American Unionization: A Regression Analysis of Union Membership on Wages

Connor Alcorn & Savannah Halvorson

Dr. Nathanael D. Peach

Fort Lewis College

American Unionization: A Regression Analysis of Union Membership on Wages

Income inequality and the distribution of wealth is a significant issue in contemporary American society. While income and wealth can be generated in various ways, from real estate to owning stocks, for most Americans, their wages are their primary source of income. Since 1980, income inequality in the U.S. has increased drastically, with the Gini Coefficient (a common measure of income inequality) increasing from 35.2 in 1980 to 41.5 in 2019 (World Bank, 2023). A Gini index of 100 would indicate that income distribution is perfectly unequal, or one individual receiving 100% of all income in a given region, which means that the United States has been moving closer to perfect wealth inequality. Wage inequality increased in this same time frame, with wages for the top 1% of earners jumping 179.3% and 389.1% for the top 0.1% of earners between 1979 and 2020. Wages of the bottom 90% of earners only saw a 28.2% increase (Mishel & Kandra, 2021). Less dynamic wage growth at the bottom of the income scale is often attributed to a few factors, including the decline in collective bargaining power and weaker labor standards (Mishel and Bivens, 2021).

Historically, unions have had large impacts on leveling the wage gap and increasing the welfare of the working class. The passing of the National Labor Relations Act (NLRA) and the creation of the National Labor Relations Board (NLRB) in 1935 gave federal protection to the organization of private-sector laborers. From 1934 to 1939, union membership of the nonfarm workforce more than doubled, rising from 11% to 28% and continued to increase to peak levels in the 1950s of around one-third of the workforce (Farber et al., 2018). During this high period of unionization, the income share of the top 1% of earners was declining and reached its lowest point, supporting the theory that unionization has beneficial impacts on income inequality historically. Income inequality was low, and the share of the U.S. population graduating high school and college was also increasing—the percentage of children entering the workforce with higher expected earnings than their parents was increasing and was at a higher level than it is today (Raj et al., 2017). Even with this trend, unionization began to decrease, and the income share of the top 1% began to rise to the levels that we see today. The question this study confronts is that of this decrease in unionization: do unions still provide the benefit to wage expectation that was seen on the tails of the NRLA, and are they a tool that can be used to lower current levels of income inequality?

The theoretical impact of unions on labor markets can be presented in two ways: monopoly and voice. The first of these is the "monopoly" power of the laborers, which concerns the strike power afforded to laborers and members of unions (U.S. Department of the Treasury, 2023). By withholding work, or striking, they can apply pressure on the employer to increase their wages by halting production and lowering revenues for the company. Unions monopolize work, as when the laborers organize, they become the sole source of labor that the company can employ—giving credibility to the threat of striking. In a non-unionized environment, employees could withhold their labor, but the firm would simply hire another laborer to replace them. This power, and its effectiveness on raising wages, depends on the industry structure the firm operates in. Perfectly competitive employers (or firms closer to perfectly competitive structure) may not be able to raise the wages should the laborers demand it, as the cost of the product would increase, and the firm would no longer be able to compete. In a perfectly competitive industry, a union would need to span the entire industry and all laborers within it to achieve the desired monopoly power and have the possibility of increasing wages through the threat of striking. The power of unions is not exclusively limited to their ability to strike; communication between the laborers and the firm is also a key power afforded to unionized laborers.

The "voice" power refers to the voice that the laborers are given concerning production (U.S. Department of the Treasury). The work of the laborers is directly concerned with production: they have the unique individual knowledge of the process of production and may provide suggestions to the firm on how to increase productivity. The “voice” power is also concerned with aspects such as workplace safety or environment, which the employee may suggest freely to the employer without fear of retaliation—the union protects the employee from being fired for such suggestions. The increase in these changes concerning the laborers will increase worker welfare, through avenues such as greater safety protections. The egalitarian nature of unions also addresses issues in the workplace such as discrimination, as wages no longer are worker specific. As such, unions may increase worker welfare through both channels, in the form of direct monetary benefits and indirect monetary benefits.

Unions also provide social structures for their members. The benefit of this social dimension extends into the broader community. Feigenbaum, Hertel-Fernandez, and Williamson (2018) describe the impacts that union membership has on the political process, specifically through the presence of right-to-work laws. Weakening unions through these right-to-work laws results in lower voter turnout, lower organized labor contributions, and fewer working-class representatives in the legislature. The increased civic engagement from the presence of unions has broad influences on policy, furthering worker welfare across classes and the benefits of unionization.

This study investigates if bargaining power, through unionization, has an impact on increasing wages and if union membership is therefore a means to decrease income inequality. It further discusses the data used to analyze union membership and the issues that present themselves regarding wage analysis and union data collection. Multiple models are constructed to explore various ways in which union membership may impact wages. Additionally, interaction terms (multiplying terms together within the model to isolate a possible relationship) are utilized to determine if the impact of unions is different across industries or between males and females. The models do provide significant evidence that union membership impacts wages, especially within the blue-collar industry.

Data Overview

The data analyzed is a cross-sectional data set from the U.S. Census Bureau and the U.S. Bureau of Labor Statistics' Current Population Survey (CPS), Annual Social and Economic Supplement from March 2022. This allows data to be pulled from the census, in which each individual (observation) has associated characteristics that facilitate an analysis of the impact of certain classifications on other reported data, such as wages or earnings. The data includes characteristics specifically relevant to union membership, wages, and the labor market. Union membership is tracked by the dummy variable Union_Member, where “Yes” is coded as (=1), and “No” is coded as (=0). The dependent variable is observed through the variable WSAL_VAL, which is an individual's total wage and salary earnings measured yearly.

Other variables relevant to wages include age of the observation, reflected in the data set as A_AGE and sq_A_AGE, the latter being age value squared. Sex is reflected through the dummy variable Male, where an observation is coded as (=1) if they identified as a male in the census. Female is coded similarly but is omitted in each model to prevent perfect multicollinearity.

Occupation is reflected through 10 different industries of occupation, reduced to three designations which may present a causal relationship with union membership – OCC_WhiteCollar, OCC_BlueCollar, and the service industry. Armed forces constituted the 11th

occupational industry designation in the census but was omitted in the sample restriction as the armed forces falls well outside the traditional labor force. OCC_WhiteCollar contains the following occupational designations of the Census: management, business, and financial occupations; professional and related occupations; sales and related occupations; and office and administrative support occupations. OCC_BlueCollar contains the following occupational designations of the Census: farming, fishing, and forestry occupations; construction and extraction occupations; installation, maintenance, and repair occupations; production occupations; and transportation and material moving occupations. The service industry contains the following: healthcare support; protective service; food service; building and grounds maintenance; and personal care occupations (occupational codes 3601 through 4655). The service industry occupational designations were contained within its own occupational variable and serves as the omitted variable to avoid perfect multicollinearity.

Education was broken down into micro-levels of educational attainment of each observation in the Census, which was condensed into a collection of dummy variables in this study. “Contains” for the sake of simplicity means that an individual is coded (=1), and coded (=0) if not within the given designation. HS_Grad contains all observations that graduated high school, Some_College contains all observations that attended some years of college but did not attain a degree. College_Grad contains all observations with a college degree, including associate degrees, vocational/occupational degrees, and bachelor's degrees. Masters_Deg contains all observations with a master's degree, and Prof_Doc_Deg contains all observations with a professional degree or a doctoral degree.

Location was accounted for by individual state in the Census and aggregated into the four census regions for this study. Reg_1 (Region 1, Northeast) contains the following states: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania. Reg_2 (Region 2, Midwest) contains the following states: Indiana, Illinois, Michigan, Ohio, Wisconsin, Iowa, Nebraska, Kansas, North Dakota, Minnesota, South Dakota, and Missouri. Reg_3 (Region 3, South) contains the District of Columbia (Washington DC) and the following states: Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, and Texas. Reg_4 (Region 4, West) contains the following states: Arizona, Colorado, Idaho, New Mexico, Montana, Utah, Nevada, Wyoming, Alaska, California, Hawaii, Oregon, and Washington. Region 1 is omitted in the analysis to avoid perfect multicollinearity. It serves as the comparison region.

The data selected and the observations taken from the CPS have been limited for relevancy to individuals who are subject to unionization or within a demographic that may be affected by unions. Observations under the age of 18 have been excluded from the data set, as they are not in the labor force full-time or would be considered an outlier if they are in the labor force full-time. This limits the survey of individuals to ages 18 to 85. The education variable also includes a classification for children, and those also have been excluded, although this likely overlapped with the age limitation considerably and had the same effect. All other grade levels for possible educational attainment range from first grade to a doctoral degree, lumped into dummy variables stated previously to see the effect of general levels of education. Wage and salary observations with a value of zero have also been excluded, as those observations are not employed and not subject to the influence or effect of a labor union; any observation coded as “not in universe” regarding industry of occupation was excluded for the same reason.

Table 1

Summary statistics of each variable in the data set

Variable	Average	Median	Minimum	Maximum	Standard Deviation
WSAL_VAL	65,088	48,000	12	1,158,000	78,859
A_AGE	43.06	42	18	85	14.096
Male	0.5105	1	0	1	0.4999
Female	0.4895	0	0	1	0.4999
Union_Member	0.0992	0	0	1	0.2990
OCC_WhiteCollar	0.6406	1	0	1	0.4798
OCC_BlueCollar	0.2170	0	0	1	0.4122
Service Industry	0.1424	0	0	1	0.3494
No HS diploma	0.0545	0	0	1	0.2270
HS_Grad	0.2664	0	0	1	0.4421
Some_College	0.1509	0	0	1	0.3580
College_Grad	0.3736	0	0	1	0.4838
Masters_Deg	0.1151	0	0	1	0.3191
Prof_Doc_Deg	0.0395	0	0	1	0.1947
Reg_1	0.1550	0	0	1	0.3619
Reg_2	0.1987	0	0	1	0.3990
Reg_3	0.3698	0	0	1	0.4828
Reg_4	0.2766	0	0	1	0.4473

Sample size = 10,312

Looking at the distribution of data points among the sample, there are issues that present themselves for a regression analysis and the ability to isolate causation. The most significant flaw to this study is the number of individuals that are members of labor unions. From the CPS, there are three responses to the question if an individual is a member of a union – “yes”, “no”, and “not in universe”. The latter response could encapsulate several meanings for each different individual, and for the sake of this study, was excluded. However, in the raw microdata sample taken from the CPS of 114,158 observations, 90.74% (103,587) of individuals responded, “not in universe”. Of the remaining individuals, 8.35% (9,532) responded “no”, leaving only 0.9066% (1,035) of individuals responding “yes”. After the exclusion of “not in universe” for the purpose of analysis, as well as the other restrictions for relevance to the labor market described above, the sample size dropped to 10,312 – 9.92% (1,023) of individuals in the restricted sample were members of a union, which is more in line with U.S. union membership at 10.1% in 2022 (U.S. Bureau of Labor Statistics, 2023). The sample size and number of union members does not directly present an issue to the regression model, but it does provide issues with the necessary assumption of randomness in the sample. Lack of unionization is not surprising, with the current downtrend of union membership.

An underlying issue with the data collection and reporting methods also causes issues with the dependent variable, wage and salary. To keep confidentiality, the Census Bureau topcodes the highest values from each source of income which “leads to artificial increases or decreases in earnings...at the top of the earnings distribution as different fractions of the

population are subject to topcoding each year" (Burkhauser & Larrimore, 2009). Topcoding specifically presents difficulties to measuring income gaps and earnings ratios, such as those between a union member and a non-union member, which may create an error that is not accounted for in the model. The other issue with wage and salary data is the number of observations with a wage and salary of zero. For this study, these observations are treated as unemployed and excluded. However, if it is assumed that these observations are unemployed, the data would present an unemployment rate of 38.7% – an unemployment rate over 10 times higher than the national rate, assuming 3.6% was the national unemployment rate in March 2022 (U.S. Bureau of Labor Statistics, 2023). It is worth noting that the national unemployment rate and the data in the study are both pulled from the CPS, so there is significant error in wage and salary reporting that is not accounted for in this study.

Methods

From education to on-the-job experience, wages and earnings are influenced by many factors. In gathering potential explanatory variables, economic theory and data availability guided which of these factors are accounted for. Education is controlled for via the previously discussed dummy variable corresponding to the individual's highest level of educational attainment. Age is also included as a proxy for experience—age being correlated with the amount of experience that one has in each given field or industry, as well as general experience and knowledge that may improve their earnings. Industry, and occupation specifically, is another large factor in earnings as individuals in certain occupations earn more than others—which also provides the opportunity to see if there is a larger effect of union membership within specific occupations. Specific occupations and industries (such as trades) may be more prone to being unionized, for historical reasons or otherwise, and could be the reason those individuals have higher earnings than in a separate occupation. Sex may be correlated with this as well, as it could be possible that more men work within an occupation which is more prone to organization or vice versa. Including sex also allows to control for possible discrimination, too.

Figure 1

Mincer earnings function with age proxy

$$\ln(w) = ß_0 + ß_1\text{School} + ß_2\text{Age} + ß_3\text{Age}^2$$

Wage estimation is a large body of knowledge, with widely used models that attempt to tease out causality in a multitude of factors that may impact wages—union membership being the focus in this study. The model in this study builds on the standard Mincer earnings function, first used by Mincer in 1958 and widely cited in empirical economics, in which the model explains earnings and income as a function of schooling and experience. The typical expression of this model is explaining the log of wages as the sum of years of education and the quadratic function of years of potential experience (figure 1). Age is used as a proxy for potential experience in this model. Education ("School" in figure 1) is broken down into individual dummy variables representing the level of educational attainment.

The logarithm of the dependent variable, wage and salary, is used to account for the possibility that the explanatory variables have a nonlinear relationship with earnings. For many individuals, wages and salary increase more during their early working years than when they near retirement. From this baseline, relevant variables are included to capture the effect of union

membership and regional variations in wage and salary, including but not limited to the cost of living or the prominence of certain industries in a designated region.

Figure 2

Estimated earnings function of the model

$$\ln(w) = ß_0 + ß_1\text{School} + ß_2\text{Age} + ß_3\text{Age}^2 + ß_4\text{Male} + ß_5\text{Union\_Member} + ß_6\text{WhiteCollar} + ß_7\text{BlueCollar} + ß_8\text{HS\_Grad} + ß_9\text{Some\_College} + ß_{10}\text{College\_Grad} + ß_{11}\text{Masters\_Deg} + ß_{12}\text{Prof\_Doc\_Deg} + ß_{13}\text{Reg2} + ß_{14}\text{Reg3} + ß_{15}\text{Reg4} + ß_{16}\text{Interaction}$$

Where Interaction represents Union_Member multiplied by BlueCollar in Table 3 and Union_Member multiplied by Male in Table 4.

The remaining variables are chosen to isolate the effect of union membership of wages. Male is included both as a control for any gender discrepancies in wages and for possible correlation with union membership. Males may be more prone to joining unions, and this possibility is isolated in Table 4 with an interaction term. Occupational industry faces these same possibilities and is the primary reason why industries were aggregated into “blue collar” and “white collar”; blue collar laborers may be more inclined to join a union due to the nature of their industry and the historical presence of unionization—the American labor movement began in factories and mines, and that sentiment may still be present today. This relationship is further investigated through an interaction term in Table 3.

Education may correlate with union membership as well. It may be true that a specific level of educational attainment is more prone to unionization, and this relationship becomes more intuitive when looking at the occupational industry of individuals with certain levels of education. Individuals who graduate high school, but do not attend college, may be more prone to working in the “blue collar” sector, and in turn be more prone to union membership. The opposite may be true for those with college and advanced degrees. Adding these variables to the baseline Mincer earnings function yields the wage estimation function in Figure 2.

Results & Interpretations

The first model, displayed in Table 2, provides the baseline function estimated. Because the dependent variable is expressed as a natural logarithm, the estimated coefficients can be interpreted as a percent change in earnings, holding all other factors constant.

With the exception of Reg_4, all the explanatory variables are found to be statistically significant. Looking comparatively at the statistical significance of all the regional variables, the model indicates that region is not indicative of income. The U.S. Census regions are extremely large areas, and region on its face may not be the predictor of income. It may be beneficial to investigate the relationship between union membership and location to more specific areas, such as rural and urban areas. Region 1 also includes the state of New York, of which New York City is considered an outlier in a majority of empirical studies and may provide error within those variables. As region was a control variable, statistical and economic significance of those variables is not relevant to the scope of this study.

Other variables, all of which are statistically significant, tend to fit within the bounds of social expectations regarding influential factors of income when looking at the economic significance of the coefficient. Age provides an interesting relationship, as it is reflected within two coefficients —both as a linear and a quadratic function. Earnings are reflected to increase 9.76% with each additional year, but this is subject to diminishing returns as indicated by the

negative coefficient on age squared. Industry also follows expected wage patterns: blue collar laborers have a higher income than the same worker in the service industry (24% increase), while white collar laborers have a higher expected wage than both the blue collar and the service sector (49.53% increase compared to the same worker in the service industry).

The education coefficients present a similar story. Wage expectation increases as one moves up the levels of educational attainment, although not as much as may be expected. The model provides that an individual with a college degree will expect a 59.29% increase in wage compared to an individual with no high-school degree, while an individual with a high school degree will earn 25.96% more than one with no high-school degree, holding all other attributes constant. These coefficients do not fit typical societal expectations of the return of a college degree with the estimated increase in earnings for college graduates being more than double that of the expected increase from a high school degree. College_Grad contains individuals with associate degrees and vocational degrees, which may introduce bias from the skill of these laborers as many skilled trade occupations may fall under these levels of attainment. College_Grad may also provide conflicting correlation with union membership, as vocational degrees may lend themselves to blue collar and traditionally unionized industry, while bachelor's degrees lend themselves to white collar and traditionally less unionized industry—all while contained within the same variable.

The male dummy variable also provides questionable results with the magnitude of the coefficient. Should there be two similar laborers, one being male and one not, the male worker would have more than a 39% higher wage expectation. The gender wage gap currently is around 18% (Aragão, 2023) and that gap is reflected at almost twice as high in the model. This discrepancy could be reflected in the issues of topcoding discussed earlier, or indicative of an omitted variable positively correlated with the dependent variable (wages) that is reflected within the male coefficient. It may also indicate the relationship between males and union membership rates, which is further investigated in Table 4.

Union membership provides promising results, with a similar worker who is a union member receiving an estimated 14.8% on average higher wage than a worker who is not. This is a significant increase in wages, especially when accounting for benefits outside of direct compensation that are provided by membership. With a 99% degree of confidence, the model places the impact of unionization between .0823 and .2137 (an 8.23% to 21.37% increase in wage and salary). This allows for the rejection of the hypothesis that union membership has no impact on wage expectation and allows for the determination that union membership indeed increases wages.

While this may be promising, union membership does not provide the same wage increase as moving industries, from blue collar to white collar. According to the model, it would be in an individual's best interest to leave a blue-collar occupation in favor of a white-collar occupation —possibly due to the labor demand for white-collar work, or that these industries have higher barriers of entry assuming many white-collar occupations require a higher degree of educational attainment compared to entry into the blue-collar workforce.

Table 2

OLS estimation of the log of wage and salary

Dependent variable (l_WSAL_VAL)

Variable	Coefficient	Standard Error	T-Ratio	P-Value
const	7.4889	0.0983	76.22	<0.001
A_AGE	0.0977	0.0041	23.71	<0.001
Sq_A_AGE	-0.001	0.0000	-21.24	<0.001
Male	0.3931	0.0174	22.59	<0.001
Union_Member	0.1480	0.0255	5.807	<0.001
OCC_WhiteCollar	0.4953	0.0267	18.57	<0.001
OCC_BlueCollar	0.2401	0.0299	8.026	<0.001
HS_Grad	0.2596	0.0388	6.686	<0.001
Some_College	0.3016	0.0426	7.075	<0.001
College_Grad	0.5929	0.0396	14.98	<0.001
Masters_Deg	0.8283	0.0460	18.02	<0.001
Prof_Doc_Deg	1.1939	0.0572	20.86	<0.001
Reg_2	-0.0862	0.0280	-3.075	0.002
Reg_3	-0.0738	0.0252	-2.932	0.003
Reg_4	-0.0185	0.0264	-0.699	0.484

S.E.R.=0.959; Adjusted R^2=0.282; P-Value of F-Statistic=<0.001

Model Two: Industry and Union Membership

Industries included in the blue-collar dummy variable are the most unionized workforces in the private sector, with utilities (19.6%), motion pictures and sound recording industries (17.3%), and transportation and warehousing (14.5%). White-collar industries are the least, with insurance (1.2%), finance (1.3%), professional and technical services (1.3%), and food services and drinking places (1.4%) (U.S. Bureau of Labor Statistics, 2023). Food services and drinking places fall under the service industry category, which further solidifies the assumption that most unionized individuals are blue-collar laborers. This relationship is investigated further in Table 3, to see if the impact of union membership is unique within blue-collar industries. To do so an interaction term, Blue×Union, where OCC_BlueCollar is multiplied by Union_Member, is utilized.

Introducing the interaction term between union membership and blue-collar designated industries has a dramatic effect on the individual coefficients of union membership and blue-collar industry, supporting the claim that unionization has larger impacts on the industries within the variable and is relatively isolated. The coefficient of OCC_BlueCollar drops from .24 to .20, about a 4% decrease in earnings for an individual in one of those industries. However, if that individual becomes a member of a union, their expected wage increases by 31%. Isolating union membership within blue-collar industries magnified the impact of unions, bringing the effect of union membership in said industries to leveling or even exceeding the impact with a white-collar job (compared to the service industry). The magnitude of union impact in this model is the sum of the coefficient on Union_Member, OCC_BlueCollar, and Blue×Union, or around a 57% effect

on wage and salary. White-collar occupation increases wages by 49.32%, meaning that unionization within a blue-collar industry may bridge the wage gap between these classes — although a significant change in the demographics of an individual makes predicting the overall impact on wages challenging.

Table 3

OLS estimation of the log of wage and salary including industry interaction term

Dependent variable (l_WSAL_VAL)

Variable	Coefficient	Standard Error	T-Ratio	P-Value
const	7.4872	0.0983	76.30	<0.001
A_AGE	0.0977	0.0041	23.72	<0.001
Sq_A_AGE	−0.001	<0.0001	−21.24	<0.001
Male	0.3909	0.0174	22.49	<0.001
Union_Member	0.0663	0.0307	2.16	0.031
OCC_WhiteCollar	0.4932	0.0267	18.46	<0.001
OCC_BlueCollar	0.2043	0.0308	6.63	<0.001
HS_Grad	0.2549	0.0390	6.54	<0.001
Some_College	0.2928	0.0428	6.84	<0.001
College_Grad	0.5883	0.0397	14.80	<0.001
Masters_Deg	0.8296	0.0461	18.01	<0.001
Prof_Doc_Deg	1.1880	0.0573	20.73	<0.001
Reg_2	−0.0884	0.0280	−3.16	0.002
Reg_3	−0.0736	0.0252	−2.92	0.004
Reg_4	−0.0153	0.0264	−0.58	0.561
Blue×Union	0.3104	0.0524	5.93	<0.001

S.E.R.=0.812; Adjusted R^2=0.284; P-Value of F-Statistic=<0.001

Union membership is extremely beneficial to wages of individuals in blue-collar industries, as shown in the model. With a 99% degree of confidence, the model places the population impact of union membership in blue collar industries (the interaction term) between 0.1754 and 0.4454 (a 17.54% to 44.54% increase in wages). This allows for the rejection of the hypothesis that union membership within these industries has no impact. The data are again limited by the lack of unionization in white-collar designated industries. It could have similar impacts that aren't reflected as there is not significant union presence, or it could be true even if there was union presence; the nature of the occupational industries is significantly different and provides environments where a union may or may not be beneficial.

Model Three: Sex and Union Membership

To investigate the causality between wages and unions, an interaction term between males and unions was added (Table 4). A possible cause of the increased male wage expectation is that more male laborers are members of unions—although current data does not support this, male membership rates are only slightly higher (10.5%) than women (9.6%) (U.S. Bureau of Labor Statistics, 2023) —and that relationship is being captured within blue-collar industry and

union membership. This interaction term additionally is utilized to address the issue of the gender wage gap explained earlier, where the model provides a wage expectation for males at almost double the established wage gap.

Table 4

OLS estimation of the log of wage and salary including male interaction term

Dependent variable (l_WSAL_VAL)

Variable	Coefficient	Standard Error	T-Ratio	P-Value
const	7.4917	0.0982	76.28	<0.001
A_AGE	0.0976	0.0041	23.72	<0.001
Sq_A_AGE	−0.001	<0.0001	−21.24	<0.001
Male	0.3861	0.0182	21.24	<0.001
Union_Member	0.1109	0.0373	2.970	0.003
OCC_WhiteCollar	0.4978	0.0267	18.64	<0.001
OCC_BlueCollar	0.2411	0.0299	8.066	<0.001
HS_Grad	0.2591	0.0388	6.671	<0.001
Some_College	0.3002	0.0426	7.043	<0.001
College_Grad	0.5924	0.0396	14.96	<0.001
Masters_Deg	0.8296	0.0460	18.02	<0.001
Prof_Doc_Deg	1.1931	0.0573	20.84	<0.001
Reg_2	−0.0867	0.0280	−3.09	0.002
Reg_3	−0.0741	0.0252	−2.94	0.003
Reg_4	−0.0181	0.0263	−0.69	0.492
Male×Union	0.0738	0.0505	1.46	0.144

S.E.R.=0.813; Adjusted R^2=0.283; P-Value of F-Statistic=<0.001

However, neither of these possibilities are supported by the model. Adding the interaction term did influence Union_Member, dropping the coefficient from 0.148 to 0.11 (a 14.8% increase in wage compared to an 11.1% increase), while increasing the total impact of male union members to 0.183 (an 18.3% increase in wage expectation), but this change is not statistically significant at the 10% level, and not economically significant at approximately a 4% increase across all relevant identifiers. The adjusted *R*-squared drops as well, even if very slightly, but affirms the assumption that the variance is better explained by the relationship between unions and industry and is the driving causal factor of the impact of unions on wage expectation. A 90% confidence interval falls between -0.0090 and 0.1566, meaning the hypothesis that the population impact of the interaction term is zero cannot be rejected, further supporting the argument that the impact of union membership on earnings is not conditional on the individual's sex.

Each model provides evidence that unionization has drastic impacts on the wage and salary of individuals, especially within unionized blue-collar laborers. However, there are many other factors which impact earnings that have not been accounted for. The adjusted *R*-squared is highest in Table 3 at only 0.2836. This signifies that the model only accounts for about 28.36% of the variation in wages, which does not prove a strong or complete relationship between the

model and estimated wage expectation. Other variables may be needed to predict wage expectations with a higher degree of accuracy. The *F*-Statistic provides that the null hypothesis of all coefficients being 0 can be rejected for all models. The standard error of regression is similar for model two and three, but significantly larger for model 1, without interaction terms. Wage expectation becomes more accurate including the interaction terms, the most accurate when accounting for the relationship between industry and unionization (Table 3) with an S.E.R. of 0.812.

Conclusion

Building from the Mincer earnings function with variables relevant to the labor force and union membership, and accounting for the relationship between unions and industry and unions and sex, union membership is shown to have a beneficial impact on earnings for individuals, especially in blue-collar occupations. The models provide promising results into the effects of union membership on wages which opens the possibility that declining union membership plays a role in the increased income inequality the United States has experienced in recent decades. Becoming a member of a union is shown to increase wage expectation by 8.23% to 21.37%, and 17.54% to 44.54% within the blue-collar industry. Although the model does not account for every factor which influences wages and is subject to weaknesses because of the data set analyzed, it provides evidence that joining a union provides increased benefits in the form of higher wages and salaries.

Wages are difficult to predict, as there are many influential factors that cannot be accounted for. Skill and experience can be difficult to measure and analyze on a large scale. Age was used as a proxy in these models but does not fully encompass experience and neglects skill in industries where skill may be crucial to wage expectation. Measuring wages in an environment absent from unionization is also unachievable, providing additional barriers to isolating the causality of unions. Additionally, union membership is not random. It is possible that workers with the highest skill set, demand union jobs more than those with lesser skills. IF this is the case, then union membership may be serving as a proxy for skill in the estimated models. To account for possible selection bias, this study would need to be expanded in significant ways.

Future studies may address these issues of estimating the impact of unionization of laborers in a few ways. Race may be an important variable to include for both wage estimation as a whole and the relationship between different industries and unionization rates within those industries. Unions also may be impactful in reducing discrimination in the workplace, and accounting for this relationship may provide support for this argument. Furthermore, the education variables in these models included vocational degrees and bachelor's degrees in the same variable. Due to the industries those levels of education lend themselves to, it would be beneficial to investigate wage expectations and unionization rates accounting for vocational and bachelor's degrees as different variables. As mentioned, attempting to account for more skilled laborers may provide for more accurate and descriptive results. These models used age as a proxy for experience and skill, investigating the possibility of other proxies may be beneficial. Public and private industries also have significant differences when it comes to unionization—unions must operate differently in the public and private sphere. Public unions may be less effective as they are bargaining with their employers while simultaneously being significantly involved with public interest. The largest labor unions today are in public industries—isolating

this relationship may provide for more direct relationships between unionization and wage expectation.

Unions don't only provide benefits for their members: they also may provide benefits for laborers in proximity and within industries. An effective union in one firm would compete with another, ununionized firm for laborers. This competition would force the ununionized firm to provide similar benefits to employees for employee retention. Not only would the unionized firm provide competition to other firms in the industry, but it would also provide competition to firms which require laborers with similar skillsets. A laborer in construction may possess similar skills utilized in another industry, such as the auto industry, and should one of those industries have higher unionization rates and therefore higher wages, the less unionized industry would similarly have to provide increased benefits to retain those employees. Analyzing spillovers such as these, while accounting for unionization rates, may provide interesting results regarding the full impact of unionization on labor markets.

There are currently institutions designed to encourage and support unionization efforts, such as the National Labor Relations Board (NLRB), but these institutions are not as effective as they once were. The Taft-Hartley Act of 1947 significantly weakened possible union actions in the bargaining process, and more recently, the United States Supreme Court chipped away at the right to strike through *Glacier Northwest, Inc. v. International Brotherhood of Teamsters* (2023) decision, narrowing the application of the NLRA and the protection that it offers to striking laborers. Public opinion of labor unions and the sentiment of laborers towards unionization seems to be becoming more favorable. Unions have made large steps back into the public eye, such as the JFK8 warehouse at Amazon—a historically anti-union corporation—successfully unionizing. While new industries are unionizing, the established unions of the United Auto Workers (UAW), the Screen Actors Guild (SAG-AFTRA), and the Writers Guild (WGA), successfully secured increased benefits for each of their members in negotiations brought about by striking. With these union successes in the public eye, the prospect of unionization will be brought back to the attention of laborers. Should the laborers begin to unionize, this study supports the argument that this trend will combat and mitigate the growing wage gap and income inequality faced by the American working class.

References

Aragão, C. (2023). Gender pay gap in U.S. hasn't changed much in two decades. Pew Research Center. Pew Research Center.

Burkhauser, R., & Larrimore, J. (2009). Using internal CPS data to reevaluate trends in labor-earnings gaps. *Monthly Labor Review, August 2009.* U.S. Census Bureau.

Raj, C., Grusky, D., Hell, M., Hendren, N., Manduca, R., & Narang, J. (2017). The fading American dream: Trends in absolute income mobility since 1940. *Science, 356* (6336): 398–406. doi:10.1126/science.aal4617

Farber, H., Herbst, D., Kuziemko, I., Naidu, S. (2018). Unions and inequality over the twentieth century: New evidence from survey data. *NBER Working Paper Series.* National Bureau of Economic Research. http://www.nber.org/papers/w24587

Feigenbaum, J., Hertel-Fernandez, A., Williamson V. (2018). From the bargaining table to the ballot box: political effects of right to work laws. *NBER Working Paper Series.* National Bureau of Economic Research. https://doi.org/10.3386/w24259.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, *66*(4): 281–302. doi:10.1086/258055.

Mishel, L., & Bivens, J. (2021). Identifying the policy levers generating wage suppression and wage inequality. Economic Policy Institute. Economic Policy Institute.

Mishel, L., & Kandra, J. (2021). Wage inequality continued to increase in 2020. *Working Economics Blog.* Economic Policy Institute. Economic Policy Institute.

U.S. Bureau of Labor Statistics, Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/UNRATE, April 28, 2023.

U.S. Bureau of Labor Statistics. (2023). Union members – 2022. *News Release.* https://www.bls.gov/news.release/pdf/union2.pdf

U.S. Department of the Treasury. (2023). Labor unions and the middle class. United States Department of the Treasury.

World Bank. (2023). Gini index – United States. Retrieved April 25, 2023, from https://data.worldbank.org/indicator/SI.POV.GINI?locations=US.