State of Maryland

Workforce Data Quality Initiative (WDQI)

DESIGN PLAN

Grantee Identifying Information

Grantee Name	State of Maryland, Department of Labor, Licensing and Regulation, 1100 N.Eutaw Street, Room 508, Baltimore, Maryland 21201
Grant Number	MI-21167-10-60-A-24
Program/Project Name	State of Maryland Workforce Data Quality Initiative
Report Submitted by	Shreya Pillai, Program Manager, DLLR
Grantee Address	Department of Labor, Licensing and Regulation, 1100 N.Eutaw Street, Room 508, Baltimore, Maryland 21201
Report Submission Date –	03/22/2011

OVERVIEW

The Division of Workforce Development & Adult Learning, Maryland Department of Labor, Licensing and Regulation (DLLR), and the Jacob France Institute, University of Baltimore (JFI), have engaged in a continuous principal-agent partnership[1] since 1991. JFI is also included in related data sharing agreements with other federal, regional, state and local government agencies.

For the Workforce Data Quality Initiative Study in Maryland, DLLR has entered into an agreement with the Jacob France Institute (JFI) at the University of Baltimore. The Principal Investigator on the study will be Dr David Stevens. The agreement will deliver new policy-relevant research findings over a three-year time span—2011 through 2013. These findings will deliver an attractive return-on-investment in two ways:

[1] The *principal-agent* partnership means that DLLR retains full and sole ownership of all DLLR administrative data delivered to and held by JFI for authorized research and evaluation uses. The *Memorandum of Understanding* (MOU) between the two organizations is renewed annually, including the agreed upon scope-of-work for the defined state fiscal year. Any within-year change to the scope-of-work is negotiated and, if affirmed, attached as an amendment to the basic MOU. Third-party requests for access to the DLLR administrative data files maintained by JFI are processed by DLLR only; JFI has no independent authority to act upon such requests.

- JFI research findings based on linkage of up to 25 years of maintained or accessible administrative records will contribute to an evidence-based foundation for making a strong national business case for sustained investment in state longitudinal workforce data systems.

- Individual JFI studies will deliver new research findings and data collection, quality control and secure processing insights. These will advance shared understanding among the WDQI states, and ultimately all states, about workforce longitudinal data system opportunities and how to respond to problems that may be encountered.

RESEARCH PROJECTS[2] AND DATA QUALITY ANALYSES

Introduction

This section begins with a succinct overview of six JFI projects that will be completed during the three years of MD WDQI funding. Each of the four research projects and two data quality analyses is then described in more detail.

Year one

JFI staff has begun two Year One[3] MD WDQI studies:

- Mapping of the frequency, timing and sequence of individual customer appearances in multiple historical administrative data files.

- Social Security Number (SSN) data quality analytics.

A Year One report from each of these studies will be delivered in November 2011, or earlier if circumstances allow more timely completion.

Year two

A research project describing the workforce and related transition paths taken by students enrolled in a Maryland public community college in the fall of 1984, will begin in Year One and be completed during Year Two (no later than November 2012).
The start date of this study is contingent upon the successful resolution and timing of any legal and/or regulatory issues that may arise from an updated and topic-specific review, and amendment if necessary, of current interagency MOUs.

---

[2] Four research projects and two data quality studies are described here. Three of the four research projects are fully consistent with the descriptions found on pp. 15-17 of DLLR's *Technical Proposal: Part I Statement of Work* responding to SGA/DFA PY-09-10. The fourth research project design has been modified based on a 3-10-11 telephone conversation with Kate Louton, ETA project officer for the WDQI awards.

[3] Year One is December 1, 2010 through November 30, 2011. Year Two is December 1, 2011 through November 30, 2012. Year Three is December 1, 2012 through November 30, 2013.

Also during Year Two, an analysis of common data field consistency among maintained administrative data sources will be completed; again no later than November 2012, or earlier if circumstances permit more timely completion.

Year three

The remaining two research studies will begin in Year Two and be completed during Year Three (no later than November 2013):

- A replicable research design and report of findings from a multi-occupation analysis of job retention durations, inter-industry transition paths and rates, and geographic mobility profiles.

- A baseline report on estimation of the impacts of the December 2007-June 2009 national recession on student enrollment in, persistence through and completion of selected Maryland public postsecondary education programs. Included in this study will be analyses of the time alignment of documented employment affiliations and earnings amounts with the enrollment profiles.

DLLR and JFI staffs will collaborate throughout the three years of MD WDQI funding to identify and carry out new field reporting practices and communication strategies that will increase the return on investment from the MD WDQI partnership.

Descriptions of each of the six projects to be undertaken by JFI staff in partnership with DLLR and other state and local government agency colleagues follow.[4]

Mapping of individual appearances in longitudinal administrative data files

Separate historical administrative data files have been acquired, maintained and updated by JFI since 1991 (with source-specific start dates)—MD UI wage records, JTPA/WIA participant records; Employment Service registrant records; AFDC/TANF recipient records; Food Stamps/SNAP recipient records; and currently in process UI benefit recipient records (less than universal coverage).

JFI has negotiated use-restricted data access agreements with three MD public school districts (Baltimore City Public Schools; Baltimore County Public Schools; and Cecil County Public Schools), MD public community colleges (all), and the University System of Maryland. The Maryland State Department of Education and the Maryland Higher Education Commission use archived JFI data resources for defined federal reporting and regulatory compliance purposes.

---

[4] The DLLR proposal responding to SGA/DFA PY 09-10 included letters of support and commitment from Governor O'Malley, DLLR Secretary Sanchez, State Superintendent of Schools Dr. Nancy S. Grasmick, and other key officials; so authorization for administrative data access and use is assumed here. Any problems encountered in achieving timely access and use will be conveyed in WDQI progress reports.

This research project will create individual longitudinal data files that draw upon maintained and authorized accessible administrative data sources. These consolidated longitudinal data files will be used to create profiles of the frequency, timing and sequence of individual use(s) of defined public services, understanding that historical time coverage will differ among the assembled data sources.

A principal goal of this research project will be to illustrate the importance of extended longitudinal data retention to support research and evaluation studies designed to deliver reliable evidence of program interdependencies. This evidence is needed to inform strategic decisions about priorities for intervention criteria and timing.

Social Security Number (SSN) quality analysis

A persistent challenge to state longitudinal data system progress continues to be the availability, accuracy and use of SSNs. The mapping project described on the previous page exemplifies this challenge—the common SSN identifier will be the starting point for assembling the individual consolidated longitudinal data files.

This SSN quality analysis project will deliver a case-study of value added when an accurate SSN is available for administrative data linkage use. The project will include replicable diagnostic steps to confirm that a nine-digit identifier has been issued by the Social Security Administration and to determine whether this SSN has been used by only one person. Additional diagnostics will include acquisition and analysis of alternative/complementary data linkage approaches that have been adopted or pilot tested (e.g., name, date-of-birth and gender).

A principal goal of this project will be to craft a strong business-case for the benefit(s) that accrue from availability of an accurate SSN identifier associated with multiple administrative data sources.

The 1984 entering cohort of MD public community college students

This research project will complement the mapping project described on the previous page. Here, too, a principal goal will be to illustrate the value of extended time coverage needed to understand life-cycle workforce (and education) flows.

There will be three stages of the research design for this project:

1. The original transcripts extract file for the 1984 entering cohort of community college students remains secure. Research of applicable law(s) and regulations will determine what steps need to be taken before this file can be used for the defined purpose.

2. An attempt will be made to document the *current* affiliations of the '84 cohort members by linkage with the available administrative data files described previously, plus inclusion in the next Federal Employment Data Exchange System (FEDES) year/quarter cycle and in the next year/quarter cycle of

authorized data exchanges among the DLLR counterpart state agencies in DC, NJ, OH, PA, VA and WV.[5]

3. The individual consolidated longitudinal data files created for the mapping project described previously will be used to document *intervening* affiliations of the '84 cohort members, understanding that this documentation will be illustrative of state longitudinal workforce data system capability, but not fully inclusive.

Common administrative record data field quality

An important unknown about individual longitudinal data file creation and use is whether consolidation of information acquired from multiple administrative data sources introduces undefined data quality differences. Such differences, if and when discovered and reported, can threaten public confidence about the integrity of state longitudinal workforce data systems.

Again drawing upon the individual longitudinal data files created for the mapping project defined previously, this administrative data quality analysis will identify opportunities to compare two or more sources of what are expected to be *the same* data field definition—such as date-of-birth and gender, or a time-contingent definition—such as educational attainment and receiving of AFDC/TANF or Food Stamps/SNAP or UI benefits.

One important facet of this project will be testing of different data updating or refreshment strategies. Updating involves absorption of cost, so a defined benefit should be identified to justify this cost. This introduces the importance of starting the design process with the anticipated uses of the data.

Occupational mobility profiles

Employers have a strategic choice to make—they can incur costs to *recruit* new hires to replace departing incumbent employees, or they can incur costs to *retain* incumbents.[6] Employee churning, or turnover, has been the focus of many research studies.

The principal goal of this research project will be to illustrate how consolidated longitudinal workforce and education administrative data files can be used to improve the types and quality of information available to multiple stakeholders—employers, workforce professionals, educators, students, and employee and employer advocacy organizations—that make strategic decisions impacting recruitment and retention costs and outcomes.

---

[5] If the WRIS 2 exchange capability becomes operational, and if this use satisfies the defined voluntary state participation criteria, in time to be drawn upon for this research study that additional source of current status information will be included.

[6] This oversimplifies an employer's actual decision process and options.

Accurate *occupational* information is often defined as the highest priority missing data element in state longitudinal data systems and system designs. With a single state exception (Alaska), state UI wage records do not contain an occupational data field. Survey data collection is costly and subject to many criticisms about representative coverage and reliability of self-reporting and subsequent coding.

This project will choose education/training programs that are known to produce completers/graduates with defined skill competencies—including teachers, selected licensed trades, and some healthcare clinical specialties, and track the subsequent sequence and duration of industry affiliation(s) and geographic movement(s) of graduates, to the extent possible relying on the available administrative data sources.

Recession impacts on student enrollment, persistence and completion rates

The design of this research project will again demonstrate that extended time coverage of integrated administrative data systems is important. A recession changes the relative weights given to immediate, intermediate and long-term considerations by graduating students, those already enrolled but not about to graduate, and those thinking about future enrollment. This impact is particularly acute at the postsecondary level.

A State longitudinal workforce administrative data system capability is needed to improve our understanding of what happened to students that experienced the misfortune of 2007-2009 recession conditions[7] impacting their decisions about whether to enter postsecondary education, to remain enrolled and to pursue an additional or different credential. Employment opportunity influences entry, retention and return decisions.

JFI staff will complete a baseline analysis of the 2007-2009 recession's impact on Maryland public postsecondary student entry, retention and return decisions, in continued collaboration with Maryland Higher Education Commission, University System of Maryland, and Maryland Association of Community Colleges public postsecondary education partners.

JFI maintains a releasable extract of the DLLR Quarterly Census of Employment and Wages (QCEW) file, which is routinely updated. Extensive research using this file and MD UI wage records has been completed and pertinent new studies are planned.

The principal goal to be achieved from the baseline longitudinal analysis of student flows into and from employment will be to advance knowledge in two ways—first, by demonstrating a practical way for other states to replicate this approach; and second, by describing the short-term baseline of employment-education interactions to which future updates can then be added beyond the three-year grant award years.

---

[7] The NBER-defined official national recession started in December 2007 and ended in June 2009. The timing and severity of this recession on Maryland and intrastate locations has been documented in recent JFI reports sponsored by DLLR.