Stata example

Thor Donsby Noe[*]

Aarhus University, Department of Economics and Business Economics

January 20, 2023

Contents

[*] *E-mail:* Thor.Noe@econ.au.dk

1 Introduction

Here I display the output of my Stata code example developed for the course 2525: Applied Economics at Aarhus University, Department of Economics and Business Economics.

I proceed with presenting the data in section 2, empirical model in section 3, and results in section 4 before discussion in section 5 and conclusion in section 6.

2 Data

2.1 Firm scrap rates and the MJOB program

> " *The Michigan Job Opportunity Bank-Upgrade program was in effect during the years 1986-90. The program was designed to provide one-time training grants to eligible firms, defined as manufacturing companies with 500 or fewer employees that were implementing some type of new technology and were not past recipients of a grant.* "
>
> - Holzer et al. (1993)

Table 1: Descriptive statistics

	mean	sd	min	p50	max
Scrap rate	3.843642	6.00777	.01	1.415	30
log(scrap rate)	.3936814	1.486471	-4.60517	.3471233	3.401197
Grant	.1790123	.3845514	0	0	1
Observations	162				

Table 2: Descriptive statistics by year

	mean	sd	min	p50	max	count
1987						
Scrap rate	4.611667	6.414963	.01	1.675	30	54
log(scrap rate)	.597434	1.594659	-4.60517	.5158087	3.401197	54
Grant	0	0	0	0	0	54
1988						
Scrap rate	3.787778	5.984144	.05	1.51	25	54
log(scrap rate)	.4284409	1.409956	-2.995732	.4120877	3.218876	54
Grant	.3518519	.4820322	0	0	1	54
1989						
Scrap rate	3.131481	5.617764	.03	1	30	54
log(scrap rate)	.1551692	1.44214	-3.506558	0	3.401197	54
Grant	.1851852	.3920952	0	0	1	54
Total						
Scrap rate	3.843642	6.00777	.01	1.415	30	162
log(scrap rate)	.3936814	1.486471	-4.60517	.3471233	3.401197	162
Grant	.1790123	.3845514	0	0	1	162

3 Empirical model

As a baseline, I first consider a simple model of the scrap rate of firm i in year t

$$\ln(scrap_{i,t}) = \beta_1 grant_{i,t} + \beta_2 grant_{i,t-1} + \boldsymbol{\delta d_t} + \epsilon_{i,t} \tag{1}$$

where $\boldsymbol{d_t}$ is a vector of dummies for each year t (excluding the first to avoid perfect collinearity) to capture time effects. In the special case of an approximately linear yearly growth rate (or in this case, a steady decline), the model could be simplified with a single *trend* variable as

$$\ln(scrap_{i,t}) = \beta_1 grant_{i,t} + \beta_2 grant_{i,t-1} + \delta trend_t + \epsilon_{i,t} \tag{2}$$

The error term, $\epsilon_{i,t} = a_i + u_{i,t}$, can be reduced to $u_{i,t}$ by explicitly including the time-constant effects as dummies for each firm, a_i, (i.e. firm-specific intercepts)

$$\ln(scrap_{i,t}) = \beta_1 grant_{i,t} + \beta_2 grant_{i,t-1} + \boldsymbol{\delta d_t} + a_i + u_{i,t} \tag{3}$$

Time-demeaning each variable eliminates the time-constant effect, a_i, and provides the Fixed Effects estimation model (4) that is applied by Wooldridge (2019, example 14.1, pp. 464-465):

$$\ln\left(sc\ddot{r}ap_{i,t}\right) = \beta_1 gr\ddot{a}nt_{i,t} + \beta_2 gr\ddot{a}nt_{i,t-1} + \boldsymbol{\delta \ddot{d}_t} + \ddot{u}_{i,t} \tag{4}$$

Model (4) is estimated under the assumption of independent and identically distributed (i.i.d.) errors, $u_{i,t}$. That is, for all years $t \neq s$, the idiosyncratic errors in two different years are uncorrelated, $\text{Cov}(u_{i,t}, u_{i,s}|\boldsymbol{X}_i, a_i) = 0$ (Assumption FE.6), and have the same variance, $\text{Var}(u_{i,t}|\boldsymbol{X}_i, a_i) = \text{Var}(u_{i,s}|\boldsymbol{X}_i, a_i) = \sigma_u^2$ (Assumption FE.5), conditional on the explanatory variables $\boldsymbol{X}_i$ and the firm-specific intercepts a_i.

In model (5), I relax Assumption FE.5 and FE.6 in allowing for heteroscedasticity over time as well as serial correlation between observations for the same firm by applying heteroscedasticity- and cluster-robust standard errors (see Wooldridge, 2019, appendix 14A.2, pp. 493-494).

4 Results

Table 3: Estimation of scrap rates

	(1) Baseline b/se	(2) Trend b/se	(3) Dummies b/se	(4) FE b/se	(5) FE cluster robust b/se
Grant	0.2000	0.2030	-0.2523*	-0.2523*	-0.2523*
	(0.3383)	(0.3254)	(0.1506)	(0.1506)	(0.1434)
Grant lagged	0.0489	0.0459	-0.4216**	-0.4216**	-0.4216
	(0.4361)	(0.4247)	(0.2102)	(0.2102)	(0.2825)
Year 1988	-0.2394		-0.0802	-0.0802	-0.0802
	(0.3109)		(0.1095)	(0.1095)	(0.0978)
Year 1989	-0.4965		-0.2472*	-0.2472*	-0.2472
	(0.3379)		(0.1332)	(0.1332)	(0.1968)
Time trend		-0.2480			
		(0.1682)			
Firm dummies	No	No	Yes	No	No
R^2	0.0173	0.0173	0.9276	0.2010	0.2010
Obs.	162	162	162	162	162
Number of firms				54	54
Obs. per firm				3	3

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

Table 3 above shows OLS estimation results for the different models outlined in section 3, which all use the same sample described in table 1 and 2 in section 2.1.

5 Discussion

5.1 Robustness

The $\hat{\beta}_2$ estimate of the Fixed Effects estimation model (4) in table 3 above (identical to table 14.1 in Wooldridge, 2019, p. 464) indicates a large effect of *grant lagged*. However, the estimated effect becomes insignificant in model (5) under heteroscedasticity- and cluster-robust inference (which by definition increases the size of the standard errors). That is, I cannot be 95% certain that receiving a grant had a statistically significant effect on scrap rates in the same year or a lagged effect the next year. On the contrary, a p-value of 0.084 for $\hat{\beta}_1$ is decent considering the weak statistical power due to the low number of firms for which *scrap* is recorded. Moreover, the large size of $\hat{\beta}_2$ and its p-value of 0.141 implies that I cannot rule out lagged effects either.

5.2 Random treatment

The validity of either model relies on the assumption of random treatment assignment, that is, the firms should be drawn from the same distribution (prior to treatment) regardless of whether they will later receive the grant (i.e. be treated) or not. In appendix A.1, I investigate this assumption by plotting the distribution of scrap rates in 1987 for the treated and untreated firms respectively. Despite the median scrap rate in 1987 was the same among the treated and untreated firms, Panel A of figure 1 suggests a selection bias as none of the firms that received the grant in 1988 had a scrap rate close to zero (smallest was 0.28%) or above 18% in 1987. On the other hand, firms that did not receive the grant in 1988 nor in 1989 had much higher variance in scrap rates in 1987 due to several extreme observations (i.e. of the 25 untreated firms, five had a scrap rate below 0.28% and three above 18%).

6 Conclusion

The verdict on whether the estimated effects are significant is very sensitive to model choice. Identification of the effects of the training grant is further challenged by a low number of observations, large heterogeneity in scrap rates, and concerns about selection bias.

References

Holzer, Harry J, Richard N Block, Marcus Cheatham, and Jack H Knott (1993). "Are training subsidies for firms effective? The Michigan experience". In: *ILR Review* 46.4.

Wooldridge, Jeffrey M (2019). *Introductory econometrics: A modern approach*. 7th ed. Cengage Learning.

A Appendix

A.1 Scrap rate distributions

Figure 1: Kernel density of scrap rates by treatment and year