ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Micro-Econometrics

Project 6

Peter Pan

How the number of workers and the stock of capital affect the amount of the product produced

Panou Panagiotis

Abstract

The theoretical model of the production process is the subject of numerous scientific papers. So and in this paper its purpose is to study how the number of workers and inventory affect the amount of product produced. Pooled ols, fixed effect and random effect models were used for the analysis of our data. A primary advantage of these models is the ability to control for time-invariant omitted variables that may bias observed relationships. From these three models, we concluded that the one with fixed effects is the most ideal for our analysis, as you justify in the rest of the paper.

Introduction

The purpose of this article is to study how employment and gross capital stockaffect the amount of product produced in industries. Human capital refers to the knowledge, skill sets, and experience that workers have in an economy. The skills provide economic value since a knowledgeable workforce can lead to increased productivity. The concept of human capital is the realization that not everyone has the same skill sets or knowledge. That is why it is very important for a business to properly invest in the right human resources. Capital stock refers to the assets—physical tools, plants, and equipment—that allow for increased work productivity. With other words capital can be any financial asset that is used. The money made from its current activities is shown as capital on a company's balance sheet. By increasing productivity through improved capital equipment, more goods can be produced and the standard of living can rise. For an industry, output is a measure of all the goods and services produced in a given time period by businesses in that industry and sold either to consumers or to businesses outside that industry. All these factors are interrelated and the purpose of this article is to study this autocorrelation. It is vital to refer that as we know from Arellano, M. and S.R. Bond (1991), and it is also valid in the case of our analysis because we have to deal with unbalanced panel data, by unbalanced panel data we refer to a sample in which consecutive observations on individual units are available, but the number of time periods available may vary from unit to unit as well as the historical points to which the observations correspond.

Data

The set of variables that will be applied throughout our analysis are the following:

year (for the survey year), ind (for the industry that we are looking at each time), emp (for the employment-workforce), wage (for the real wage), cap (for the gross capital stock), indoutpt (for the industry outputs) and lastly id (for the Firm id for each industry). Then we logarithmize the variables indoutput, cap, and emp and there are derived the variables lindoutput, lcap, and lemp correspondingly for which the summary statistics appear in table 2.

After mentioning all the variables what they represent in our sample, from table 1 we can explain the following conclusions for the sample of 1037 observations: The variable Year has an average of 1979.65 years, the minimum and maximum values are 1976 and 1984 years, which means that we start our survey at 1976 and we have the latest information's for industries data at 1984 the standard deviation is 2.22. The mean of industry outputs is 103,80 which means that half of companies produce more than 103.80 production units when the other 50% of industries in our sample produce lower than 103.80 production units. The standard deviation is 9.94, and the range is from 86.9 to 128.37 which means that the lowest productivity which mentioned in our sample is 86.9 and the highest is 128.37. From the variable id we can mention that each industry has a uniwue id and this is a formatting attribute. The mean of the gross capital stock in the sample is 2.51 units, with standard deviation 6.25, minimum value corresponding to 0.01 and the maximum value which is 47.11. Finally, employment index (emp) has an average of 7.90 units, its minimum and maximum values are 0.1 and 108.56 units respectively while standard deviation is 15.93.

Table 1. Summary Statistics

Variable	Mean	Standard deviation	Minimum	Maximum	Observations
Id	73.20	41.23	1	140	1031
Year	1979.65	2.22	1976	1984	1031
Indoutpt	103.80	9.94	86.9	128.37	1031
Cap	2.51	6.25	0.01	47.11	1031
Emp	7.90	15.93	0.1	108.56	1031

Source: Datastream panel, U.K.

After setting our data as panel data we have the following summarized statistics which are presented in table 2. As we have already mentioned according to the description of the identification number variable of each company in the sample, 140 companies are observed. For the variable year which refers to the survey years we observe that according to the time separation the standard deviation is 2.13 and the sample is observed between 1975 and 1983. Based on the separation according to the firm's id the standard deviation is equal to 0.6 and 140 companies are observed where the first observation date is 1979 while the last one is 1981. Therefore a fairly large difference between the standard deviations is observed.

For the variable Indoutpt which represents industrial production, the average output equals 103.8. A separation is made based on firm's id and time. In the first case the lowest outflow value is 96.53 and the highest 115.06. While in the second case it is 85.6 and 121.94 respectively. We notice a difference between the values, since chronologically more extreme values are observed. The standard deviation in the between separation is 4.36 and in the within almost double 8.95.

The average gross capital equals 2.51 and the variable that expresses it (cap) takes values from 0.01 to 47.11. There is a large difference in standard deviations when disaggregating by follow-up year of firms and between firms, with 6.10 and 1.06 standard deviations respectively. The same is observed in the emp variable (between std. dev.: 15.17 and within std. dev.: 2.21) which expresses the current number of employees in each company with an average number of employees of 7.89. The differences express that specifically there is greater variation between companies (between) than over time (within), so the within estimate can lead to a significant loss of efficiency of the coefficients.

Table 2. Summary Statistics for panel data

Variable		Mean	Std. Dev.	Min	Max	Observations
id	overall	**73.20**	**41.23**	**1**	**140**	N = 1031
	between		**40.56**	**1**	**140**	n = 140
	within		**0**	**73.20**	**73.20**	T-bar = 7.36
year	overall	**1979.65**	**2.22**	**1976**	**1984**	N = 1031
	between		**0.60**	**1979**	**1981**	n = 140
	within		**2.13**	**1975.65**	**1983.65**	T-bar = 7.36
indoutpt	overall	**103.80**	**9.94**	**86.90**	**128.37**	N = 1031
	between		**4.36**	**96.53**	**115.07**	n = 140
	within		**8.95**	**85.60**	**121.94**	T-bar = 7.36
cap	overall	**2.51**	**6.25**	**0.01**	**47.11**	N = 1031
	between		**6.10**	**0.03**	**38.49**	n = 140
	within		**1.06**	**-12.42**	**12.48**	T-bar = 7.36
emp	overall	**7.89**	**15.93**	**0.10**	**108.56**	N = 1031
	between		**16.17**	**0.13**	**102.19**	n = 140
	within		**2.21**	**-14.81**	**34.76**	T-bar = 7.36

Source: Datastream panel, U.K.

Empirical model

For our empirical assessment, in how the number of workers and the stock of capital affect the amount of the produced product, the following three models were used: pooled ols, fixed effects and random effects. Pooled regression is standard ordinary least squares (OLS) regression without any cross-sectional or time effects. The error structure is simply, where the are independently and identically distributed with zero mean and variance. Generally we use pooled OLS to derive unbiased and consistent estimates of parameters even when time constant attributes are present, but random effects will be more efficient. Panel data models examine cross-sectional (group) and/or time-series (time) effects. These effects may be fixed and/or random. Fixed effects assume that individual group/time have different intercept in the regression equation, while random effects hypothesize individual group/time have different disturbance. If the study effect sizes are seen as having been sampled from a distribution of effect sizes, then the random-effects model, which reflects this idea, is the logical one to use. If the between-studies variance is substantial (and statistically significant) then the fixed-effect model is inappropriate.

For estimating the industry output we worked with a simple two input Cobb-Douglas production function: $y_{it} = e^{b0} * K_{it}^{b1} * L_{it}^{b2} * e_{it}$ (1)

where, i indexes firms, t indexes years, Kit is units of capital, Lit is units of labor, and Yit is units of output. b0, b1, b2 areparameters and eit captures unobservables that afects output. By taking natural logs we can create the next equation: $y_{it} = b_0 + b_{1t} * k_{it} + b_{2t} * l_{it} + e_{it}$ (2)

But we have to deal with the unobserved heterogeneity to obtain valid inference on the common parameters b. As a result of kit and lit correlation with eit we expect the coefficients from OLS method to be positively biased. We will divide eit in two components, i.e.

$$y_{it} = b_0 + b_{1t} * k_{it} + b_{2t} * l_{it} + h_i + u_{it} \qquad (3)$$

where, hi is unobservable firm-speciffc effect.

Estimation results

Table 3 shows the results from the three subsamples used for our analysis. For all subsamples we have the same number of observations 1031. Initially for the pooled ols we draw the following conclusions: the coefficient for lemp variable (employment) is 0.01 is statistically significant at all significance levels. The coefficient means that if the industry hire 1 more employee in its workforce, its hourly total industry output will increase by 1%. Through the coefficient R^2 = 0.01 we see that the variability of the industry output is explained by 1% by the change in employment.

For the second subsample (fixed effects) we can observe that the number of groups is 140 which means that we examine 140 industries during the period of 9 years. We observe that coefficient for lcap is 0.05 which means that if 1 unit of gross capital stock added we will have an increase of 5%. Also for lemp we observe that if 1 more employee be hired in the industry we will have an increase of 20% for the industry output. We observe that all coefficients are statistically significant at all levels of significance. For the industries which are examined through time we observe that we have minimum statistics for some industries for 7 periods, maximum periods are 9, and the average observations for our sample is

7.4. If we examine every time period separately (within) we can observe that R-squared is higher than the other two occasions (between, overall). This means that. R^2 explains that 29,1% of the variation in industry output is explained by variation in gross capital stock and employment.

For the third subsample (fixed effects) we can observe again that the number of groups is 140 which means as we have already refere that we examine 140 industries during the period of 9 years. We observe that coefficient for lemp is 0.002 wich means that if 1 more employee be hired in the industry we will have an increase of 2% for the industry output. This coefficient is statistically significant at level of significance 5%. For the industries which are examined through time we observe that we have minimum statistics for some industries for 7 periods, maximum periods are 9, and the average observations for our sample is 7.4. If we examine every time period separately (within) we can observe that R-squared is higher than the other two occasions (between, overall). This means that. R^2 explains that 20% of the variation in industry output is explained by variation in employment.

Table 3. Industry output estimates (Pooled OLS, Fixed Effects, Random Effects)

Variable	Pooled OLS	Fixed effects	Random effects
lcap	-0.01(0.00)	0.05*** (0.02)	-0.01 (0.01)
lemp	0.01*** (0.01)	0.20*** (0.20)	0.02** (0.01)
Sigma u	-	0.33	0.03
Sigma e	-	0.08	0.08
Rho	-	0.95	0.11
R-squared:	0.01	-	-
Within	-	0.29	0.20
Between	-	0.00	0.00
Overall	-	0.01	0.01
Observations	1031	1031	1031
Number of groups		140	140
Obs per group	-	-	-
minimum	-	7	7
average	-	7.4	7.4
maximum	-	9	9

Source: Datastream panel, U.K.

Notes: Standard errors on parenthesis. *** at 1%, ** at 5%, * at 10%

Conclusions

Main purpose of this paper is to investigate the impact of the total number of employees and gross capital stock in the industry output of 1031 observations of 140 firms which included in 9 industries for the period 1976-1984. The conclusion we reached is that for the empirical analysis of our panel data the most ideal method of analysis is fixed effects compared to pooled ols and random effects as it has the highest R squared as we observe in table 3 which is equal to 0.29 in the coefficient within which it concerns each time period separately for each industry. Therefore, this method can explain most of the variability of our dependent variable, which is industry outputs.

References

Arellano, M. and S.R. Bond (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, Review of Economic Studies, 58, 277-297.

Cameron, A. C., & Trivedi, P. K. (2010). Microeconometrics using stata (Vol. 2). College Station, TX: Stata press.