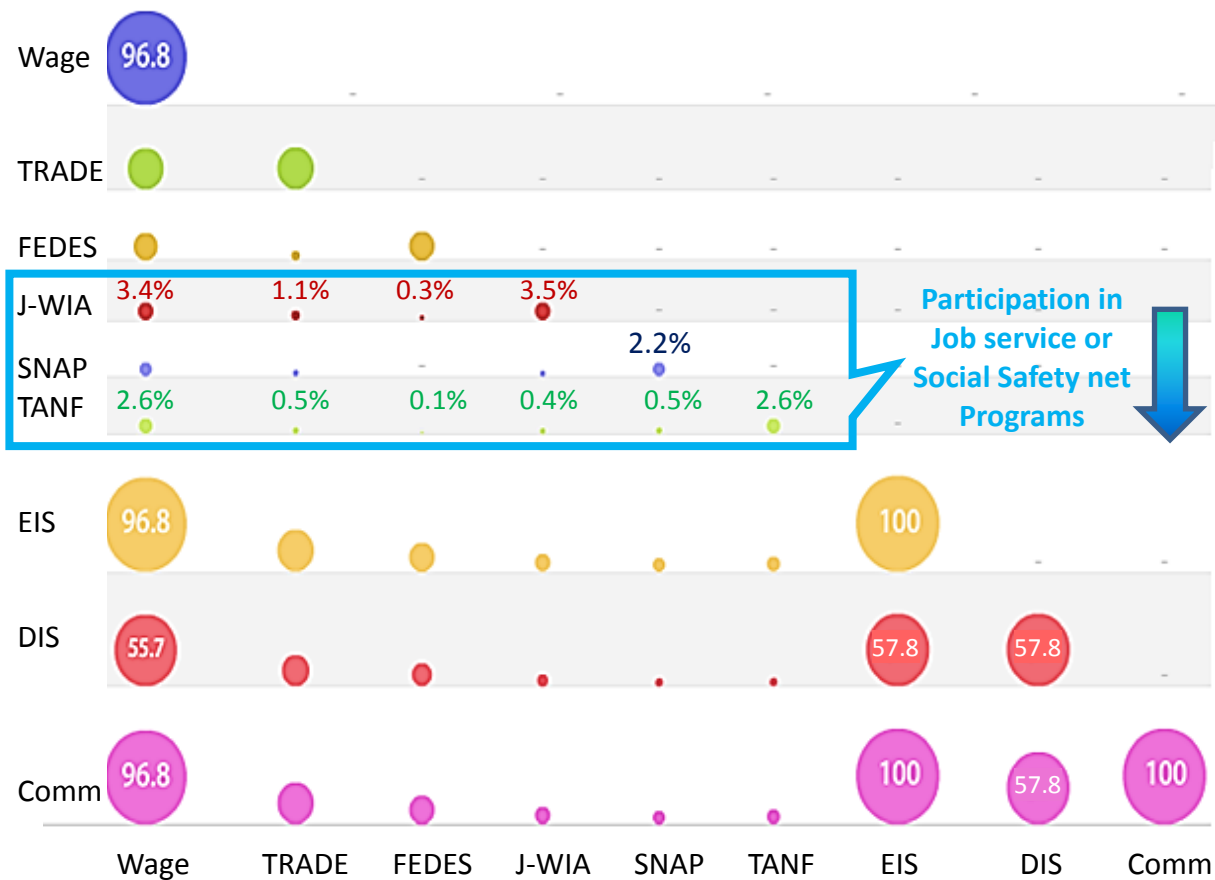
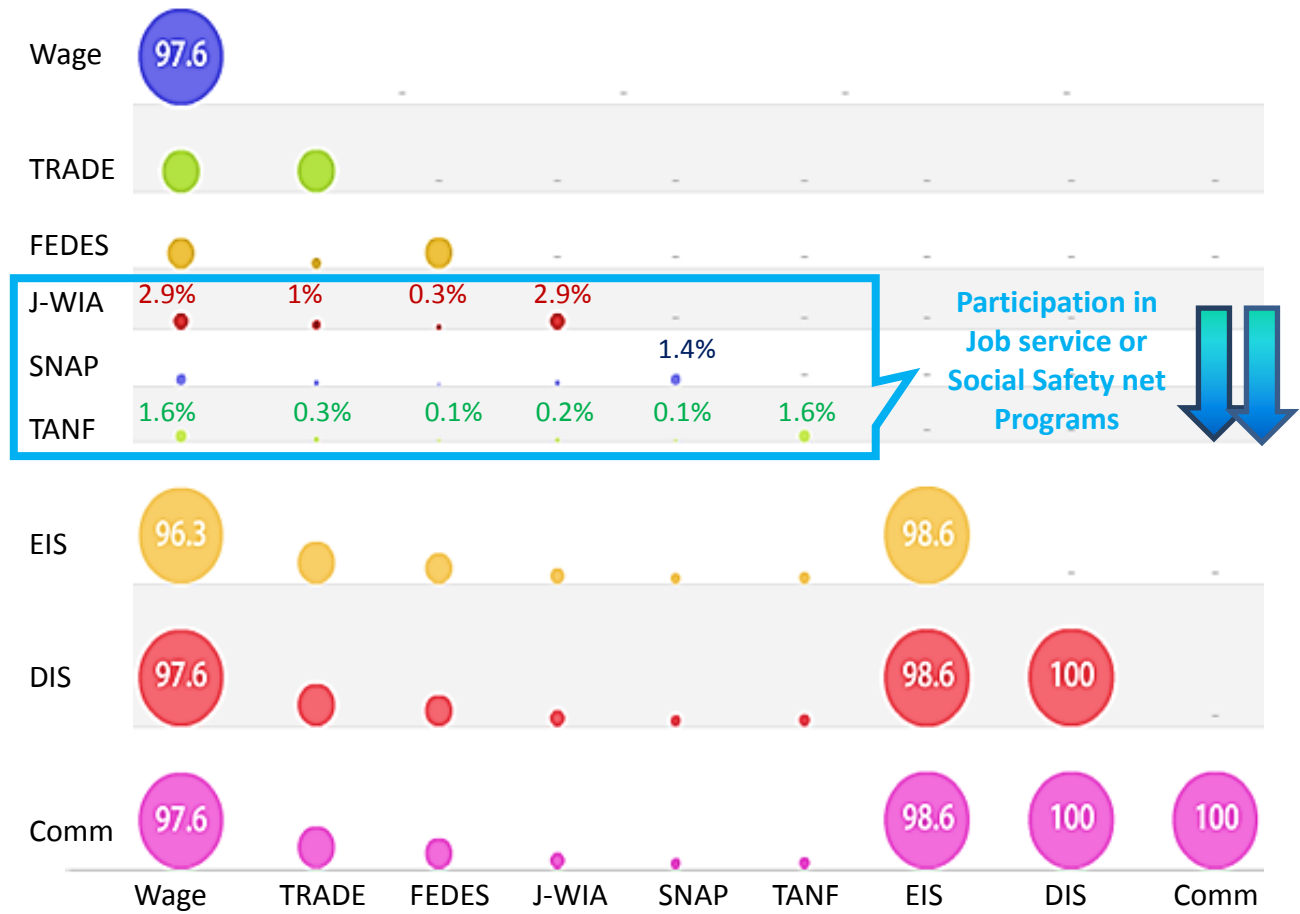


FIGURE 9: Appearance in Other Sources for the University System of Maryland Degree Holders who were the 1984 Community College Enrollees



CONCLUSIONS

It should be clear by now why we have described this as an *interim* report. Our longitudinal integrated data system remains a work in progress.

We return to the report's title one more time to set the stage for drawing conclusions from what we have learned to date—*Neglecting the “L” in a Longitudinal Integrated Data System Can be a Costly Mistake*. Our conclusions focus on making a business case for maintenance of strategically constructed integrated data systems over extended time spans; spans that are long enough to observe downstream events and interactions that constitute teachable lessons pointing the way to actions that can be expected to have positive individual and social impacts.

High on our list of long time span coverage and robust data source content benefits that are derivable from a longitudinal integrated data system is improvement of rate of return and outcome estimates associated with education and training investments. We expect to be able to make substantial progress on this front in 2013, even though some relevant data gaps will remain in our assembled administrative databases.

We have shown in the previous section that 1984 Maryland public community college enrollees appear in multiple available administrative databases over the next 27 years. Some of these engagements have been with social safety-net programs. Some appearances indicate movement out-of-state. Others document participation in *workforce* programs, undefined further up to this point. And we are anxious to fill-in the gap of Enrollment Information System and Degree Information System appearances since current coverage ended in 1997.

As we probe beneath the surface of the appearances documented to date we expect to arrive at actionable recommendations. Likely examples include:

- The 1984 community college enrollees traveled different human-capital investment paths, starting at the community college level, but then branching off along diverse formal education, on-the-job training, and other advancement stepping stones. Some were more directly vulnerable to recession jeopardy than others. Some absorbed UI benefit, food assistance and temporary cash assistance resources. We hope to be able to define high impact actions and timing that can improve former student prospects for remaining on a positive trajectory of stable employment and rising earnings.
- Strong interest continues to be expressed in two topics that we can, and will, address using our updated longitudinal integrated data system:

3. Public, business and educator concern is expressed that public postsecondary education entities in the U.S. are not preparing enough science, technology, engineering and mathematics (STEM) graduates to satisfy current and projected demand for these competencies.
4. Widespread concern is expressed that students unfortunate enough to have graduated since 2006 will suffer lasting untoward employment and earnings effects traceable to their immediate struggle to establish an appropriate entry-level affiliation.

Neither of the two topics described above can be adequately studied in the absence of a robust longitudinal integrated data system capability covering a time span of at least 5 years; and preferably more years as time passes.

We are enthusiastic about new data acquisition possibilities that may happen in 2013. Maryland has been invited to join a six state pilot initiative that will replicate and improve upon a one state pilot completed in 2012 for Illinois. The design involves negotiation of appropriate *Memoranda of Understanding* between participating state education entities and CompTIA and the Manufacturing Institute, with a goal of adding a former student's reliable skill certification accomplishments to enhance longitudinal integrated data system capabilities. This pilot offers hope that well known, but poorly documented, skill advances can be captured to improve return on investment estimates. These improvements, in turn, should be translatable into better information flowing to tomorrow's students about how their choices might be expected to impact their long-term employment and earnings prospects.