

Merging UMETRICS Data with Census Employee Data to Assess Human Capital Outcomes

Nikolas Zolas and Nathan Goldschlag
Center for Economic Studies
U.S. Census Bureau

Food Safety Workshop
December 1, 2015

Disclaimer: Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.



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



Overview

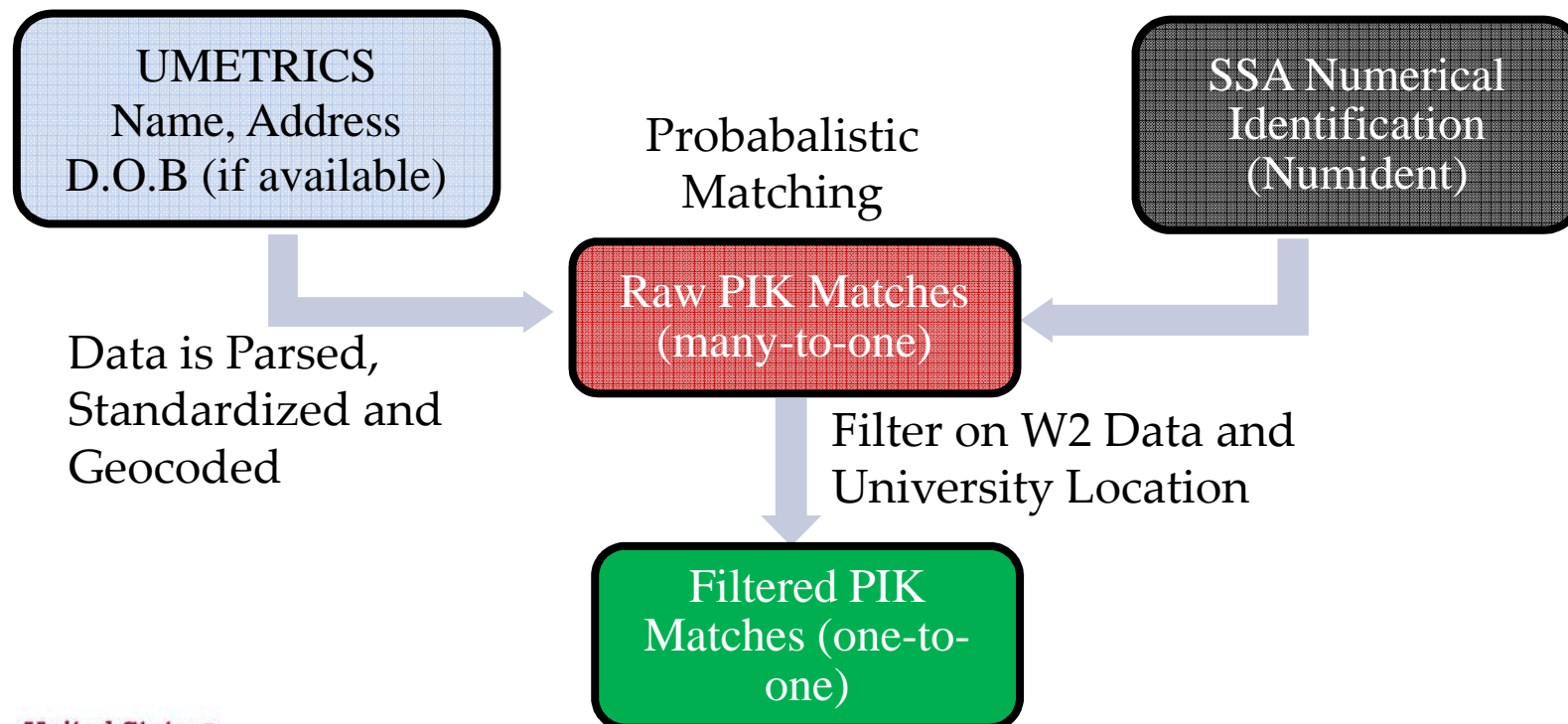
- UMETRICS provides us with a host of employee characteristics not found anywhere else
- Merging UMETRICS with Census confidential microdata allows us to track the outcomes of employees
 - Wage Outcomes
 - Location Outcomes
 - Sector and Industry Outcomes
 - Entrepreneurial Outcomes
- As well as outcomes of the firms and establishments they work for
 - Employment Growth
 - Patents
 - Employment Networks

Data Description

- **Business Register (BR)**
 - Universe of U.S. non-agricultural businesses and the source of data from which all other economic data are ultimately created
 - Key data provided: Industry Classification (NAICS), Geographic data, Employment, Payroll, EIN Codes, Available from 2002-2013
- **Longitudinal Business Database (LBD)**
 - Universe of employer businesses, unique establishments, the LBD covers all industries and all U.S. States linked over time
 - Key data provided: Industry Classification (NAICS), Geographic data, Employment, Payroll, Firm Age, Available from 2002-2013
- **Integrated Longitudinal Business Database (iLBD)**
 - Universe of non-employer businesses with links to employer universe
 - iLBD records are identified by either PIKs or EINS, 85-88% are PIKs and 12-15% are EINS
 - Key data provided: Industry Classification, Gross Receipts, Geographic data, Available from 2002-2010
- **Longitudinal Employer-Household Dynamics (LEHD)**
 - Employee-Employer linked dataset
 - Key data provided: EIN-Geocode Linkage, Wage Data, Available from 2002-2010
- **W2 Data**
 - Key data provided: PIK, Wage Data, Available from 2005-2014

PVS Process

- Before any matching can take place, each employee must be assigned an anonymous, unique person identifier through the US Census Bureau's Personal Identification Validation System (PVS)



PVS Process (Cont.)

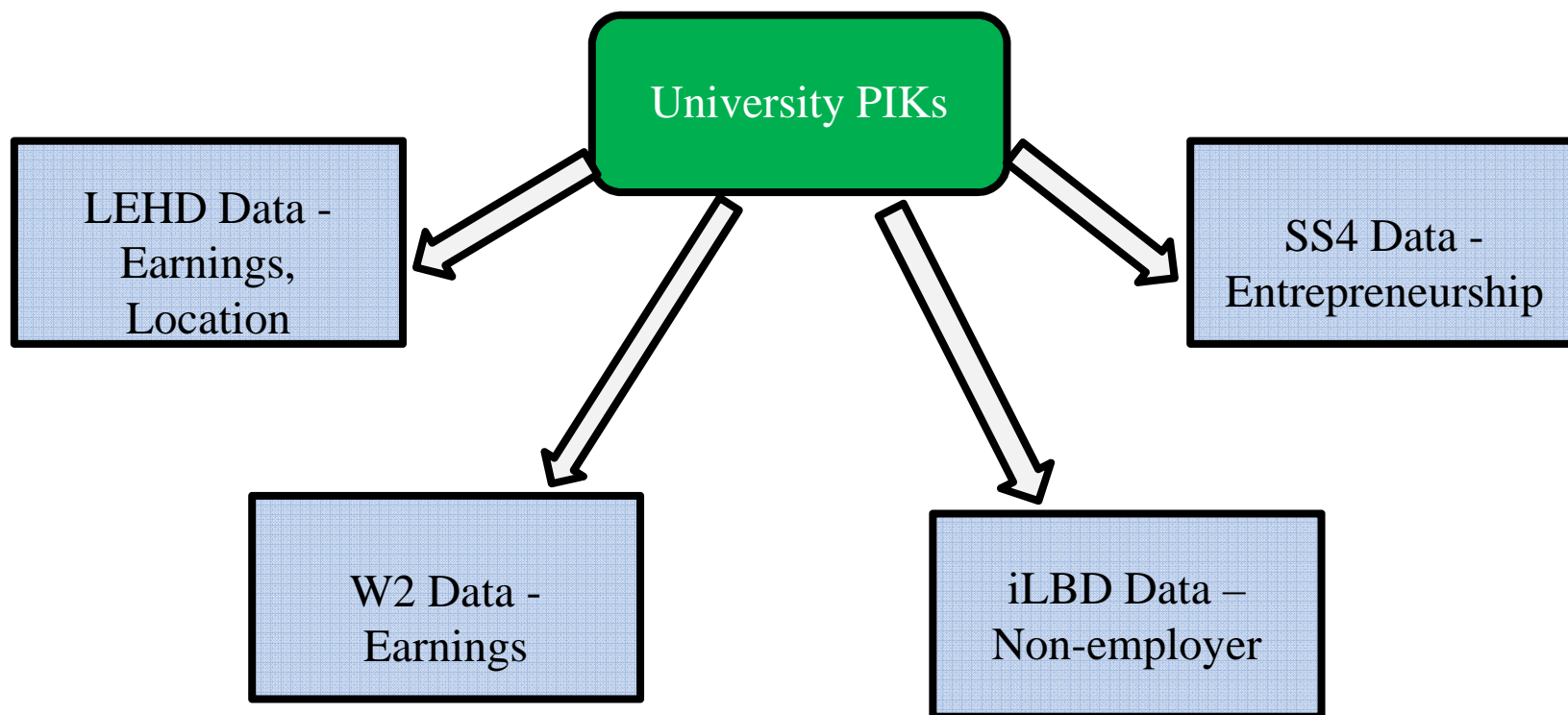
- PIK rates vary by University and what details of the University Employees we receive (i.e. birthdates, W2 Data, etc...)

PIK Rate by University

Last Year	A	B	C	D	E	F	G	H	I
2008	65.6%	82.6%	99.1%	74.3%	-	60.6%	63.2%	-	99.3%
2009	64.1%	80.2%	99.0%	75.5%	(D)	62.9%	61.0%	-	99.7%
2010	66.1%	79.6%	99.1%	72.4%	75.1%	62.3%	64.7%	70.7%	99.5%
2011	67.7%	78.6%	95.8%	73.3%	74.9%	60.0%	70.1%	67.5%	99.9%
2012	60.9%	77.8%	97.7%	71.6%	73.1%	59.2%	63.5%	69.5%	99.6%
Total	63.4%	79.6%	97.8%	73.2%	74.4%	60.9%	65.2%	69.0%	99.6%

Matching Process – Stage 1

Once we have obtained PIKs, we can match University PIKs to four different Census Databases



Preliminary Findings

Match Rates

Final Year at University	University Employees	PIK-ized	Total Matched	Total Matched to Non-University Firm
2009	13,702	11,137	10,729	8,259
2010	23,463	16,231	15,516	11,884
2011	23,382	18,553	17,679	12,294

Sectoral Breakdown

Year	Private	Education	Public
2010	52.7%	36.7%	10.2%
2011	57.7%	34.2%	7.7%
2012	57.8%	34.7%	7.3%
Total	56.1%	35.2%	8.4%

Faculty

Year	Private	Education	Public
2010	41.7%	42.4%	14.6%
2011	41.8%	45.6%	12.7%
2012	36.8%	53.0%	9.7%
Total	39.6%	48.0%	11.9%

Post-Doc

Year	Private	Education	Public
2010	30.7%	59.6%	8.9%
2011	41.1%	49.8%	9.1%
2012	35.4%	59.0%	5.6%
Total	35.9%	56.3%	7.6%

Graduate Students

Year	Private	Education	Public
2010	46.6%	42.7%	10.4%
2011	52.9%	38.7%	8.2%
2012	55.0%	36.5%	8.2%
Total	51.7%	39.2%	8.9%

Undergraduates

Year	Private	Education	Public
2010	67.3%	23.8%	8.5%
2011	70.1%	22.9%	6.5%
2012	70.5%	22.9%	6.4%
Total	69.7%	23.1%	6.9%

Job Placements - 1 Year After Leaving Institution

By University (Graduate Students)

A

Last Year	Proportion Matched (At Least 6 Months)		
	Industry	Academia	Government
2010	49.0%	41.2%	9.8%
2011	42.9%	46.0%	9.5%
2012	50.9%	43.6%	5.5%

B

Last Year	Proportion Matched (At Least 6 Months)		
	Industry	Academia	Government
2010	48.3%	40.7%	10.7%
2011	58.9%	34.9%	5.9%
2012	59.6%	34.3%	5.7%

C

Last Year	Proportion Matched (At Least 6 Months)		
	Industry	Academia	Government
2010	35.9%	55.6%	8.3%
2011	46.7%	44.4%	8.9%
2012	49.1%	42.3%	8.0%

D

Last Year	Proportion Matched (At Least 6 Months)		
	Industry	Academia	Government
2010	-	-	-
2011	62.5%	28.1%	9.4%
2012	57.1%	32.1%	10.7%

E

Last Year	Proportion Matched (At Least 6 Months)		
	Industry	Academia	Government
2010	55.9%	27.3%	15.9%
2011	58.0%	28.0%	13.0%
2012	56.8%	27.1%	15.6%

F

Last Year	Proportion Matched (At Least 6 Months)		
	Industry	Academia	Government
2010	45.1%	47.4%	7.5%
2011	57.5%	39.2%	3.3%
2012	45.7%	49.4%	4.9%

Industry Breakdown

NAICS	NAICS Description	National Sample	All Universities
11	Forestry, Fishing, Hunting, and Agriculture Support	1.12%	0.92%
21	Mining	0.59%	0.41%
22	Utilities	0.72%	0.34%
23	Construction	4.64%	1.31%
31-33	Manufacturing	9.75%	13.35%
42	Wholesale Trade	4.97%	4.58%
44-45	Retail Trade	13.03%	8.20%
48-49	Transportation and Warehousing	3.95%	0.99%
51	Information	2.85%	4.17%
52	Finance and Insurance	5.57%	4.08%
53	Real Estate and Rental and Leasing	1.79%	0.85%
54	Professional, Scientific, and Technical Services	7.09%	15.06%
55	Management of Companies and Enterprises	2.64%	3.07%
56	Administrative and Support and Waste Management and Remediation Services	8.17%	9.25%
62	Health Care and Social Assistance	16.24%	18.52%
71	Arts, Entertainment, and Recreation	1.83%	2.40%
72	Accommodation and Food Services	10.25%	6.36%
81	Other Services (except Public Administration)	4.81%	6.11%

Most Grant Recipients are concentrated in Manufacturing, Professional Services and Healthcare

Industry Breakdown (Cont.)

NAICS	NAICS Description	LBD	All Universities
11	Forestry, Fishing, Hunting, and Agriculture Support	1.12%	0.92%
21	Mining	0.59%	0.41%
22	Utilities	0.72%	0.34%
23	Construction	4.64%	1.31%
31-33	Manufacturing		
42	Wholesale Trade		
44-45	Retail Trade		
48-49	Transportation and Warehousing		
51	Information		
52	Finance and Insurance		
53	Real Estate and Rental and Leasing		
54	Professional, Scientific, and Technical Services		
55	Management of Companies and Enterprises		
56	Administrative and Support and Waste Management and Remediation Services		
62	Health Care and Social Assistance		
71	Arts, Entertainment, and Recreation		
72	Accommodation and Food Services		
81	Other Services (except Public Administration)		

3-digit NAICS Breakdown

NAICS	NAICS Description	LBD	All Universities
330	Primary Metal Manufacturing	0.00%	0.01%
331	Primary Metal Manufacturing	0.33%	0.28%
332	Fabricated Metal Product Manufacturing	1.18%	1.01%
333	Machinery Manufacturing	0.55%	0.38%
334	Computer and Electronic Product Manufacturing		
335	Electrical Equipment, Appliance, and Component Manufacturing		
336	Transportation Equipment Manufacturing		
337	Furniture and Related Product Manufacturing		
339	Miscellaneous Manufacturing		
541	Professional, Scientific, and Technical Services		
621	Ambulatory Health Care Services		
622	Hospitals		
623	Nursing and Residential Care Facilities		
624	Social Assistance		

4-digit NAICS Breakdown

NAICS	NAICS Description	LBD	All Universities
3341	Computer and Peripheral Equipment Manufacturing	0.06%	0.26%
3342	Communications Equipment Manufacturing	0.10%	0.17%
3343	Audio and Video Equipment Manufacturing	0.01%	0.02%
3344	Semiconductor and Other Electronic Component Manufacturing	0.25%	0.54%
3345	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	0.34%	0.74%
3346	Manufacturing and Reproducing Magnetic and Optical Media	0.01%	0.00%
5411	Legal Services	1.02%	1.23%
5412	Accounting, Tax Preparation, Bookkeeping, and Payroll Services	1.15%	1.29%
5413	Architectural, Engineering, and Related Services	1.13%	1.92%
5414	Specialized Design Services	0.09%	0.04%
5415	Computer Systems Design and Related Services	1.30%	1.99%
5416	Management, Scientific, and Technical Consulting Services	0.86%	1.67%
5417	Scientific Research and Development Services	0.63%	0.00%

Over/Under-Represented Industries

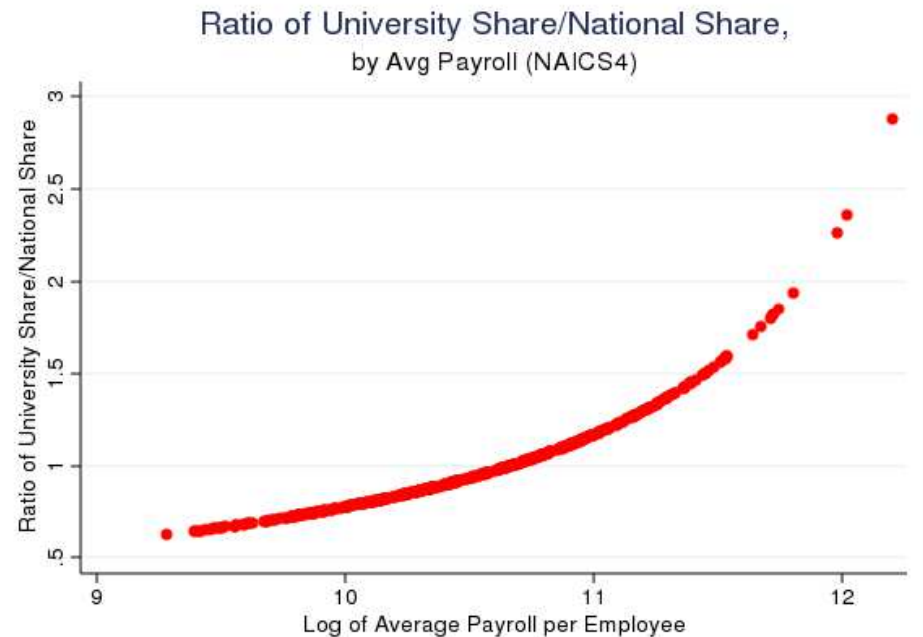
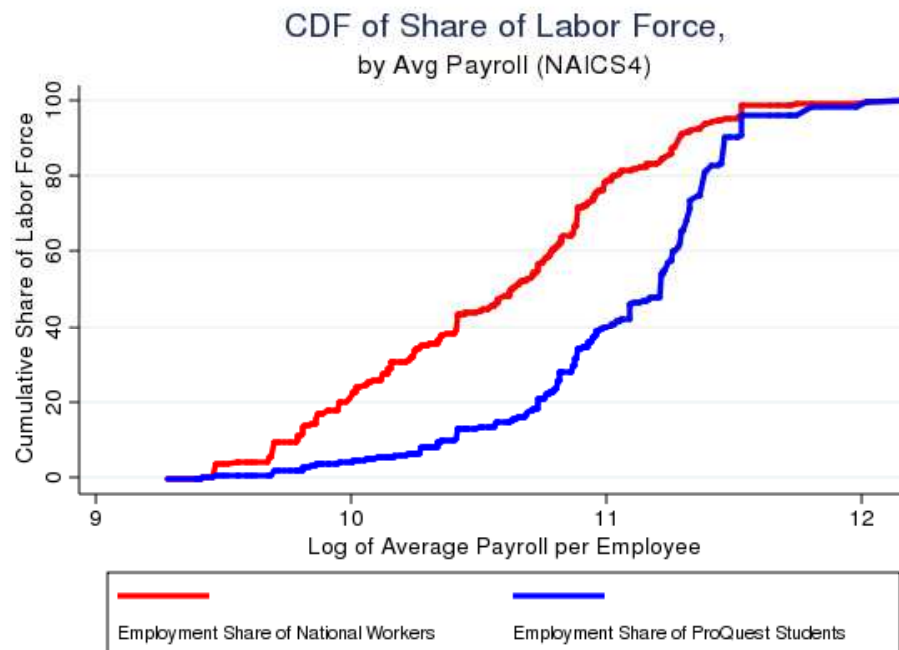
Most Overrepresented 4-digit NAICS

	NAICS	NAICS Description	U.S.	Univs.	Dif
1	5413	Architectural, Engineering, and Related Services	1.13%	4.34%	3.21%
2	5415	Computer Systems Design and Related Services	1.30%	3.97%	2.68%
3	5613	Employment Services	3.87%	6.26%	2.39%
		Management, Scientific, and Technical Consulting			
4	5416	Services	0.86%	2.67%	1.82%
5	6221	General Medical and Surgical Hospitals	4.63%	5.96%	1.33%
6	4236	Electrical and Electronic Goods Merchant Wholesalers	0.43%	1.72%	1.28%
7	6214	Outpatient Care Centers	0.69%	1.82%	1.12%
8	8132	Grantmaking and Giving Services	0.17%	1.25%	1.08%
9	5112	Software Publishers	0.32%	1.35%	1.03%
10	5191	Other Information Services	0.23%	1.25%	1.02%

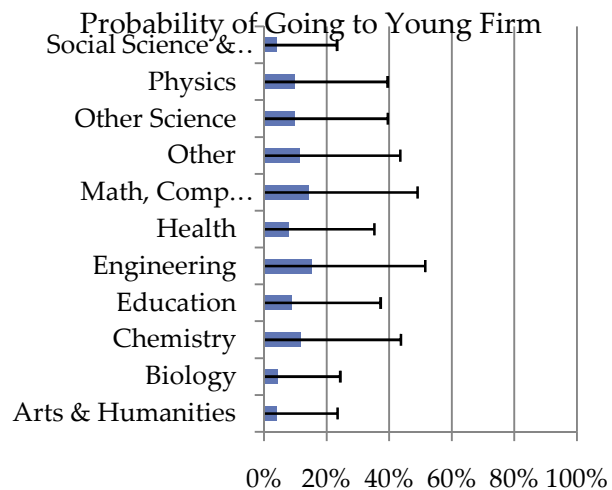
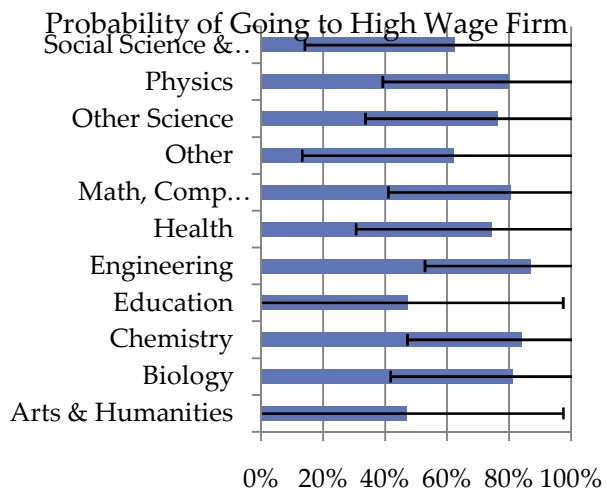
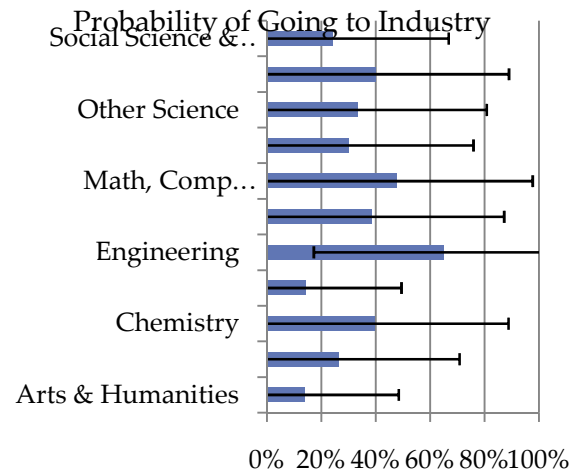
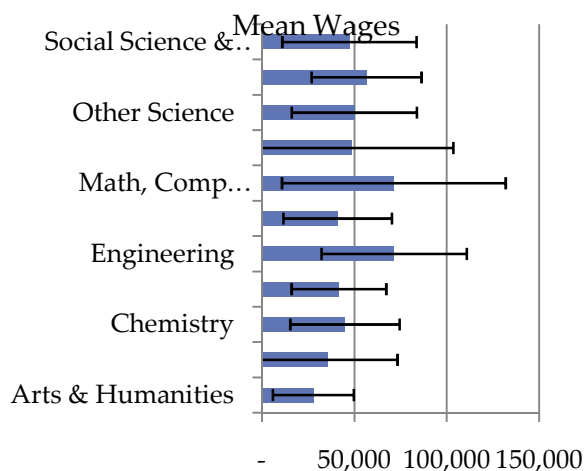
Most Underrepresented 4-digit NAICS

1	7222	Limited-Service Eating Places	3.63%	1.84%	-1.79%
2	4451	Grocery Stores	2.26%	0.69%	-1.58%
3	2382	Building Equipment Contractors	1.39%	0.27%	-1.12%
4	5221	Depository Credit Intermediation	1.80%	0.71%	-1.09%
5	4529	Other General Merchandise Stores	1.51%	0.44%	-1.07%
6	7211	Traveler Accommodation	1.66%	0.66%	-1.00%
7	7221	Full-Service Restaurants	4.03%	3.12%	-0.91%
8	8131	Religious Organizations	1.47%	0.56%	-0.90%
9	5617	Services to Buildings and Dwellings	1.46%	0.56%	-0.89%
10	6231	Nursing Care Facilities	1.46%	0.64%	-0.82%

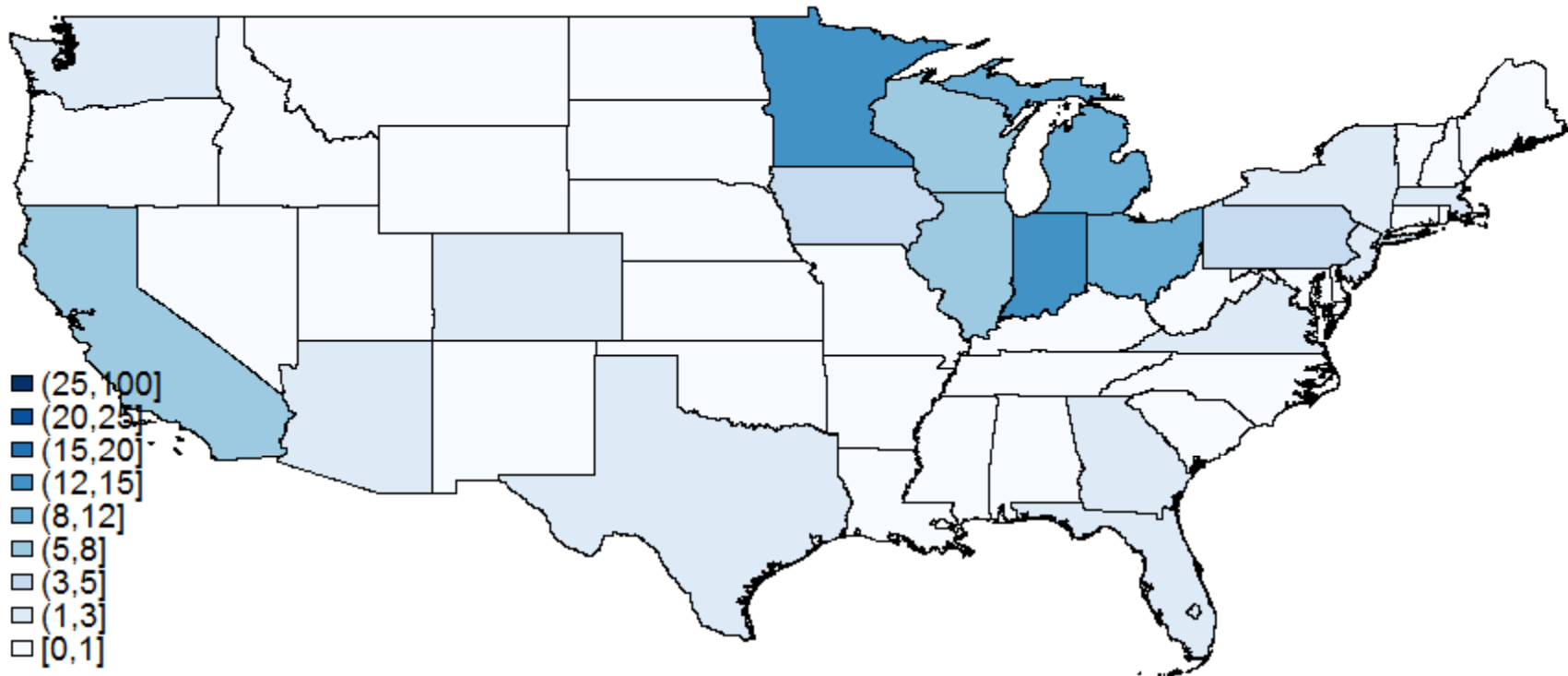
Explanatory Factors for Employment Share Differences



Explanatory Factors for Employment Share Differences

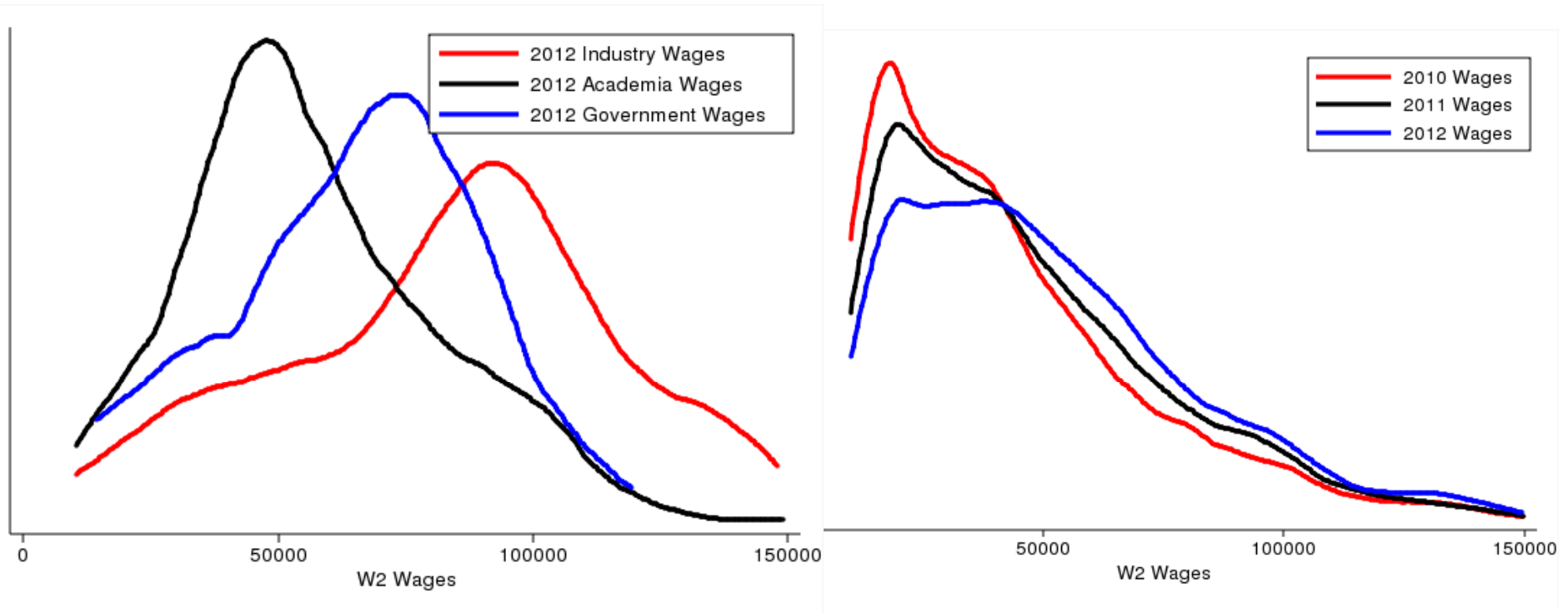


Geographic Outcomes



- Midwestern University sample shows great concentration in geographic outcomes
- Can drill down at further levels of detail (MSA, County, 5-digit Zip, etc...) to assess local economic impacts

Wage Differences and Growth



- Wage Differences Across Sectors fits our priors

- Can also track wage outcomes, differences and growth among employees

Destination Firm Characteristics

Variable	2010		2011		2012	
	All US Firms	Employers*	All US Firms	Employers*	All US Firms	Employers*
Median Employment**	109.3	260.0	110.5	281.3	113.9	296.2
Median Age**	17.4	11.0	17.4	12.0	17.8	13.0
Median Payroll**	3,596.3	13,828.5	3,728.9	15,187.0	3,914.6	16,883.2
Median Average Payroll**	33,860.2	58,310.0	34,558.3	60,186.3	35,019.3	62,882.8
Median YoY Emp Growth**	0.00%	-0.01%	0.18%	0.76%	0.34%	2.55%

* Employment Weighted

** Fuzzy Medians (Mean of Firms in between the 45th and 55th Percentile)

- Grant recipients tend to get jobs in larger, older, higher growth and higher wage firms
 - Median size is 2.5x greater
 - Median Age is more than 6 years younger
 - Median Payroll is more than 3x greater
 - Median Average Payroll is nearly 2x higher
- Within High-Tech Industries (R&D Intensive Industries), the differences are even more stark
 - Median size is nearly 5x greater
 - Median Payroll is more than 6x greater
 - Growth Rates are much faster

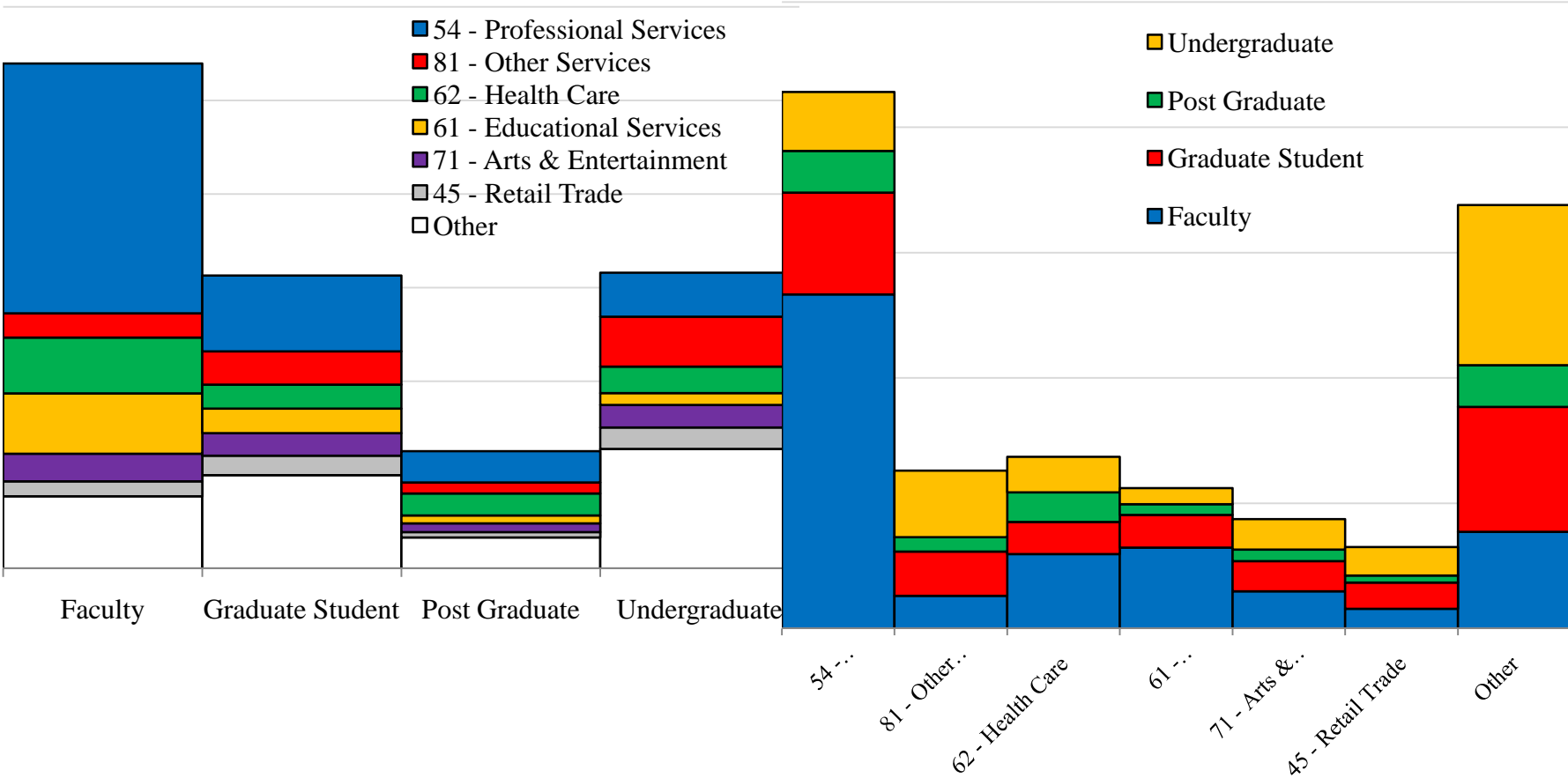
Entrepreneurial Activity

Sole Prop/Non-Employer Match Statistics

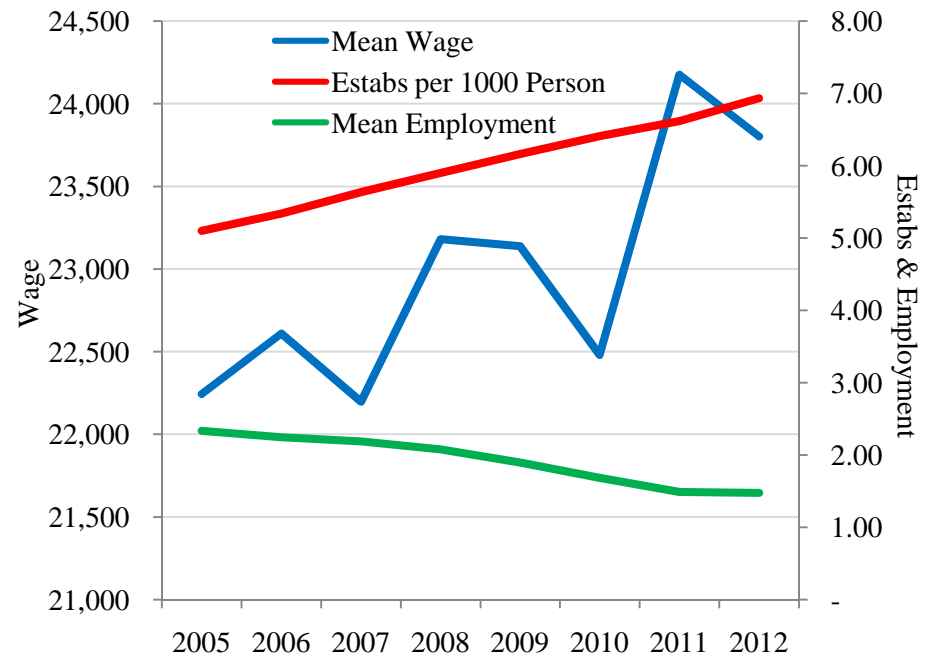
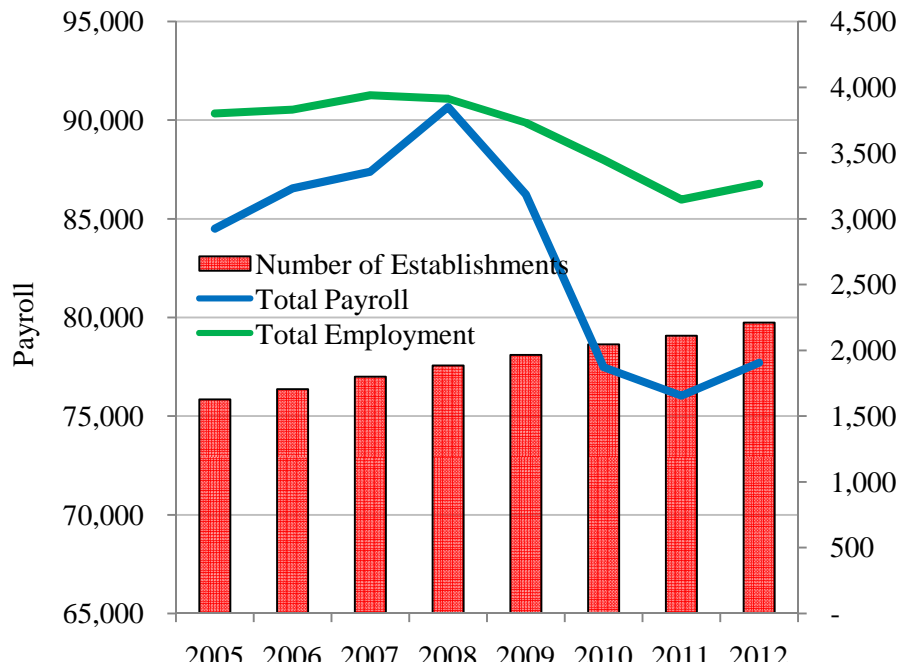
Year	All	Faculty	Grad Student	Post Graduate	Undergrad	Other
2005	14,051	2,818	1,001	351	691	9,190
2006	12,688	2,523	931	381	658	8,195
2007	16,504	3,066	1,257	511	1,007	10,663
2008	17,375	3,156	1,466	539	1,137	11,077
2009	17,937	3,145	1,547	569	1,249	11,427
2010	18,147	3,133	1,561	600	1,323	11,530
Total	46,986	5,880	4,977	1,678	4,013	30,438
Possible	222,622	10,958	29,759	8,531	34,119	139,255
Match Rate	21.1%	53.7%	16.7%	19.7%	11.8%	21.9%

Non-Employer Outcomes

Sole Prop Industries (by occupation)



Startup Business Dynamics (Matched through SS-4)



- Number of startups has been steadily increasing, although the cumulative size of these firms has been somewhat flat

Discussion

Two Key Issues to Proceed with Analysis:

Discussion

Two Key Issues to Proceed with Analysis:

- How to define individuals associated with Food Safety Research?

- How to define Food Safety Outcomes?

Discussion

Two Key Issues to Proceed with Analysis:

- How to define individuals associated with Food Safety Research?
 - Identification of Food Safety Grants
 - Identification of Food Safety Dissertations (ProQuest)
 - Identification of Known PI's in Food Safety
 - Research Papers and Patent Output
- How to define Food Safety Outcomes?

Discussion

Two Key Issues to Proceed with Analysis:

- How to define individuals associated with Food Safety Research?
 - Identification of Food Safety Grants
 - Identification of Food Safety Dissertations (ProQuest)
 - Identification of Known PI's in Food Safety
 - Research Papers and Patent Output
- How to define Food Safety Outcomes?
 - Industry Codes of Establishments?
 - 6-digit NAICS - Nearly 1,100 categories. Is it disaggregated enough?
 - Product Codes of Imports and Exports?
 - 10-digit Harmonized System – Nearly 40,000 categories
 - Other identifiers?