To: Commission on Evidence-Based Policymaking  
From: Workforce Data Quality Campaign  
Re: Docket ID USBC-2016-0003-0001  
Date: November 14, 2016

Workforce Data Quality Campaign (WDQC), a project of National Skills Coalition, is a non-profit initiative that promotes inclusive, aligned and market-relevant education and workforce data. Guided by a diverse group of national partners and state officials representing stakeholders across the education and workforce spectrum, WDQC encourages the use of data to ensure that all of our nation’s education and training programs are preparing students and workers to succeed in a changing economy.

The Commission on Evidence-Based Policymaking’s examination of federal administrative and survey data provides an exceptional opportunity to address the management and use of data for measuring postsecondary education and workforce outcomes. The Commission’s findings could go far beyond enabling the provision of better information for researchers and policymakers. It could also help educators, students, employers, and workers all make more informed decisions.

Overarching Challenge

The collection and use of education and workforce data can improve human capital policies that play a part in strengthening the national economy, but information still remains too disconnected to realize its full potential for improving workforce and economic development.

We advise the Commission to consider the creation of a national clearinghouse that would facilitate linkages between employment data and data about postsecondary and workforce program participants. The system should enable public access to data aggregated or anonymized sufficiently to prevent tracing back information to individuals, but disaggregated sufficiently to allow useful program analysis and research. Absent the creation of a clearinghouse, further developing the use and linkages of these data can still improve program reporting, evaluation, and research. We also recommend aligning federal laws like the Workforce Innovation and Opportunity Act (WIOA), The Perkins Career and Technical Education Act (Perkins), and the Higher Education Act (HEA) to improve data collection. Finally, we recommend providing assistance and incentives for states to strengthen their roles in the collection, use, and sharing of data.

The Commission also should issue an admonition, possibly in its final report, to remind audiences that data alone will not drive policy improvements, and that human factors will influence how data are interpreted and used. While policy improvements should become more likely because of better data, linkages, and access, we must continually work to create a culture of responsible data use. Federal data policies should be regularly assessed and refined as thinking and technology change.
Below are recommendations organized according to questions that are the most pertinent to our remit of improving the collection and use of data for education and workforce policies and programs. Our comments incorporate input from multiple national and state experts.

COMMISSION QUESTION #1: Are there successful frameworks, policies, practices, and methods to overcome challenges related to evidence-building from state, local, and/or international governments the Commission should consider when developing findings and recommendations regarding Federal evidence-based policymaking? If so, please describe.

Through our work with states and federal agencies over the past several years, we have observed the following effective practices for using data to inform workforce development policy:

- Convene relevant stakeholders to identify high-priority policy questions, then develop a longitudinal data infrastructure to enable meaningful analysis to answer these questions.
- Allow states (and their agents) to access this information to assist with required program reporting and evaluation, and further tailor information for their specific needs.
- Outline the importance of data linkages, privacy, and security. Select systems that ensure privacy, security, and confidentiality.
- Educate the public and policymakers on the merits of data collection and use, and particularly on the importance of labor market information.
- Facilitate and encourage the incorporation of data into easy-to-use tools that fulfill specific stakeholder needs.
- Invest in technical assistance and professional development to ensure that data managers know how to use data to inform policy choices.

More detailed examples of these practices and proposals are included in answers to the questions below.

COMMISSION QUESTION #2: Based on identified best practices and existing examples, what factors should be considered in reasonably ensuring the security and privacy of administrative and survey data?

Ensuring that privacy and security are integrated with the development of new data linkages or the creation of a national clearinghouse will be essential for building trust and support for the Commission’s recommendations.

Policies established to guide system development and use should always address privacy and security. Moreover, these policies should be transparent, meaning that they should be published, and also communicated in a way that will be understood by relevant stakeholders.¹

¹ Several important federal laws and rules applicable to privacy and confidentiality were covered during the Commission’s meeting of September 9, 2016, including The Office of Management and Budget Circular A-130, the 1974 Privacy Act, the Confidential Information Protection and Statistical Efficiency Act (CIPSEA), and the Family Educational Rights and Privacy Act (FERPA).
To reduce risk, the following basic practices should be followed:

- Only collect data that are needed. This limits the opportunity for revealing personal information, and has the added benefit of reducing burden on data providers, collectors, and systems.
- Ensure that publicly available data are aggregated or otherwise stripped of all information that could be used to identify particular individuals or employers.
- Avoid willful disclosure or inappropriate use. Convey consequences to all relevant parties.

We also encourage the use of the following best practices:

- Promote research: Maintaining privacy, confidentiality and security is essential, but it should not be construed as a reason to pull back on research. Data policies should explicitly allow research, policy analysis, and program evaluation, in addition to publication of aggregate performance outcomes.
- Transparency: Each agency that has administrative data should publish their process for data-sharing with researchers and policymakers; including identifying what data the agency has and are available for sharing, the mechanism for requesting access, specific timelines for making decisions, and the specific individuals who are authorized to make decisions about access. The process should also identify who is eligible to access data and for what purposes, what security measures are required for maintaining the data that are ultimately received, and any fees imposed for preparing the data.
- Create Memoranda of Understanding: Craft standard templates for memoranda of understanding (MOUs) to ensure that personnel are sharing and using information according to consistent guidelines.
- Take into account different cultural perspectives: Residents have different levels of comfort about government and other entities having access to personal data. Sensitivity to different perspectives on privacy should shape how recommendations are made and how privacy is addressed in data collection.

Technological advances should be explored and used, with any plan including provisions for regularized assessments to adopt improved practices and adaptation to evolving security threats. Currently, two technological innovations may be worth exploring in particular:

**Tiered access** allows layers of access and aggregation using the same foundational data sets. This technological solution could also eliminate the need to hold different data sets in different locations or among disparate systems with varied levels of security.

California provides a good example of tiered access. The California Community Colleges Chancellor’s Office career and technical education (CTE) “LaunchBoard” tool, which links education data and wage records, provides community college employees with login access to detailed dashboards that analyze outcomes of CTE programs. The same underlying data is used to create “Salary Surfer,” which is a publicly-available tool that allows users to see the median salaries of postsecondary programs at California’s community colleges.

---

Texas has made some of the biggest strides in data integration, and provides different types of users with access to specific portions of its workforce database. The U.S. Government Accountability Office (GAO) described the Texas data infrastructure in its recent review of Workforce Innovation and Opportunity Act (WIOA) reporting. Texas agencies draw on multiple data systems that share information through regular, automated exchanges. The data from these systems are compiled in a separate system, called The Workforce Information System of Texas (TWIST) web reports, where UI wage data and other data are merged with participant information and then sent to the U.S. Department of Labor (DOL). TWIST uses a system providing different levels of restricted access so agencies can see individual-level outcomes for only their own participants.

**Synthetic data** are artificial data sets that are similar to, but are not the raw, confidential data from which they are derived. A statistical model is created that adjust specific data fields so that individual-level data no longer corresponds to real people, but the data set as a whole maintains the characteristics of the original. Because the data are not tied to real people, confidentiality rules do not apply.

Analysts can access these data that have the properties of the real data, without facing the same restrictive barriers in addressing their research questions. The types of inferences that may be drawn from synthetic data, however, will be more limited. Therefore, the fact that these data sets are synthetic must be made clear so researchers and other users can take into account the limitations of these data sets for appropriate use.

The Maryland Longitudinal Data System (MLDS) Center has embarked on a multi-year project to create synthetic data sets, a project which was funded by a portion of a grant from the U.S. Department of Education (ED). The U.S. Census Bureau’s Longitudinal Employment-Household Dynamics (LEHD) program uses partially synthetic data to examine workers’ residential patterns.

**COMMISSION QUESTION #3: Based on identified best practices and existing examples, how should existing government data infrastructure be modified to best facilitate use of, and access to, administrative and survey data?**

Rationalizing the federal data landscape would require significant changes to multiple laws, but it could cut costs and reduce security risks by cutting down on the number of redundant data sets scattered across government, and could also make the creation of a national clearinghouse more manageable. In the answer to Question #7, we identify federal data sets that deserve consideration. In

---

particular, revising access policies for wage record sources would greatly improve the quality of information for performance reporting and analysis. For example, access to the Department of Health and Human Services’ National Directory of New Hires (NDNH) data, which includes actual hire dates in addition to quarterly wages, would help to inform workforce development programs on earnings and indicate how long it takes people to return to work.

In addition to revising the organizational structure of data systems, the federal government should align data definitions and performance metrics across key U.S. laws that authorize education and workforce development programs. Consistent definitions would streamline data collection and provide comparable information, facilitating program collaboration and allowing policymakers to have a holistic view of human capital development.

WIOA passed in 2014 with wide bipartisan approval to help job seekers succeed in the labor market by better coordinating access to employment, education, training, and support services. Since then, government agencies significantly increased their collaboration to issue rules and guidance on performance reporting, making WIOA a useful guidepost by which to align reporting requirements with other programs related to those services. We advise the Commission to take a closer look at the following laws and programs for alignment of definitions and metrics:

**Perkins** – This law, which governs career and technical education, is currently due for reauthorization. The revised law should use the WIOA definition of “postsecondary recognized credential,” which includes degrees, licenses, and industry-recognized certificates and certifications. This inclusive definition reflects growing evidence that many types of credentials can give students career opportunities. In addition, Perkins should include employment outcome metrics aligned with WIOA for postsecondary students.

**HEA** - While this law drives a large portion of postsecondary funding, the current law, last reauthorized in 2008, does not require performance reporting to address employment outcomes. ED currently publishes a College Scorecard with employment outcome information, using annual data on employment from the Department of the Treasury. Legislative language formalizing the ED College Scorecard should consider how that information aligns with other federally-mandated scorecards, especially training provider reporting required by WIOA. WIOA performance reporting requires quarterly metrics based primarily on UI wage records held by states, so the data source is different. Furthermore, WIOA focuses on short-term outcomes of less than one year, whereas the College Scorecard’s figures show median earnings 10 years after graduation. Having comparable data on education and training providers, with interval snapshots showing short-term and long-term outcomes, would provide a clearer, broader view for consumers.

**Supplemental Nutrition and Assistance Program Education and Training (SNAP E&T)** - As an example of agency progress toward alignment, the Department of Agriculture’s Food and Nutrition Service (FNS) recently established four national metrics that all states must report for SNAP E&T participants, which closely align with core measures in WIOA. For example, outcomes metrics must

---

now include the number and percentage of E&T participants and former participants who are in unsubsidized employment during the second and fourth quarters after completion of participation in E&T; and the median quarterly earnings of all current and former participants who are in unsubsidized employment during the second quarter after completion of participation in E&T.⁹ FNS consulted with DOL when it created the new metrics.

**Temporary Assistance for Needy Families (TANF)** - WIOA requires that TANF be a partner program in the federally-supported One-Stop work center system (if a governor does not opt out), and encourages states to include TANF as part of the comprehensive state workforce plans required under WIOA. The TANF program currently lacks performance requirements relating to participant outcomes. TANF reauthorization is overdue and will present an opportunity for alignment with WIOA. The Commission should recommend that states be authorized to at least have the option of negotiating performance rates for credential attainment, skills gains, employment, and earnings for work-eligible individuals to collect better data on outcomes and align these data with WIOA.¹⁰

**COMMISSION QUESTION #5: What challenges currently exist in linking state and local data to federal data? Are there successful instances where these challenges have been addressed?**

Although having data linked across stages of education, training and employment holds an incredible amount of promise for informing education and workforce development, a number of persistent hurdles still prevent linkages of state and federal data. Legal barriers and lack of capacity have presented challenges for government agencies and institutions to collect, process, and use data in ways that make it possible to match data on a systematic basis. Moreover, with an inconsistent application of definitions and metrics across laws and programs, such as those discussed above, it will be hard to develop a set of common standards that could support an integrated clearinghouse where data could be linked.

Two barriers worth revising through legislative action are the Family Educational Rights and Privacy Act (FERPA) and the HEA ban on a federal student record system.

**Family Educational Rights and Privacy Act (FERPA)** – (20 U.S.C. 1232g; 34 CFR Part 99) – This law to protect student education records should be amended for clarity. Current federal regulations elaborating on statutory FERPA language allow education entities to designate an “authorized representative” to evaluate education (including job training) programs. Authorized representatives, such as workforce agencies, may receive individual-level student data. This regulation is crucial for allowing education and workforce data linkages.

Although state education and workforce agencies are commonly able to collaborate and produce longitudinal analysis of education through to workforce outcomes, misinterpretation of FERPA

---


continues to cause roadblocks. Amending FERPA to incorporate the regulatory language into the statute itself would help to reassure public officials and facilitate data linkages.

The Higher Education Act (HEA) ban on a federal student record system - (P.L. 110-315, Sec. 113). The ban on a federal student record system took effect with the last reauthorization of HEA in 2008. A federal student record system used for all postsecondary students would allow for building a more complete picture of postsecondary paths, especially for transfer students and working adults who attend school part-time. Through linkages with employment data, a federal student record system also would allow more accurate aggregate reporting on graduates’ employment and wages.  

Although the federal government is banned from retaining individual-level data, most of this information is already collected and held by a range of agencies and institutions. States have made strides in developing similar systems for their students, but with much effort and cost. More than half of states have mechanisms in place to show student progress in education and careers, but they lose precious information when these students go onto schools or find employment in other states.

Given that states have faced hurdles in accounting for information on students who leave their state borders, and also for those who do not appear in state-held UI wage records (e.g. self-employed and military), several initiatives have helped to fill in these gaps. Examples of systems for states to directly share wage records include:

- **The Wage Record Interchange System I and II (WRIS/2)** allows states to exchange quarterly UI wage records through an automated clearinghouse system that directs queries and manages flow of data between states. The DOL operates this exchange system through a cooperative agreement with Maryland. The system has been challenged by slow response times from states in returning queried data.

  Because WRIS/WRIS2 is not a central database, but rather a system that facilitates transfer of requested individual-level UI records, state officials must manually gather the wage records to respond to queries. This is instructive for the Commission when considering different models for a federal clearinghouse. A process that requires individual agencies to manually respond to data requests, even if those requests come through a central clearinghouse, may pose a capacity challenge and result in delays in data exchange.

  - **The Western Interstate Commission for Higher Education’s (WICHE) Multistate Longitudinal Data Exchange (MLDE)** created a data sharing arrangement between Hawaii, Idaho, Oregon, and Washington. Each state contributed to a data set containing information on over 190,000 students who graduated from a public high school and/or began at a public

---

11 HEA: Sec. 113. Database of Student Information Prohibited. Part C of title I (20 U.S.C. 1015) is further amended by adding after section 133 (as added by section 112 of this Act) the following: “Sec. 134. Database of Student Information Prohibited. (a) Prohibition.—Except as described in subsection (b), nothing in this Act shall be construed to authorize the development, implementation, or maintenance of a Federal database of personally identifiable information on individuals receiving assistance under this Act, attending institutions receiving assistance under this Act, or otherwise involved in any studies or other collections of data under this Act, including a student unit record system, an education bar code system, or any other system that tracks individual students over time.”
postsecondary institution in the state. The data set included information on enrollments, postsecondary credential attainment, and wage records so researchers could look at employment outcomes and subsequent enrollments of students after completing a program of study. About half of graduates had employment records in the same state where they earned a credential. MLDE enabled states to find wage records for an additional 9 percent to 22 percent of students, depending on the state.\(^\text{12}\) The project is in the process of expanding to include over 10 states.

In addition, there are two notable processes that allow states to access employment data held by federal agencies:

- **Federal Employment Data Exchange System (FEDES)** - DOL operates an exchange system through a cooperative agreement with Maryland that allows states to query relevant wage records from the U.S. Department of Defense and Office of Personnel Management files. This system has helped states fill in UI wage record gaps for mandated performance reporting, but the participating federal agencies prohibit use of this data for broader research purposes.

- **LEHD data pilot** – This Census program — which holds UI wage records, Office of Personnel Management files, and tax data for self-employed workers — has nationwide coverage except for Wyoming. Key data has been covered for over ten years and the collection continues to grow. Census recently agreed to conduct a pilot program with the University of Texas (UT) System to collect student records from UT and link them with LEHD employment data, to produce aggregate employment outcomes. UT already collects employment outcomes information about students who remain in Texas. This project is intended to show what happens with the students who move out of state, and will provide more information on the utility of harnessing LEHD data for larger-scale research projects on postsecondary and workforce outcomes.

COMMISSION QUESTION #7: What data should be included in a potential U.S. government data clearinghouse(s)? What are the current legal or administrative barriers to including such data in a clearinghouse or linking the data?

The streamlining of information systems through the creation of a national clearinghouse could reduce the burden on agencies that have had to submit the same data multiple times, and strengthen data security and privacy by using a system with better privacy protection and security protocols.

A national clearinghouse(s) should include employment data that allows for the calculation of reliable and comparable employment outcomes for students across education and workforce programs. It

also should include information on a wide array of postsecondary students and jobseekers – including part-time students, and all students at education and training providers that receive federal aid. Because of the changing nature of education and career paths, the clearinghouse should include information on all meaningful credentials — certificates, licenses, certifications — not only two- or four-year degrees. The Commission should recommend overturning the HEA ban on a federal student record system and the ban on a federal database of WIOA data, so that we can have a nationwide, inclusive data set to show how people are moving through a variety of education pathways.

A number of the most relevant, currently-available data collections include:

- **LEHD** at the U.S. Census Bureau, which includes UI wage records submitted by all states, except for Wyoming, Office of Personnel Management files, and tax data for self-employed workers.

- **NDNH**, under the Department of Health and Human Services, which also contains UI wage records submitted by states, employment data for most federal workers, and data that includes employees’ dates of hire.

- **Tax records from the Social Security Administration and the Treasury Department** - in limited instances, agencies have found ways to use these data to show employment outcomes on programs. The LEHD program at Census has used records to fill in gaps on information for the self-employed. In an example provided during the Commission’s inaugural public event, Prof. Raj Chetty shared how he used Treasury data to show correlation between social mobility and where people grew up.

- **Student record information** - ED’s Integrated Postsecondary Education Data System (IPEDS) collects information nationwide by institution, but has limited utility for linking with other data sets when students (not institutions) are the unit of analysis. The LEHD program at Census has used records to fill in gaps on information for the self-employed. Individual-level student data is kept at ED’s Federal Student Aid office, but only includes students receiving Title IV federal aid and is maintained in antiquated data systems that inhibit meaningful analysis. The College Scorecard links this financial aid student data with employment data from Treasury to calculate institutional employment outcomes.

The National Student Clearinghouse (NSC), a non-profit organization, illustrates the potential for a federal student record system. It collects individual-level student data on a voluntary basis from institutions that use it to facilitate required reporting. The NSC operates under contract as an agent of each school to allow for FERPA-compliant data exchanges and has received approval to operate from ED’s Federal Student Aid office. NSC is not, however,

obligated to meet the same privacy standards as ED or other federal agencies,¹⁴ or a national clearinghouse that would be created by the federal government. In addition, NSC primarily serves institutions. Its central mission does not include offering transparency about higher education.

Additional data collection would enhance the abovementioned resources, such as data on hours worked and occupation. DOL has been running a pilot program to assess the utility of adding hours worked and job titles to wage records. If successful, this additional information could fill in knowledge gaps in administrative data: whether people are earning at a full-time or part-time rate, and whether they are working in occupations related to their program of study.

Some legal/administrative and technical barriers would need to be overcome to include data in a national clearinghouse:

- **Legal/administrative barriers** – Several key legal barriers, or clarifications, would be necessary before incorporating data into a national system:
  
  o Lifting the ban on a federal student record system (2008 reauthorization of HEA), Higher Education Opportunity Act, P.L. 110-315, Sec. 113 currently prohibits the provision of a national database on student information, and would need to be overturned or amended to allow for inclusion of such information in a national clearinghouse.

  o Linkages with employment data in LEHD or NDNH are currently limited. States voluntarily provide wage data to LEHD through a data sharing agreement that restricts how data may be used. Wyoming recently withdrew from the arrangement, illustrating that renegotiating the agreement with states to allow incorporation into a federal clearinghouse would be challenging. NDNH data may only be used for purposes specifically enumerated in the Social Security Act.

  o Lifting the ban on a national database under WIOA (Pub. L. 113-128, Sec. 501 (b) would be necessary to include data on workforce program participants in a national clearinghouse.

- **Informational barriers for linking data** – linking employment data with student records usually requires matching personally identifiable information, such as Social Security Numbers (SSNs), and first and last names. Tax and wage records usually SSNs, but many student records do not. Some states have taken on the challenge of trying to match student

---

information to wage records without SSNs, for example, by using department of motor vehicle records to serve as a check that improves the probability of matching other data sources.15

COMMISSION QUESTION #9: What specific administrative or legal barriers currently exist for accessing survey and administrative data?

Legal and administrative barriers that obstruct the inclusion of employment and education data into a federal clearinghouse also restrict access to individual-level data by external entities for program evaluation and research. These issues include some of those addressed above in question 7.

Linkages with employment data in LEHD or NDNH are currently limited. States voluntarily provide wage data to LEHD through a data sharing agreement that restricts how data may be used. Wyoming recently withdrew from the arrangement, illustrating that renegotiating the agreement with states to allow incorporation into a federal clearinghouse would be challenging. The original process of negotiating the sharing of state wage record data with Census took years, and in some instances required legislative changes in states. Substantial staff resources in federal and state government are dedicated to re-negotiating those arrangements as they expire in each state – a process which can be challenging. NDNH data may only be used for purposes specifically enumerated in the Social Security Act.

State laws on UI confidentiality may make it challenging for states to contribute data to NDNH or LEHD if they are accessed for broader purposes through a federal clearinghouse.16 The Center for Regional Economic Competitiveness conducted a survey of all states and found variance in their interpretation of various federal and state laws applicable to wage record information.17 The variance in accessing information for matching data causes hurdles for reporting on employment and wage outcomes within states, and between states where certain agreements have been established for matching data across borders.18

It would also be useful to amend FERPA to incorporate current regulatory language that allows an “authorized representative” to evaluate education programs, broadly defined as including job

16 See Title 20, Section 603, of the U.S. Code of Federal Regulations (20 CFR § 603) — The Federal State Unemployment Compensation Program; Confidentiality and Disclosure of State Unemployment Compensation Information. State laws vary on how they address the sharing of education and workforce-related information.
18 Ibid. The Center for Regional Economic Competitiveness is currently assessing state laws and regulations that govern business revenue and UI wage data disclosure. The project will determine best practices for maintaining the safety of data and for expanding access to confidential data for policymakers and researchers.
training. The “authorized representative” clause is important for allowing education and workforce data linkages.

**COMMISSION QUESTION #14: What incentives may best facilitate interagency sharing of information to improve programmatic effectiveness and enhance data accuracy and comprehensiveness?**

The federal government uses incentives to promote data sharing within states and between states and the federal government. These practices could inform further recommendations of the Commission, especially on the subject of creating a national clearinghouse.

**Tie greater funding to the provision of data** – Within the context of greater funding for the development of state longitudinal data systems, the Commission could recommend that funding be tied to states providing data to a national clearinghouse. For example, states that apply for DOL’s Workforce Data Quality Initiative (WDQI) grants must agree to participate in the WRIS 2 system that facilitates wage record matching across states for education and workforce performance reporting. This incentive in the existing WDQI grant program has proven effective in motivating some states to join, so that 42 states, plus the District of Columbia and Puerto Rico, now participate in WRIS2.

**Invest in training data stewards and users** – While tying funding to the use of data for policy, it will be important to invest in technical assistance and professional development to ensure that data managers know how to collect quality data, follow privacy and security rules, and effectively use data to inform policy choices.

**Promote public facing information** – Encourage agencies to provide user-friendly information to help customers select the training and education programs that best suit their needs. Public and political support would be helpful in generating the will and resources for creating a national clearinghouse. Support for data collection and system development will strengthen as employers, workers, students, and policymakers increasingly benefit from the use of this information and as long as highly sensitive information remains confidential and secure.

**Conclusion**

We would like to thank the Commissioners for their work on this subject, and the opportunity to comment through this written submission. Should you have questions or want to request additional information, please contact:

Rachel Zinn, Director
rachelz@workforcedqc.org
202-223-8355

Workforce Data Quality Campaign – a project of National Skills Coalition
1730 Rhode Island Ave. NW, Suite 712
Washington, DC 20036
www.workforcedqc.org