Econometrics Assignment

Matteo Bertasio[1]

March 2025

1 Describe the data

The data set includes 3138 observations in 4 sectors. We analyze the data considering the mean value for all variables within each sector, the mean value across all sectors, and the standard deviation across all sectors. We present the mean values of the variables divided by sector (1-4). Natural logarithm of employment (n): 10.8, -0.7, 3.77, 25.24. Natural logarithm of the relative wage (w): 3.59, 3.58, 3.47, 3.39. Natural logarithm of physical capital (k): -0.03, 0.03, 0.05, -0.01. Sectoral output (ys): 4.23, 1.52, 3, 5.72. For sectoral output we are provided with one single value for each sector. Taking the four sectors together, the mean value of employment is 16.76, for the relative wage is 3.46, for physical capital is 0.00, and for sectoral output is 4.66. Considering the four sectors, the standard deviation of employment is 28.6, for relative wage is 1.51, for physical capital is 1.99, and for sectoral output it is 1.32.

2 Estimate an employment equation through a OLS regression with dependent variable n and regressor w and k

The OLS regression shows that employment is negatively correlated with wage at the 1 percent significance level, and positively correlated with physical capital at the same significance. The R-squared=0.0075 and the Adj R-squared=0.0068 indicate that wage and capital do not explain much of the employment's variation, suggesting that other variables not included here may influence it.

3 Significance of coefficients

We employed two F-tests with single linear restrictions and an F-test with two joint linear restrictions to account for the significance of the coefficients. The F-test with a single linear restriction tests the null of the individual coefficient equal to zero, while the one with two joint linear restriction tests the linear combination of both coefficients equal to zero at the same time. The F-test can be expressed as the ratio between the difference in the residual sum of squares of the restricted and unrestricted model and the residual sum of squares of the unrestricted model, multiplied by their appropriate degress of freedom. The F distribution depends on these parameters: the degrees of freedom of the numerator, i.e. the number of restrictions, and the degrees of freedom of the denominator, i,e. the df of the residuals. The following table reports the results of the three tests, which allow us to reject the null hypothesis of non-significance of the two coefficients.

H_0	F-statistic	p-value
$w = 0, k = 0$	$F(2, 3135) = 11.77$	0.0000
$w = 0$	$F(1, 3135) = 10.25$	0.0014
$k = 0$	$F(1, 3135) = 13.52$	0.0002

Table 1: F-test results

4 Testing the null of homogeneity across sectors

To test the null of homogeneity across sectors, we ran an auxiliary regression involving the dummies for each sector (with sector 1 as baseline) using as dependent variables ys, w, k and n. Since it assumes

only four values assigned to each sector, the variable ys shows perfect heterogeneity with an R-squared of 1. For the subsequent regressions we performed the F-tests defining the null as if the effects across sectors were homogeneous (`test [sector_2] = [sector_3] = [sector_4]`). The p-values of the F-tests allowed us to reject the hypothesis of sectoral homogeneity for the regressions over wage and capital (0.166 and 0.864), while for employment (0.000) the test suggests the existence of sectoral heterogeneity.

5 Heterogeneity

As heterogeneity is observed in employment, we here use 3 ways to accommodate it in our regression model. Firstly, sectoral output is added as the interaction term. The result is that the effect of relative wage on employment is 1.72 higher if the sectoral output increase by 1 unit and this is significant at 99% confidence level. No significant effect is observed on the interaction term of k and ys. Another way is to do the subgroup regression. When conducting regression separately on each sector, we only observe significant effect of relative wage on employment in sector 4 (95% CI), which means the effect is mostly contributed by the single one sector. And the effect of physical capital on employment is significant in sector 1 and 4. At last, we simply add the dummy of sectors as independent variables. After controlling for sectors, the coefficient of w and k on n remain their sign and significance.

VARIABLES	(1) employment	(2) employment
ln of relative wage	-8.882***	-0.793**
	(0.547)	(0.318)
ln of capital	-0.0506	0.970***
	(0.934)	(0.241)
sector_2		-11.56***
		(2.117)
sector_3		-7.191***
		(1.560)
sector_4		14.26***
		(1.183)
c.w#c.ys	1.717***	
	(0.0973)	
c.k#c.ys	0.208	
	(0.190)	
Constant	19.94***	13.67***
	(1.217)	(1.508)
Observations	3,138	3,138
R-squared	0.098	0.121

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

Table 2: Regression result accounting for heterogeneity.

6 Heteroskedasticity

In the last part, we use both Breusch-Pagan and White methods to test for heteroskedasticity. The p value in both tests is less than 1%, which means there is sufficient evidence for heteroskedasticity. Therefore, robust standard deviation is used to correct for it. The result is that relative wage remains to have significant negative effect on employment at the 95% confidence level but the effect of physical capital is no longer significant.

---

[1] In collaboration with Anna Puricelli, Kexin Xu