Proving Income Inequality in the Field of Psychology

Jacob Willoughby

December 2023

Abstract:

This econometrics research paper investigates the persistent wage gap experienced by women, and demonstrating its existence within the field of psychology. It does this by employing advanced statistical techniques to eliminate doubt on presence of this gender disparity. Drawing on comprehensive data sets spanning diverse demographic parameters, the study employs rigorous econometric models to isolate and quantify the impact of variables such as education, experience, and sex on psychologists’ wages. The findings contribute nuanced insights to the ongoing discourse on gender pay inequity within the psychological profession, informing policymakers, practitioners, and educators about targeted interventions to address and mitigate these disparities. This research aims to facilitate evidence-based strategies for fostering gender equity in compensation within the psychology workforce, promoting a more inclusive and equitable professional landscape.

Introduction:

Gender income inequality remains a persistent and widely discussed issue across various professions and industries. I was curious about the extent that this issue affects the field of psychology as it is an industry that is quite significantly dominated by woman. This econometrics paper delves into the complex landscape of possible gender-based wage disparities within the realm of psychology. By employing rigorous statistical analysis and econometric techniques, I aim to uncover the underlying factors contributing to the pay gap between male and female psychologists. Understanding the intricacies of this phenomenon is not only essential for promoting fairness and equality in the workplace but also for fostering a more inclusive and diverse community within the field of psychology. This study endeavors to shed light on the extent of the pay gap, its determinants, and potential policy implications to address this critical issue in the profession.

Data Selection:

The data is sourced from the IPUMS ACS and HigherEd data bases. The population of interest is woman working in the field of psychology. The data is coming from the years 2003, 2006, 2008, 2010, and 2013. The data should be a random sample since IPUMS follows procedures to insure that. The outcome variable of interest is total yearly income measured in dollars. The key independent variable is sex of worker. Other variables used are: age, race, highest level of education, field of major for highest degree, hours worked per week, employer sector, and size of employer. These other control variables should help to eliminate/limit the omitted variable bias creating a more accurate model. To back up the claim of the APA that the field of psychology is woman dominated (or at least majority), I ran summary statistics on the sex/gender variable for my two data sets. Results are as follows:

ACS:

Sex	Frequency	Percent	Cumulative
Male	3,127	32.33%	32.33
Female	6,545	67.67%	100.00
Total	9,672	100.00%	

HigherEd:

Sex	Frequency	Percent	Cumulative
Male	4,913	37.75%	37.75
Female	8,101	62.25%	100.00
Total	13,014	100.00%	

We can conclude that in these data sets about 2/3 of the observations are from women, backing up the claim of the APA. Because the study was randomly assigned, we can assume that this is representative of the population. Observations where salary or total income below 5,000 was dropped to remove outliers.

Method:

For my initial model I used natural log of income with dummy variable for being a female (1 = female, 0 = male) to make a Log-linear model.

$$\ln(inctot) = B_0 + B_1(female) + e$$

To remove omitted variable bias I add in other control variables that would change the expected total income amount. Variables include education, field of degree, sector of work (private or public/governmental), race, age, and hours worked over the year (usual hours worked per week * weeks worked last year)

$$\ln(inctot) = B_0 + B_1(female) + B_{2-4}(EducationLevels) + B_5(age) + B_{6-10}(races) + B_{11}(usualweeklyhours) + e$$

This model shows the percent of total income for females compared to males (B1 for females) All of the variables within the model are linear, and they are not perfectly correlated between themselves. The model also follows strict exogeneity where income doesn't predict any of the independent variables.

Results:

ACS Data:

"White males with bachelor degrees" is the comparative group. Model 5 is most robust model because the change in coefficients between model 3 and 5 is quite low. Model 6 uses the same independent variables as model 5 but uses total income instead of the natural log of total income. The model shows statistical significance at the 99% level for the impact of being a woman in the field of psychology compared to being a man while controlling for demographic variables, such as race and age, and level of educational attainment. Being a female has a decrease of about 18%. Sex is not the only variable to make a statistically significant change to income at the 99% level. Hours worked over the year, age, and different educational attainment all impact income. Race seems to have no statistically significant change in income besides races that are not listed above, but only at the 95% level.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	ln_inctot	ln_inctot	ln_inctot	ln_inctot	ln_inctot	inctot
female	-0.366***	-0.285***	-0.179***	-0.283***	-0.182***	-15,365***
	[0.0153]	[0.0148]	[0.0141]	[0.0147]	[0.0141]	[1,218]
uhrswork		0.0166***	0.0214***	0.0165***	0.0212***	1,186***
		[0.000541]	[0.000523]	[0.000536]	[0.000521]	[45.05]
age			0.0184***		0.0177***	1,225***
			[0.000511]		[0.000515]	[44.53]
black			-0.0307		-0.00848	-4,905
			[0.0347]		[0.0347]	[2,997]
nativeamerican			-0.0593		-0.0395	-7,006
			[0.124]		[0.123]	[10,651]
asian			0.0297		0.0328	1,204
			[0.0409]		[0.0407]	[3,522]
otherrace			-0.110**		-0.104**	-6,924*
			[0.0431]		[0.0430]	[3,718]
lessHS				0.375**	0.131	-5,816
				[0.181]	[0.170]	[14,684]
highSchool				0.423***	0.286**	9,897
				[0.129]	[0.121]	[10,461]
someCol				-0.0210	-0.0895	-4,250
				[0.0904]	[0.0850]	[7,346]
highered				0.411***	0.250***	9,238***
				[0.0317]	[0.0302]	[2,609]
Constant	11.20***	10.53***	9.370***	10.15***	9.179***	-30,810***
	[0.0126]	[0.0249]	[0.0399]	[0.0389]	[0.0464]	[4,008]
Observations	8,993	8,993	8,993	8,993	8,993	8,993
R-squared	0.060	0.149	0.258	0.166	0.265	0.164
Standard errors in brackets						
*** p<0.01, ** p<0.05, * p<0.1						

Figure 1 [Regression data from ACS]

A residual vs fitted values plot was run to test for homoscedasticity. The plot below shows that our model passes this test.

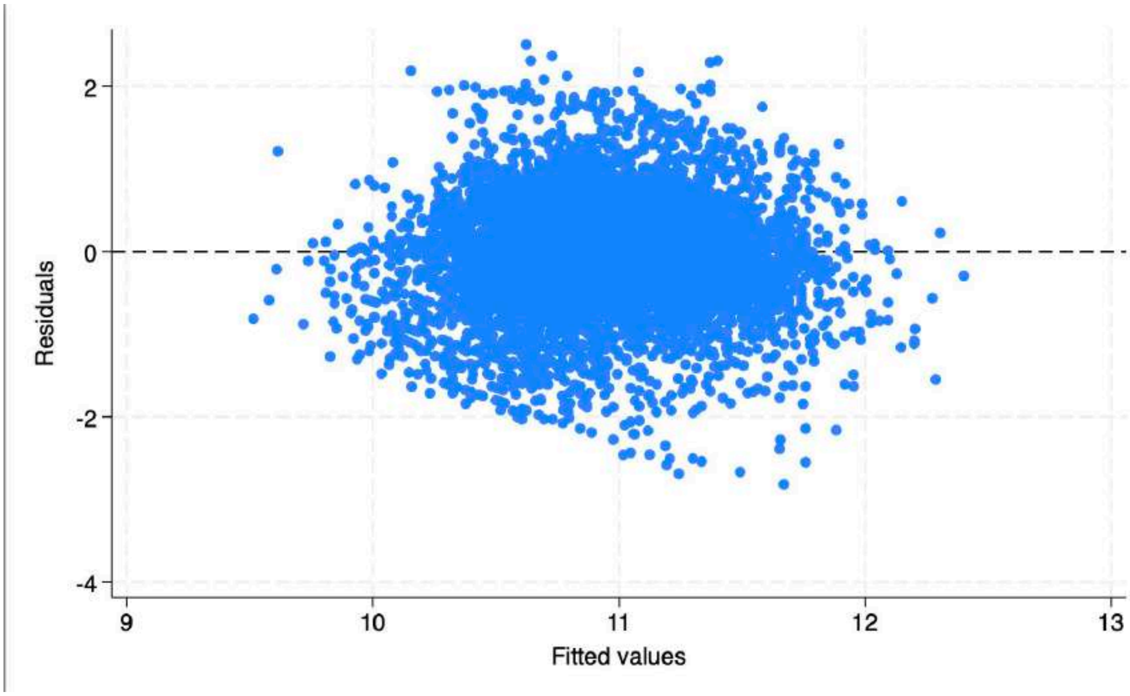

Figure 2 [Residual vs Fitted plot of ACS data]

HigherEd Data:

Using the data from the IPUMs HigherEd section the variables don't match up perfectly so instead the model was recreated as:

$$\ln(salary) = B_0 + B_1(female) + B_{2-4}(EducationLevels) + B_5(age) + B_6(minority) + B_{7-9}(Differenthours) + e$$

A base model where natural log of salary was regressed on female, and then a model that added demographic variables (age and minority) were used to test robustness of our main model.

Doing a standard OLS regression came back with a residual vs fitted plot that shows indications of heteroscedasticity (See figure 3). So instead of a OLS model I used a weighted least squares regression model. For the model I weighed proportional to age, proportional to log of residuals squared, and used no constant. After checking the new residual vs fitted plot for the WLS model, it passed for homoscedasticity. (See figure 4.)

Regression results are shown in figure 5. With non-minority males that work 21 to 35 hours weekly as the comparison group: being female has about a -3% change in salary which is quite reduced from the previous dataset. This model shows robustness in the low change from Model 1 to Model 2 (adding college degrees). All of the variable coefficients are statistically significant at the 99% level except for minority which is not statistically significant. The R-squared value is 0.41, indicating 41% of the variation of income is explained by our model.

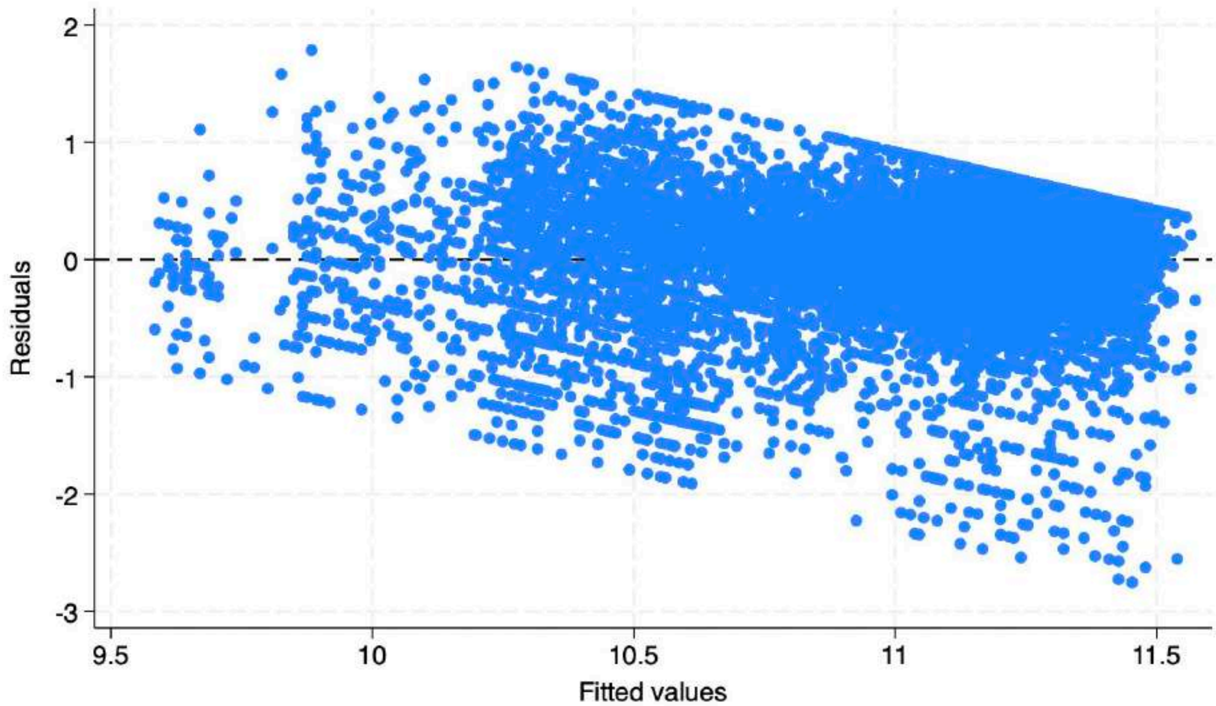

Figure 3. [Residuals vs Fitted plot for OLS]

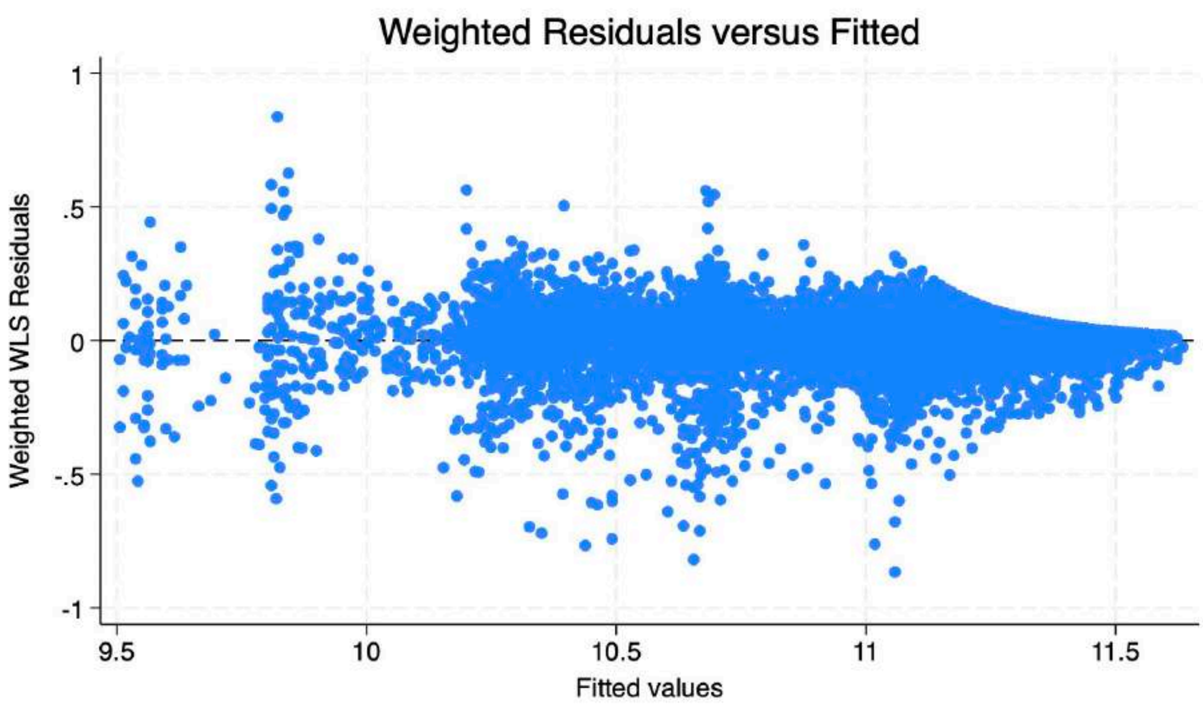

Figure 4. [Residuals vs Fitted plot for WLS]

	(1)	(2)	(3)
VARIABLES	ln_salary	ln_salary	ln_salary
female	-0.218***	-0.0267***	-0.0291***
	[0.0118]	[0.0101]	[0.00960]
age		0.0183***	0.0122***
		[0.000400]	[0.000409]
minrty		-0.0136	-0.00709
		[0.0113]	[0.0107]
mastersDg			0.260***
			[0.0208]
doctorateDg			0.602***
			[0.0204]
professionalCert			0.574***
			[0.0318]
hours20orLess		-0.675***	-0.641***
		[0.0157]	[0.0148]
hours36_40		0.163***	0.172***
		[0.0131]	[0.0124]
hours40plus		0.244***	0.217***
		[0.0132]	[0.0124]
Constant	11.06***	10.13***	9.903***
	[0.00968]	[0.0239]	[0.0268]
Observations	12,809	12,809	12,809
R-squared	0.026	0.340	0.412
Standard errors in brackets			
*** p<0.01, ** p<0.05, * p<0.1			

Figure 5. [HigherEd WLS regression results]

Comparison of fields:

With these results I wanted to compare the gender-based wage gap for psychologists to the gap for the total of occupational observations and also against another woman dominated field. I used the HigherEd dataset for this test due to the high number of observations. The following regression results use the same models as in figure 5.

	(1)	(2)	(3)
VARIABLES	ln_salary	ln_salary	ln_salary
female	0.0141***	-0.150***	-0.151***
	[0.00498]	[0.00362]	[0.00363]
age		0.0205***	0.0191***
		[0.000151]	[0.000160]
minrty		-0.0454***	-0.0411***
		[0.00422]	[0.00422]
mastersDg			-0.0103**
			[0.00427]
doctorateDg			0.0997***
			[0.00479]
professionalCert			0.281***
			[0.00930]
hours21_35		-3.703***	-3.716***
		[0.00787]	[0.00787]
hours36_40		-3.278***	-3.285***
		[0.00521]	[0.00521]
hours40plus		-3.142***	-3.165***
		[0.00506]	[0.00512]
Constant	11.57***	13.58***	13.61***
	[0.00333]	[0.00787]	[0.00798]
Observations	521,153	521,153	521,153
R-squared	0.000	0.505	0.507
Standard errors in brackets			
*** p<0.01, ** p<0.05, * p<0.1			

Figure 6. [WLS Regression on all occupations in dataset]

As we can see from model 3, women in the field of Psychology have a much smaller pay gap (3%) than for average women (15% difference). Then for comparing against another woman dominated field, summary statistics were run to see which fields had more woman than men. Only 4 other occupations of the possible choices listed in the dataset were woman dominated: “Other social scientists”, “health-related occupations”, “Science and engineering pre-college teachers”, and “Non-science and engineering pre-college and post-secondary teachers”. Health-related occupations was chosen as the comparison group due to the large number of observations compared to the other groups.

Gender	Frequency	Percent	Cumulative
Male	18,957	38.40%	38.40%
Female	30,409	61.60%	100.00%
Total	49,366	100.00%	

As we can see the ratio of woman to men in the health-related field is nearly identical to the ratio in the Psychology field. This helps reduce any bias that would have been caused by demand of a particular gender due to equal ratios.

Figure 7 shows the regression results of repeating the same 3 models as in figure 5 for both fields. The first three models (1, 2, 3) are for Health-related occupations and the following three models (4, 5, 6) are for Psychology for ease of comparison.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	ln_salary	ln_salary	ln_salary	ln_salary	ln_salary	ln_salary
female	-0.375***	-0.187***	-0.0937***	-0.218***	-0.0267***	-0.0291***
	[0.00599]	[0.00529]	[0.00508]	[0.0118]	[0.0101]	[0.00960]
age		0.0178***	0.0145***		0.0183***	0.0122***
		[0.000217]	[0.000208]		[0.000400]	[0.000409]
minrty		-0.0751***	-0.0556***		-0.0136	-0.00709
		[0.00585]	[0.00543]		[0.0113]	[0.0107]
mastersDg			0.232***			0.260***
			[0.00588]			[0.0208]
doctorateDg			0.341***			0.602***
			[0.00756]			[0.0204]
professionalCert			0.567***			0.574***
			[0.00641]			[0.0318]
hours21_35		0.718***	0.692***		0.675***	0.641***
		[0.0107]	[0.00995]		[0.0157]	[0.0148]
hours36_40		0.926***	0.932***		0.838***	0.813***
		[0.00954]	[0.00885]		[0.0144]	[0.0137]
hours40plus		1.092***	0.961***		0.919***	0.858***
		[0.00972]	[0.00918]		[0.0145]	[0.0138]
Constant	11.21***	9.536***	9.419***	11.06***	9.453***	9.262***
	[0.00486]	[0.0133]	[0.0125]	[0.00968]	[0.0240]	[0.0265]
Observations	48,569	48,569	48,569	12,809	12,809	12,809
R-squared	0.075	0.346	0.438	0.026	0.340	0.412
Standard errors in brackets						
*** p<0.01, ** p<0.05, * p<0.1						

Figure 7. [WLS regression for health-related occupations]

As we can see from the results, the health-related occupations group also has a wage gap between genders by about 9%, and is significant at the 99% confidence level, with all other variables held constant. This percentage also significantly lower than in Figure 3, consistent with Psychology. This model may not be robust for health-related occupations due to the large difference in coefficients from model 2 to model 3. The coefficient for female is greater than Psychology's by about triple but I wanted to test to see if the difference of these coefficients was statistically different from zero.

To do this test I had to use our main model to regress salary for both fields, but created dummy variables corresponding to: being female and in psychology field (femaleandPsy), and being female and in the health-related occupations field (femaleandNotPsy). A linear combinations of parameters test was performed on these two dummy variables (felmaleandPsy – femaleandNotPsy = 0). FemaleandPsy is the omitted category. The results of the regression and linear combinations of parameters test is on the next page. (Figure 8)

	(1)	(2)
VARIABLES	ln_salary	ln_salary
female	-0.0792***	-0.155***
	[0.00451]	[0.00722]
age	0.0141***	0.0140***
	[0.000186]	[0.000186]
minrty	-0.0500***	-0.0483***
	[0.00487]	[0.00486]
mastersDg	0.193***	0.205***
	[0.00539]	[0.00545]
doctorateDg	0.336***	0.371***
	[0.00580]	[0.00634]
professionalCert	0.572***	0.581***
	[0.00612]	[0.00615]
hours21_35	0.685***	0.680***
	[0.00838]	[0.00837]
hours36_40	0.918***	0.909***
	[0.00747]	[0.00748]
hours40plus	0.952***	0.940***
	[0.00766]	[0.00770]
femaleandPsy		0
		[0]
femaleandNotPsy		0.0948***
		[0.00702]
Constant	9.424***	9.422***
	[0.0110]	[0.0110]
Observations	61,378	61,378
R-squared	0.426	0.427

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Linear combination test of FemaleandPsy - FemaleandNotPsy = 0

ln_salary	Difference	Std. err.	t	P>|t|	[95% conf. interval]	
(1)	-.0421088	.0083443	-5.05	0.000	-.0584636	-.0257539

Figure 8 [WLS Regression of both fields with Linear Combinations test]

As we can see, the difference between the coefficients (about 4%) is statistically different from zero at the 99% confidence level. This means that the wage gap for psychologists is lower than other woman dominated fields.

I would like to include the demographics of each person's patients, particularly the sex. This would show any correlation between if patient prefer to pick therapists or psychologists as the same sex as them and impact on salary. However, I have yet to be able to acquire such a dataset but if one is found I can easily include a percentage of clients that are female variable into the model. This missing variable could create a bias, hypothesized to be a negative bias due to the demographics of patients being a majority female (SAMHSA data), and them requesting to have female psychologists, therefore impacting the supply and demand in a way that changes income for female psychologists.

Conclusion:

Women are making less money than men in the same field of work, even when controlling for other variables. This is consistent in the field of Psychology but the percentage of salary gap is significantly lower than other occupational fields. It is important that people get paid similarly for similar work. Being a female dominated field is not enough to reduce these differences. Governments may want to consider making policies that would help reduce these pay gaps in order to incentivize people of all demographics to have an occupation.

Works Cited:

Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality, 2021. Mental Health Client-Level Data: Annual Detailed Tables, 2016-2021. Rockville, MD: Substance Abuse and Mental Health Services Administration, 2023. https://www.samhsa.gov/data/sites/default/files/reports/rpt42719/2021-mhcld-annual-detailed-tables.pdf

Canal, Emily. “These Are The Gender Wage Gaps In 25 Popular Jobs.” *Forbes.Com*, May 2016, p. 1. *EBSCOhost*, research.ebsco.com/linkprocessor/plink?id=0a15732b-b5ae-3dbd-a625-837f95907e88.

Folbre, Nancy, et al. “Gender Inequality, Bargaining, and Pay in Care Services in the United States.” *ILR Review*, vol. 76, no. 1, Jan. 2023, pp. 86–111. *EBSCOhost*, https://doi-org.proxy.li.suu.edu:2443/10.1177/00197939221091157.

Clay, R. A. (2017, August). *Women outnumber men in psychology, but not in the field’s top echelons*. Monitor on Psychology. https://www.apa.org/monitor/2017/07-08/women-psychology

Nigrinis, A., et al. (2014). *Does the gender pay gap in psychology differ by work setting?*. Monitor on Psychology. https://www.apa.org/monitor/2014/12/datapoint

Blau, Francine & Kahn, Lawrence. “Analyzing the gender pay gap” *The Quarterly Review of Economics and Finance,* vol 39, no. 5, 1999, pp. 625-646. https://doi.org/10.1016/S1062-9769(99)00021-6.

(https://www.sciencedirect.com/science/article/pii/S1062976999000216)

PSYCHOLOGIST DEMOGRAPHICS AND STATISTICS IN THE US. Zippia. (2022). https://www.zippia.com/psychologist-jobs/demographics/

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry. Integrated Public Use Microdata Series, Current Population Survey: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2023. https://doi.org/10.18128/D030.V11.0

Minnesota Population Center. IPUMS Higher Ed: Version 1.0 [dataset]. Minneapolis, MN: University of Minnesota, 2016. https://doi.org/10.18128/D100.V1.0