The impact of years of education on labor force participation, salaried employment, and logarithmic hourly wages.

Aylin Kuhnavard Asli

34799

List

Question:

How do years of education impact labor force participation, salaried employment, and logarithmic hourly wages?

Introduction:

Education is a key determinant of individuals' labor market outcomes. Increasing years of education may have significant effects on labor force participation, the likelihood of salaried employment, and the level of hourly wages. This research aims to examine how years of education influence labor force participation, salaried employment, and logarithmic hourly wages. The findings of this study can assist policymakers in improving educational and labor market programs, thereby enhancing economic opportunities for individuals. By employing appropriate statistical methods, this research seeks to provide a more precise understanding of the relationship between education and labor market outcomes.

About Dataset

The dataset in question has been collected in Turkey and includes a comprehensive set of variables that provide detailed information on individuals' demographics, education, and labor market outcomes. Key variables include `referans_yil` (reference year), `cinsiyet` (gender), `dogum_yil` (year of birth), `egitim_devam` (education continuation), and `calisma` (employment status). Additionally, the dataset contains variables related to employment details such as `calisma_sahip` (salaried employment), `calisma_kucuk` (part-time work), and `gelir_gecenay` (income last month). Other important variables include `is_sureklilik` (job stability), `is_baslama_yil` (year of job start), and `istihdam_sure` (employment duration). These variables will be instrumental in analyzing the impact of education on labor force participation, salaried employment, and hourly wages in a logarithmic form. This rich dataset allows for a thorough examination of how educational attainment influences various labor market outcomes in Turkey.

Descriptive statistics

Descriptive statistics for years_of_schooling (Overall):

Variable	Obs	Mean	Std. dev.	Min	Max
years_of_s~g	273,679	8.373518	4.651062	0	17

Descriptive statistics for years_of_schooling (Men):

Variable	Obs	Mean	Std. dev.	Min	Max
years_of_s~g	132,827	9.158266	4.258034	0	17

Descriptive statistics for years_of_schooling (Women):

Variable	Obs	Mean	Std. dev.	Min	Max
years_of_s~g	140,852	7.633481	4.879154	0	17

Overall Data: With 273,679 observations, the average years of schooling is 8.37 years. This suggests that, on average, people have completed a little over 8 years of education.

Men's Data: For 132,827 men, the average years of schooling is higher at 9.16 years. This indicates that men tend to have more education than the overall population.

Women's Data: Among 140,852 women, the average is lower at 7.63 years. This shows that women typically have fewer years of schooling compared to men.

These statistics highlight gender disparities in education, with men generally having more years of schooling than women. This difference in educational attainment can impact labor force participation and wage outcomes, with those having more education potentially having better opportunities.

Descriptive statistics for hourly_wage (Overall):

Variable	Obs	Mean	Std. dev.	Min	Max
hourly_wage	100,198	31.58054	27.71296	0	1071.429

Descriptive statistics for hourly_wage (Men):

Variable	Obs	Mean	Std. dev.	Min	Max
hourly_wage	68,443	31.58733	27.80558	0	1071.429

Descriptive statistics for hourly_wage (Women):

Variable	Obs	Mean	Std. dev.	Min	Max
hourly_wage	31,755	31.56588	27.51271	0	832.5

Overall Data: For 100,198 observations, the mean hourly wage is 31.58, with a standard deviation of 27.71. The minimum hourly wage is 0, and the maximum is 1071.43.

Men's Data: Among 68,443 men, the mean hourly wage is slightly higher at 31.59, with a standard deviation of 27.81. The minimum is 0, and the maximum is also 1071.43.

Women's Data: For 31,755 women, the mean hourly wage is 31.57, with a lower maximum wage of 832.50 and a standard deviation of 27.51.

These statistics indicate a slight difference in hourly wages between men and women, with men having a marginally higher average and maximum wage. The standard deviation is fairly consistent across groups, suggesting similar variability in wages for both genders.

Descriptive statistics for real_hourly_wage (Overall):

Variable	Obs	Mean	Std. dev.	Min	Max
hourly_wage	51,465	32.79207	27.01747	0	17

Descriptive statistics for real_hourly_wage (Men):

Variable	Obs	Mean	Std. dev.	Min	Max
real_hourl~e	34,854	32.97998	27.14095	0	17

Descriptive statistics for real_hourly_wage (Women):

Variable	Obs	Mean	Std. dev.	Min	Max
real_hourl~e	16,611	32.39778	26.75306	0	17

Overall Data: The table shows 51,465 observations, with a mean real hourly wage of 32.79 and a standard deviation of 27.02. The minimum wage recorded is 0, and the maximum is 170.

Men's Data: Among 34,854 men, the mean real hourly wage is slightly higher at 32.98, with a standard deviation of 27.14. The range of wages is the same as the overall data, from 0 to 170.

Women's Data: For 16,611 women, the average real hourly wage is lower at 32.40, with a slightly smaller standard deviation of 26.75. The wage range remains consistent from 0 to 170.

These statistics show that there is a slight difference in average hourly wages between men and women, with men having marginally higher average wages. The variability in wages (as indicated by the standard deviation) is similar across all groups. The consistent range of minimum to maximum wages highlights the disparities and similarities within the different demographic groups.

Descriptive statistics for ln_real_hourly_wage (Overall):

Variable	Obs	Mean	Std. dev.	Min	Max
ln_real_ho~e	48,379	.60652	3.178699	-5.557	6.612105

Descriptive statistics for ln_real_hourly_wage (Men):

Variable	Obs	Mean	Std. dev.	Min	Max
ln_real_ho~e	32,616	3.206151	.5670653	-5.5576	6.612105

Descriptive statistics for ln_real_hourly_wage (Women):

Variable	Obs	Mean	Std. dev.	Min	Max
ln_real_ho~e	15,763	3.121895	.6774424	-1.016968	6.35979

The image contains three tables showing descriptive statistics for the variable "ln_real_hourly_wage." The tables are divided into three categories: Overall, Men, and Women. Each table includes the following columns: Variable, Obs (Observations), Mean, Std. dev. (Standard Deviation), Min (Minimum), and Max (Maximum).

These statistics indicate that men have a slightly higher average logarithmic real hourly wage compared to women. The standard deviation is slightly higher for women, suggesting more variability in their wages. The range of wages, from minimum to maximum, is similar across both groups, highlighting some consistency in wage distribution.

Descriptive statistics for lfp (Overall):

Variable	Obs	Mean	Std. dev.	Min	Max
lfp	132,846	. 0741611	.2620338	0	1

Descriptive statistics for lfp (Men):

Variable	Obs	Mean	Std. dev.	Min	Max
lfp	38,103	.1541348	.3610827	0	1

Descriptive statistics for lfp (Women):

Variable	Obs	Mean	Std. dev.	Min	Max
lfp	94,743	.0419978	.2005852	0	1

The image contains three tables showing descriptive statistics for the variable "lfp" (labor force participation). The tables provide the number of observations (Obs), mean, standard deviation (Std. dev.), minimum (Min), and maximum (Max) values for three different groups: overall, men, and women.

These statistics indicate that men have a higher mean labor force participation rate compared to women. The standard deviation for men is also higher, suggesting more variability in their participation. The maximum value of 1 indicates that there are individuals in each group who are fully participating in the labor force.

Descriptive statistics for wage_employee (Overall):

Variable	Obs	Mean	Std. dev.	Min	Max
lfp	132,846	.0741611	.2620338	0	1

Descriptive statistics for wage_employee (Men):

Variable	Obs	Mean	Std. dev.	Min	Max
lfp	38,103	.1541348	.3610827	0	1

Descriptive statistics for wage_employee (Women):

Variable	Obs	Mean	Std. dev.	Min	Max
lfp	94,743	.0419978	.2005852	0	1

The image contains three tables showing descriptive statistics for the variable "lfp" (labor force participation). The tables are divided into three categories: Overall, Men, and Women. Each table includes the number of observations (Obs), mean, standard deviation (Std. dev.), minimum (Min), and maximum (Max) values.

These statistics indicate that men have a higher mean labor force participation rate compared to women. The standard deviation for men is also higher, suggesting more variability in their participation. The maximum value of 1 indicates that there are individuals in each group who are fully participating in the labor force.

Collinearity

(obs=273,679)	years_~g dogum_~l
years_of_s~g	1.0000
dogum_yil	0.3770 1.0000

The image shows a correlation matrix for two variables: "years_of_s~g" and "dogum_yil," based on 273,679 observations. The matrix includes the correlation coefficients, which measure the strength and direction of the linear relationship between the variables.

Years of Schooling and Itself: The correlation coefficient is 1.0000, indicating a perfect positive correlation, as expected for a variable with itself.

Dogum Yil and Itself: Similarly, the correlation coefficient is 1.0000, indicating a perfect positive correlation.

Years of Schooling and Dogum Yil: The correlation coefficient is 0.3770, indicating a moderate positive correlation. This means that as the years of schooling increase, the birth year (dogum_yil) also tends to increase.

These correlations help in understanding the relationships between education level and birth year. A moderate positive correlation suggests that younger individuals might have slightly higher educational attainment. This analysis can be useful for policymakers and researchers studying educational trends and their implications.

Results for lfp (Men) – OLS

Number of obs = 38,103

F(1, 38101) = 936.13

Prob > F = 0.0000

R-squared = 0.0240

Adj R-squared = 0.0240

Root MSE = .35673

Source	SS	df	MS
Model Residual	119.129196 4848.63686	1 38,101	119.129196 .12725747
Total	4967.76606	38,102 .	130380716

lfp	Coefficient	Std. err.	z	P>|z|	[95% conf . interval]	
years_of_schooling	.0129171	.0004222	30.60	0.000	.0120896	.0137446
_cons	.0467374	.0039574	11.81	0.000	.0389808	.054494

Here is a detailed analysis of the image showing the results of an Ordinary Least Squares (OLS) regression analysis for the labor force participation (lfp) of men:

Observations: The analysis is based on 38,103 observations. This large sample size provides robust results for the regression analysis.

F-statistic: The F-statistic is 936.13, indicating that the overall regression model is statistically significant.

Probability > F: The probability of the F-statistic being significant is 0.0000, confirming that the model's explanatory variables have a statistically significant relationship with labor force participation.

R-squared: The R-squared value is 0.0240, indicating that the model explains approximately 2.4% of the variance in labor force participation.

Adjusted R-squared: The adjusted R-squared value is also 0.0240, suggesting that the model's explanatory power remains consistent after adjusting for the number of predictors.

Root Mean Squared Error (MSE): The root MSE is 0.35673, providing a measure of the average deviation of the observed values from the fitted values.

The ANOVA table shows the sources of variation:

Model: The sum of squares (SS) for the model is 119.129196 with 1 degree of freedom (df) and a mean square (MS) of 119.129196.

Residual: The residual sum of squares is 4848.63866 with 38,101 degrees of freedom and a mean square of 0.12725747.

Total: The total sum of squares is 4967.76606 with 38,102 degrees of freedom and a mean square of 0.130380716.

The regression coefficients table indicates:

Years of Schooling: The coefficient for years of schooling is 0.0129171 with a standard error of 0.0004222, a z-value of 30.60, and a p-value of 0.000, indicating a statistically significant positive relationship with labor force participation. The 95% confidence interval is [0.0120896, 0.0137446].

Constant: The constant (intercept) coefficient is 0.0467374 with a standard error of 0.0039574, a z-value of 11.81, and a p-value of 0.000, indicating statistical significance. The 95% confidence interval is [0.0389808, 0.054494].

These results suggest that years of schooling positively impact labor force participation for men.

Results for wage_employee (Men) - OLS:

Number of obs = 38,103

F(1, 38101) = 936.13

Prob > F = 0.0000

R-squared = 0.0240

Adj R-squared = 0.0240

Root MSE = .35673

Source	SS	df	MS
Model Residual	119.129196 4848.63686	1 38,101	119.129196 .12725747
Total	4967.76606	38,102 .	130380716

wage_employee	Coefficient	Std. err.	z	P>|z|	[95% conf	. interval]
years_of_schooling	.0129171	.0004222	30.60	0.000	.0120896	.0137446
_cons	.0467374	.0039574	11.81	0.000	.0389808	.054494

The image shows the results of an Ordinary Least Squares (OLS) regression analysis for the variable "wage_employee" (Men). Here are the key details:

Number of Observations: The analysis includes data from 38,103 observations, providing a robust sample for the analysis.

F-statistic: The F-statistic is 936.13, indicating that the regression model is statistically significant.

Probability > F: The p-value is 0.0000, confirming the statistical significance of the model's explanatory variables.

R-squared: The R-squared value is 0.0240, suggesting that the model explains 2.4% of the variance in the dependent variable.

Adjusted R-squared: The adjusted R-squared is also 0.0240, consistent with the R-squared value.

Root Mean Squared Error (MSE): The root MSE is 0.35673, which measures the average deviation of observed values from the fitted values.

The analysis includes the following sources of variation:

Model: The sum of squares for the model is 119.129196 with 1 degree of freedom.

Residual: The residual sum of squares is 4848.63866 with 38,101 degrees of freedom.

Total: The total sum of squares is 4967.76606 with 38,102 degrees of freedom.

The regression coefficients table shows:

Years of Schooling: The coefficient is 0.0129171, with a standard error of 0.0004222 and a z-value of 30.60. The p-value is 0.000, indicating statistical significance.

Constant: The constant (intercept) is 0.0467374 with a standard error of 0.0039574 and a z-value of 11.81. The p-value is also 0.000, indicating statistical significance.

Results for ln_real_hourly_wage (Men) – OLS

Number of obs = 32,616

F(1, 32614) = 12889.95

Prob > F = 0.0000

R-squared = 0.2833

Adj R-squared = 0.2832

Root MSE = .48008

Source	SS	df	MS
Model	2970.88309	1	2970.88309
Residual	487516.8959	32,614	.23048065
Total	10487.779	32,615	.321563054

ln_real_hourly_w~e	Coefficient	Std. err.	z	P>|z|	[95% conf . interval]	
years_of_schooling	.0733084	.0006457	113.53	0.000	.0720428	.074574
_cons	. 2.440586	.0072481	336.72	0.000	2.42638	2.454793

The image shows the results of an Ordinary Least Squares (OLS) regression analysis for the natural logarithm of real hourly wage (ln_real_hourly_wage) for men. Here are the key details:

Number of Observations: The analysis is based on data from 32,616 observations, which provides a substantial sample size for reliable results.

F-statistic: The F-statistic is 12,889.95, indicating that the overall regression model is highly statistically significant.

Probability > F: The p-value is 0.0000, confirming that the model's explanatory variables have a statistically significant relationship with the dependent variable.

R-squared: The R-squared value is 0.2833, suggesting that the model explains about 28.33% of the variance in the natural logarithm of real hourly wages.

Adjusted R-squared: The adjusted R-squared is also 0.2832, consistent with the R-squared value, which accounts for the number of predictors in the model.

Root Mean Squared Error (MSE): The root MSE is 0.48008, indicating the average deviation of observed values from the fitted values.

The regression coefficients table shows:

Years of Schooling: The coefficient for years of schooling is 0.0733084, with a standard error of 0.0006457 and a z-value of 113.53. The p-value is 0.000, indicating a highly statistically significant positive relationship between years of schooling and ln_real_hourly_wage. The 95% confidence interval is [0.0720428, 0.074574].

Constant: The constant (intercept) coefficient is 2.440586, with a standard error of 0.0072481 and a z-value of 336.72. The p-value is also 0.000, indicating statistical significance. The 95% confidence interval is [2.42638, 2.454793].

These results suggest that years of schooling have a positive and statistically significant impact on the natural logarithm of real hourly wages for men.

Results for lfp (Men) - IV

(obs=273,679)	years_~g dogum_~l
years_of_s~g dogum_yil	1.0000 0.3770 1.0000

Number of obs = 38,103

Wald chi2(1) = 515.22

Prob > chi2 = 0.0000

R-squared = .

Root MSE = .37096

lfp	Coefficient	Std. err.	z	P>|z|	[95% conf . interval]	
years_of_schooling	.0364289	.0016049	22.70	0.000	.0332833	.0395744
_cons	-.1487489	.0134784	-11.04	0.000	-.1751661	-.1223316

Variable	R-sq.	Adjusted R-sq.	Partial R-sq.	F(1,38101)	Prob > F
years_of_s~g	0.0748	0.0748	0.0748	3081.58	0.0000

2SLS relative bias	5%	10%	20%	30%
	(not available)			
	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	16.38	8.96	6.66	5.53
LIML size of nominal 5% Wald test	16.38	8.96	6.66	5.53

The image provides the results of an Ordinary Least Squares (OLS) regression analysis for labor force participation (lfp) among men. Here are the key points:

The analysis shows a statistically significant model, with labor force participation positively influenced by years of schooling.

The F-statistic indicates that the explanatory variables in the model are significant.

The coefficient for years of schooling is positive, suggesting that higher levels of education lead to increased labor force participation.

The constant term is statistically significant, indicating the baseline level of labor force participation when other variables are zero.

The R-squared value suggests that the model explains a small but notable portion of the variance in labor force participation.

The standard error and confidence intervals provide a measure of the precision of the estimated coefficients.

Overall, the regression analysis supports the conclusion that education has a positive and statistically significant impact on labor force participation among men.

These findings emphasize the importance of education in enhancing labor force engagement, highlighting a clear relationship between higher educational attainment and increased participation in the workforce.

Results for wage_employee (Men) - IV

Number of obs = 38,103

Wald chi2(1) = 515.22

Prob > chi2 = 0.0000

R-squared = .

Root MSE = .37096

wage_employee	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
years_of_schooling	.0364289	.0016049	22.70	0.000	.0332833	.0395744
_cons	-.1487489	.0134784	-11.04	0.000	-.1751661	-.1223316

Variable	R-sq.	Adjusted R-sq.	Partial R-sq.	F(1,38101)	Prob > F
years_of_s~g	0.0748	0.0748	0.0748	3081.58	0.0000

2SLS relative bias	5%	10%	20%	30%
	(not available)			
	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	16.38	8.96	6.66	5.53
LIML size of nominal 5% Wald test	16.38	8.96	6.66	5.53

The image provides the results of an instrumental variable (IV) regression analysis for the variable "wage_employee" among men. Here are the key points:

Statistical Significance: The model is highly statistically significant, indicating that the explanatory variables have a strong relationship with the dependent variable.

Positive Impact of Education: The coefficient for years of schooling is positive and statistically significant, suggesting that more years of education are associated with higher employment in wage jobs.

Baseline Employment: The constant term is statistically significant, representing the baseline level of wage employment when other variables are zero.

Explained Variance: The R-squared values indicate how well the model explains the variance in wage employment, with partial R-squared values showing the contribution of individual predictors.

Model Precision: The standard errors and confidence intervals provide measures of the precision of the estimated coefficients, with narrow intervals indicating high precision.

Model Performance: The Root Mean Squared Error (MSE) indicates the average deviation of the predicted values from the observed values.

ANOVA Table: The analysis includes detailed sources of variation, helping to understand the overall fit of the model.

Labor Economics Insight: This regression highlights the importance of education in increasing wage employment among men.

Policy Implications: These findings suggest that investing in education can lead to better employment opportunities in wage jobs, underscoring the value of educational policies.

Overall Conclusion: The regression analysis confirms that years of schooling have a positive and statistically significant impact on wage employment for men, highlighting the role of education in enhancing job prospects.

Results for ln_real_hourly_wage (Men) - IV

Number of obs	=	32,616
Wald chi2(1)	=	28.96
Prob > chi2	=	0.0000
R-squared	=	.
Root MSE	=	.60843

ln_real_hourly_w~e	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
years_of_schooling	-.0174868	.0032493	-5.38	0.000	-.0238553	-.0111183
_cons	3.388767	.0340995	99.38	0.000	3.321933	3.455601

Variable	**R-sq.**	Adjusted **R-sq.**	Partial **R-sq.**	**F(1,38101)**	**Prob > F**
years_of_s~g	0.0634	0.0634	0.0634	2208.66	0.0000

2SLS relative bias	**5%**	**10%**	**20%**	**30%**
	(not available)			
	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	16.38	8.96	6.66	5.53
LIML size of nominal 5% Wald test	16.38	8.96	6.66	5.53

The image shows the results of an instrumental variable (IV) regression analysis for the natural logarithm of real hourly wage (ln_real_hourly_wage) for men. Here are the key points:

Statistical Significance: The model is highly statistically significant, as indicated by the Wald chi-squared test and its p-value. This means that the explanatory variables in the model significantly explain the variation in ln_real_hourly_wage.

Negative Coefficient for Years of Schooling: The coefficient for years of schooling is negative and statistically significant, suggesting that more years of schooling are associated with lower real hourly wages. This unexpected result may require further investigation.

Baseline Wage Level: The constant term in the model is statistically significant, representing the baseline level of real hourly wage when years of schooling is zero.

Explained Variance: The adjusted R-squared value indicates the proportion of the variance in the dependent variable that is explained by the independent variable, though it is relatively low in this model.

Precision of Estimates: The standard errors and confidence intervals provide measures of the precision of the estimated coefficients, with narrow intervals indicating higher precision.

Model Performance: The Root Mean Squared Error (MSE) provides an indication of the model's accuracy in predicting the dependent variable.

Further Considerations: The negative coefficient for years of schooling could suggest complex interactions or other factors affecting real hourly wages that are not captured by the model.

Policy Implications: These findings might imply that simply increasing years of schooling may not necessarily lead to higher wages and that the quality and relevance of education might also play critical roles.

Overall Conclusion: The regression analysis highlights the complex relationship between education and real hourly wages for men, suggesting the need for a deeper understanding of the underlying factors.

Further Analysis: Additional research could explore other variables and factors that might explain the negative association between schooling and wages, such as labor market conditions, types of education, or other demographic factors.

These points summarize the key findings and implications of the regression analysis.

Results for lfp (Women) - OLS

Number of obs	= 94,743
F(1, 94741)	= 4129.24
Prob > F	= 0.0000
R-squared	= 0.0418
Adj R-squared	= 0.0418
Root MSE	= .19635

Source	SS	df	MS
Model	159.200561	1	159.200561
Residual	3652.69009	94,741	.038554481
Total	3811.89065	94,742	.040234433

lfp	Coefficient	Std. err.	z	P>|z|	[95% conf	. interval]
years_of_schooling	.008873	.0001381	64.26	0.000	.0086024	.0091437
_cons	-.0198945	.0011553	-17.22	0.000	-.0221588	-.0176302

The image presents the results of an Ordinary Least Squares (OLS) regression analysis for labor force participation (lfp) among women. Key points include:

Model Significance: The model is statistically significant, as indicated by the high F-statistic and the p-value of 0.0000. This means the explanatory variables significantly predict labor force participation.

Positive Impact of Schooling: The coefficient for years of schooling is positive and statistically significant, suggesting that higher levels of education lead to increased labor force participation among women.

Baseline Participation: The constant term is statistically significant, representing the baseline level of labor force participation when other variables are zero.

Explained Variance: The R-squared value indicates that the model explains about 4.18% of the variance in labor force participation, which, while modest, is still notable.

Model Precision: The standard errors and confidence intervals provide measures of the precision of the estimated coefficients, with narrow intervals indicating high precision.

Model Performance: The Root Mean Squared Error (MSE) suggests the average deviation of observed values from the fitted values.

Importance of Education: The positive and significant coefficient for years of schooling underscores the critical role of education in enhancing labor force participation among women.

Policy Implications: These results suggest that investing in women's education could lead to higher labor force participation rates, highlighting the value of educational policies.

Overall Conclusion: The regression analysis confirms that years of schooling have a positive and statistically significant impact on labor force participation for women.

Further Research: Additional studies could explore other factors influencing labor force participation to build a more comprehensive understanding of the dynamics involved.

Results for wage_employee (Women) - OLS

Number of obs	=	94,743
F(1, 94741)	=	4129.24
Prob > F	=	0.0000
R-squared	=	0.0418
Adj R-squared	=	0.0418
Root MSE	=	.19635

Source	**SS**	**df**	**MS**
Model	159.200561	1	159.200561
Residual	3652.69009	94,741	.038554481
Total	3811.89065	94,742	.040234433

lfp	**Coefficient**	**Std. err.**	**z**	**P>|z|**	**[95% conf**	**. interval]**
years_of_schooling	.008873	.0001381	64.26	0.000	.0086024	.0091437
_cons	-.0198945	.0011553	-17.22	0.000	-.0221588	-.0176302

The image provides the results of an Ordinary Least Squares (OLS) regression analysis for wage employment among women. Here are the key points:

Model Significance: The analysis shows a highly significant model, as evidenced by the F-statistic and its p-value. This means the variables used in the model significantly predict wage employment among women.

Positive Impact of Education: The coefficient for years of schooling is positive and statistically significant, indicating that higher levels of education are associated with higher wage employment among women.

Baseline Employment Level: The constant term is statistically significant, representing the baseline level of wage employment when years of schooling is zero.

Explained Variance: The R-squared value suggests that the model explains a small but notable portion of the variance in wage employment, indicating that education plays a role but other factors may also be important.

Model Precision: The narrow confidence intervals for the coefficients indicate high precision in the estimates.

Model Performance: The Root Mean Squared Error (MSE) provides a measure of the average deviation of observed values from the fitted values, indicating the model's predictive accuracy.

Education's Role: The positive relationship between years of schooling and wage employment emphasizes the importance of education in enhancing women's job prospects.

Policy Implications: These findings suggest that improving educational opportunities for women could lead to better employment outcomes, highlighting the importance of educational policies.

Overall Conclusion: The regression analysis confirms that higher educational attainment positively impacts wage employment for women, supporting the need for continued investment in women's education.

Future Considerations: Further research could explore additional factors influencing wage employment to provide a more comprehensive understanding of the dynamics at play.

Results for ln_real_hourly_wage (Women) – OLS

Number of obs = 15,763

F(1, 15761) = 10022.56

Prob > F = 0.0000

R-squared = 0.3887

Adj R-squared = 0.3887

Root MSE = .52967

Source	SS	df	MS
Model	2811.84805	1	2811.84805
Residual	4421.77923	15,761	.280551946
Total	7233.62728	15,762	.458928263

lfp	Coefficient	Std. err.	z	P>|z|	[95% conf	. interval]
years_of_schooling	.0900813	.0008998	100.11	0.000	.0883176	.091845
_cons	2.117114	.0108871	194.46	0.000	2.095774	2.138454

The image shows the results of an Ordinary Least Squares (OLS) regression analysis for the natural logarithm of real hourly wage for women. Here are the key points:

Model Significance: The model is highly statistically significant, indicating that the explanatory variables have a strong relationship with the dependent variable.

Positive Impact of Schooling: The coefficient for years of schooling is positive and statistically significant, suggesting that higher levels of education are associated with higher real hourly wages.

Baseline Wage Level: The constant term is also statistically significant, representing the baseline level of real hourly wage when years of schooling is zero.

Explained Variance: The R-squared value indicates that the model explains a substantial portion of the variance in real hourly wages.

Precision of Estimates: The narrow confidence intervals for the coefficients suggest high precision in the estimates.

Model Performance: The Root Mean Squared Error (MSE) provides an indication of the model's accuracy in predicting the dependent variable.

Impact of Schooling: The significant positive coefficient for years of schooling emphasizes the importance of education in increasing women's real hourly wages.

Policy Implications: These findings suggest that investing in women's education can lead to higher wages, underlining the value of educational policies.

Overall Conclusion: The regression analysis confirms that years of schooling have a positive and statistically significant impact on the real hourly wages of women.

Further Analysis: Additional research could explore other factors that might influence wages, such as labor market conditions or the quality of education, to provide a more comprehensive understanding of the dynamics at play

Results for lfp (Women) - IV

Number of obs	=	94,743
Wald chi2(1)	=	661.47
Prob > chi2	=	0.0000
R-squared	=	0.0417
Root MSE	=	.19636

lfp	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
years_of_schooling	.0092503	.0003597	25.72	0.000	.0085453	.0099552
_cons	-.022526	.0025886	-8.70	0.000	-.0275996	-.0174524

Variable	**R-sq.**	Adjusted **R-sq.**	Partial **R-sq.**	**F(1,38101)**	**Prob > F**
years_of_s~g	0.1474	0.1474	0.1474	16379.2	0.0000

2SLS relative bias	**5%**	**10%**	**20%**	**30%**
	(not available)			
	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	16.38	8.96	6.66	5.53
LIML size of nominal 5% Wald test	16.38	8.96	6.66	5.53

The image shows the results of an instrumental variable (IV) regression analysis for labor force participation (lfp) among women. Here are the key points:

Statistical Significance: The model is highly statistically significant, indicating that the explanatory variables significantly predict labor force participation.

Positive Impact of Education: The coefficient for years of schooling is positive and statistically significant, suggesting that more years of education are associated with higher labor force participation.

Baseline Participation: The constant term is statistically significant, representing the baseline level of labor force participation when years of schooling is zero.

Explained Variance: The R-squared value indicates that the model explains a notable portion of the variance in labor force participation.

Precision of Estimates: The standard errors and confidence intervals provide measures of the precision of the estimated coefficients, with narrow intervals indicating high precision.

Model Performance: The Root Mean Squared Error (MSE) provides an indication of the model's accuracy in predicting the dependent variable.

Impact of Schooling: The significant positive coefficient for years of schooling emphasizes the importance of education in increasing labor force participation among women.

Policy Implications: These findings suggest that investing in women's education can lead to higher labor force participation rates, highlighting the value of educational policies.

Overall Conclusion: The regression analysis confirms that years of schooling have a positive and statistically significant impact on labor force participation for women.

Further Research: Additional studies could explore other factors influencing labor force participation to build a more comprehensive understanding of the dynamics involved.

These points summarize the key findings and implications of the regression analysis.

Results for wage_employee (Women) - IV

Number of obs	=	94,743
Wald chi2(1)	=	661.47
Prob > chi2	=	0.0000
R-squared	=	0.0417
Root MSE	=	.19636

lfp	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
years_of_schooling	.0092503	.0003597	25.72	0.000	.0085453	.0099552
_cons	-.022526	.0025886	-8.70	0.000	-.0275996	-.0174524

Variable	**R-sq.**	Adjusted **R-sq.**	Partial **R-sq.**	**F(1,38101)**	**Prob > F**
years_of_s~g	0.1474	0.1474	0.1474	16379.2	0.0000

2SLS relative bias	**5%**	**10%**	**20%**	**30%**
	(not available)			
	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	16.38	8.96	6.66	5.53
LIML size of nominal 5% Wald test	16.38	8.96	6.66	5.53

The image provides the results of an Ordinary Least Squares (OLS) regression analysis for labor force participation (lfp) among women. Here are the key insights:

Model Significance: The model is highly statistically significant, indicating that the explanatory variables have a strong predictive power for labor force participation.

Positive Impact of Schooling: The coefficient for years of schooling is positive and statistically significant, suggesting that higher levels of education are associated with increased labor force participation.

Baseline Participation: The constant term is also statistically significant, representing the baseline level of labor force participation when years of schooling is zero.

Explained Variance: The R-squared value indicates that the model explains a meaningful portion of the variance in labor force participation.

Precision of Estimates: The standard errors and confidence intervals for the coefficients are narrow, indicating high precision in the estimates.

Model Performance: The Root Mean Squared Error (MSE) provides an indication of the model's accuracy in predicting labor force participation.

Impact of Education: The significant positive coefficient for years of schooling highlights the importance of education in enhancing labor force participation among women.

Policy Implications: These findings suggest that investing in women's education can lead to higher labor force participation rates, emphasizing the value of educational policies.

Overall Conclusion: The regression analysis confirms that years of schooling have a positive and statistically significant impact on labor force participation for women.

Further Research: Additional studies could explore other factors influencing labor force participation to build a more comprehensive understanding of the dynamics involved.

These key points summarize the findings and implications of the regression analysis for women's labor force participation.

Results for ln_real_hourly_wage (Women) - IV

Number of obs	=	15,763
Wald chi2(1)	=	307.46
Prob > chi2	=	0.0000
R-squared	=	0.2838
Root MSE	=	.57328

ln_real_hourly_w~e	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
years_of_schooling	.0432868	.0024687	17.53	0.000	.0384483	.0481253
_cons	2.639067	.027912	94.55	0.000	2.584361	2.693774

Variable	**R-sq.**	Adjusted **R-sq.**	Partial **R-sq.**	**F(1,38101)**	**Prob > F**
years_of_s~g	0.1556	0.1556	0.1556	2904.93	0.0000

2SLS relative bias	**5% 10% 20% 30% (not available)**
	10% 15% 20% 25%
2SLS size of nominal 5% Wald test	16.38 8.96 6.66 5.53
LIML size of nominal 5% Wald test	16.38 8.96 6.66 5.53

The image displays the results of an instrumental variable (IV) regression analysis for women's real hourly wages. The analysis indicates the following insights:

Statistical Significance: The model is highly statistically significant, showing that the explanatory variables significantly predict real hourly wages.

Positive Impact of Education: The coefficient for years of schooling is positive and statistically significant, suggesting that more years of education are associated with higher real hourly wages.

Baseline Wage Level: The constant term is also statistically significant, representing the baseline wage level when years of schooling is zero.

Explained Variance: The R-squared value indicates that the model explains a substantial portion of the variance in real hourly wages.

Precision of Estimates: The standard errors and narrow confidence intervals for the coefficients indicate high precision in the estimates.

Model Performance: The Root Mean Squared Error (MSE) provides an indication of the model's accuracy in predicting wages.

Impact of Education: The positive relationship between years of schooling and wages highlights the importance of education in enhancing women's earnings potential.

Policy Implications: These findings suggest that investing in women's education can lead to higher wages, underlining the value of educational policies.

Overall Conclusion: The regression analysis confirms that years of schooling have a positive and statistically significant impact on women's real hourly wages.

Further Analysis: Additional research could explore other factors influencing wages, such as labor market conditions or the quality of education, for a more comprehensive understanding.

These points summarize the key findings and implications of the regression analysis presented in the image.

Difference between ols and iv methode

Differences between OLS and IV results:

Coefficient Estimates

For men, the OLS estimate for the effect of years of schooling on wage_employee is 0.0129, while the IV estimate for the same variable is 0.0364. This indicates that the IV estimate is significantly higher than the OLS estimate.

For women, the OLS estimate for the effect of years of schooling on wage_employee is 0.0089, whereas the IV estimate is slightly higher at 0.0093.

Standard Errors

The standard errors for the IV estimates are generally higher than those for the OLS estimates. For example, in the case of men's wage_employee, the standard error for the IV estimate is 0.0016 compared to 0.0004 for the OLS estimate. This suggests that the IV estimates are less precise.

Significance Levels

Both OLS and IV estimates show statistically significant results for the years_of_schooling variable across all models, as indicated by the p-values being 0.000 in all cases.

R-squared Values

The R-squared values for the OLS models are available and indicate the proportion of variance explained by the model. For example, the R-squared for the OLS model for men's wage_employee is 0.0240.

IV models do not report R-squared values in the same way, making direct comparisons of model fit more challenging.

Root MSE

The Root Mean Squared Error (Root MSE) is higher for the IV models compared to the OLS models, indicating that the IV models have a higher prediction error. For instance, the Root MSE for IV on ln_real_hourly_wage (men) is 0.60843 compared to 0.48008 for the OLS model.

Direction of Effect

In the case of ln_real_hourly_wage for men, the OLS estimate suggests a positive effect of years of schooling (0.0733), whereas the IV estimate suggests a negative effect (-0.0175). This discrepancy indicates potential endogeneity issues that IV estimation addresses.

Instrument Relevance

The F-statistics for the first stage of the IV regressions are very high (e.g., 3081.58 for men's wage_employee), indicating strong instruments. This supports the validity of the IV approach in addressing endogeneity.

These differences highlight the importance of considering potential endogeneity in the analysis and suggest that IV estimates may provide a more accurate reflection of the causal impact of years of schooling on labor market outcomes.

Conclusion discussion

Discussion and Conclusion:

The analysis of the impact of years of schooling on labor market outcomes, specifically wage employment and hourly wages, reveals significant differences between Ordinary Least Squares (OLS) and Instrumental Variables (IV) estimation methods.

Discussion:

Coefficient Estimates

The IV estimates for the effect of years of schooling on wage employment are consistently higher than the OLS estimates for both men and women. This suggests that OLS may underestimate the true effect of education on employment outcomes due to potential endogeneity issues, such as omitted variable bias or measurement error. The higher IV estimates indicate a stronger positive impact of education on wage employment when endogeneity is addressed.

Standard Errors and Precision

The IV estimates exhibit larger standard errors compared to OLS estimates, reflecting reduced precision. This is a common characteristic of IV estimation, as it relies on instruments that may introduce additional variability. Despite this, the IV estimates remain statistically significant, underscoring the robustness of the findings.

Significance Levels

Both OLS and IV estimates demonstrate statistical significance for the years of schooling variable across all models. This consistency reinforces the importance of education in influencing labor market outcomes, though the magnitude of the effect differs between the two methods.

R-squared Values and Model Fit

The R-squared values for OLS models indicate that a small proportion of the variance in labor market outcomes is explained by years of schooling. The absence of R-squared values for IV models limits direct comparisons of model fit, but the higher prediction errors (Root MSE) in IV models suggest greater variability in the outcomes.

Direction of Effect

Notably, the IV estimate for the effect of years of schooling on men's real hourly wages is negative, contrasting with the positive OLS estimate. This discrepancy highlights the potential for endogeneity to bias OLS results and suggests that the true causal effect of education on wages may differ when using IV estimation.

Instrument Relevance

The high F-statistics for the first stage of the IV regressions confirm the strength and relevance of the instruments used. This supports the validity of the IV approach in addressing endogeneity and providing more reliable estimates of the causal impact of education.

Conclusion:

The findings from this analysis underscore the critical role of addressing endogeneity when estimating the impact of years of schooling on labor market outcomes. The differences between OLS and IV estimates indicate that failing to account for endogeneity may lead to biased and potentially misleading conclusions. The IV estimates suggest that the true effect of education on wage employment is more substantial than indicated by OLS, while the effect on hourly wages may be more complex and warrants further investigation.

Overall, the results highlight the importance of education in enhancing labor market prospects, particularly in terms of wage employment. Policymakers should consider these findings when designing educational policies and interventions aimed at improving labor market outcomes. Future research could explore additional instruments and methods to further refine the estimates and provide a more comprehensive understanding of the relationship between education and labor market success.