

# CT Preschool through Twenty & Workforce Information Network (P20 WIN)

## DATA REQUEST DOCUMENT

This form and all attachments are to be submitted to the P20 WIN Data Governing Board for consideration and approval before the execution of each Data Request.

Information provided here is required to fulfill the mandatory provisions for written agreements according to the Family Educational Rights and Privacy Act (FERPA), 20 U.S.C. §1232g and to comply with Unemployment Compensation (UC) considerations under state (CGS § 31-254) and federal (20 CFR 603) law. PERSONS WHO ARE NOT STATE GOVERNMENT EMPLOYEES OR OFFICIALS MAY NOT HAVE ACCESS TO PERSONALLY IDENTIFIABLE INFORMATION (Pii) FOR DOL UC RECORDS.

SECTION 1 – BASIC DATA REQUEST INFORMATION	
<b>1.a Application Submission Date: December 20, 2018</b>	
<b>1.B TYPE OF SUBMISSION:</b> <input checked="" type="checkbox"/> <b>Initial Application</b> <input type="checkbox"/> <b>Continuation*</b> <input type="checkbox"/> <b>Modification**</b>	
*Initial applications are approved for a given duration. Select 'continuation' if this request is to extend that timeframe. ** If you are requesting a modification to a request that has already been approved and is underway, select 'Modification' and address only the sections of this form which are being adjusted from the prior request.	
<b>1.c P20 WIN Data Request Number: P20W_1807_0_0019</b>	
<b>1.d Data Request Title:</b> College Performance Metrics	
<b>1.e Agencies from which data is being requested (Check all that apply):</b>	
<input type="checkbox"/> OEC <input checked="" type="checkbox"/> SDE <input checked="" type="checkbox"/> CSCU <input checked="" type="checkbox"/> Department of Labor <input checked="" type="checkbox"/> Cicu-IRPS <input checked="" type="checkbox"/> UCONN	
<b>1.f Has this Data Request been discussed with program employees at the involved agencies?:</b> <input checked="" type="checkbox"/> <b>Yes</b> <input type="checkbox"/> <b>No</b> If yes please indicate with whom:	
Agency: CSCU	Name: Jan Kiehne
Agency: SDE	Name: Chuck Martie
Agency: CCIC	Name: Maura Provencher
Agency: UConn	Name: Laura Yahn
SECTION 2– REQUESTOR INFORMATION	
<b>2.a Name (last, first) Chingos, Matthew</b>	
<b>2.b Director, Education Policy Program</b>	
<b>2.c Urban Institute</b>	
<b>2.d 2100 M St NW, Washington, DC 20037</b>	
<b>2.e <a href="mailto:mchingos@urban.org">mchingos@urban.org</a></b>	
<b>2.f 202-261-5664</b>	<b>Fax Number</b>
<b>2.g Provide Full name, organization, email address, and project role for ALL other persons within the requestor's organization who need to work with raw data, derived data or data output for this project before the P20 WIN Data Governing Board Participating Agencies involved have approved the data as appropriately aggregated for public release. Add additional space as necessary. Please print and attach a signed Confidentiality Agreement for each individual needing access to the resultant data set. (See Attachment C) Note: Persons who are not state government employees or officials may not have access to Pii for DOL UC records.</b>	

**Matthew Chingos, Urban Institute**

[mchingos@urban.org](mailto:mchingos@urban.org)

Principal Investigator

**Kristin Blagg, Urban Institute**

[kblagg@urban.org](mailto:kblagg@urban.org)

Research Associate

**Erica Blom, Urban Institute**

[eblom@urban.org](mailto:eblom@urban.org)

Research Associate

**Tomas Monarrez, Urban Institute**

[tmonarrez@urban.org](mailto:tmonarrez@urban.org)

Research Associate

**Victoria Rosenboom, Urban Institute**

[vrosenboom@urban.org](mailto:vrosenboom@urban.org)

Research Analyst

**Victoria Lee, Urban Institute**

[vlee@urban.org](mailto:vlee@urban.org)

Research Assistant

**Macy Rainer, Urban Institute**

[mrainer@urban.org](mailto:mrainer@urban.org)

Research Assistant

**2.h Provide full name, organization, email address and project role for all other persons from the Participating Agencies or Contributing Institutions who have been identified as having legitimate interests in the evaluation who need to access raw data, derived data or any data output for this project before the P20 WIN Data Governing Board Participating Agencies involved have approved the data as appropriately aggregated for public release.** For these individuals, access to the unit record data is limited. For example, named individuals may a) need access to data to support the data matching process or b) function as a subject matter experts from the Participating Agencies to support the data matching process or authorized requestor's analysis, or c) they may only have access to their own organization's data linked to other non-education records. After each name, please include an explanation as to why these individuals need access to the unit record data. **If not already signed, please print and attach a signed Confidentiality Agreement for each individual needing access to the resultant data set. (See Attachment C). Note: Persons who are not government employees or officials may not have access to Pii for DOL UC records.**

The following individuals will only have access to aggregate data (i.e. they will not have access to unit-record data that they do not already have access to):

- Dana Placzek, Research Analyst, Department of Labor; [Dana.placzek@ct.gov](mailto:Dana.placzek@ct.gov)
- Liam McGucken, Research Analyst, Department of Labor; [liam.mcgucken@ct.gov](mailto:liam.mcgucken@ct.gov)

- Jan Kiehne, Institutional Research Specialist, CSCU; [kiehnej@ct.edu](mailto:kiehnej@ct.edu)
- Laura Yahn, Research Analyst, UCONN; [laura.yahn@uconn.edu](mailto:laura.yahn@uconn.edu)
- Maura Provencher, Vice President of Research & Administration, CCIC; [provencher@theccic.org](mailto:provencher@theccic.org)
- Chuck Martie, Analyst, State Department of Education; [charles.martie@ct.gov](mailto:charles.martie@ct.gov)
- Bill Gammell, Director, Office of Research & System Effectiveness, CSCU, [gammellw@ct.edu](mailto:gammellw@ct.edu)
- Elsa Nunez, President, Eastern CT State University, [Nunez@EasternCT.edu](mailto:Nunez@EasternCT.edu)
- Lloyd Blanchard, Assoc. VP for Budget, Management & IR, UConn, [lloyd.blanchard@uconn.edu](mailto:lloyd.blanchard@uconn.edu)
- Jennifer Widness, President, CT Conference of Independent Colleges, [widnessj@theccic.org](mailto:widnessj@theccic.org)

### SECTION 3 –ALIGNMENT OF PURPOSE& SCOPE

**3.a Provide a brief summary of the Data Request.** Include a description of the categories of data, years or data and cohorts of students – if applicable. Provide enough detail so that the Participating Agencies can understand what they need to provide.

The Urban Institute aims to understand and further develop the next generation of college performance metrics through a collaborative effort bringing together policymakers and researchers. While we anticipate conducting a number of analyses, our main focus will likely be on conducting “return-on-investment” (ROI) calculations for colleges and universities in Connecticut. In order to undertake this, we need detailed student information prior to, during, and after college. In particular, we are seeking:

- demographic information, high school information (high school name/identifier, test scores, attendance, family income (if possible) or proxy (EFC from FAFSA, zip code or other geographic information)
- college attended, including length of time and intensity (course load), attempted (if possible) and completed majors, tuition and financial aid information, living circumstances during college, work during college (if possible), degrees/credentials acquired, dates of stop- or drop-out
- labor market information, including wages and industry.

In terms of cohorts of students, we would like:

- the cohort of students who graduated from high school in the 2009-2010 academic year, the 2010-11 academic year, and the 2011-12 academic year. We would also like their subsequent college enrollment (regardless of when they enroll) and labor market data, both before (if applicable), during, and after college. That is, we would like data for all their postsecondary enrollment(s) and each quarter of wage data that is available for them (regardless of whether or not they attended a postsecondary institution), through present.
- In addition, we would like data for all students who began college in Connecticut in the 2009-10, 2010-11, 2011-12, or 2012-13 academic year, regardless of whether (or when) they attended high school in Connecticut.

**3.b Clearly state the purpose of this request and describe how the purpose is an audit or evaluation of federal or state supported education program(s)** (See Audit or Evaluation Exception 20 U.S.C. 1232g(b)(1)(C), (b)(3), and (b)(5) and §§99.31(a)(3) and 99.35).

The ultimate goal of this project is to help state policymakers develop the capacity to evaluate postsecondary education programs. We will accomplish this by producing college performance metrics that best leverage Connecticut’s underlying state longitudinal data system (P20 WIN) while remaining simple enough for users to understand and incorporate into

their decision-making. This project will identify *what* college performance metrics are of greatest interest to state policymakers and *how* to enable the state to produce this information.

**3.c Describe the benefit that this audit/evaluation will provide to a local or State Education Authority or Agency (34 C.F.R. 99.1) and the state of Connecticut.** Note requests that only provide audit or evaluation of non-state Participating Agencies will not be approved.

Our analysis will leverage Connecticut's current data to provide nuanced insights into the performance of Connecticut's colleges and universities, in particular by incorporating background (high school) characteristics into our analyses. This will enable policymakers and state agencies to more accurately assess the performance of various institutions and to design or amend policies to achieve their desired outcomes.

**3.d Provide a brief description of the method for analysis.**

As described in question 3a above, we plan to undertake an ROI-style analysis. In order to calculate the needed counterfactual scenarios (i.e., wages or other outcomes if students did not attend the college/program in question), we will use a variety of common econometric techniques, including but not limited to linear regression, panel regression, fixed effects, propensity score matching, and hierarchical linear models.

**3.e Provide a description of the documents or reports that will reference data from this Data Request. For each report indicate the audience for the report and expected indicators or measures to be included in each.**

We will produce the following:

- A report on the results of our efforts to use student-level, longitudinal data to produce performance metrics geared towards policymakers. This report will draw on data from specific states but be geared toward a national audience and focus on broader lessons learned. We may also make recommendations for how states (and, if appropriate, the federal government) should change their data collection practices to address shortcomings identified through this work. The goal of this report is to provide a model for how states can make better use of their longitudinal data to production actionable information.
- A technical toolkit that helps state policymakers produce performance metrics from individual-level data. This toolkit will report the progress we make on the methodological issues described above, and provide practical guidance for states regarding subjects such as methods for adjusting metrics for differences in student characteristics, minimum sample sizes needed to obtain reliable data (especially at the program level), and the timing of measurement (e.g., when earnings are measured).
- A data dashboard tool that enables policymakers to explore a somewhat complicated dataset (e.g., program-level performance data with a wide variety of metrics). This dashboard will most likely be based on data from a single state and serve as "proof of concept" that complements the technical toolkit described above. We will make the source code (but not the underlying data) for the dashboard publicly available.

**3.f In addition to utilizing a minimum cell size as documented by the Data Governance Procedure, identify/describe the statistical methods that will be used to minimize the risk of re-identification of PII for data to be published. Some typical methods are listed below; however, this list is not comprehensive. Options should be used in combination for maximum security. See the [USED Technical Brief 3](#) as a resource.**

- **Suppression by:** cell, row, sampling (present data for a portion of students: e.g. 80%)
- **Blurring through:** aggregation of groups, rounding, use of ranges, use of top/bottom categories (e.g. 'less than 5%, greater than 95%),
- **Perturbation:** data swapping, adding noise, use of synthetic data

- We plan to use most of these techniques, including but not limited to suppression of data for small cell sizes, aggregation across cohorts or programs, the creation of synthetic cohorts, and the use of categorical rather than continuous outcome measures (ie, presenting “binned” data).

#### SECTION 4 – PROTECTION OF CONFIDENTIALITY

##### **4.a By what date will the datasets and all paper or electronic copies will be destroyed by the requestor?**

*After the Approved Requestor has received data files and the matrix of generic unique identifiers, the Approved Requestor has 12 months to complete analysis and destroy the data files. Note, that the Data Governing Board may set a shorter or longer time frame before the data files must be destroyed, and this 12 month limitation may be extended with written approval from the Data Governing Board. Requestors need to submit the Query Management Document as a ‘continuation’ for approval of an extension.*

Data files will be destroyed within twelve months of receipt of data.

##### **4.b With the addition of your signature at the bottom of this section, indicate that you agree to each of the following statements:**

- I will only utilize the data received through this Data Request to meet the purpose as described. The approval given to receive data through this Data Request does not confer approval to use it for another purpose.
- I will not re-disclose the data received through this Data Request approval process to any public official who has not been authorized by the Data Governing Board to receive it, and who has not also signed a Personal Statement of Confidentiality and Non-Disclosure.
- I will take the necessary and appropriate precautions to safeguard personal information and will comply with all state and federal laws concerning the safeguarding and disclosure of such information.
- I will not use the datasets to re-identify individuals.
- I agree that approval to receive data does not convey ownership of the data.
- I agree that prior to the public release of any documents or reports generated from this Data Request I will supply all reports or documents to the Data Governing Board for review and verification that the intended purpose has been adhered to.
- I agree to store all resultant data, in print or electronic form, in a locked receptacle that can be accessed by authorized persons only.
- I agree to store all resultant data on secure desktop computers and in secure files to which access is restricted to authorized persons only.
- I agree that no resultant data may be transmitted via email or placed or stored on a mobile computing or storage device. For purposes of this agreement, a definition of “mobile computing device” includes, but is not limited to, notebooks, palmtops, PDAs, iPods®, Blackberry® devices, and cell phones with internet browsing capability. A “mobile storage device” includes but is not limited to, mobile computing devices, diskettes, magnetic tapes, external/removable hard drives, flash cards (e.g., SD, Compact Flash), thumb drives (USB keys), jump drives, compact disks, digital video disks, etc.
- I agree to destroy all confidential information obtained through this agreement as soon as such information is no longer needed.
- I agree that the Data Governing Board will be allowed access to monitor all authorized users to ensure such users adhere to the confidentiality requirements of the information obtained under this Agreement.
- I agree to comply with all provisions of the P20 WIN Data Request Management Procedure.

**I, Matthew Chingos, agree with each of the statements in section 4.b.**

#### SECTION 5—ADDITIONAL INFORMATION

5.a Provide any additional comments that would be useful to the Data Governing Board in considering this request.

**SECTION 6 – SIGNATURE OF REQUESTOR AND DATE**

**6.a Signature of Requestor**

I understand that the entities that are providing data to P20 WIN (hereinafter Participating Agencies) have made reasonable efforts to ensure that the data available through P20 WIN are up-to-date, accurate, complete and comprehensive at the time of disclosure. These records reflect data as reported to the Participating Agencies by their data-submitting organizations for the reporting period indicated. Changes or updates to the data may occur after the time of disclosure and may impact data that have previously been made available. The Participating Agencies are not responsible for data that are misinterpreted or altered in any way. Derived conclusions and analyses generated from this data are not to be considered attributable to the Participating Agencies or the participating organization(s) from which the data originated.

I certify that the information supplied in this form, with attachments, is complete, accurate. The analysis will be conducted according to the protocol approved by the Data Governing Board, applicable federal, state and local laws regarding the protection of education records and unemployment insurance records. I will ensure that all protocol changes will be prospectively reviewed by the Data Governing Board. I will request approval from the Data Governing Board for changes to the Data Request and will not implement proposed changes until I receive Data Governing Board approval. I will promptly report to the Data Governing Board any related complaints, problems, and/or breaches of confidentiality.

Signature Walter Ching Date December 20, 2018