Returns to education and your education data warehouse

Greg Weeks

Forecasting Division

Office of Financial Management

Pacific Northwest Regional Economic Conference

Vancouver, Washington

May 2016

EDUCATION RESEARCH
&
DATA CENTER

Outline

1. Overview and brief discussion of non-experimental evaluation research.

2. Description of Propensity Score Matching (PSM)

3. Impacts of PSM

4. STEM Results (with and without PSM)

5. Discussion/Next steps.

But first, a word from our sponsor… ERDC

Education Research & Data Center

Washington state Education Research & Data Center

- **ERDC created in 2007 to:**
  - Act as objective broker for education and workforce data
  - Assemble, link and analyze education and workforce data
  - Provide research focusing on student transitions
  - Make data available to the education agencies and institutions
- **Located in Governor's budget agency** (Office of Financial Management)
- **Work closely with State Education Agency (OSPI)**
- **Working on second SLDS and WDQI grants**
  - Focus on research and reporting projects
  - Broadening subject areas to human services, corrections and data visualization
  - Continue to operate the ERDC data warehouse

ERDC
Education Research & Data Center

Rigorous evaluation studies matter for programs that enrich human capital

- Often required for US Department of Labor grant funded programs

- Help define evidence-based approaches that work

- Best practice

- Efficiency

- Help target audiences

- Often analytically challenging

The problem

With random assignment (such as clinical trials and experimental evaluation designs):

1. the outcome of the treatment is conditionally independent from the treatment.
   1. Chosen at random, the treatment and control groups are statistically identical
   2. *The only difference is one group has the treatment, the other does not*.
2. In observational non-experimental studies this assumption is invalid, resulting in "selection bias."
   1. *The treatment group may have better outcome measures even in the absence of a treatment*.
   2. *Measured outcomes reflect both the differences in the groups and the differences attributable to the treatment.*

Selection bias

- Selection bias occurs when observable or unobservable factors influence both the decision to participate in the treatment and the outcomes.

- For example, our Bachelor's degree study assumes that college graduates differ from high school graduates in ways that affect both the likelihood of attending and completing college, *and* post-graduation earnings.

- *Simple (unadjusted) comparisons of earnings by educational attainment lead to selection-biased (over-stated) estimates of the earnings premium associated with a college degree.*

ERDC
Education Research & Data Center

Propensity score matching (PSM)

- Propensity score matching is utilized to develop a closely matched comparison group and correct selection bias.
- A propensity score is the estimated probability that an individual from the treatment or comparison group will participate in the treatment.
- This single measure indexes all the variables in the characteristics vector and provides a selection corrected comparison of the outcomes between the two groups.
- Estimated propensity scores allow individual treatment group members to be matched with and compared to individual comparison group members.

ERDC
Education Research & Data Center

PSM- the counterfactual

- "PSM uses information from a pool of units that do not participate in the intervention to identify what would have happened to participating units in the absence of the intervention"

    – Heinrich, C., Maffioli, A. and Vazquez, G. "A Primer for Applying Propensity Score Matching". Office of Strategic Planning and Development Effectiveness. Inter-American development Bank. 2010. Retrieved from: http://publications.iadb.org/bitstream/handle/11319/1681/A%20Primer%20for%20Applying%20Propensity-Score%20Matching.pdf?sequence=1

Requirements for PSM

- Comparison group roughly equivalent in size to treatment group.
  - Applicants for the training or educational program not accepted into the program.
  - SLDS educational data warehouse – may be able to provide an anonymized comparison group from same high school classes, or by gender or age.
- Clearly defined treatment(s) – start date, end date, time for follow up in UI wage record (often a six month lag).
- Pre-treatment descriptive data – the SLDS educational data warehouse may be able to help with this.
- Clearly defined outcomes/effects – often UI wage data.

ERDC
Education Research & Data Center

Basic PSM Process

Once data is assembled for both treatment and comparison groups:

1. Use logistic regression using pre-treatment variables to predict the probability (propensity score) of participating in the treatment (using both groups together)

2. Match comparison group members to treatment group members based on this propensity score.

   There are several matching approaches including with or without replacement, nearest neighbor, weighted, …)

3. The difference in outcome measures of the treatment group and the matched comparison group is the measure of program net impact or effect.
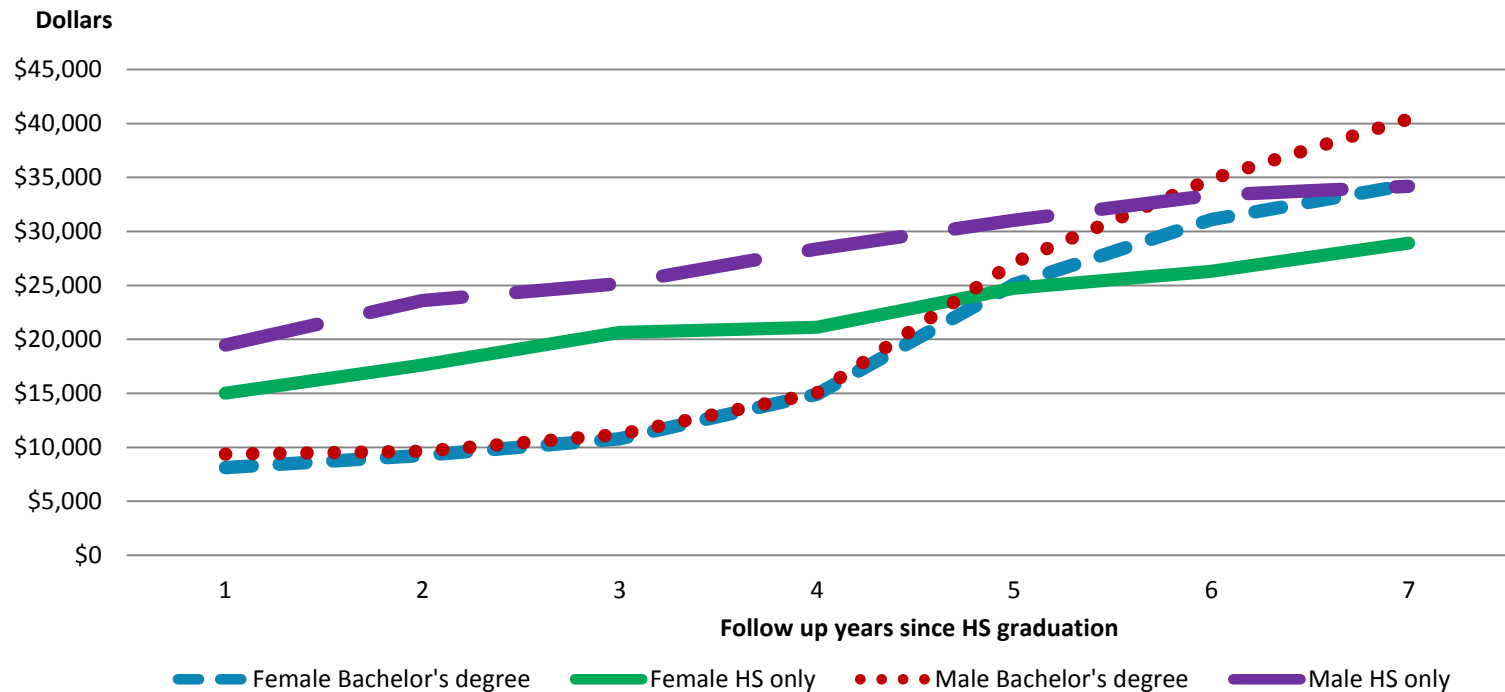
There is a substantial literature on PSM. I would recommend starting with:
http://publications.iadb.org/bitstream/handle/11319/1681/A%20Primer%20for%20Applying%20Propensity-Score%20Matching.pdf?sequence=1

Some examples from our ERDC research

- Returns to a Bachelor's degree by gender: (http://www.erdc.wa.gov/sites/default/files/publications/201403_0.pdf)

- Returns to STEM degrees by gender and race categories: (http://www.erdc.wa.gov/sites/default/files/publications/EarningsPremiums-STEMBachelorDegrees.pdf)

- Returns to an associate degree by gender: (http://www.erdc.wa.gov/sites/default/files/publications/201501.pdf)

- Comparison of pathways to a bachelor's degree: http://www.erdc.wa.gov/sites/default/files/publications/CompPathwaysBachDegreesWashState.pdf

ERDC
Education Research & Data Center

Female and male earnings trajectory, bachelor's degree and high school only, PSM, 2012 dollars, follow up years 1-7.

Dollars

$45,000
$40,000
$35,000
$30,000
$25,000
$20,000
$15,000
$10,000
$5,000
$0

Follow up years since HS graduation

- - - Female Bachelor's degree   —— Female HS only   ••••• Male Bachelor's degree   —— Male HS only

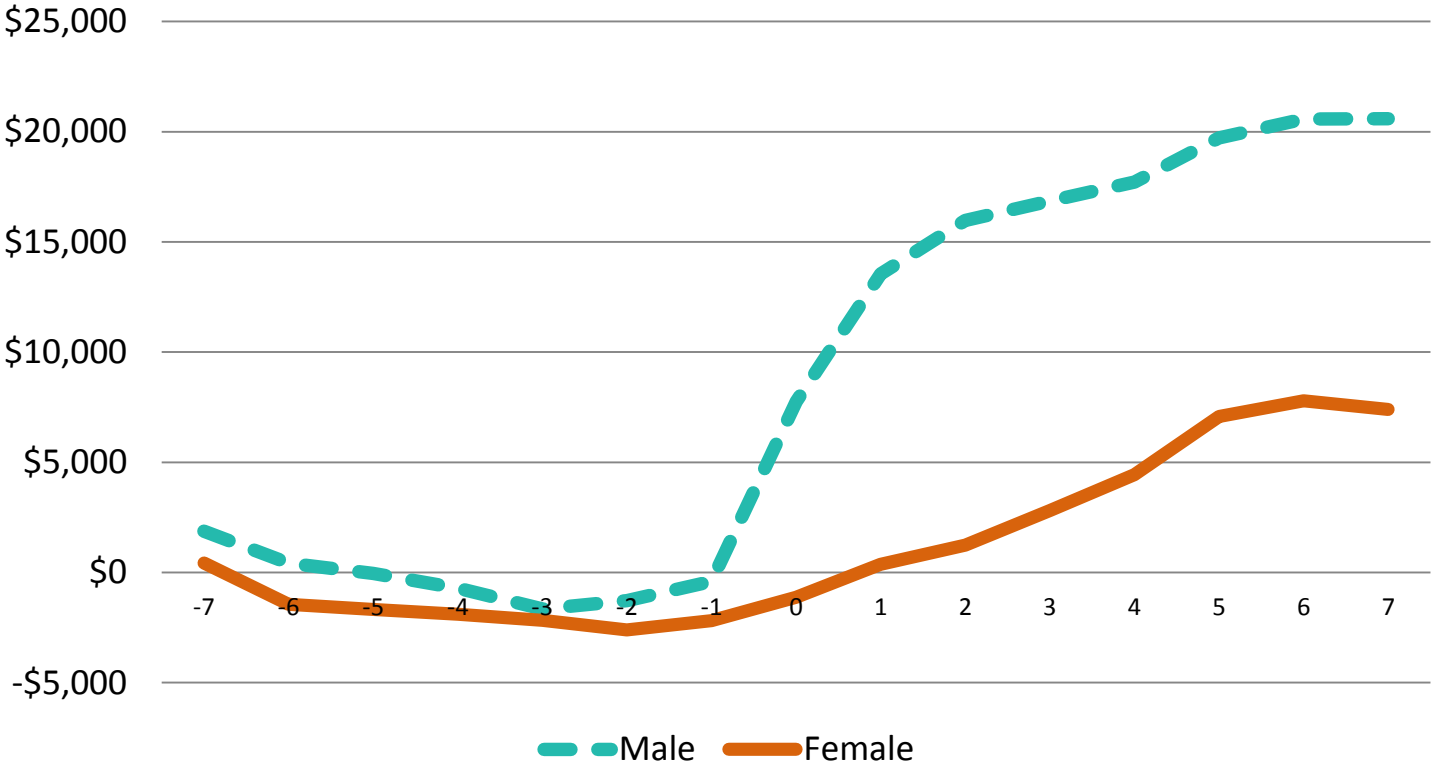**ERDC**
Education Research & Data Center

Female bachelor's degree earnings differentials, with and without PSM; current dollars
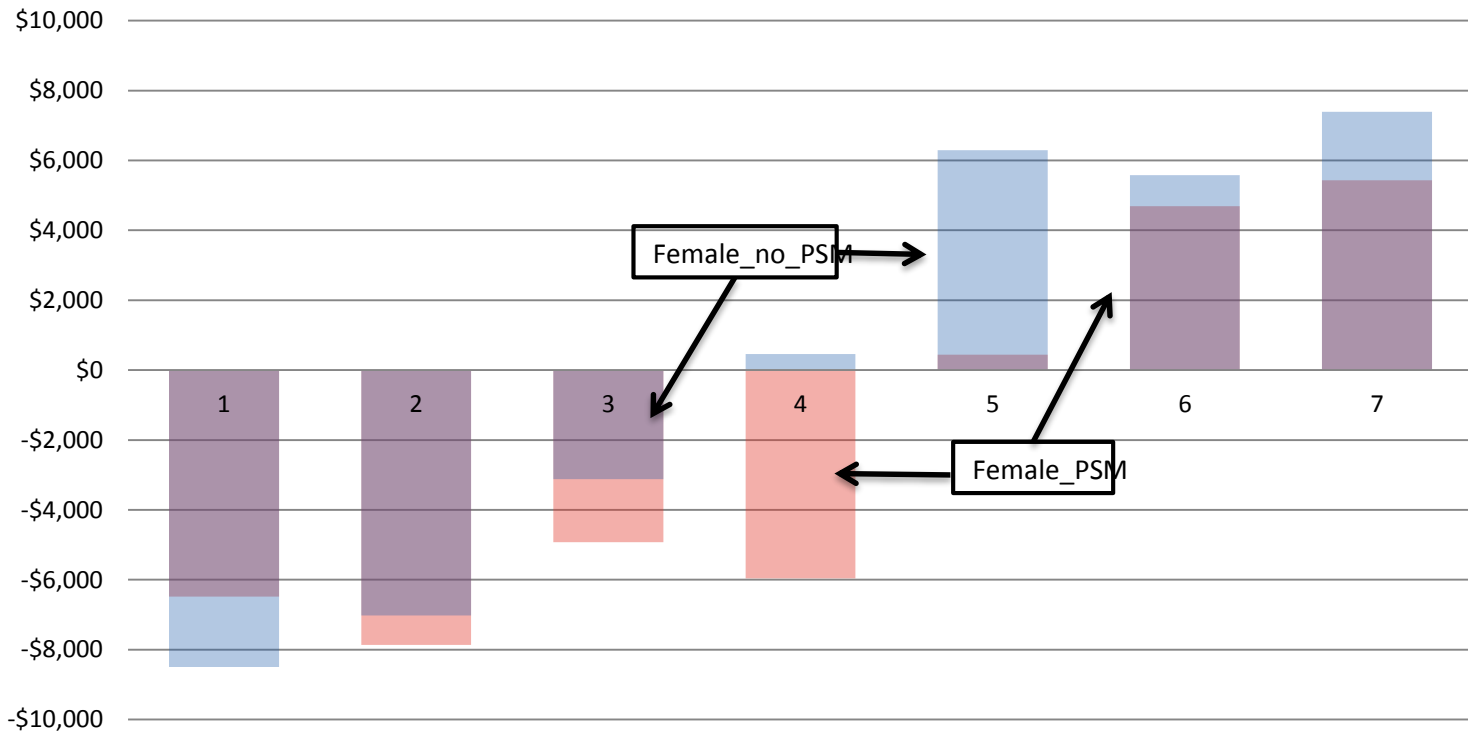
Male bachelor's degree earnings differentials, with and without PSM; current dollars
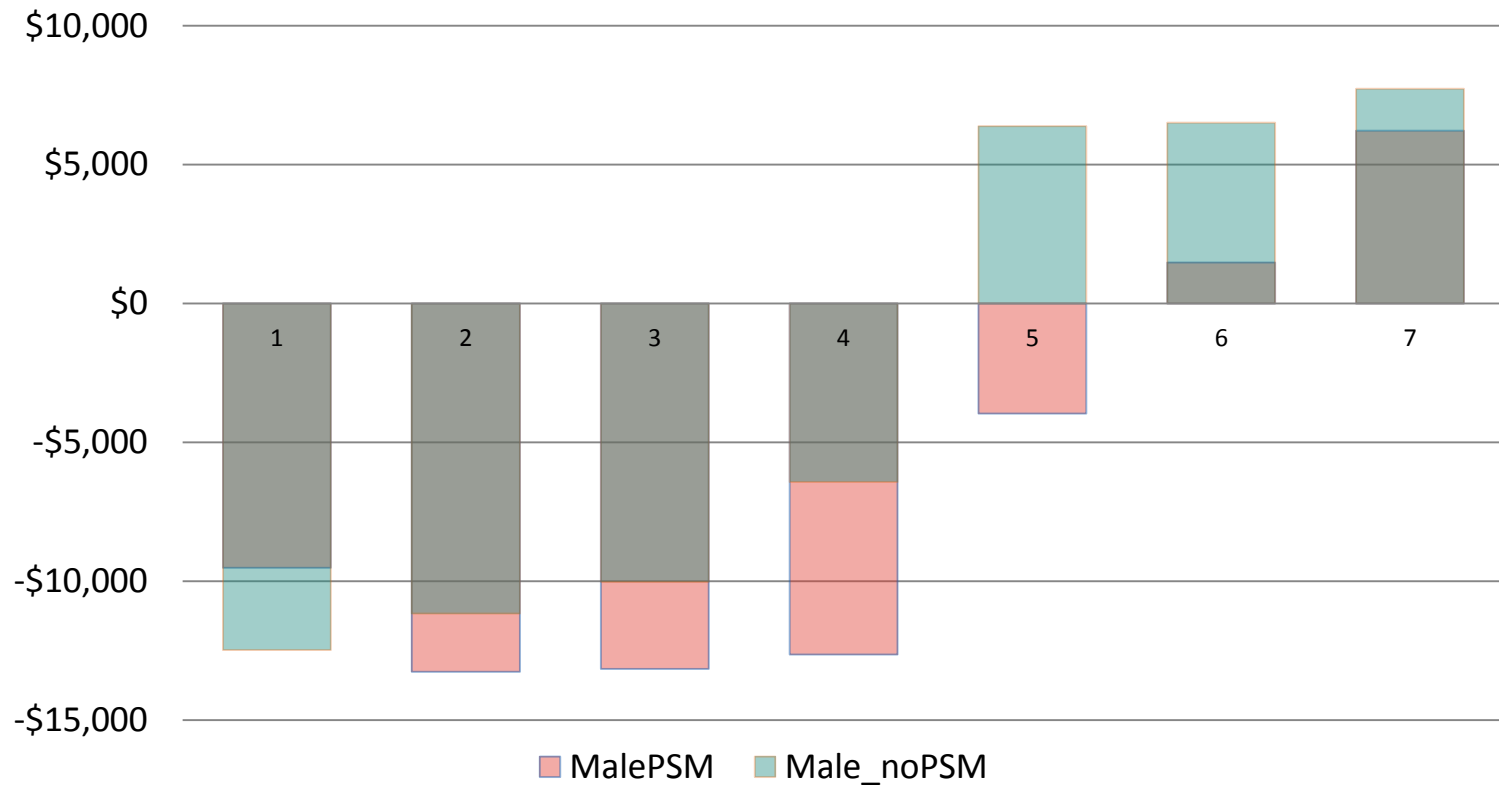
Female and male STEM earnings premium in current dollars for years before and after graduation (year 0), no PSM adjustment

Female and male STEM earnings premium, 2013 dollars, years after HS graduation (year 0), with PSM adjustment

Legend: Male STEM earnings premium (dashed teal line), Female STEM earnings premium (orange line)
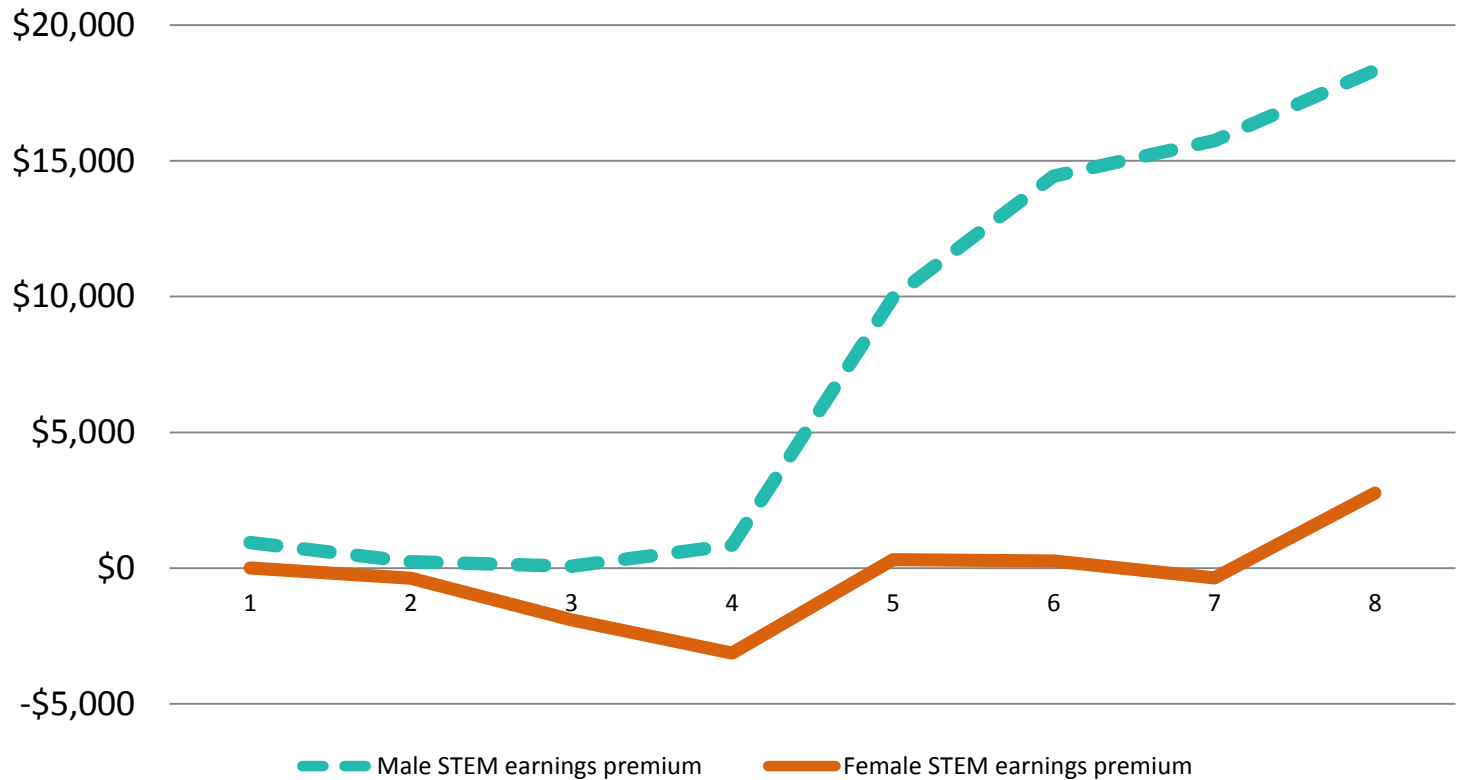
Proportion of STEM graduates in occupations by gender

Male		female	
Software developers, applications and system software	20.85%	Registered nurses	9.00%
Civil engineers	3.92%	No occupation	5.01%
Postsecondary teachers	3.30%	Customer service representatives	3.46%
Accountants and auditors	2.69%	Software developers, applications and system software	3.23%
Computer programmers	2.54%	Postsecondary teachers	3.21%
Computer support specialists	2.38%	Counselors	3.11%
Managers, all other	2.37%	Secretaries and administrative assistants	2.82%
Sailors and marine oilers	2.32%	Miscellaneous life, physical and social science technicians	2.49%
Mechanical engineers	2.24%	Managers, all other	2.38%
Carpenters	2.21%	Physical therapists	2.16%
Total top ten	44.83%	Total top ten	36.86%

ERDC
Education Research & Data Center

Median overall wage rates for top occupations of STEM graduates by gender

Male		Female	
Software developers, applications and system software	$52.70	Registered nurses	$36.74
Civil engineers	$39.15	No occupation	---------
Postsecondary teachers	$39.09	Customer service representatives	$17.48
Accountants and auditors	$32.16	Software developers, applications and system software	$52.70
Computer programmers	$53.66	Postsecondary teachers	$39.09
Computer support specialists	$25.32	Counselors	$21.47
Managers, all other	$50.48	Secretaries and administrative assistants	$17.75
Sailors and marine oilers	$22.77	Miscellaneous life, physical and social science technicians	$30.25
Mechanical engineers	$42.68	Managers, all other	$50.48
Carpenters	$22.68	Physical therapists	$17.48
Weighted average hourly wage rate	$44.23	Weighted average hourly wage rate	$28.12

Takeaways

- Rigorous evaluations of job skills training and job search assistance programs are more often possible than sometimes assumed.

- Selection bias should be taken into account whenever possible (PSM)

- A rigorous PSM study implies rigorous data requirements (treatment and comparison groups)

- Use your SLDS education data warehouse as a source of data and as a partner.

Thank you for your time and attention. Questions?

Greg Weeks, Ph.D.
Greg.weeks@ofm.wa.gov
(360) 902-0660