Perceived Wages

John J. Horton
NYU Stern*

November 18, 2013

Abstract

Using a convenience sample of workers, I test their perceptions of hourly wages and actual hourly wages for the top one hundred BLS occupations, as measured by employment. Estmation errors are decreasing in the prevelance of an occupation. Knowing someone who has that occupation also decreases error rates. There appears to be a U-shaped pattern in knowledge about occupations Workers do not appear to know random subsets of people with an occupation: knowledge is substantially clustered.

JEL J01, J24, J3

1 Introduction

Workers need know what various occupations pay–and will pay in the future—to plan their careers. Systematic informational gaps could be consequential, leading to misallocation of job search efforts and human capital investments. However, to the extent that these gaps exist, they might be cheaply fixable: several studies have highlighted the powerful effects obtainable from purely informational interventions (e.g., [2], [1], [**?**]). At least part of the motivation of government statistical efforts like the OES is precisely this allocative efficiency goal.

The purpose of this study is to measure information gaps and explore whether any patterns emerge across occupations. This was done using a convenience sample of US-based workers on their knowledge about the top (by May 2013 employment figures) 100 BLS occupations. For each occupation, participants were asked:

1. Whether they know what the job consists of
2. Whether they know anyone with that job
3. The hourly wage for that job
4. Whether wages and employment for that occupation will rise or fall in the future.

For each occupation, 30 evaluations were performed. The short time taken to answer questions suggests that workers are not doing research. There was no incentive for correct answers.

*Author contact information, datasets and code are currently or will be available at http://www.john-joseph-horton.com/.

The focus on the analysis was to: -characterize performance and see whether the "wisdom of crowds" hypotheses holds. -find poorly labeled occupations in the BLS/OES -see what occupations are misperceived. -see how total employment and social knowledge mediate performance in estimating hourly wages -test whether social knowledge is "segmented" by wage

There is substantial variation across occupations in whether they are known by the sample. First, individual error rates (MSE between predicted and actual hourly wage) are decreasing the total employment in that occupation. This is both re-assuring and unsurprising. Suggesting a mechanism, the more people workers know with that occupation, the better their estimates. Knowledge of an occupation seems to be substantially wage-biased, in that workers are more accurate when predicting the hourly wage of low-wage jobs. This is true in percentages. The wisdom of crowds hypotheses suggests that combining the noisy signals of a large number of individuals can lead to very accurate predictions. This is more or less the case here: the mean worker rating for each occupation explains 70% of the variation in actual log hourly wages.

2 Results

The sample of workers answering questions on MTurk. What we lack in representativeness, we gain in numbers.

If a pattern or causal effect exists in one population, there is always a question of generalizability.

2.1 Discussion

What if the sample of respondents is just a sample of low-wage workers?

- We can ask workers for their incomes - We can show in the literature that this is not a particularly biased group - We can use Google Survey, a nationally representative sample, to adjust

What if high-way jobs are just less common? Controlling for the total employment in that category does not make the effect go away.

What if the hourly/annual salary throws things off?

Do high-way jobs just have more variance?

2.2 Knowledge of what different occupations are

In Figure **??**, we plot the mean knowledge index by occupations.

In Figure 2 we show the fraction of respondents reporting that they know what an occupation is, binned by hourly wage. A clear U-shaped pattern is apparent: respondnents know more about low- and high-wage occupations, but less about the middle. This probably reflects universal experience with low-wage jobs like waitressing and cleaning and high-wage professional jobs like doctor and lawyer.

2.3 Personal knowledge of someone working with an occupation

2.4 Prediction

In Table 1 we regress MSE in hourly wage estimates on a number of predictors.

Figure 4 shows that workers generally underestimat the returns at high end of the labor market.

Figure 5 shows the scatter plot of predicted wages verus actual hourly wages.

Figure 1: Self-reported knowledge by BLS occupation title

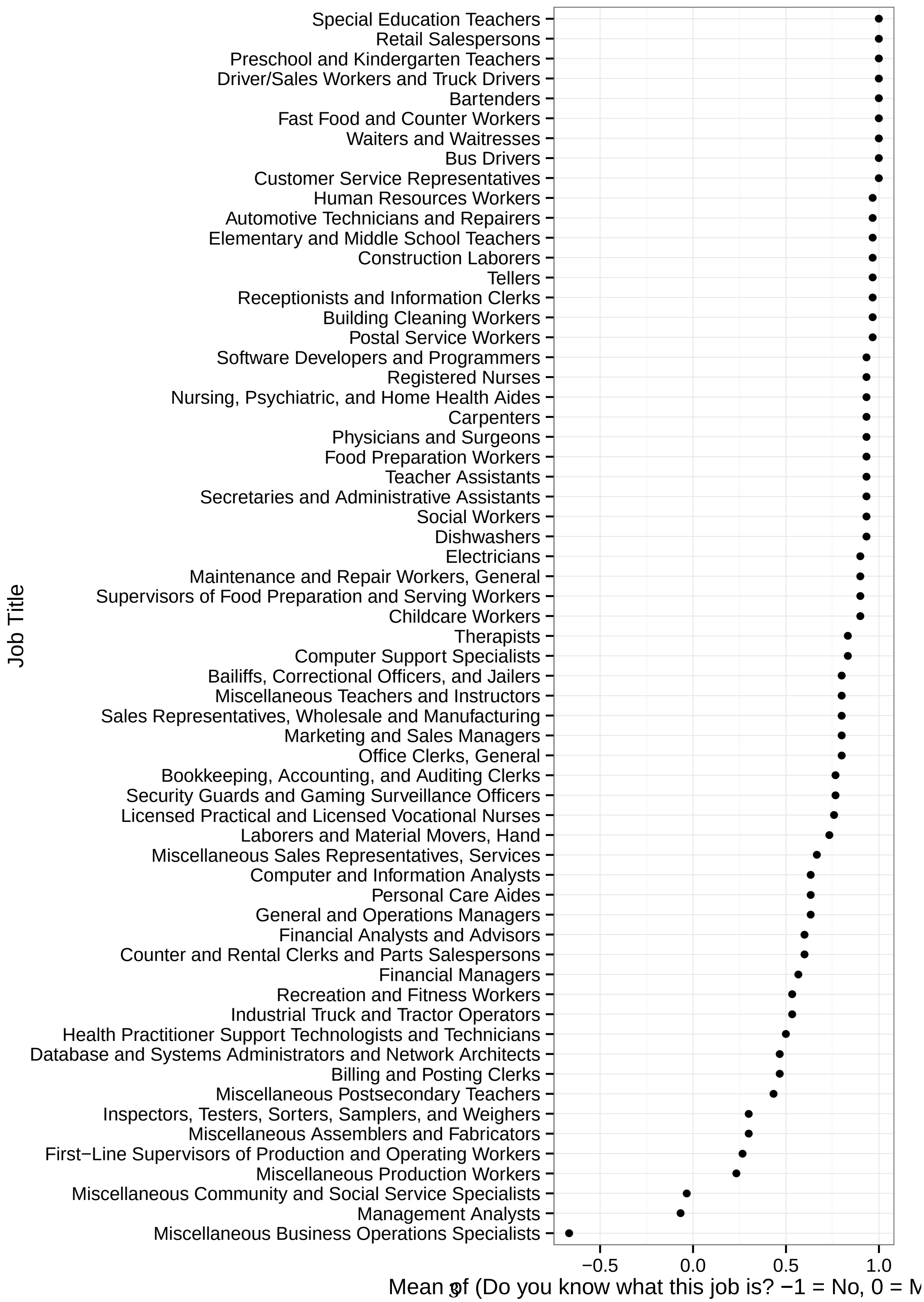


Table 1: Mean square error of hourly wage predictions

	(1)	(2)	(3)
(Intercept)	0.627^{***}	0.363^{***}	0.369^{***}
	(0.110)	(0.019)	(0.020)
social	-0.008	-0.011^{*}	-0.010^{*}
	(0.005)	(0.005)	(0.005)
know	-0.035^{***}	-0.010	-0.009
	(0.010)	(0.009)	(0.010)
log(TOT$_E MP$)	-0.018^{*}		
	(0.008)		
Var((Intercept)|Input.Title)		0.024	0.024
Var(|Residual)		0.061	0.058
Var((Intercept)|WorkerId)			0.004
R-squared	0.011		
adj. R-squared	0.010		
sigma	0.290		
F	10.679		
p	0.000		
Log-likelihood	-535.826	-201.886	-163.483
Deviance	248.486	403.771	326.965
AIC	1081.652	413.771	338.965
BIC	1111.603	443.723	374.907
N	2952	2952	2952

Figure 2: Hourly wage and occupational knowledge

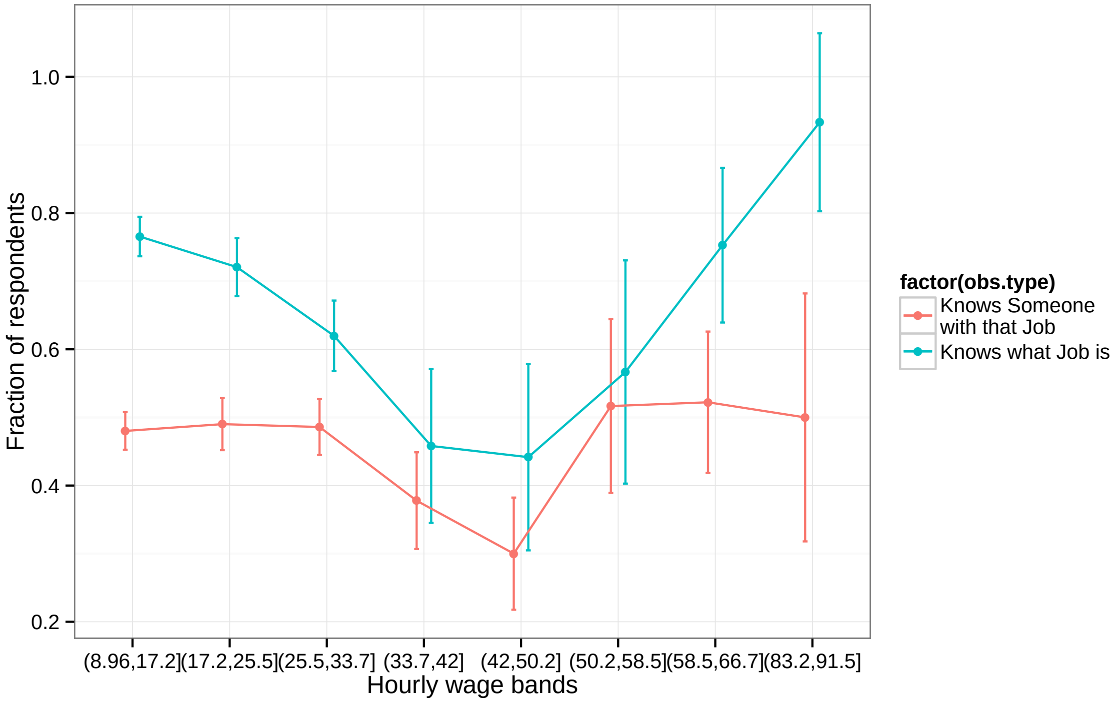


Figure 3: Social index by BLS occupation title

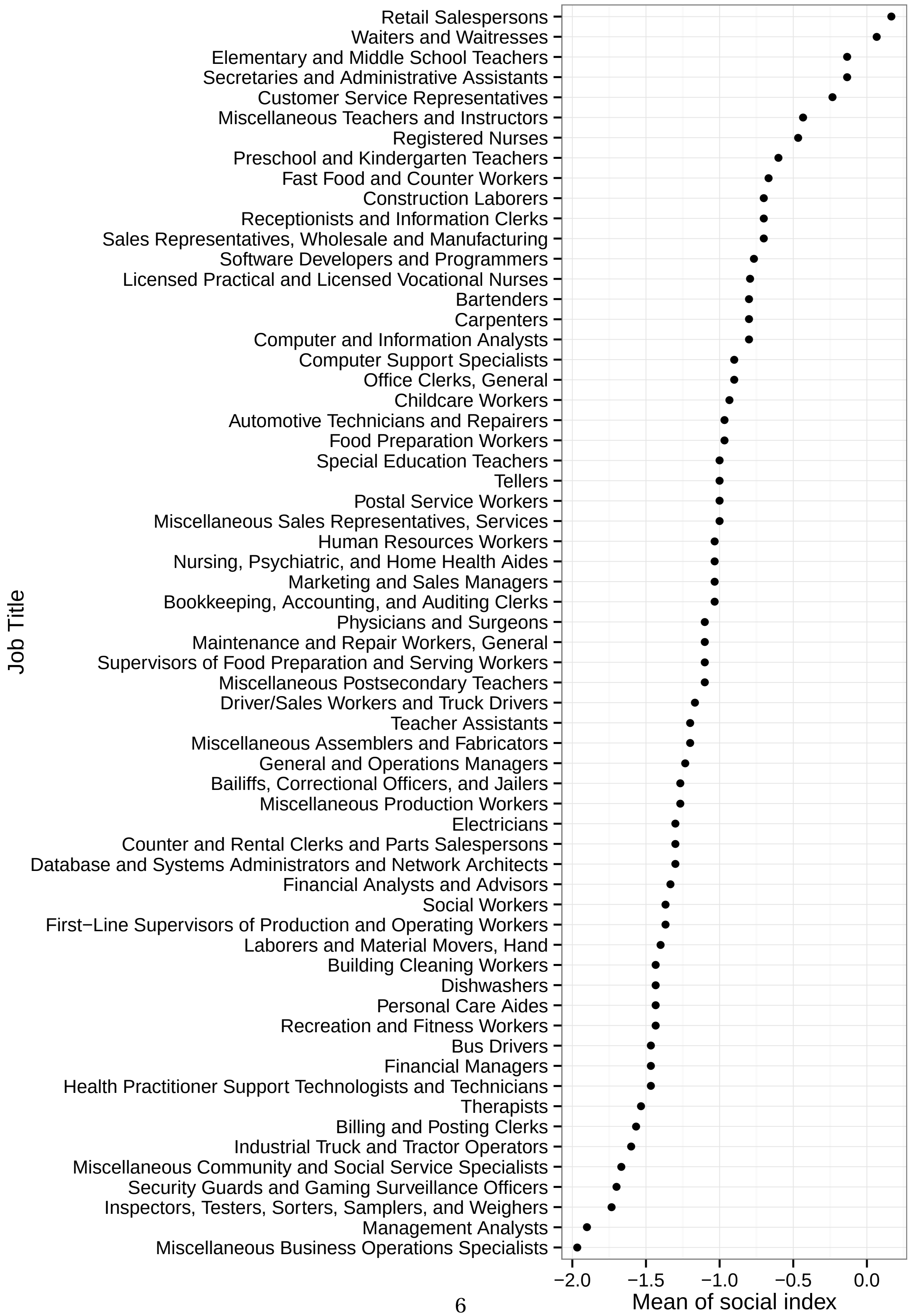


Figure 4: Box plots showing distribution of respondent hourly wage perceptions by occupation

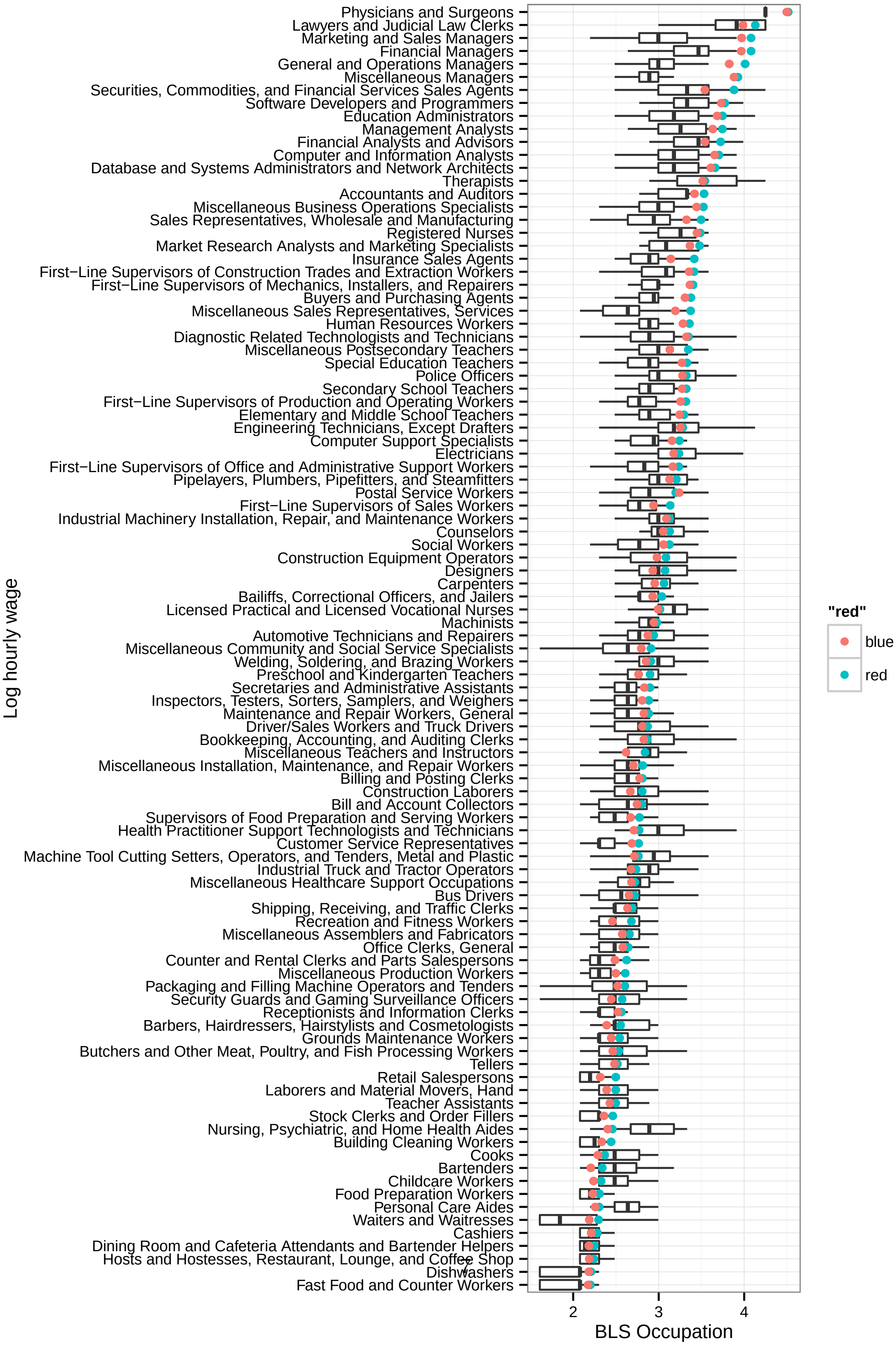
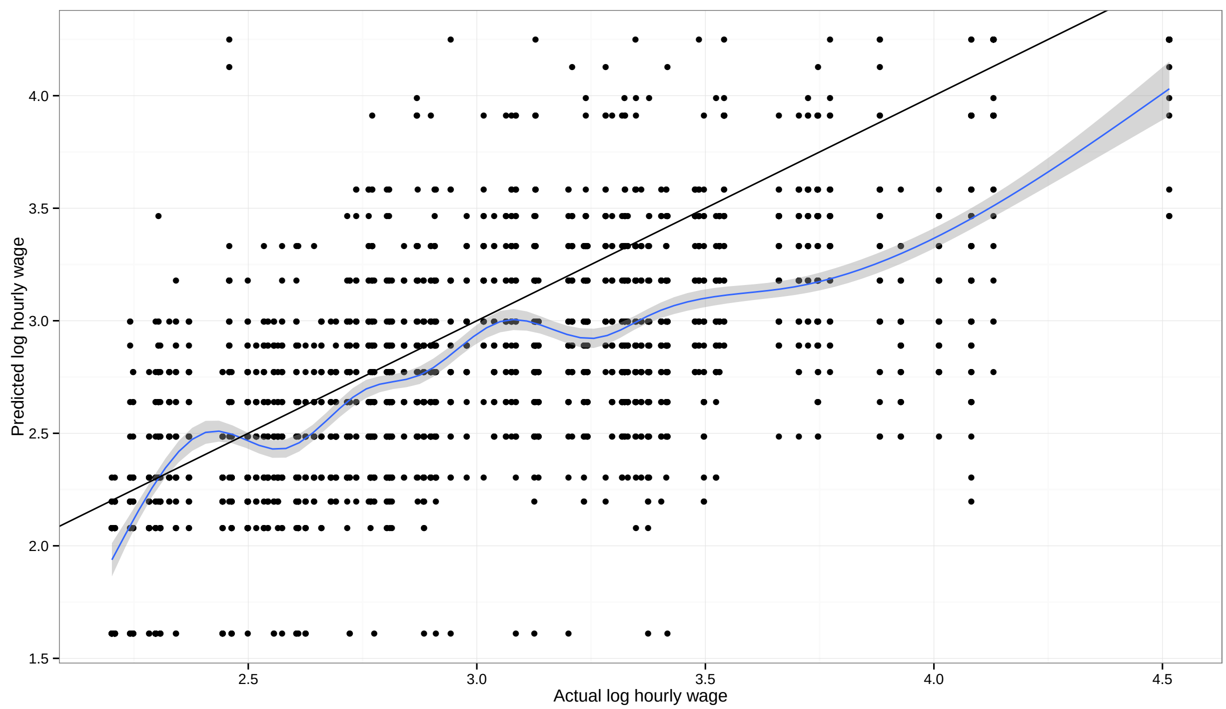


Table 2: Clustering of social knowledge by wage

	(1)	(2)
(Intercept)	3.574^{***}	2.699^{**}
	(1.015)	(0.977)
log($\mathrm{H}_W AGE$)	-1.510^{***}	-1.244^{***}
	(0.327)	(0.306)
mean.wage.others.social	-1.626^{***}	-1.318^{***}
	(0.332)	(0.320)
log($\mathrm{TOT}_E MP$)	0.131^{***}	0.131^{***}
	(0.015)	(0.014)
log($\mathrm{H}_W AGE$) × mean.wage.others.social	0.504^{***}	0.415^{***}
	(0.109)	(0.102)
Var((Intercept)|WorkerId)		0.045
Var(|Residual)		0.198
R-squared	0.040	
adj. R-squared	0.039	
sigma	0.489	
F	26.252	
p	0.000	
Log-likelihood	−1754.623	−1609.371
Deviance	596.067	3218.743
AIC	3521.245	3232.743
BIC	3556.182	3273.503
N	2497	2497

Figure 5: Perceived hourly wages versus actual hourly wages



2.5 Clustering

	1	2	3	4	5
(Intercept)	-0.635^{***}	-0.291^{***}	-0.594^{***}	-0.617^{*}	-0.783^{***}
	(0.122)	(0.044)	(0.119)	(0.291)	(0.118)
know	0.041	0.045		0.029	0.039
	(0.032)	(0.032)		(0.030)	(0.030)
social	-0.039^{*}	-0.038^{*}	-0.012^{**}	-0.030^{*}	-0.031^{*}
	(0.016)	(0.016)	(0.005)	(0.015)	(0.015)
log($\text{TOT}_E MP$)	0.024^{**}		0.023^{**}	0.023	0.020^{**}
	(0.008)		(0.008)	(0.019)	(0.008)
log($\text{H}_W AGE$)	0.209^{***}	0.200^{***}	0.211^{***}	0.209^{***}	0.293^{***}
	(0.010)	(0.010)	(0.010)	(0.026)	(0.011)
know × social	0.031	0.032^{*}		0.020	0.027
	(0.016)	(0.016)		(0.015)	(0.016)
Var((Intercept)|Input.Title)				0.013	
Var(|Residual)				0.061	
I(log(Answer.wage) > 3)					-0.206^{***}
					(0.013)
Aldrich-Nelson R-sq.	0.011	0.011	0.011		0.017
McFadden R-sq.	0.132	0.129	0.129		0.203
Cox-Snell R-sq.	0.011	0.011	0.011		0.017
Nagelkerke R-sq.	0.136	0.134	0.134		0.210
phi	0.074	0.074	0.074		0.068
Likelihood-ratio	33.039	32.365	32.566		51.068
p	0.000	0.000	0.000		0.000
Log-likelihood	−343.627	−348.183	−348.287	−183.829	−216.304
Deviance	218.147	218.822	219.290	367.658	200.118
AIC	701.254	708.367	706.574	383.658	448.608
BIC	743.186	744.308	736.538	431.580	496.530
N	2952	2952	2960	2952	2952

Figure 6: Wage trends

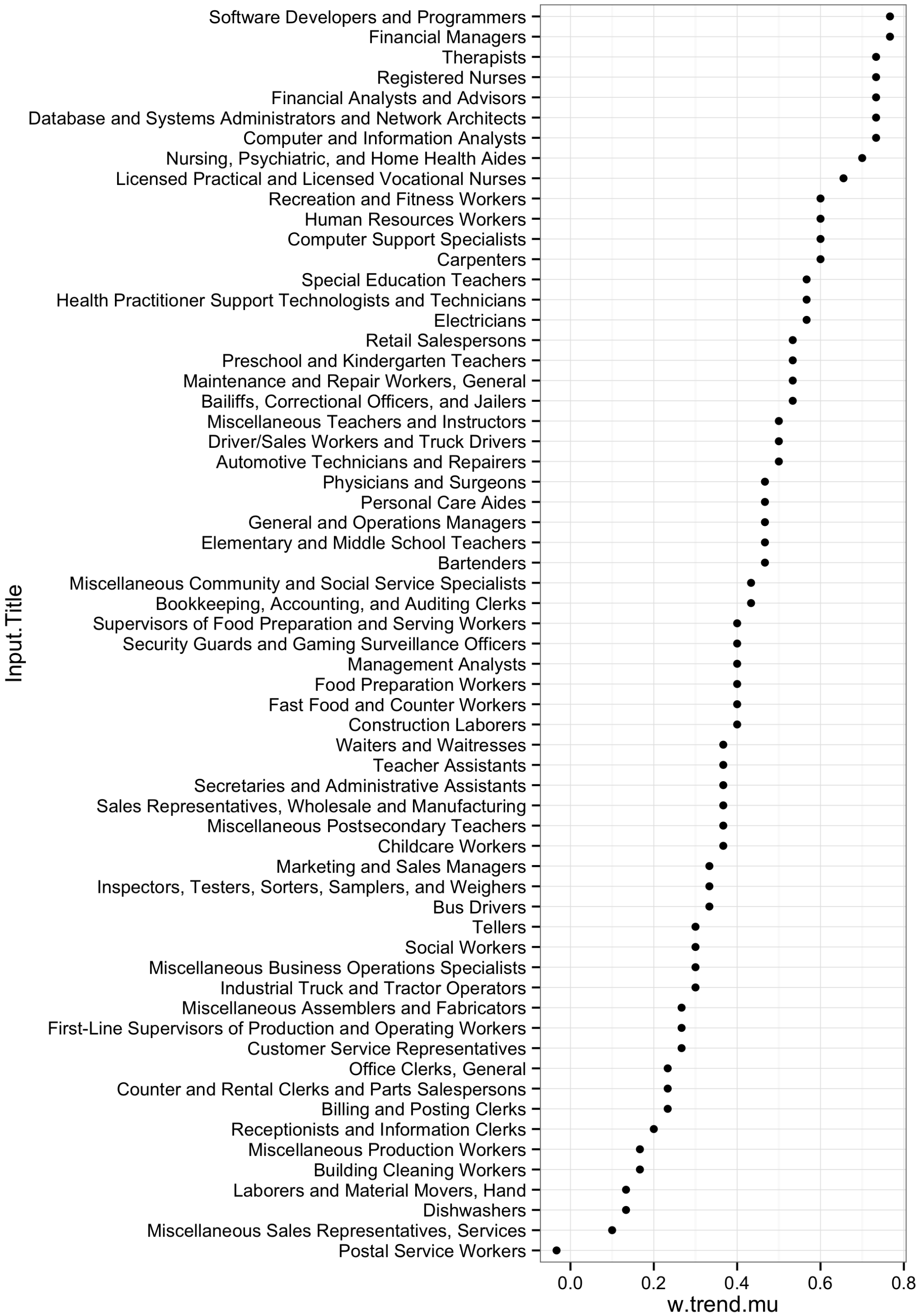
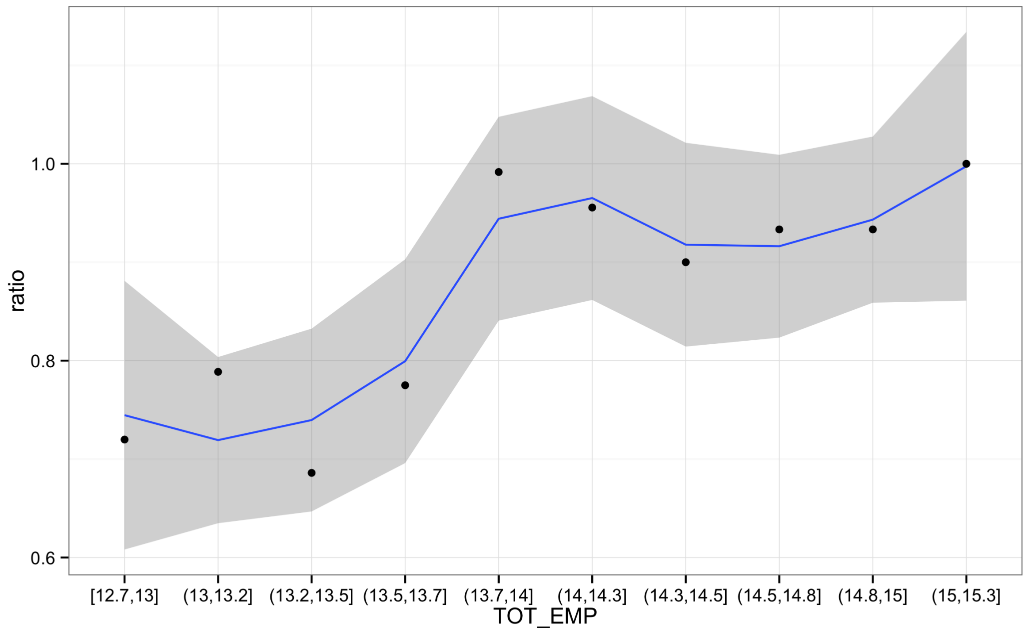


Figure 7: Knowledge by total employment



		1	2	3
	(Intercept)	-0.501^{***}	0.423	-0.566^{*}
		(0.057)	(0.223)	(0.252)
	$\mathrm{TOT}_E MP$	0.000^{***}		0.000^{***}
		(0.000)		(0.000)
	$\log(\mathrm{H}_W AGE)$		-0.177^{*}	0.020
			(0.073)	(0.077)
	Aldrich-Nelson R-sq.	0.027	0.002	0.027
Know anyone	McFadden R-sq.	0.020	0.001	0.020
	Cox-Snell R-sq.	0.027	0.002	0.027
	Nagelkerke R-sq.	0.036	0.003	0.036
	phi	1.000	1.000	1.000
	Likelihood-ratio	81.499	5.950	81.570
	p	0.000	0.015	0.000
	Log-likelihood	−2012.617	−2050.392	−2012.582
	Deviance	4025.234	4100.784	4025.164
	AIC	4029.234	4104.784	4031.164
	BIC	4041.226	4116.776	4049.152
	N	2969	2969	2969

References

[1] Pascaline Dupas. Do teenagers respond to hiv risk information? evidence from a field experiment in kenya. Technical report, National Bureau of Economic Research, 2009.

[2] Robert Jensen. The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2):515–548, 2010.

[3] Andrew I Kohen and Susan C Breinich. Knowledge of the world of work: A test of occupational information for young men. *Journal of Vocational Behavior*, 6(1):133 – 144, 1975.

[4] A. Smith. *Wealth of nations*. Wiley Online Library, 1999.

[5] Erik Snowberg, Justin Wolfers, and Eric Zitzewitz. How prediction markets can save event studies. Technical report, National Bureau of Economic Research, 2011.