Determining the Magnitude of the Gender Pay Gap, and how It Has Changed over the Course of the Pandemic

Yushan Xie, Isaiah Sinclair and Jack Ronson

Professor Plesca

ECON*3740

10 April 2023

Work Distribution: distributed evenly between group members.

Changes since the proposal: Changed the project title, literature section, and the regression on the change in the gender wage gap to only include 'sex * covid' and 'covid' as the additional regressors, as opposed to using a fully interacted model. Added a new sensitivity test conducting this same regression, but with weekly wages. A new weekly wages and log weekly wages variable was constructed for this sensitivity test.

Introduction

The “gender pay gap” has long been the most prevalent topic regarding gender differences in the workplace. There has been a gender pay gap in Canada for a long time, but the gap has narrowed over the years. A study found that the gender wage gap decreased from 18.8% in 1998 to 13.3% in 2018 (Pelletier et al., 2019). However, the COVID-19 pandemic dealt a devastating blow to the world economy, resulting in the first global wage decline since the inception of the Global Wage Report in 2008 (Global Wage Report 2022–23, 2022). This means that employment circumstances have been seriously affected, not only due to COVID-19 infections among the working population but also from the policies and measures decreed by governments to control the pandemic (e.g. ‘lockdowns’, enforced teleworking, etc.) (Moncada et al., 2021). Therefore, the COVID-19 pandemic continues to pose a major challenge to wage growth in the Canadian economy. Specifically, many of the industries that have struggled the most during the pandemic have a high proportion of female employees (e.g., tourism, services), putting many women out of work. As a result, many women were put in a bad predicament where they were forced to make tough decisions about their careers, sometimes choosing lower-pay positions or opting out of the labour force entirely. As a result of some of these unfortunate circumstances, many have argued that the gender wage gap has widened during the pandemic, setting women back further.

Our paper seeks to answer the following questions: What is the true magnitude of the gender pay gap in Canada? To determine the magnitude of the gender pay gap we regressed ‘mincer’ regressions on the December 2019 and December 2022 Labour Force Survey (LFS) data, holding variables (sex, education, actual hours worked, etc.) constant. We also asked: how has the pay gap changed during the pandemic? To answer this question, we combined the December 2019 and December 2022 LFS datasets for testing and will interact sex with a ‘covid’ variable that indicates the period (post-COVID/pre-COVID). This will allow us to see the percentage change in wages pre-pandemic and post-pandemic and the relative effect of the pandemic on hourly earnings for women. Using this regression, we conducted a t-test to see whether there is statistically significant evidence that the pay gap has improved or worsened throughout of the pandemic. We found that the gender wage gap still exists but is moving in a downward trajectory. In fact, between December 2019 to December 2022, we found that the gender pay gap has improved slightly, putting women in a better position than they were before the pandemic.

Literature

Allard (2020), Moncada et al. (2021) and Dubois et al. (2022) examine changes in the gender wage gap during the pandemic and factors that may have influenced the change. They found that the gender wage gap may have widened further during the pandemic. According to their research, social factors, marital status, job industry and educational attainment all play a role in wage determination. We added factors in our analysis to try to get a different perspective on the change in the gender pay gap and the contribution of each factor. We also added them to partial-out the effects of other determinants of earnings. These studies analyzed changes in the gender pay gap from the start of the pandemic to 2020.

Therefore, our findings will complement these studies to find whether there has been any change in the gender pay gap from 2019 to 2022. Unlike their work on American, German, and Spanish data, we observe whether different factors that may contribute to the gender pay gap have changed over the pandemic *in Canada.*

Data

Data Sources

The Labour Force Survey was used as the data source. The Labour Force Survey is a monthly survey that measures the current state of the Canadian labour market. According to the *Statistics Act*, it is a mandatory survey, making 'invited' Canadians complete the survey no matter what. This should reduce sampling bias since it is not necessarily a voluntary response survey.

Two versions Labour Force Survey datasets used:

1. Labour Force Survey, December 2019
2. Labour Force Survey, December 2022

Data from 2019 was taken to use as pre-COVID data, and data from 2022 was taken to be used as post-COVID data.

Data Cleaning, Manipulation, and Generation of New Variables

The following adjustments were made to the datasets:

1. Created a log hourly wages variable with hourly earnings.
2. Removed all rows with missing hourly earnings.
3. A new sex variable was created which takes the SEX variable and subtracts the categorical variable indicators by 1.
4. A new covid variable was created for each data set. For the December 2019 data set, the covid variable will be set to 0 to indicate the start of the pandemic, and for the December 2022 data set, the covid variable will be set to 1 to indicate after the pandemic.
5. A new weekly wages variable was constructed multiplying usual hours worked at main job by hourly earnings.
6. A new log weekly wages variable was constructed, taking the log of each of the values in weekly wages.

Description of Variables

Hourly earnings: hourly earnings of someone, measured in Canadian dollars.

Log hourly wages: measures the percentage change in hourly earnings.

Actual Hours Worked at Main Job: measures the self-reported hours worked of an individual at their main job (what they actually worked, not necessarily only the hours they were paid to work for).

Sex: female or male.

Age: age group of a respondent, in years.

Marital status: factor variable that has shows the marital status of a respondent (married, single, divorced, etc.).

Education: Education level (bachelor's degree, 0-8 years of high school, etc.).

Industry of Main Job: factor variable for the industry of main job (agriculture, utilities, etc.).
Occupation at Main Job: factor variable for occupation at main job (health occupations, management, etc.).
Weekly Wages: the usual hours that someone works in a week * hourly wage, measured in Canadian dollars.
Log weekly wages: measures the percentage change in hourly earnings.

Summary Statistics

Recorded after creating log hourlywages (lwage) and removing all rows where lwage = N/A.

# of Observations	
December 2019	December 2022
49502	55752

Means of Continuous Variables

The hourly earnings variable is not used explicitly in the project (log of hourly earnings is used instead), but this table is included to provide more information about what is being studied.

Hourly Earnings (CAD)		
Male/Female	December 2019	December 2022
Both Sexes	27.57	31.61 **(14.7%)**
Male	29.46	33.52 **(+13.8%)**
Female	25.71	29.72 **(+15.6%)**

Log hourly wages (% change in CAD)	
December 2019	December 2022
3.207	3.346

Actual hours worked, hours	
December 2019	December 2022
33.22	32.84

Tabulations for Categorical Variables

Sex		
	December 2019	December 2022
Male	24499	27718
Female	25003	28034

Age		
Age Group	December 2019	December 2022
15 to 19 years	2563	296
20 to 24 years	4145	4293
25 to 29 years	5046	5372
30 to 34 years	5434	6097

35 to 39 years	5682	6249
40 to 44 years	5423	6363
45 to 49 years	5288	5991
50 to 54 years	5221	5974
55 to 59 years	5263	5566
60 to 64 years	3510	4272
65 to 69 years	1331	1722
70 and over	596	890

Marital Status		
Marital Status	**December 2019**	**December 2022**
Married	22916	26278
Living in common-law	8156	8636
Widowed	528	572
Separated	1286	1386
Divorced	2075	2437
Single, never married	14551	16453

Education		
Education Level	**December 2019**	**December 2022**
0 to 8 years	648	640
Some high school	3672	3747
High school graduate	9558	10126
Some postsecondary	2994	2944
Postsecondary certificate or diploma	19033	20307
Bachelor's degree	9371	12109
Above bachelor's degree	4226	5879

Class of Worker		
Class of Worker	**December 2019**	**December 2022**
Public sector employees	13394	15540
Private sector employees	36108	40212
Self-employed incorporated, with paid help	0	0
Self-employed incorporated, no paid help	0	0
Self-employed unincorporated, with paid help	0	0
Self-employed unincorporated, no paid help	0	0
Unpaid family worker	0	0

Industry of Main Job		
Industry	**December 2019**	**December 2022**
Agriculture	518	483

Forestry and logging and support activities for forestry	198	122
Fishing, hunting and trapping	52	52
Mining, quarrying, and oil and gas extraction	1180	1108
Utilities	437	550
Construction	3320	3716
Manufacturing - durable goods	2771	2901
Manufacturing - non-durable goods	2125	2487
Wholesale trade	1528	1749
Retail trade	6194	6670
Transportation and warehousing	2488	2685
Finance and insurance	1896	2634
Real estate and rental and leasing	613	712
Professional, scientific and technical services	2490	3661
Business, building and other support services	1514	1600
Educational services	4553	5091
Health care and social assistance	7376	8266
Information, culture and recreation	1703	1921
Accommodation and food services	3449	3190
Other services (except public administration)	1755	1913
Public administration	3339	4241

Occupation at Main Job		
Occupation	**December 2019**	**December 2022**
Management occupations	2775	4154
Business, finance and administration occupations	8081	9798
Natural and applied sciences and related occupations	3457	4474
Health occupations	4033	4766
Occupations in education, law and social, community and government	6400	4288
Occupations in art, culture, recreation and sport	865	1023
Sales and service occupations	12602	12422
Trades, transport and equipment operators and related occupations	7586	7870
Natural resources, agriculture and related production occupations	1261	1099
Occupations in manufacturing and utilities	2442	2558

Before performing our tests and regressions, we can see preliminary evidence that perhaps women's earnings were not affected as much as some have suggested. Specifically, among both sexes, wages have gone up 14.7% over this 3-year period. Women's hourly earnings increased by 15.6% on average, while men's wages increased by 13.8% on average. While we do not know the exact reasoning for these findings, it is apparent that women (on average) had a larger increase in wages than men, at least from this sampled data.

From the hourly earnings data, we can also see the prevalence of a gender wage gap, by purely comparing the mean wages for women and men (and not holding any factors constant). In the 2019 sample data, men earned $29.46 an hour, while women earned $25.71 an hour on average. In the 2022 sample data, men earned $33.52 an hour while women earned $29.72. While there is a chance that this difference in wages could be in part due to other factors such as education level or occupation, it is reasonable to hypothesize that some of the wage gap is driven by the effect of the sex variable.

Methodology

Assumptions

For our analysis with each regression, we have made a variety of assumptions that are used in many multiple linear regression models, specifically using the six classical linear model assumptions:

1. Linear in Parameters
2. Random Sampling
3. No *perfect* collinearity between independent variables
4. Zero Conditional Mean
5. Homoscedasticity
6. Normality of unobserved error

The random sampling assumption is a safe assumption to make as our sample size is very large given that we're using LFS data. Respondents are required to answer the survey by Canadian law, which reduces the possibility of voluntary response bias.

For perfect collinearity, base categories were dropped since including them would cause perfect collinearity between each of the variables.

Zero conditional mean assumes that the error term has an expected value of zero given the values of any independent variables, implying that there is no correlation between independent variables and the value of the error term. By including a wide variety of independent variables that account for educational differences, actual hours worked, age, etc., we have attempted to mitigate issues with this assumption. However, there may be other independent variables that were omitted that could be heavily correlated with our existing independent variables. This could cause an upward/downward bias on our overall results and coefficients.

The homoscedasticity assumption assumes that the variance of the error terms is the same, given any value of the explanatory variables. This is an important assumption as it simplifies varying statistical formulas that would otherwise become much more complex in the case of heteroskedasticity. This assumption along with the other Gauss-Markov assumptions allows its coefficients to be the best linear unbiased estimators of its corresponding population coefficients.

Our final assumption, the normality assumption of errors is used for statistical inference. Under the first 5 assumptions (Gauss-Markov), the coefficient values' sampling distributions could take on a wide variety of shapes. To restrict this for the sake of statistical inference, an assumption is made that the unobserved error is normally distributed in the population, given that the sampling distributions of OLS estimators depend on the distribution of errors. Since the LFS is a very large sample size, it is reasonable to invoke the central limit theorem to assume that the unobserved errors roughly have a normal distribution.

General Estimation Equation

We will use the LFS to create a model that tests the impact of sex on the change in hourly earnings. We will hold age, education level, marital status, class of worker (e.g. public vs private sector), industry of main job, occupation at main job, and the actual hours worked at main job constant. We will use an Ordinary Least Squares (OLS) regression model. Two regressions were created and reported, one for the December 2019 LFS, and one for the December 2022 LFS, so that we can see how the coefficients have changed.

Here is our regression model:

$$\begin{aligned}\log(\widehat{hourly}\, wages) &= \widehat{\beta_0} + \widehat{\beta_1} age + \widehat{\beta_2} marital\ status + \widehat{\beta_3} education + \widehat{\beta_4} class\ of\ worker \\ &+ \widehat{\beta_5} industry\ of\ main\ job + \widehat{\beta_6} occupation\ at\ main\ job + \widehat{\beta_7} actual\ hours\ worker\end{aligned}$$

Note that this model is written in a simplified form. All the categorical variables (age, marital status, education, class of worker, industry of main job and occupation at main job) have their categories listed in the 'data' section of this report. Aside from the base categories, all the categories for each categorical variable are included in the above regression model. The base categories are cited below in a table as a reference.

Categorical Variable	**Base Category**
Sex	Male
Age	15 to 19 years
Marital Status	Married
Education	0 to 8 years
Class of Worker	Public sector employees
Industry of Main Job	Agriculture
Occupation at Main Job	Management occupation

We decided to deploy a multiple regression model because choosing sex as the only regressor on log wages would bias the results of our analysis, as we know that age, education, class of worker, etc. all have a noticeable effect on hourly wages, and are often correlated with gender. For example, women graduate from university at higher rates than men, which would imply that the effects of omitting the education variable could provide a bias on the overall results for log wages.

With our choices of specific regressors, we picked ones that we thought would affect wages but had a disproportionate effect on women or men. Education is an example of this, which was just cited above. With other variables, we thought that industry of main job, occupation at main job and actual hours worked were important to include as it is often argued that women tend to make less than men in large part due to working fewer hours and choosing lower-paying careers – so these variables were used to partial out most of these effects. Class of worker was included as one's wage can be largely affected based on if they work in the public or private sector, and women tend to work more in the public sector than men so they may be disproportionately affected. Marital status and age were included as these two variables may affect how many hours women work, and there may be some discrepancies between the make-up of ages between men and women in the workforce because of labour force exits due to varying reasons such as having children.

Analyzing the Change in the Gender Wage Gap (hourly) Over the Pandemic

The December 2019 and December 2022 LFS data sets were analyzed together with the December 2019 data entries having a 'covid' variable equal to zero, and December 2022 data having a 'covid' variable equal to 1. This denotes the pre-pandemic data differently from the post-pandemic data to perform the comparison of before and after the pandemic.

We then conducted a similar regression to our main regression using the combined dataset of values from 2019 and 2022. We analyzed how sex interacted on the covid variable to see if there has been a statistically significant effect. This is shown in relation to the single variable and shows the general increase in wages during the pandemic.

$$
\begin{aligned}
&log\ (\widehat{hourly}\ wages) \\
&\qquad = \widehat{\beta_0} + \widehat{\beta_1} sex + \widehat{\beta_2} age + \widehat{\beta_3} marital\ status + \widehat{\beta_4} education \\
&\qquad + \widehat{\beta_5} class\ of\ worker + \widehat{\beta_6} industry\ of\ worker\ at\ main\ job \\
&\qquad + \widehat{\beta_7} occupation\ of\ worker\ at\ main\ job + \widehat{\beta_8} actual\ hours\ worked + \widehat{\delta_1} sex \\
&\qquad * covid + \hat{\beta}_9 covid
\end{aligned}
$$

This equation is different from that of the main regression model as it does not show all the categories for the individual categorical variables. The base categories for each of the categorical variables are shown earlier in this section and are represented in this equation.

A t-test was conducted for the coefficient on the interacted variable between sex and covid, which shows if there has been a statistically significant change in the gender pay gap since the start of the pandemic.

Sensitivity Analysis Tests

Analyzing the Change in the Gender Wage Gap (weekly) over the Course of the Pandemic

$$\begin{aligned} log\ (\widehat{weekly}\ wages) &= \widehat{\beta_0} + \widehat{\beta_1} sex + \widehat{\beta_2} age + \widehat{\beta_3} marital\ status + \widehat{\beta_4} education \\ &+ \widehat{\beta_5} class\ of\ worker + \widehat{\beta_6} industry\ of\ worker\ at\ main\ job \\ &+ \widehat{\beta_7} occupation\ of\ worker\ at\ main\ job + \widehat{\beta_8} actual\ hours\ worked + \widehat{\delta_1} sex \\ &* covid + \hat{\beta}_9 covid \end{aligned}$$

A sensitivity test was conducted using the same regression that was used to analyze the change in the gender wage gap (hourly) during the pandemic, however, this uses the log(weekly wages) as the dependent variable. This regression test was used in consideration of the fact that women might still retain similar hourly wages but may be working reduced hours from men in part due to 'pandemic-type' effects. Thus, their weekly wages would be affected more than their hourly wages.

Removing Industry from December 2022 Regression

This sensitivity test will take the original December 2022 data set regression performed with log wages as the dependent variable and will remove the industry of main job as one of the regressors. The aim of this is to see how to omission of the industry variable biases the results, and in what direction.

Results

General 'Mincer' Regression on Log Hourly Earnings

Name of Variable	2022 Estimated Coefficient (standard error)	2019 Estimated Coefficient (standard error)
Female	-0.096***(0.003)	-0.101*** (-0.004)
20 to 24 years	-0.014* (0.008)	0.008 (0.009)
25 to 29 years	0.100***(0.009)	0.089*** (0.009)
30 to 34 years	0.161***(0.009)	0.164***(0.009)
35 to 39 years	0.199***(0.009)	0.192*** (0.009)
40 to 44 years	0.207***(0.009)	0.217*** (0.009)
45 to 49 years	0.213***(0.009)	0.227*** (0.009)
50 to 54 years	0.216***(0.009)	0.230*** (0.009)
55 to 59 years	0.205***(0.009)	0.200*** (0.009)
60 to 64 years	0.173***(0.009)	0.177*** (0.010)
65 to 69 years	0.114***(0.011)	0.115*** (0.012)
70 and over	0.072***(0.013)	0.042*** (0.016)
Living in common-law	0.001 (0.004)	-0.009** (0.004)
Widowed	-0.035**(0.014)	-0.031**(0.015)
Separated	-0.033***(0.009)	-0.008 (0.009)

Divorced	-0.009 (0.007)	-0.016** (0.008)
Single, never married	-0.060***(0.004)	-0.066***(0.004)
Some high school	0.047***(0.014)	0.049***(.014)
High school graduate	0.078***(0.013)	0.090***(0.013)
Some postsecondary	0.100***(0.014)	0.108***(0.014)
Postsecondary certificate or diploma	0.136***(0.013)	0.148***(0.013)
Bachelor's degree	0.240***(0.013)	0.268***(0.014)
Above bachelor's degree	0.309***(0.014)	0.347*** (.014)
Private sector employees	-0.159***(0.005)	-0.190***(0.006)
Forestry and logging and support activities for forestry	0.211***(0.033)	0.259***(0.028)
Fishing, hunting and trapping	-0.022 (0.047)	0.023 (0.048)
Mining quarrying, and oil and gas extraction	0.527***(0.019)	0.590***(0.019)
Utilities	0.411***(0.024)	0.476***(0.025)
Construction	0.291***(0.020)	0.320*** (0.020)
Manufacturing - durable goods	0.282***(0.021)	0.296*** (0.020)
Manufacturing - non-durable goods	0.217***(0.021)	0.217***(0.021)
Wholesale trade	0.231***(0.021)	.266*** (0.021)
Retail trade	0.062***(0.020)	0.058*** (0.020)
Transportation and warehousing	0.152***(0.021)	0.159*** (0.020)
Finance and insurance	0.344***(0.021)	0.349***(0.021)
Real estate and rental and leasing	0.152***(0.023)	0.158*** (0.023)
Professional, scientific and technical services	0.293***(0.020)	0.280***(0.020)
Business, building and other support services	0.089***(0.021)	0.081***(0.021)
Educational services	0.122***(0.021)	0.141***(0.021)
Health care and social assistance	0.052**(0.020)	0.066***(0.020)
Information, culture and recreation	0.154***(0.021)	0.167***(0.021)
Accommodation and food services	0.054***(0.021)	0.043** (0.020)
Other services (except public administration)	0.121***(0.021)	0.132***(0.021)
Public administration	0.236***(0.021)	0.232***(0.021)
Business, finance and administration occupations	-0.413***(0.006)	-0.386***(0.007)
Natural and applied sciences and related occupations	-0.209***(0.007)	-0.191***(0.009)
Health occupations	-0.204***(0.008)	-0.178***(0.009)
Occupations in education, law and social, community and government services	-0.288***(0.007)	-0.291*** (0.008)

Occupations in art, culture, recreation and sport	-0.436***(0.012)	-0.391*** (0.014)
Sales and service occupations	-0.575***(0.006)	-0.531***(0.008)
Trades, transport and equipment operators and related occupations	-0.408***(0.007)	-0.367***(0.008)
Natural resources, agriculture and related production occupations	-0.423***(0.015)	-0.407***(0.015)
Occupations in manufacturing and utilities	-0.556***(0.010)	-0.486***(0.011)
Actual hours worked	0.002***(0.0001)	0.002** (0.0001)
Constant	3.366***(0.025)	3.217***(0.026)

December 2019: $R^2 = 0.501, n = 49502$
December 2022: $R^2 = 0.505, n = 55752$

Statistical significance can be shown denoted by the *p<0.1; **p<0.05; ***p<0.01 on the coefficient estimations. The coefficient of 0.590 for mining quarrying, and oil and gas extraction shows that any individual working in this sector is more likely to earn 59% more than someone working in agriculture, holding all else constant. The same is shown as the level of education increases.

In the December 2019 data set, holding all else constant, a woman earns 10.1% less than a man. In the December 2022 data set, holding all else constant, a woman earns 9.6% less than a man. Even after we included multiple wage-effecting independent variables in our regression model to hold constant, it appears that a wage gap is still prevalent. However, it is notable that with the same regression model, this coefficient has reduced in magnitude over 3 years. The next regression (below) computes a t-test to see if this change is statistically significant.

Change in Gender Wage Gap Over the Pandemic (hourly earnings and weekly earnings)

Name of Variable	Estimated Coefficient (standard error) with log(hourly earnings)	Estimated Coefficient (standard error) with log(weekly earnings)
Female	-0.104*** (0.003)	-0.140*** (0.005)
20 to 24 years	-0.004 (0.006)	0.353*** (0.009)
25 to 29 years	0.094*** (0.006)	0.612*** (0.009)
30 to 34 years	0.162*** (0.006)	0.702*** (0.009)
35 to 39 years	0.195*** (0.006)	0.722*** (0.009)
40 to 44 years	0.211*** (0.006)	0.726*** (0.009)
45 to 49 years	0.219*** (0.006)	0.738*** (0.010)
50 to 54 years	0.221*** (0.006)	0.739*** (0.010)
55 to 59 years	0.202*** (0.006)	0.709*** (0.010)
60 to 64 years	0.174*** (0.007)	0.638*** (0.010)
65 to 69 years	0.114*** (0.008)	0.435*** (0.012)
70 and over	0.061*** (0.010)	0.251*** (0.015)
Living in common-law	-0.004 (0.003)	0.016*** (0.004)
Widowed	-0.033*** (0.010)	-0.044*** (0.015)
Separated	-0.021*** (0.006)	-0.011 (0.010)

Divorced	-0.012** (0.005)	0.004 (0.008)
Single, never married	-0.063*** (0.003)	-0.082*** (0.004)
Some high school	0.048*** (0.010)	-0.016 (0.015)
High school graduate	0.084*** (0.009)	0.090*** (0.014)
Some postsecondary	0.103*** (0.010)	0.046*** (0.015)
Postsecondary certificate or diploma	0.141*** (0.009)	0.143*** (0.014)
Bachelor's degree	0.253*** (0.010)	0.255*** (0.014)
Above bachelor's degree	0.326*** (0.010)	0.328*** (0.015)
Private sector employees	-0.174*** (0.004)	-0.201*** (0.006)
Forestry and logging and support activities for forestry	0.239*** (0.021)	0.302*** (0.031)
Fishing, hunting and trapping	0.0003 (0.034)	0.140*** (0.049)
Mining quarrying, and oil and gas extraction	0.561*** (0.013)	0.641*** (0.020)
Utilities	0.442*** (0.017)	0.422*** (0.026)
Construction	0.307*** (0.014)	0.344*** (0.021)
Manufacturing - durable goods	0.291*** (0.014)	0.327*** (0.021)
Manufacturing - non-durable goods	0.220*** (0.015)	0.249*** (0.022)
Wholesale trade	0.250*** (0.015)	0.303*** (0.022)
Retail trade	0.063*** (0.014)	0.054*** (0.021)
Transportation and warehousing	0.157*** (0.014)	0.173*** (0.021)
Finance and insurance	0.350*** (0.015)	0.398*** (0.021)
Real estate and rental and leasing	0.158*** (0.016)	0.165*** (0.024)
Professional, scientific and technical services	0.292*** (0.014)	0.308*** (0.021)
Business, building and other support services	0.088*** (0.015)	0.109*** (0.021)
Educational services	0.134*** (0.015)	0.014 (0.022)
Health care and social assistance	0.061*** (0.014)	0.045** (0.021)
Information, culture and recreation	0.163*** (0.015)	0.100*** (0.022)
Accommodation and food services	0.051*** (0.014)	0.016 (0.021)
Other services (except public administration)	0.129*** (0.015)	0.084*** (0.022)
Public administration	0.236*** (0.015)	0.244*** (0.022)
Business, finance and administration occupations	-0.402*** (0.005)	-0.435*** (0.007)
Natural and applied sciences and related occupations	-0.202*** (0.005)	-0.206*** (0.008)
Health occupations	-0.193*** (0.006)	-0.225*** (0.009)
Occupations in education, law and social, community and government services	-0.291*** (0.005)	-0.336*** (0.008)
Occupations in art, culture, recreation and sport	-0.418*** (0.009)	-0.623*** (0.013)
Sales and service occupations	-0.556*** (0.005)	-0.628*** (0.007)
Trades, transport and equipment operators and related occupations	-0.391*** (0.005)	-0.367*** (0.008)
Natural resources, agriculture and related production occupations	-0.416*** (0.011)	-0.380*** (0.016)
Occupations in manufacturing and utilities	-0.524*** (0.007)	-0.504*** (0.010)
Actual hours worked	0.002*** (0.0001)	0.015*** (0.0001)
Covid	0.102*** (0.003)	0.112*** (0.004)
Covid * Sex	0.011*** (0.004)	0.011* (0.006)
Constant	3.242*** (0.018)	5.901*** (0.027)

For the regression that estimates log(wages): $R^2 = 0.513, n = 105254$

For the regression that estimates log(weekly earnings): $R^2 = 0.583, n = 105254$

Two regressions were conducted to analyze the change in the gender wage gap throughout the pandemic (from December 2019 to December 2022). One was for the percentage change in hourly wages (log hourly earnings), and the other was for the percentage change in yearly wages (log yearly earnings). The log yearly wages regression was conducted as a sensitivity test.

With the log hourly wages regression, the coefficient on female was -0.104. We can interpret this as holding all else constant, a female earns 10.4% less than a man in our combined time periods. The covid coefficient was equal to 0.102, which can be interpreted as holding all else constant, someone working in December 2022 earns 10.2% more than someone working in December 2019. This makes sense given the general increase in nominal wages coming mainly as a result of inflation.

Notably, the covid * sex coefficient was 0.011. This is statistically significant at a 99% confidence level, like the previously cited coefficients from this regression. This can be interpreted as holding all else constant, the return on the December 2022 time period (relative to December 2019) is 1.1% higher for women in comparison to men. We are unable to hold numerous factors constant that may have generated this result so it is hard to come to a causal conclusion (e.g. "the COVID-19 pandemic has benefited women more than men"), but it does appear that women have faired better than men over this December 2019-2022 time period then men have, with hourly earnings increases.

In our sensitivity analysis, we performed the same regression just with weekly hourly wages, and we got slightly different results. The coefficient on female was of higher magnitude in the negative direction. Holding all else constant, the percentage change in weekly wages was 14% lower for Canadian women in this combined data set vs men. Notably, the covid * sex coefficient was the same but reported at a lower confidence level (90% confidence) due to a higher standard error associated with the coefficient. So at a 90% confidence level and holding all else constant, a Canadian woman had a 1.1% better return on the December 2022 period (in comparison to December 2019) than a man did. Once again, it appears that women tended to fair better regarding their change in earnings then men did over these 3 years.

Industry Sensitivity Test

Name of Variable	**December 2019 - Estimated Coefficient (standard error)**	**December 2022 - Estimated Coefficient (standard error)**
Female	-0.126*** (0.004)	-0.116*** (0.003)
20 to 24 years	0.020** (0.009)	0.0003 (0.009)
25 to 29 years	0.112*** (0.009)	0.123*** (0.009)
30 to 34 years	0.191*** (0.009)	0.188*** (0.009)
35 to 39 years	0.218*** (0.010)	0.230*** (0.009)
40 to 44 years	0.244*** (0.010)	0.233*** (0.009)
45 to 49 years	0.251*** (0.010)	0.238*** (0.009)
50 to 54 years	0.251*** (0.010)	0.240*** (0.009)
55 to 59 years	0.220*** (0.010)	0.225*** (0.009)
60 to 64 years	0.192*** (0.010)	0.191*** (0.009)
65 to 69 years	0.126*** (0.013)	0.124*** (0.011)
70 and over	0.040** (0.016)	0.078*** (0.014)

Living in common-law	-0.011** (0.005)	0.001 (0.004)
Widowed	-0.036** (0.015)	-0.036** (0.014)
Separated	-0.015 (0.010)	-0.038*** (0.009)
Divorced	-0.021*** (0.008)	-0.011 (0.007)
Single, never married	-0.079*** (0.005)	-0.069*** (0.004)
Some high school	0.061*** (0.015)	0.059*** (0.015)
High school graduate	0.110*** (0.014)	0.095*** (0.014)
Some postsecondary	0.128*** (0.015)	0.120*** (0.015)
Postsecondary certificate or diploma	0.181*** (0.014)	0.165*** (0.014)
Bachelor's degree	0.307*** (0.014)	0.278*** (0.014)
Above bachelor's degree	0.385*** (0.015)	0.346*** (0.014)
Private sector employees	-0.183*** (0.004)	-0.147*** (0.004)
Business, finance and administration occupations	-0.362*** (0.008)	-0.402*** (0.006)
Natural and applied sciences and related occupations	-0.129*** (0.009)	-0.165*** (0.007)
Health occupations	-0.285*** (0.009)	-0.331*** (0.007)
Occupations in education, law and social, community and gov -ernment services	-0.333*** (0.008)	0.348*** (0.007)
Occupations in art, culture, recreation and sport	-0.391*** (0.014)	-0.454*** (0.012)
Sales and service occupations	-0.600*** (0.008)	-0.653*** (0.006)
Trades, transport and equipment operators and related occu pations	-0.320*** (0.008)	-0.390*** (0.007)
Natural resources, agriculture and related production occupa -tions	-0.329*** (0.012)	-0.399*** (0.012)
Occupations in manufacturing and utilities	-0.410*** (0.010)	-0.498*** (0.009)
Actual hours worked	0.002*** (0.0001)	0.002*** (0.0001)
Constant	3.345*** (0.018)	3.501*** (0.017)

2019 Data: $R^2 = 0.455, n = 49502$

2022 Data: $R^2 = 0.464, n = 55752$

From this regression (using the December 2022 LFS data set), holding all else constant, a woman earns 11.6% less than a man. We can be 99% confident that the null hypothesis is rejected, meaning that the effect of being a female on wages is non-zero. In comparison to the original December 2022 regression where a woman was estimated to earn 9.6% less than a man, this regression coefficient is of larger magnitude in the negative direction by 2%. This implies that the omission of the industry variable would negatively bias the results, making the gender wage gap look larger than it is.

Using the December 2019 data, holding all else constant, a woman earns 12.6% less than a man. Again, this represents a larger gender pay gap than what was originally stated in the regression that included industry as an independent variable.

Overview of Results

Focusing on the gender pay gap and the factors that influence wages, the Labour Force Surveys from December 2019 and 2022 were used to assess the magnitude of differences in income between gender. Holding education and hours worked constant, our analysis still showed women earning less than men. During the pandemic, women saw a greater increase in income and a smaller wage gap in comparison to pre-pandemic levels in 2019. This is a positive change for women’s earnings over the three-year period that is trending towards eliminating a wage gap between genders.

After analyzing this data, we can see that wages are predicted by education, experience, and field of work. Knowing this, we still see a wage gap between genders that could be attributed to the factors of education, experience, or type of work – while there may be societal issues affecting the balance of income between genders, progress has been made to narrow this wage gap and we can see this narrowing over the pandemic years. Men and women are represented differently in the labour force and many factors influence wages for different individuals.

Conclusion

From our analysis, it appears that the gender pay gap persists, through the regressions performed to showcase the effects that age, marital status, education, industry, occupation, and other variables have on the gender pay gap when only comparing the differences in the mean wage of men and women (and holding no variable constant). Our analysis also concluded that the gender pay gap has shrunk in Canada from December 2019 to December 2022.

The gender pay gap is largely a societal issue at its core, and this is an element that cannot be exactly quantified to use in this analysis. So, although there are many personal and employment factors used, attitudes of the Canadian population are an important factor that this analysis does not consider. If further research could be used to expand this topic, including an anthropological element would help to include the societal aspect of the gender pay gap. If research were continued on a more specified basis, a psychological aspect could also be included for the possibility of more detailed findings. It is also difficult to fully determine all the independent variables that should be included in the regression, which could lead to biased results and coefficients in our regression results.

References

Allard, S. (2020). Placements and salaries 2020: Salaries and full-time employment are up, but so are unemployment and the gender gap; 2019 graduates faced A mixed job market even before the pandemic. *Library Journal, 145*(10), 27. Retrieved from https://subzero.lib.uoguelph.ca/login?URL=?url=https://www.proquest.com/trade-journals/placements-salaries-2020/docview/2446961211/se-2

Salas-Nicás, S., Moncada, S., Llorens, C., & Navarro, A. (2021). Working conditions and health in Spain during the COVID-19 pandemic: Minding the gap. *Safety Science*, *134*(Complete). https://doi.org/10.1016/j.ssci.2020.105064

Dubois, C., Lambertini, L., & Wu, Y. (2022). Gender effects of the COVID-19 pandemic in the Swiss labor market. *Swiss Journal of Economics and Statistics*, *158*(1), NA. https://link-gale-com.subzero.lib.uoguelph.ca/apps/doc/A729450127/AONE?u=guel77241&sid=bookmark-AONE&xid=05cc7db8

Global Wage Report 2022–23: The impact of inflation and COVID-19 on wages and purchasing power. Geneva: International Labour Office, 2022.

Pelletier, R., Patterson, M., & Moyser, M. (2019). *The Gender Wage Gap in Canada: 1998 to 2018*. Statistics Canada. https://books-scholarsportal-info.subzero.lib.uoguelph.ca/uri/ebooks/ebooks5/cpdc5/2020-03-12/1/10102538

R Output

```
#set-up
rm(list=ls())
setwd("C:/Users/isss1/Documents/University/Semester 4/ECON-3740/Group Project
/Data and Code/")

#loading libraries
library(haven)
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.
3.2 ──
## ✓ ggplot2 3.4.0      ✓ purrr   1.0.1
## ✓ tibble  3.1.8      ✓ dplyr   1.0.10
## ✓ tidyr   1.2.1      ✓ stringr 1.5.0
## ✓ readr   2.1.3      ✓ forcats 0.5.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflict
s() ──
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()    masks stats::lag()

library(stargazer)

##
## Please cite as:
##
##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary St
atistics Tables.
##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

# loading up the two data sets
dec2019 <- read_stata("C:/Users/isss1/Documents/University/Semester 4/ECON-37
40/Group Project/Data and Code/LFS-71M0001-E-2019-December/LFS-71M0001-E-2019
-December_F1.dta")
dec2022 <- read_stata("C:/Users/isss1/Documents/University/Semester 4/ECON-37
40/Group Project/Data and Code/LFS-71M0001-E-2022-December/LFS-71M0001-E-2022
-December_F1.dta")

# making wage adjustments
# creating a log wage
dec2019 <- dec2019 %>% mutate(lwage=log(HRLYEARN))
dec2022 <- dec2022 %>% mutate(lwage=log(HRLYEARN))
dec2019 <- dec2019 %>% mutate(sex = SEX - 1)
dec2022 <- dec2022 %>% mutate(sex = SEX - 1)
```

```
# filter out all rows with missing values for lwage
dec2019 <- dec2019 %>% filter(!is.na(lwage))
dec2022 <- dec2022 %>% filter(!is.na(lwage))

# creating a data set where we can create a covidTime variable to run an unre
stricted model
# creating a covid indicator to be able to know what is post-covid and pre-co
vid when we created combinedData
dec2019 <- dec2019 %>% mutate(covid=0)
dec2022 <- dec2022 %>% mutate(covid=1)
combinedData <- bind_rows(dec2019, dec2022)

## Warning: `..1$NOC_10` and `..2$NOC_10` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$NOC_10`.
## ✖ Values: 5, 8, and 9

## Warning: `..1$NOC_40` and `..2$NOC_40` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$NOC_40`.
## ✖ Values: 3, 4, 6, 7, 9, 11, 17, 18, …, 37, and 38

## Warning: `..1$WHYPT` and `..2$WHYPT` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$WHYPT`.
## ✖ Values: 6 and 7

## Warning: `..1$UNION` and `..2$UNION` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$UNION`.
## ✖ Values: 2

## Warning: `..1$FLOWUNEM` and `..2$FLOWUNEM` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$FLOWUNEM`.
## ✖ Values: 4

## Warning: `..1$YNOLOOK` and `..2$YNOLOOK` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$YNOLOOK`.
## ✖ Values: 3

## Warning: `..1$EFAMTYPE` and `..2$EFAMTYPE` have conflicting value labels.
## ℹ Labels for these values will be taken from `..1$EFAMTYPE`.
## ✖ Values: 1, 2, 3, 4, 5, 6, 7, 8, …, 16, and 17

# filtering out data where we cannot get a weekly wage using usual hours work
ed, so we can compute weekly wage (usual hours is non-zero for wage reporters
)
# implies that everyone in our data set must have worked
combinedData <- combinedData %>% mutate(weeklywage = UHRSMAIN * HRLYEARN)
combinedData <- combinedData %>% mutate(lweeklywage = log(weeklywage))

# getting means and proportions of variables for proposal
summary(dec2019$HRLYEARN)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3.00   17.00   24.00   27.57   35.00  110.19

summary(dec2022$HRLYEARN)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3.00   20.00   27.50   31.61   40.00  115.38

# getting specific means for female wages and men by creating new variables
dec2019Men <- dec2019 %>% filter(sex==0)
dec2022Men <- dec2022 %>% filter(sex==0)
dec2019Women <- dec2019 %>% filter(sex==1)
dec2022Women <- dec2022 %>% filter(sex==1)

summary(dec2019Men$HRLYEARN)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3.00   18.38   26.00   29.46   37.50  108.17

summary(dec2022Men$HRLYEARN)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3.00   21.00   30.00   33.52   42.00  115.38

summary(dec2019Women$HRLYEARN)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3.08   16.00   22.00   25.71   31.85  110.19

summary(dec2022Women$HRLYEARN)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    3.08   19.00   25.54   29.72   37.50  112.82

summary(dec2019$lwage)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.099   2.833   3.178   3.207   3.555   4.702

summary(dec2022$lwage)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.099   2.996   3.314   3.346   3.689   4.748

table(dec2019$SEX)

##
##     1     2
## 24499 25003

table(dec2022$SEX)
```

```
##
##     1     2
## 27718 28034
```

```
table(dec2019$AGE_12)
```

```
##
##    1    2    3    4    5    6    7    8    9   10   11   12
## 2563 4145 5046 5434 5682 5423 5288 5221 5263 3510 1331  596
```

```
table(dec2022$AGE_12)
```

```
##
##    1    2    3    4    5    6    7    8    9   10   11   12
## 2963 4293 5372 6097 6249 6363 5991 5974 5566 4272 1722  890
```

```
table(dec2019$MARSTAT)
```

```
##
##     1     2     3     4     5     6
## 22916  8156   528  1276  2075 14551
```

```
table(dec2022$MARSTAT)
```

```
##
##     1     2     3     4     5     6
## 26278  8636   572  1376  2437 16453
```

```
table(dec2019$EDUC)
```

```
##
##     0     1     2     3     4     5     6
##   648  3672  9558  2994 19033  9371  4226
```

```
table(dec2022$EDUC)
```

```
##
##     0     1     2     3     4     5     6
##   640  3747 10126  2944 20307 12109  5879
```

```
table(dec2019$NAICS_21)
```

```
##
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
16
##  518  198   52 1180  437 3320 2771 2128 1528 6194 2488 1896  613 2490 1514
4553
##   17   18   19   20   21
## 7376 1703 3449 1755 3339
```

```
table(dec2022$NAICS_21)
```

```
##
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
16
##  483  122   52 1108  550 3716 2901 2487 1749 6670 2685 2634  712 3661 1600
5091
##   17   18   19   20   21
## 8266 1921 3190 1913 4241
```

```
table(dec2019$NOC_10)
```

```
##
##     1     2     3     4     5     6     7     8     9    10
##  2775  8081  3457  4033  6400   865 12602  7586  1261  2442
```

```
table(dec2022$NOC_10)
```

```
##
##     1     2     3     4     5     6     7     8     9    10
##  4154  9798  4774  4766  7288  1023 12422  7870  1099  2558
```

```
summary(dec2019$AHRSMAIN)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   26.00   37.50   33.22   40.00   99.00
```

```
summary(dec2022$AHRSMAIN)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   25.00   37.50   32.84   40.00   99.00
```

```
# more detailed regression
r2022multiple <- lm(lwage ~ sex + factor(AGE_12) + factor(MARSTAT)
                 +factor(EDUC) + factor(COWMAIN) + factor(NAICS_21) + factor(
NOC_10) + AHRSMAIN, data=dec2022)

stargazer(r2022multiple,type="text")
```

```
##
## =================================================
##                          Dependent variable:
##                     -----------------------------
##                                 lwage
## -------------------------------------------------
## sex                           -0.096***
##                                (0.003)
##
## factor(AGE_12)2                -0.014*
##                                (0.008)
##
## factor(AGE_12)3               0.100***
##                                (0.009)
##
```

```
## factor(AGE_12)4                0.161***
##                                 (0.009)
##
## factor(AGE_12)5                0.199***
##                                 (0.009)
##
## factor(AGE_12)6                0.207***
##                                 (0.009)
##
## factor(AGE_12)7                0.213***
##                                 (0.009)
##
## factor(AGE_12)8                0.216***
##                                 (0.009)
##
## factor(AGE_12)9                0.205***
##                                 (0.009)
##
## factor(AGE_12)10               0.173***
##                                 (0.009)
##
## factor(AGE_12)11               0.114***
##                                 (0.011)
##
## factor(AGE_12)12               0.072***
##                                 (0.013)
##
## factor(MARSTAT)2                  0.001
##                                 (0.004)
##
## factor(MARSTAT)3               -0.035**
##                                 (0.014)
##
## factor(MARSTAT)4               -0.033***
##                                 (0.009)
##
## factor(MARSTAT)5                -0.009
##                                 (0.007)
##
## factor(MARSTAT)6               -0.060***
##                                 (0.004)
##
## factor(EDUC)1                  0.047***
##                                 (0.014)
##
## factor(EDUC)2                  0.078***
##                                 (0.013)
##
## factor(EDUC)3                  0.100***
##                                 (0.014)
```

```
##
## factor(EDUC)4                    0.136***
##                                   (0.013)
##
## factor(EDUC)5                    0.240***
##                                   (0.013)
##
## factor(EDUC)6                    0.309***
##                                   (0.014)
##
## factor(COWMAIN)2                -0.159***
##                                   (0.005)
##
## factor(NAICS_21)2                0.211***
##                                   (0.033)
##
## factor(NAICS_21)3                 -0.022
##                                   (0.047)
##
## factor(NAICS_21)4                0.527***
##                                   (0.019)
##
## factor(NAICS_21)5                0.411***
##                                   (0.024)
##
## factor(NAICS_21)6                0.291***
##                                   (0.020)
##
## factor(NAICS_21)7                0.282***
##                                   (0.021)
##
## factor(NAICS_21)8                0.217***
##                                   (0.021)
##
## factor(NAICS_21)9                0.231***
##                                   (0.021)
##
## factor(NAICS_21)10               0.062***
##                                   (0.020)
##
## factor(NAICS_21)11               0.152***
##                                   (0.021)
##
## factor(NAICS_21)12               0.344***
##                                   (0.021)
##
## factor(NAICS_21)13               0.152***
##                                   (0.023)
##
## factor(NAICS_21)14               0.293***
```

```
##                                      (0.020)
##
## factor(NAICS_21)15                 0.089***
##                                      (0.021)
##
## factor(NAICS_21)16                 0.122***
##                                      (0.021)
##
## factor(NAICS_21)17                  0.052**
##                                      (0.020)
##
## factor(NAICS_21)18                 0.154***
##                                      (0.021)
##
## factor(NAICS_21)19                 0.054***
##                                      (0.021)
##
## factor(NAICS_21)20                 0.121***
##                                      (0.021)
##
## factor(NAICS_21)21                 0.236***
##                                      (0.021)
##
## factor(NOC_10)2                   -0.413***
##                                      (0.006)
##
## factor(NOC_10)3                   -0.209***
##                                      (0.007)
##
## factor(NOC_10)4                   -0.204***
##                                      (0.008)
##
## factor(NOC_10)5                   -0.288***
##                                      (0.007)
##
## factor(NOC_10)6                   -0.436***
##                                      (0.012)
##
## factor(NOC_10)7                   -0.575***
##                                      (0.006)
##
## factor(NOC_10)8                   -0.408***
##                                      (0.007)
##
## factor(NOC_10)9                   -0.423***
##                                      (0.015)
##
## factor(NOC_10)10                  -0.556***
##                                      (0.010)
##
```

```
## AHRSMAIN                      0.002***
##                               (0.0001)
##
## Constant                      3.366***
##                                (0.025)
##
## ---------------------------------------------------
## Observations                    55,752
## R2                               0.505
## Adjusted R2                      0.504
## Residual Std. Error      0.322 (df = 55697)
## F Statistic         1,051.701*** (df = 54; 55697)
## ===================================================
## Note:                   *p<0.1; **p<0.05; ***p<0.01
```

```
r2019multiple <- lm(lwage ~ factor(SEX) + factor(AGE_12) + factor(MARSTAT)
                  +factor(EDUC) + factor(COWMAIN) + factor(NAICS_21) + factor
(NOC_10) + AHRSMAIN, data=dec2019)

stargazer(r2019multiple, type="text")
```

```
##
## =================================================
##                          Dependent variable:
##                      ---------------------------
##                                 lwage
## -------------------------------------------------
## factor(SEX)2                  -0.101***
##                               (0.004)
##
## factor(AGE_12)2                0.008
##                               (0.009)
##
## factor(AGE_12)3              0.089***
##                               (0.009)
##
## factor(AGE_12)4              0.164***
##                               (0.009)
##
## factor(AGE_12)5              0.192***
##                               (0.009)
##
## factor(AGE_12)6              0.217***
##                               (0.009)
##
## factor(AGE_12)7              0.227***
##                               (0.009)
##
## factor(AGE_12)8              0.230***
##                               (0.009)
```

```
##
## factor(AGE_12)9              0.200***
##                              (0.009)
##
## factor(AGE_12)10             0.177***
##                              (0.010)
##
## factor(AGE_12)11             0.115***
##                              (0.012)
##
## factor(AGE_12)12             0.042***
##                              (0.016)
##
## factor(MARSTAT)2             -0.009**
##                              (0.004)
##
## factor(MARSTAT)3             -0.031**
##                              (0.015)
##
## factor(MARSTAT)4              -0.008
##                              (0.009)
##
## factor(MARSTAT)5             -0.016**
##                              (0.008)
##
## factor(MARSTAT)6             -0.066***
##                              (0.004)
##
## factor(EDUC)1                0.049***
##                              (0.014)
##
## factor(EDUC)2                0.090***
##                              (0.013)
##
## factor(EDUC)3                0.108***
##                              (0.014)
##
## factor(EDUC)4                0.148***
##                              (0.013)
##
## factor(EDUC)5                0.268***
##                              (0.014)
##
## factor(EDUC)6                0.347***
##                              (0.014)
##
## factor(COWMAIN)2             -0.190***
##                              (0.006)
##
## factor(NAICS_21)2            0.259***
```

```
##                                  (0.028)
##
## factor(NAICS_21)3                 0.023
##                                  (0.048)
##
## factor(NAICS_21)4               0.590***
##                                  (0.019)
##
## factor(NAICS_21)5               0.476***
##                                  (0.025)
##
## factor(NAICS_21)6               0.320***
##                                  (0.020)
##
## factor(NAICS_21)7               0.296***
##                                  (0.020)
##
## factor(NAICS_21)8               0.217***
##                                  (0.021)
##
## factor(NAICS_21)9               0.266***
##                                  (0.021)
##
## factor(NAICS_21)10              0.058***
##                                  (0.020)
##
## factor(NAICS_21)11              0.159***
##                                  (0.020)
##
## factor(NAICS_21)12              0.349***
##                                  (0.021)
##
## factor(NAICS_21)13              0.158***
##                                  (0.023)
##
## factor(NAICS_21)14              0.280***
##                                  (0.020)
##
## factor(NAICS_21)15              0.081***
##                                  (0.021)
##
## factor(NAICS_21)16              0.141***
##                                  (0.021)
##
## factor(NAICS_21)17              0.066***
##                                  (0.020)
##
## factor(NAICS_21)18              0.167***
##                                  (0.021)
##
```

```
## factor(NAICS_21)19            0.043**
##                              (0.020)
##
## factor(NAICS_21)20           0.132***
##                              (0.021)
##
## factor(NAICS_21)21           0.232***
##                              (0.021)
##
## factor(NOC_10)2             -0.386***
##                              (0.007)
##
## factor(NOC_10)3             -0.191***
##                              (0.009)
##
## factor(NOC_10)4             -0.178***
##                              (0.009)
##
## factor(NOC_10)5             -0.291***
##                              (0.008)
##
## factor(NOC_10)6             -0.391***
##                              (0.014)
##
## factor(NOC_10)7             -0.531***
##                              (0.008)
##
## factor(NOC_10)8             -0.367***
##                              (0.008)
##
## factor(NOC_10)9             -0.407***
##                              (0.015)
##
## factor(NOC_10)10            -0.486***
##                              (0.011)
##
## AHRSMAIN                     0.002***
##                             (0.0001)
##
## Constant                     3.217***
##                              (0.026)
##
## -------------------------------------------------
## Observations                  49,502
## R2                             0.501
## Adjusted R2                    0.501
## Residual Std. Error     0.328 (df = 49447)
## F Statistic         920.962*** (df = 54; 49447)
## =================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

```r
# doing unrestricted test to see relative effect of COVID on wages
rSensitivitySex <- lm(lwage ~ sex + factor(AGE_12) + factor(MARSTAT)
                    +factor(EDUC) + factor(COWMAIN) + factor(NAICS_21) + fact
or(NOC_10) + AHRSMAIN + sex:covid + covid, data=combinedData)

stargazer(rSensitivitySex, type="text")

## 
## ==================================================
##                          Dependent variable:
##                      -----------------------------
##                                 lwage
## --------------------------------------------------
## sex                           -0.104***
##                                (0.003)
## 
## factor(AGE_12)2                 -0.004
##                                (0.006)
## 
## factor(AGE_12)3                0.094***
##                                (0.006)
## 
## factor(AGE_12)4                0.162***
##                                (0.006)
## 
## factor(AGE_12)5                0.195***
##                                (0.006)
## 
## factor(AGE_12)6                0.211***
##                                (0.006)
## 
## factor(AGE_12)7                0.219***
##                                (0.006)
## 
## factor(AGE_12)8                0.221***
##                                (0.006)
## 
## factor(AGE_12)9                0.202***
##                                (0.006)
## 
## factor(AGE_12)10               0.174***
##                                (0.007)
## 
## factor(AGE_12)11               0.114***
##                                (0.008)
## 
## factor(AGE_12)12               0.061***
##                                (0.010)
## 
## factor(MARSTAT)2                -0.004
```

```
##                                      (0.003)
##
## factor(MARSTAT)3                    -0.033***
##                                      (0.010)
##
## factor(MARSTAT)4                    -0.021***
##                                      (0.006)
##
## factor(MARSTAT)5                     -0.012**
##                                      (0.005)
##
## factor(MARSTAT)6                    -0.063***
##                                      (0.003)
##
## factor(EDUC)1                        0.048***
##                                      (0.010)
##
## factor(EDUC)2                        0.084***
##                                      (0.009)
##
## factor(EDUC)3                        0.103***
##                                      (0.010)
##
## factor(EDUC)4                        0.141***
##                                      (0.009)
##
## factor(EDUC)5                        0.253***
##                                      (0.010)
##
## factor(EDUC)6                        0.326***
##                                      (0.010)
##
## factor(COWMAIN)2                    -0.174***
##                                      (0.004)
##
## factor(NAICS_21)2                    0.239***
##                                      (0.021)
##
## factor(NAICS_21)3                     0.0003
##                                      (0.034)
##
## factor(NAICS_21)4                    0.561***
##                                      (0.013)
##
## factor(NAICS_21)5                    0.442***
##                                      (0.017)
##
## factor(NAICS_21)6                    0.307***
##                                      (0.014)
##
```

```
## factor(NAICS_21)7                0.291***
##                                  (0.014)
##
## factor(NAICS_21)8                0.220***
##                                  (0.015)
##
## factor(NAICS_21)9                0.250***
##                                  (0.015)
##
## factor(NAICS_21)10               0.063***
##                                  (0.014)
##
## factor(NAICS_21)11               0.157***
##                                  (0.014)
##
## factor(NAICS_21)12               0.350***
##                                  (0.015)
##
## factor(NAICS_21)13               0.158***
##                                  (0.016)
##
## factor(NAICS_21)14               0.292***
##                                  (0.014)
##
## factor(NAICS_21)15               0.088***
##                                  (0.015)
##
## factor(NAICS_21)16               0.134***
##                                  (0.015)
##
## factor(NAICS_21)17               0.061***
##                                  (0.014)
##
## factor(NAICS_21)18               0.163***
##                                  (0.015)
##
## factor(NAICS_21)19               0.051***
##                                  (0.014)
##
## factor(NAICS_21)20               0.129***
##                                  (0.015)
##
## factor(NAICS_21)21               0.236***
##                                  (0.015)
##
## factor(NOC_10)2                 -0.402***
##                                  (0.005)
##
## factor(NOC_10)3                 -0.202***
##                                  (0.005)
```

```
##
## factor(NOC_10)4                -0.193***
##                                 (0.006)
##
## factor(NOC_10)5                -0.291***
##                                 (0.005)
##
## factor(NOC_10)6                -0.418***
##                                 (0.009)
##
## factor(NOC_10)7                -0.556***
##                                 (0.005)
##
## factor(NOC_10)8                -0.391***
##                                 (0.005)
##
## factor(NOC_10)9                -0.416***
##                                 (0.011)
##
## factor(NOC_10)10               -0.524***
##                                 (0.007)
##
## AHRSMAIN                        0.002***
##                                 (0.0001)
##
## covid                           0.102***
##                                 (0.003)
##
## sex:covid                       0.011***
##                                 (0.004)
##
## Constant                        3.242***
##                                 (0.018)
##
## --------------------------------------------------
## Observations                    105,254
## R2                               0.513
## Adjusted R2                      0.513
## Residual Std. Error      0.325 (df = 105197)
## F Statistic         1,982.102*** (df = 56; 105197)
## ==================================================
## Note:                  *p<0.1; **p<0.05; ***p<0.01

# sensitivity test to see if weekly wage differences are different in the
rSensSexWeeklyWage <- lm(lweeklywage ~ sex + factor(AGE_12) + factor(MARSTAT)
                         +factor(EDUC) + factor(COWMAIN) + factor(NAICS_21) +
factor(NOC_10) + AHRSMAIN + sex:covid + covid, data=combinedData)

stargazer(rSensSexWeeklyWage, type="text")
```

```
##
## ===================================================
##                          Dependent variable:
##                     -----------------------------
##                              lweeklywage
## ---------------------------------------------------
## sex                           -0.140***
##                                (0.005)
##
## factor(AGE_12)2                0.353***
##                                (0.009)
##
## factor(AGE_12)3                0.612***
##                                (0.009)
##
## factor(AGE_12)4                0.702***
##                                (0.009)
##
## factor(AGE_12)5                0.722***
##                                (0.009)
##
## factor(AGE_12)6                0.726***
##                                (0.009)
##
## factor(AGE_12)7                0.738***
##                                (0.010)
##
## factor(AGE_12)8                0.739***
##                                (0.010)
##
## factor(AGE_12)9                0.709***
##                                (0.010)
##
## factor(AGE_12)10               0.638***
##                                (0.010)
##
## factor(AGE_12)11               0.435***
##                                (0.012)
##
## factor(AGE_12)12               0.251***
##                                (0.015)
##
## factor(MARSTAT)2               0.016***
##                                (0.004)
##
## factor(MARSTAT)3              -0.044***
##                                (0.015)
##
## factor(MARSTAT)4                -0.011
##                                (0.010)
```

```
##
## factor(MARSTAT)5                   0.004
##                                   (0.008)
##
## factor(MARSTAT)6                 -0.082***
##                                   (0.004)
##
## factor(EDUC)1                      -0.016
##                                   (0.015)
##
## factor(EDUC)2                     0.090***
##                                   (0.014)
##
## factor(EDUC)3                     0.046***
##                                   (0.015)
##
## factor(EDUC)4                     0.143***
##                                   (0.014)
##
## factor(EDUC)5                     0.255***
##                                   (0.014)
##
## factor(EDUC)6                     0.328***
##                                   (0.015)
##
## factor(COWMAIN)2                 -0.201***
##                                   (0.006)
##
## factor(NAICS_21)2                 0.302***
##                                   (0.031)
##
## factor(NAICS_21)3                 0.140***
##                                   (0.049)
##
## factor(NAICS_21)4                 0.641***
##                                   (0.020)
##
## factor(NAICS_21)5                 0.422***
##                                   (0.026)
##
## factor(NAICS_21)6                 0.344***
##                                   (0.021)
##
## factor(NAICS_21)7                 0.327***
##                                   (0.021)
##
## factor(NAICS_21)8                 0.249***
##                                   (0.022)
##
## factor(NAICS_21)9                 0.303***
```

```
##                                   (0.022)
##
## factor(NAICS_21)10                0.054***
##                                   (0.021)
##
## factor(NAICS_21)11                0.173***
##                                   (0.021)
##
## factor(NAICS_21)12                0.398***
##                                   (0.021)
##
## factor(NAICS_21)13                0.165***
##                                   (0.024)
##
## factor(NAICS_21)14                0.308***
##                                   (0.021)
##
## factor(NAICS_21)15                0.109***
##                                   (0.021)
##
## factor(NAICS_21)16                 0.014
##                                   (0.022)
##
## factor(NAICS_21)17                0.045**
##                                   (0.021)
##
## factor(NAICS_21)18                0.100***
##                                   (0.022)
##
## factor(NAICS_21)19                 0.016
##                                   (0.021)
##
## factor(NAICS_21)20                0.084***
##                                   (0.022)
##
## factor(NAICS_21)21                0.244***
##                                   (0.022)
##
## factor(NOC_10)2                  -0.435***
##                                   (0.007)
##
## factor(NOC_10)3                  -0.206***
##                                   (0.008)
##
## factor(NOC_10)4                  -0.225***
##                                   (0.009)
##
## factor(NOC_10)5                  -0.336***
##                                   (0.008)
##
```

```
## factor(NOC_10)6               -0.623***
##                                (0.013)
##
## factor(NOC_10)7               -0.628***
##                                (0.007)
##
## factor(NOC_10)8               -0.367***
##                                (0.008)
##
## factor(NOC_10)9               -0.380***
##                                (0.016)
##
## factor(NOC_10)10              -0.504***
##                                (0.010)
##
## AHRSMAIN                       0.015***
##                                (0.0001)
##
## covid                          0.112***
##                                (0.004)
##
## sex:covid                       0.011*
##                                (0.006)
##
## Constant                       5.901***
##                                (0.027)
##
## ---------------------------------------------------
## Observations                   105,254
## R2                              0.583
## Adjusted R2                     0.583
## Residual Std. Error      0.480 (df = 105197)
## F Statistic         2,629.013*** (df = 56; 105197)
## ===================================================
## Note:                   *p<0.1; **p<0.05; ***p<0.01
```

```
# sensitivity test for 2022 data set, taking out industry
rSensIndustry2019 <- lm(lwage ~ sex + factor(AGE_12) + factor(MARSTAT)
                        +factor(EDUC) + factor(COWMAIN) + factor(NOC_10) + AH
RSMAIN, data=dec2019)

rSensIndustry2022 <- lm(lwage ~ sex + factor(AGE_12) + factor(MARSTAT)
                    +factor(EDUC) + factor(COWMAIN) + factor(NOC_10) + AHRSMA
IN, data=dec2022)

stargazer(rSensIndustry2022, type="text")
```

```
##
## ===================================================
##                             Dependent variable:
```

```
##                      -----------------------------
##                                  lwage
## ---------------------------------------------------
## sex                            -0.116***
##                                 (0.003)
##
## factor(AGE_12)2                 0.0003
##                                 (0.009)
##
## factor(AGE_12)3                0.123***
##                                 (0.009)
##
## factor(AGE_12)4                0.188***
##                                 (0.009)
##
## factor(AGE_12)5                0.230***
##                                 (0.009)
##
## factor(AGE_12)6                0.233***
##                                 (0.009)
##
## factor(AGE_12)7                0.238***
##                                 (0.009)
##
## factor(AGE_12)8                0.240***
##                                 (0.009)
##
## factor(AGE_12)9                0.225***
##                                 (0.009)
##
## factor(AGE_12)10               0.191***
##                                 (0.009)
##
## factor(AGE_12)11               0.124***
##                                 (0.011)
##
## factor(AGE_12)12               0.078***
##                                 (0.014)
##
## factor(MARSTAT)2                 0.001
##                                 (0.004)
##
## factor(MARSTAT)3               -0.036**
##                                 (0.014)
##
## factor(MARSTAT)4               -0.038***
##                                 (0.009)
##
## factor(MARSTAT)5                -0.011
##                                 (0.007)
```

```
##
## factor(MARSTAT)6              -0.069***
##                                (0.004)
##
## factor(EDUC)1                 0.059***
##                                (0.015)
##
## factor(EDUC)2                 0.095***
##                                (0.014)
##
## factor(EDUC)3                 0.120***
##                                (0.015)
##
## factor(EDUC)4                 0.165***
##                                (0.014)
##
## factor(EDUC)5                 0.278***
##                                (0.014)
##
## factor(EDUC)6                 0.346***
##                                (0.014)
##
## factor(COWMAIN)2              -0.147***
##                                (0.004)
##
## factor(NOC_10)2               -0.402***
##                                (0.006)
##
## factor(NOC_10)3               -0.165***
##                                (0.007)
##
## factor(NOC_10)4               -0.331***
##                                (0.007)
##
## factor(NOC_10)5               -0.348***
##                                (0.007)
##
## factor(NOC_10)6               -0.454***
##                                (0.012)
##
## factor(NOC_10)7               -0.653***
##                                (0.006)
##
## factor(NOC_10)8               -0.390***
##                                (0.007)
##
## factor(NOC_10)9               -0.399***
##                                (0.012)
##
## factor(NOC_10)10              -0.498***
```

```
##                                (0.009)
##
## AHRSMAIN                      0.002***
##                               (0.0001)
##
## Constant                      3.501***
##                                (0.017)
##
## ---------------------------------------------------
## Observations                   55,752
## R2                              0.464
## Adjusted R2                     0.464
## Residual Std. Error       0.335 (df = 55717)
## F Statistic         1,420.064*** (df = 34; 55717)
## ===================================================
## Note:                   *p<0.1; **p<0.05; ***p<0.01
```

```
stargazer(rSensIndustry2019, type="text")
```

```
##
## ===================================================
##                          Dependent variable:
##                     -------------------------------
##                                lwage
## ---------------------------------------------------
## sex                           -0.126***
##                                (0.004)
##
## factor(AGE_12)2                0.020**
##                                (0.009)
##
## factor(AGE_12)3               0.112***
##                                (0.009)
##
## factor(AGE_12)4               0.191***
##                                (0.009)
##
## factor(AGE_12)5               0.218***
##                                (0.010)
##
## factor(AGE_12)6               0.244***
##                                (0.010)
##
## factor(AGE_12)7               0.251***
##                                (0.010)
##
## factor(AGE_12)8               0.251***
##                                (0.010)
##
## factor(AGE_12)9               0.220***
```

```
##                                    (0.010)
##
## factor(AGE_12)10                  0.192***
##                                    (0.010)
##
## factor(AGE_12)11                  0.126***
##                                    (0.013)
##
## factor(AGE_12)12                   0.040**
##                                    (0.016)
##
## factor(MARSTAT)2                  -0.011**
##                                    (0.005)
##
## factor(MARSTAT)3                  -0.036**
##                                    (0.015)
##
## factor(MARSTAT)4                   -0.015
##                                    (0.010)
##
## factor(MARSTAT)5                  -0.021***
##                                    (0.008)
##
## factor(MARSTAT)6                  -0.079***
##                                    (0.005)
##
## factor(EDUC)1                     0.061***
##                                    (0.015)
##
## factor(EDUC)2                     0.110***
##                                    (0.014)
##
## factor(EDUC)3                     0.128***
##                                    (0.015)
##
## factor(EDUC)4                     0.181***
##                                    (0.014)
##
## factor(EDUC)5                     0.307***
##                                    (0.014)
##
## factor(EDUC)6                     0.385***
##                                    (0.015)
##
## factor(COWMAIN)2                  -0.183***
##                                    (0.004)
##
## factor(NOC_10)2                   -0.362***
##                                    (0.008)
##
```

```
## factor(NOC_10)3              -0.129***
##                              (0.009)
##
## factor(NOC_10)4              -0.285***
##                              (0.009)
##
## factor(NOC_10)5              -0.333***
##                              (0.008)
##
## factor(NOC_10)6              -0.391***
##                              (0.014)
##
## factor(NOC_10)7              -0.600***
##                              (0.008)
##
## factor(NOC_10)8              -0.320***
##                              (0.008)
##
## factor(NOC_10)9              -0.329***
##                              (0.012)
##
## factor(NOC_10)10             -0.410***
##                              (0.010)
##
## AHRSMAIN                     0.002***
##                              (0.0001)
##
## Constant                     3.345***
##                              (0.018)
##
## ---------------------------------------------------
## Observations                  49,502
## R2                             0.455
## Adjusted R2                    0.454
## Residual Std. Error      0.342 (df = 49467)
## F Statistic         1,213.367*** (df = 34; 49467)
## ===================================================
## Note:                  *p<0.1; **p<0.05; ***p<0.01
```