# Post-Secondary Employment Outcomes

## Using National Jobs Data to Measure Graduate Impact

Maryland SDC
October 26, 2023
Andrew Foote
Principal Economist, Longitudinal Employer-Household Dynamics (LEHD) Research
Center for Economic Studies, Census Bureau

Any opinions and conclusions expressed herein are those of the author and do not represent the views of the U.S. Census Bureau. All tables and figures use publicly available PSEO data. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7517475: CBDRB-FY18-103)

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

# Outline

- What is the Post-Secondary Employment Outcomes (PSEO)?
- What does it measure?
- How do we create the PSEO?
- How are the data protected?

# What is PSEO?

An experimental data product based on partnerships between university systems, state higher education systems, and the Census Bureau, the **Post-Secondary Employment Outcomes (PSEO)** are tabulations providing national earnings and employment statistics for graduates of post-secondary institutions.

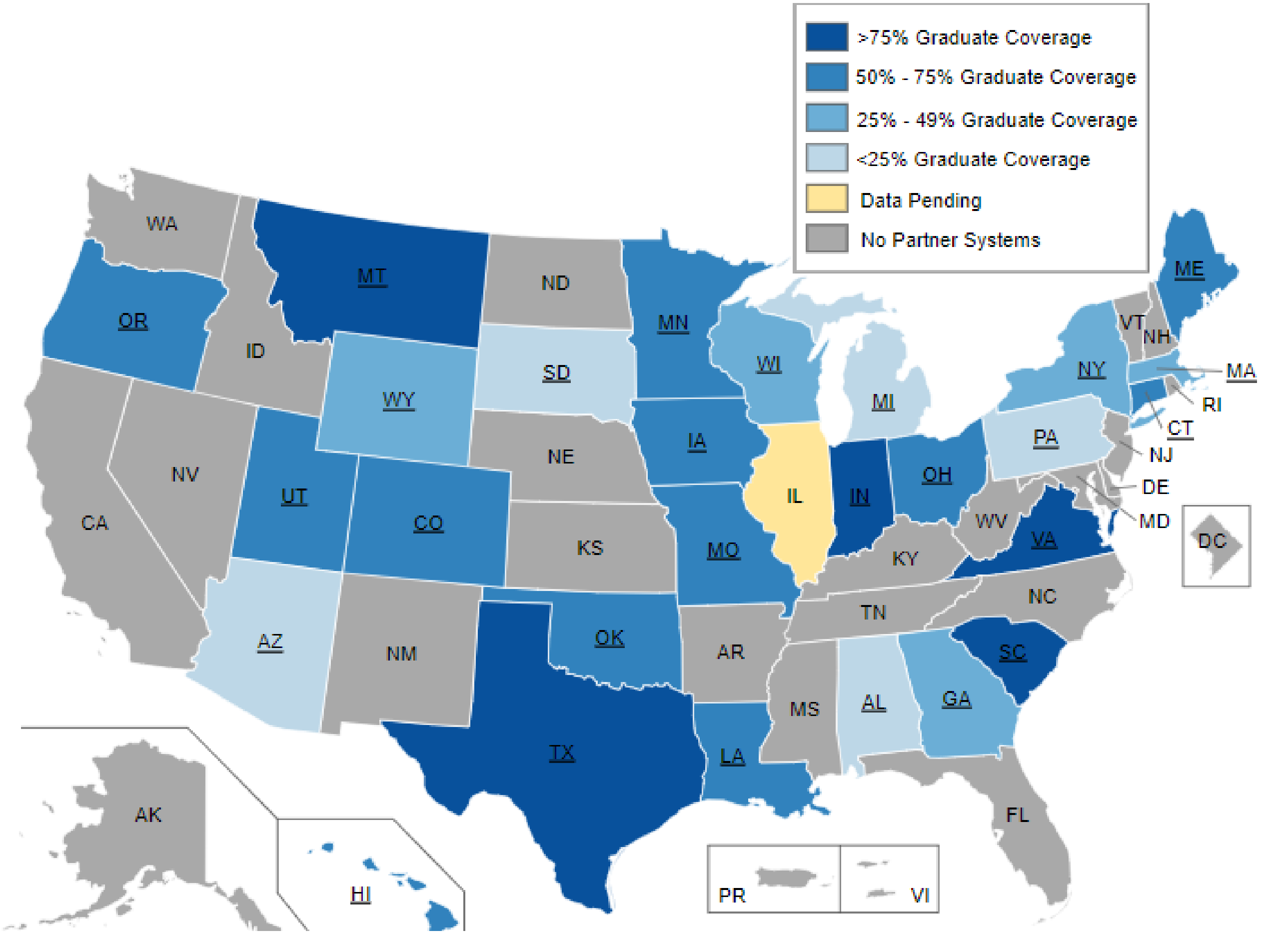
**Goal:** Provide students, parents, institutions and workforce agencies better data on the return on investment to post-secondary degrees and the flows of graduates across the country.

# What does PSEO currently release?

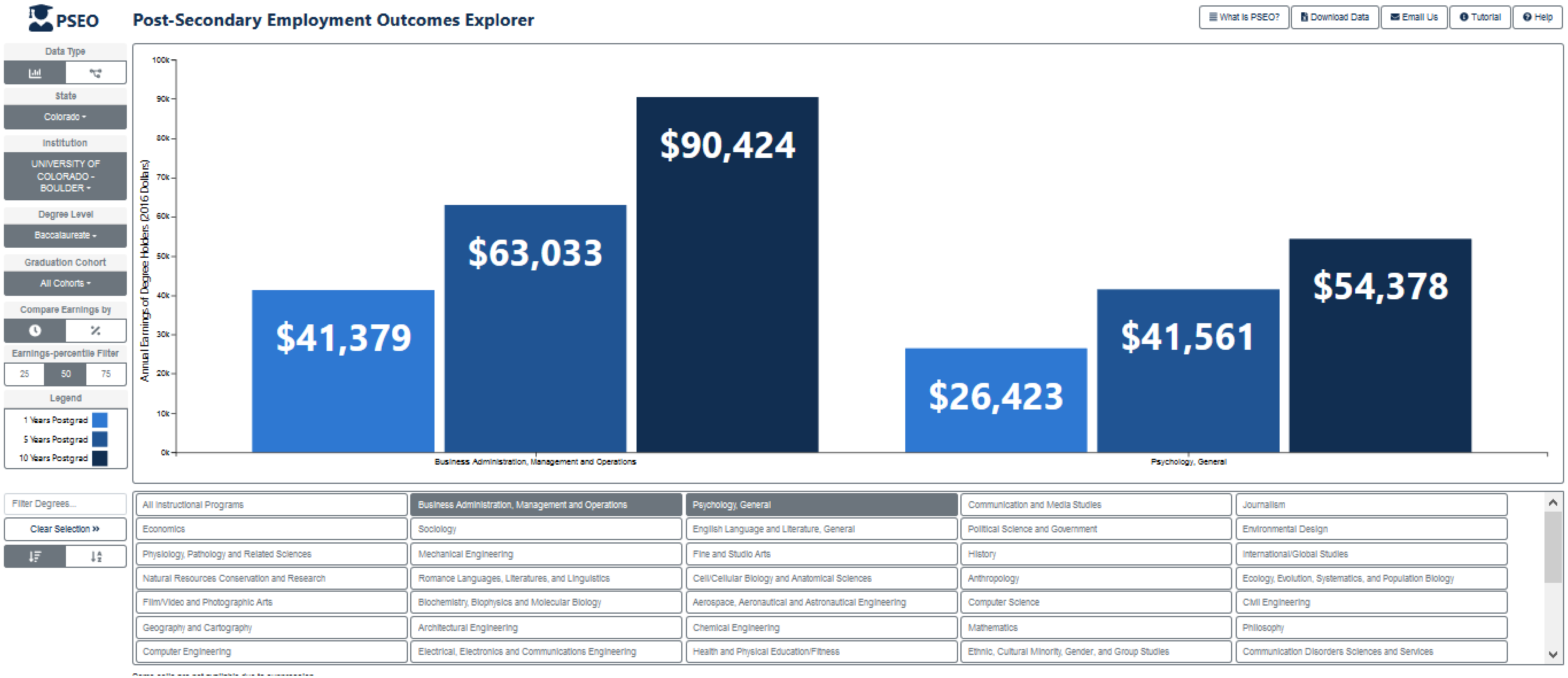
- **Earnings**: 25th, 50th, and 75th percentiles of annual earnings for college and university graduates
- **Employment Flows**: Employment within industry sector and Census division
- Statistics available by institution, degree level, graduation cohort, and field of degree, one year, five years, and 10 years after graduation

# PSEO coverage

- Current coverage
  - 825 public and private four-year and two-year institutions
  - ~29% post-secondary awards (using 2015 AY as baseline)



# Post-Secondary Employment Outcomes (PSEO) Explorer



# How do we create the PSEO?

# How are the PSEO data created?

- Census enters into legal agreements with state departments of higher education (or similar entities) and receives records of graduates from all covered institutions
- Using a masked version of the SSN, we match these data with the unemployment insurance wage records, which we obtain from state employment security agencies
- These longitudinal job histories allow us to measure earnings for graduates over a long time horizon

# Jobs data

- Census maintains a national database of jobs data, which includes information about employers (establishments, industry, location) and employees (demographics)
- Enables us to measure earnings (from quarterly wages) and employer characteristics (industry sector and employment location)
- Longer term, there is a broad effort at Census to create an integrated jobs frame using UI, W2 and Schedule SE records (prototype due at end of FY2024)

# Looking forward

- In PSEO 2.0, we are developing three major updates to the data we release:
  - Enhanced wage records (using new jobs frame, which includes W2 and 1099 records)
  - Earnings outcomes by gender and race/ethnicity
  - Release of state by industry flows, which would be incredibly helpful from a workforce perspective

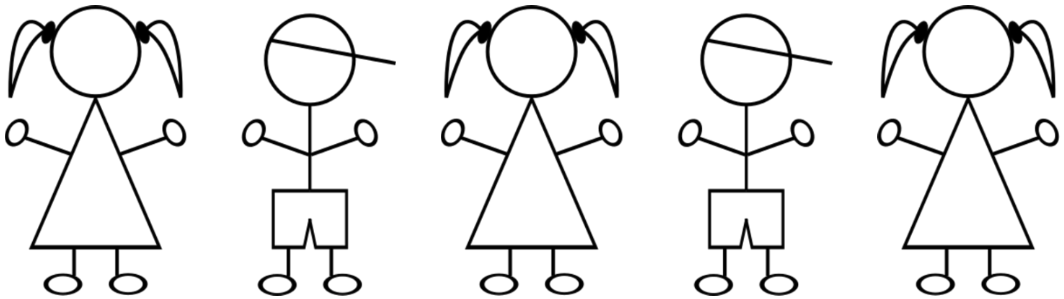
# How do we protect student privacy?

# Protecting the microdata

- Title 13 requirement:
  - The existence of a job held by an individual is confidential
- **We do not have a monopoly on microdata** – and most of our partners have access to the frame (all graduates) and most of the earnings data we use to produce the statistics

# Example disclosure risk

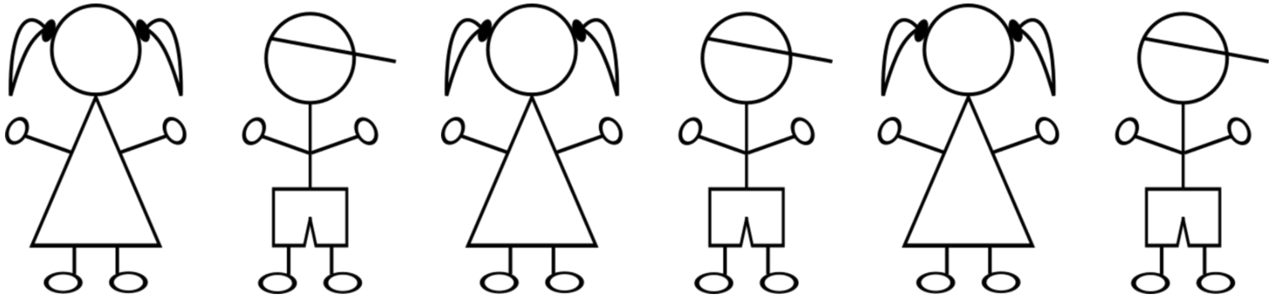
Data owner (A) releases median earnings for *in-state* graduates



Median earnings (in-state):
$80,000

# Example disclosure risk

Data owner (B) releases median earnings for *all* graduates, nationally

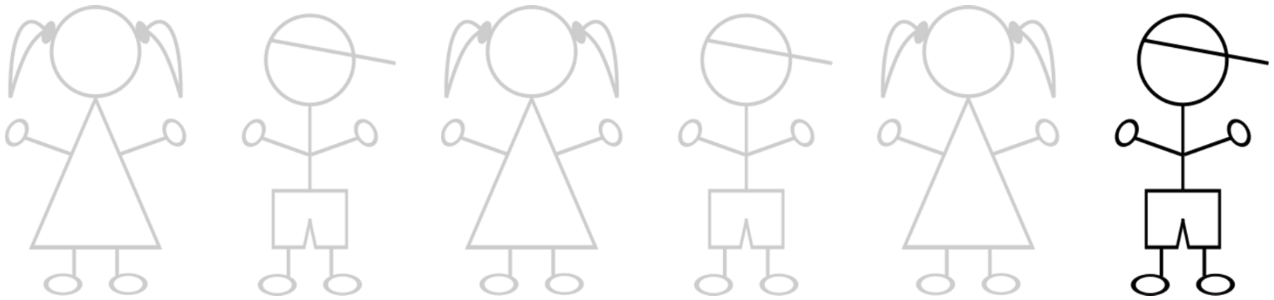


Median earnings (in-state): $80,000

Median earnings (national): $85,000

# Example disclosure risk

The difference in earnings reveals the existence of a job for the out-of-state graduate!



Median earnings (in-state):
$80,000

Median earnings (national):
$85,000

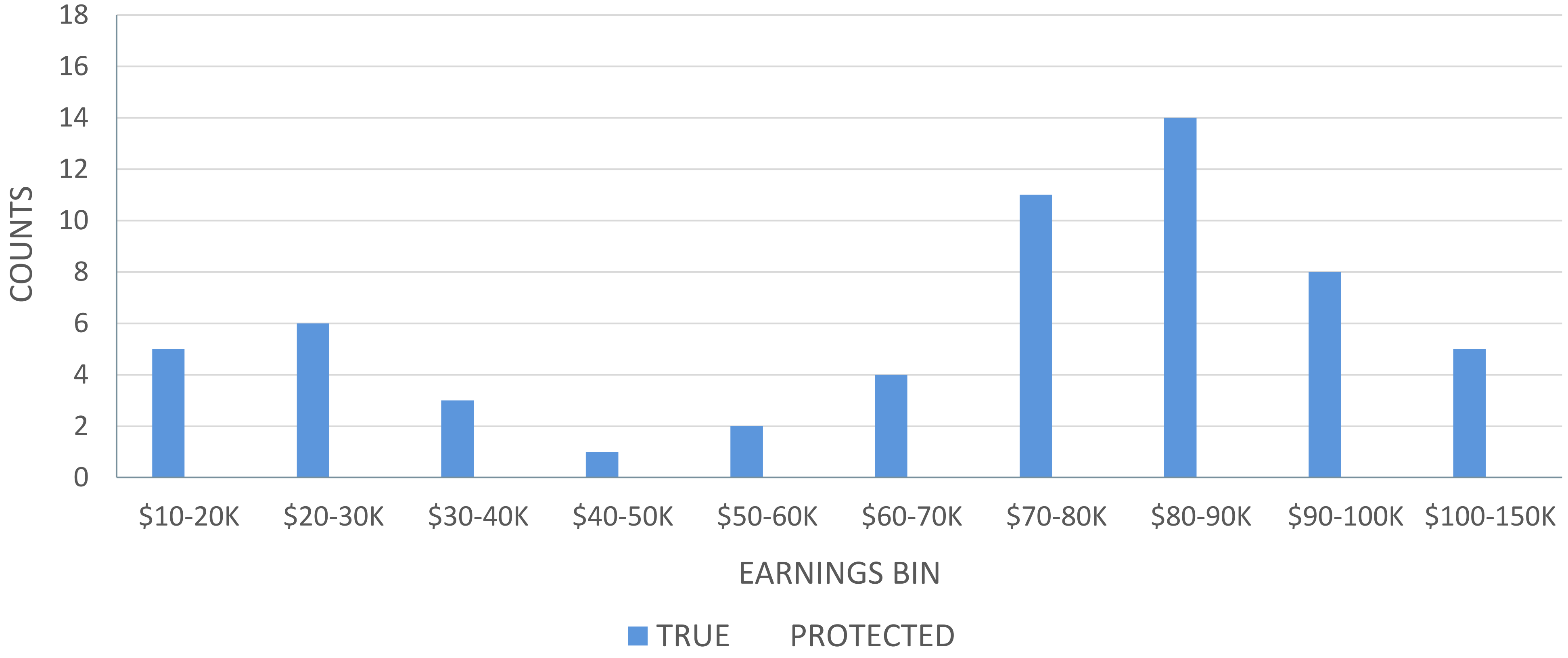
# Our solution: differential privacy

- Student data is de-identified and PII data are destroyed; anonymous identifiers are used to link graduates to jobs data
- All tabulations are protected using state-of-the-art **formal privacy** disclosure avoidance methods
- Mathematically provable information/privacy trade-offs
  - Any one individual in a database used for analysis should be indifferent between being in the database or not
- Methodology for earnings described in Foote, Machanavajjhala, and McKinney (JPC, 2019)

# How we implement differential privacy

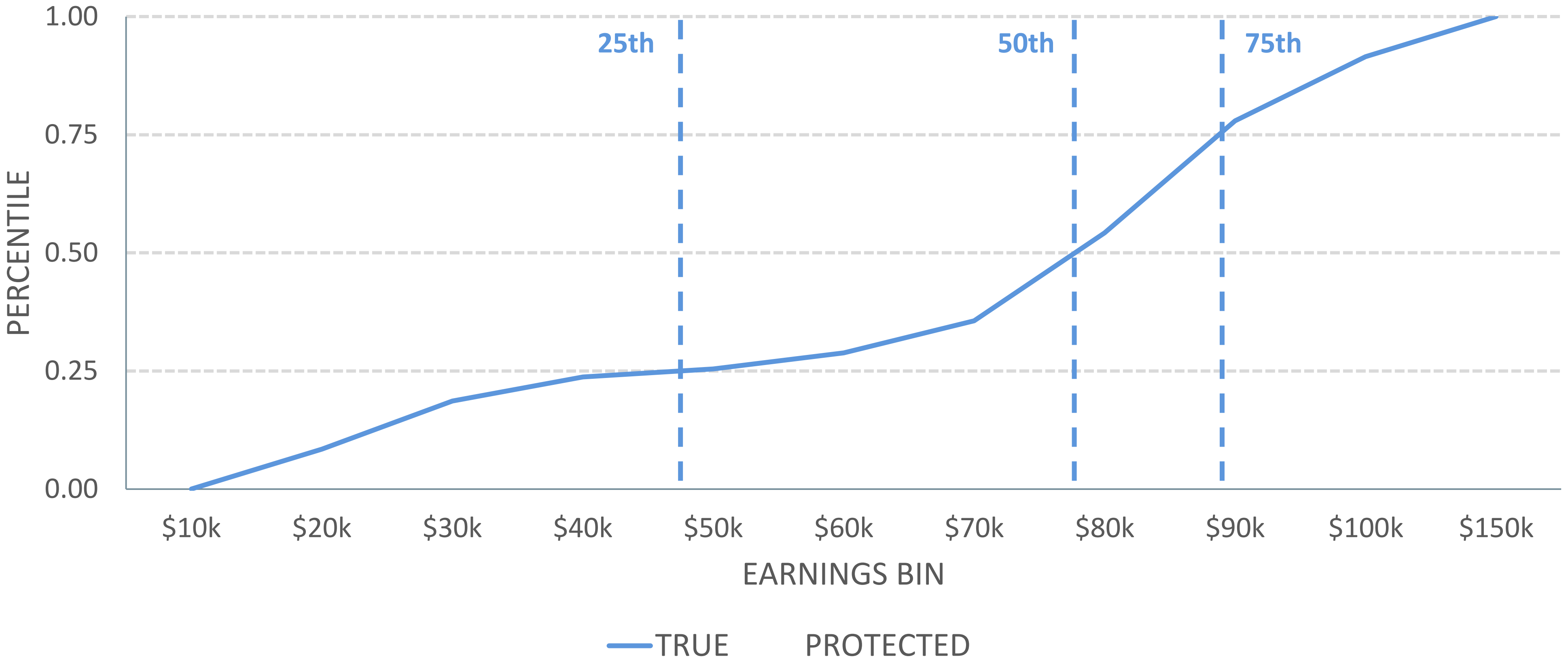
- Construct a histogram of earnings
  - Log-normal based on ACS public-use sample
  - Histogram bin ranges are public information
- Add noise to each histogram bin
- Extract percentiles from the resulting protected CDF

# Protecting histogram cells with noise

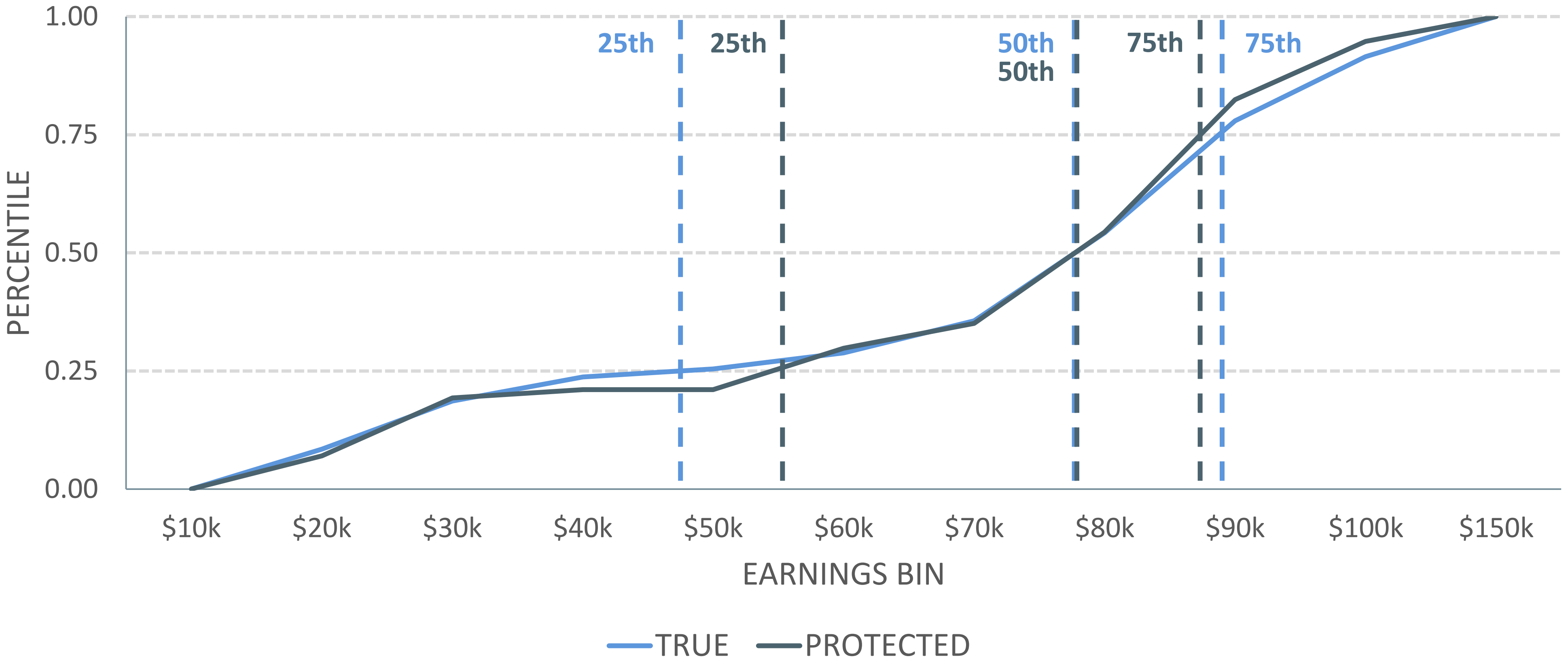


# Protecting histogram cells with noise

# Extracting percentiles from histogram



# Extracting percentiles from histogram



# Advantages of this DP implementation:

- We can aggregate to higher levels and calculate percentiles at those levels (for instance, median earnings at an institution)
- Easily implemented in other situations – Veterans Employment Outcomes used the same method, but with different histogram cutoffs
- Protection is intuitive for general public
- Code available for broader use