Analysis of Gender Wage Gap Across Educational Levels within Management Occupations

Scope and Data Filtering This analysis specifically targets the management occupations classified within the occupation codes 0010 to 0430. The dataset was filtered to include only individuals within these occupation codes, ensuring that the findings are relevant to the management sector.

Unconditional Gender Gap The initial analysis showed an unconditional gender gap in wages. This gap was quantified by adding a binary variable indicating gender (female = 1, male = 0) in a simple regression model where the dependent variable is the logarithm of hourly wages (lnw). The coefficient for the female variable was statistically significant and negative, suggesting that women earn less than men regardless of educational level.

Gender Gap Variation with Education Level To understand how the gender gap varies with educational attainment, the analysis expanded to include educational level variables. These variables were incorporated as dummy variables to capture the incremental effect of each educational level on the wage gap. The results showed that as the educational level increases, the gender wage gap changes, with the gap narrowing at higher education levels.

Inclusion of Marital Status

Further analysis incorporated marital status interaction terms to investigate their influence on the gender wage gap. The inclusion of these terms aimed to capture the combined effects of marital status and gender on wages. The model incorporating marital status interaction terms suggests that marital status also influences the gender wage gap, but like education, the significance of this influence varies. For example, coefficients for certain marital statuses, such as 'Widowed_divorced_88' and 'Divorced', were positive and statistically significant, indicating that women with these marital statuses may experience a different wage gap compared to never-married women.

Statistical Inference and Key Coefficients Interpretation The regression analysis provided the following key insights:

- The coefficient for the female variable in the simple model without education level controls was -0.186, with a p-value less than 0.05, indicating that women's earnings are about 18.6% less than men's when not accounting for education level.
- After including educational level controls, the coefficients for interaction terms between female and education levels showed differentials in the gender gap across educational categories. The coefficients for the interaction terms were positive but varied in significance. For instance, the coefficient for female*Post_secondary was 0.1738, suggesting a reduced gender wage gap at the post-secondary education level, but this coefficient was not statistically significant at conventional levels.
- The coefficient for female*Graduate was 0.9471, significant at the 10% level, indicating a substantial narrowing of the gender gap for those with graduate-level education.
- The model incorporating marital status interaction terms suggests that marital status also influences the gender wage gap, but like education, the significance of this influence varies. For example, coefficients for certain marital statuses, such as 'Widowed_divorced_88' and

'Divorced', were positive and statistically significant, indicating that women with these marital statuses may experience a different wage gap compared to never-married women.

- Regarding the R-squared value in the marital status regression output, it indicates that the model explains approximately 16.8% of the variability in the logarithm of hourly wages. When compared to earlier models that did not include marital status interaction terms, this value suggests a slight improvement in the model's ability to account for the variation in wages. This increase implies that the inclusion of marital status, alongside education level and gender, adds some explanatory power to the wage gap analysis, capturing more nuances of the underlying wage determinants.

Summary of Findings The analysis indicates that there is a persistent gender wage gap across all levels of education within management occupations. However, the gap diminishes with higher levels of educational attainment, being least pronounced among individuals with graduate degrees. This suggests that education may play a role in mitigating wage disparities between genders. The inclusion of marital status adds depth to the analysis, revealing that certain marital statuses, such as being widowed or divorced, are associated with a different wage pattern compared to never-married women, potentially indicating that marital status also has an influence on wage dynamics. Nevertheless, even after controlling for education and marital status, a statistically significant gap remains, which could be indicative of other underlying factors such as occupational segregation, work experience, or discrimination.

Tables and Graphs:

Regressions for unconditional Gender Gap

	Dependent variable: lnw	
	(1)	(2)
female	-0.186***	-0.186***
	(0.010)	(0.011)
Constant	3.399***	3.399***
	(0.007)	(0.007)
Observations	14885	14885
R^2	0.021	0.021
Adjusted R^2	0.021	0.021
Residual Std. Error	0.633 (df=14883)	0.633 (df=14883)
F Statistic	318.849*** (df=1; 14883)	314.895*** (df=1; 14883)
Note:		*p<0.1; **p<0.05; ***p<0.01

Robust SE slope: 0.06, CI $[-0.234, 0.008]$ (wide)

-> in 2014 in the USA we can be 95% confident that the average difference between hourly earnings of female market analysts versus male was -23% to 1%

-> the CI *includes* zero -> we cannot rule out with 95% confidence that their average earnings (female and male) are the same

-> $|t| = 1.8 < 1.96$ cannot reject H_0

-> it can be seen also by p-value> 0.05

-> the coefficient cannot be considered statistically significant at 5% (significant at 10%)

Regression results for weekly earnings by education levels

Interaction terms

Out[21]:

OLS Regression Results

Dep. Variable:	lnw	**R-squared:**	0.140
Model:	OLS	**Adj. R-squared:**	0.140
Method:	Least Squares	**F-statistic:**	526.9
Date:	Fri, 08 Dec 2023	**Prob (F-statistic):**	0.00
Time:	18:01:51	**Log-Likelihood:**	-13339.
No. Observations:	14885	**AIC:**	2.669e+04
Df Residuals:	14879	**BIC:**	2.674e+04
Df Model:	5		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	2.6638	0.069	38.398	0.000	2.528	2.800
C(female)[T.True]	-0.1903	0.010	-19.319	0.000	-0.210	-0.171
C(Secondary)[T.1]	0.4190	0.071	5.940	0.000	0.281	0.557
C(Post_secondary)[T.1]	0.5055	0.070	7.185	0.000	0.368	0.643
C(Undergraduate)[T.1]	0.8172	0.070	11.715	0.000	0.680	0.954
C(Graduate)[T.1]	1.0156	0.070	14.522	0.000	0.879	1.153

Omnibus:	13993.999	**Durbin-Watson:**	1.874
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	2279850.610
Skew:	-4.054	**Prob(JB):**	0.00
Kurtosis:	63.085	**Cond. No.**	41.4

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)

The results indicate increase in weekly earnings as education levels go higher.

Out[22]:

OLS Regression Results

Dep. Variable:	lnw	**R-squared:**	0.140
Model:	OLS	**Adj. R-squared:**	0.140
Method:	Least Squares	**F-statistic:**	295.2
Date:	Fri, 08 Dec 2023	**Prob (F-statistic):**	0.00
Time:	18:01:51	**Log-Likelihood:**	-13337.
No. Observations:	14885	**AIC:**	2.669e+04
Df Residuals:	14875	**BIC:**	2.677e+04
Df Model:	9		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	2.7176	0.079	34.521	0.000	2.563	2.872
female[T.True]	-0.3618	0.156	-2.314	0.021	-0.668	-0.055
Secondary	0.3707	0.081	4.599	0.000	0.213	0.529
female[T.True]:Secondary	0.1586	0.158	1.001	0.317	-0.152	0.469
Post_secondary	0.4507	0.080	5.610	0.000	0.293	0.608
female[T.True]:Post_secondary	0.1738	0.158	1.099	0.272	-0.136	0.484
Undergraduate	0.7689	0.079	9.693	0.000	0.613	0.924
female[T.True]:Undergraduate	0.1593	0.157	1.014	0.310	-0.149	0.467
Graduate	0.9471	0.079	11.914	0.000	0.791	1.103
female[T.True]:Graduate	0.2055	0.157	1.305	0.192	-0.103	0.514

Omnibus:	14002.539	**Durbin-Watson:**	1.874
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	2286373.667
Skew:	-4.058	**Prob(JB):**	0.00
Kurtosis:	63.171	**Cond. No.**	97.3

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)

There is a upward increasing trend in the weekly earnings as the education levels go higher considering for females in the above table.

Regression with marital status included (interaction terms)

Out[27]:

OLS Regression Results

Dep. Variable:	lnw	R-squared:	0.168
Model:	OLS	Adj. R-squared:	0.167
Method:	Least Squares	F-statistic:	155.6
Date:	Fri, 08 Dec 2023	Prob (F-statistic):	0.00
Time:	18:02:10	Log-Likelihood:	-13095.
No. Observations:	14885	AIC:	2.623e+04
Df Residuals:	14863	BIC:	2.640e+04
Df Model:	21		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	2.4689	0.083	29.850	0.000	2.307	2.631
female[T.True]	-0.2771	0.162	-1.710	0.087	-0.595	0.041
Secondary	0.3716	0.083	4.479	0.000	0.209	0.534
female[T.True]:Secondary	0.1513	0.163	0.928	0.353	-0.168	0.471
Post_secondary	0.4517	0.083	5.461	0.000	0.290	0.614
female[T.True]:Post_secondary	0.1708	0.163	1.050	0.294	-0.148	0.490
Undergraduate	0.7677	0.082	9.390	0.000	0.607	0.928
female[T.True]:Undergraduate	0.1636	0.162	1.013	0.311	-0.153	0.480
Graduate	0.9235	0.082	11.269	0.000	0.763	1.084
female[T.True]:Graduate	0.2218	0.162	1.369	0.171	-0.096	0.539
Married_civilian	0.3148	0.016	19.231	0.000	0.283	0.347
female[T.True]:Married_civilian	-0.1032	0.025	-4.172	0.000	-0.152	-0.055
Married_AF	-0.3228	0.169	-1.909	0.056	-0.654	0.009
female[T.True]:Married_AF	0.4442	0.196	2.266	0.023	0.060	0.829
Married_absent	0.3166	0.054	5.827	0.000	0.210	0.423
female[T.True]:Married_absent	-0.2075	0.113	-1.838	0.066	-0.429	0.014
Widowed_divorced_88	0.2915	0.062	4.688	0.000	0.170	0.413
female[T.True]:Widowed_divorced_88	-0.0656	0.078	-0.839	0.401	-0.219	0.088
Divorced	0.2614	0.026	10.144	0.000	0.211	0.312
female[T.True]:Divorced	-0.0354	0.035	-1.016	0.309	-0.104	0.033
Separated	0.2020	0.062	3.238	0.001	0.080	0.324
female[T.True]:Separated	-0.0192	0.080	-0.239	0.811	-0.177	0.138

Omnibus:	14491.247	Durbin-Watson:	1.860
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2685320.684
Skew:	-4.268	Prob(JB):	0.00
Kurtosis:	68.244	Cond. No.	150.

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)