

STEM Training and Early Career Outcomes of Female and Male Graduate Students: Evidence from UMETRICS Data linked to the 2010 Census

By Catherine Buffington, Benjamin Cerf Harris, Christina Jones, and Bruce A. Weinberg

Online Appendix

Appendix Table 1. Description of Sample Construction

Data Descriptions	Observations	Notes
UMETRICS Payroll Records linked to 2010 Census	3,551,730	This is the employee file from 4 universities linked to the 2010 Census. There is a record for every payment to every employee.
Linked to ProQuest	11,773	ProQuest data include all doctoral recipients from the 4 universities. ProQuest files do not include UMETRICS employees who did not obtain doctorates from the university in question (e.g., undergraduates, masters students, post doctoral researchers, faculty, and staff).
2007-2010 Graduating Cohort	3,837	Based on ProQuest year of degree.
Linked to placement and earnings composite	2,008	Not all matches have earnings information in every year.
Sample Restrictions	1,237	Keep those with PIKs; keep only those with STEM dissertation topics; keep those aged 24 to 40; drop students whose main awards do not include faculty members; keep those appearing in UMETRICS after 2000 and before 2010. Drop observations with missing values of analysis variables. These missing values are entirely due to missing demographic information. Some matches to the 2010 Census may have missing demographic information because we are not using imputed or proxy responses.

Source: Author calculations. UMETRICS linked to 2010 Census, ProQuest, LEHD, W2, LBD, BR, and iLBD

Appendix Table 2. Summary Statistics

	Females		Males	
	Mean	SD	Mean	SD
2010 Census Variables				
White Alone	0.57	0.50	0.57	0.50
Black Alone	0.03	0.18	0.02	0.14
Asian Alone	0.39	0.49	0.40	0.49
Neither White Alone, Black Alone, nor Asian Alone	0.01	0.10	0.01	0.10
Hispanic	0.03	0.16	0.03	0.16
Age	30.40	2.90	30.42	2.93
Married or Partnered	0.62	0.49	0.63	0.48
Presence of Children Under 17	0.19	0.39	0.24	0.43
ProQuest Variables				
Dissertation in Biology, Chemistry, or Health	0.59	0.49	0.27	0.45
Dissertation in Engineering	0.21	0.41	0.45	0.50
Dissertation in Computer Science, Math, or Physics	0.19	0.40	0.28	0.45
Year of Degree	2008.61	1.11	2008.72	1.07
UMETRICS Variables				
First Year in Data	2005.41	1.80	2005.49	2.04
Left Censored in Data	0.29	0.46	0.34	0.47
Observations	370		867	

Source: Author calculations. UMETRICS linked to 2010 Census, ProQuest, LEHD, W2, LBD, BR, and iLBD