

Homework 4

Environmental Economics II
Maghira Ramadhani

February 12, 2024

Problem 1 Python

- Figure 1 shows that there's a parallel trend before the treatment started in Jan 2018. This is important to ensure that the treatment and control group are comparable before the treatment started.

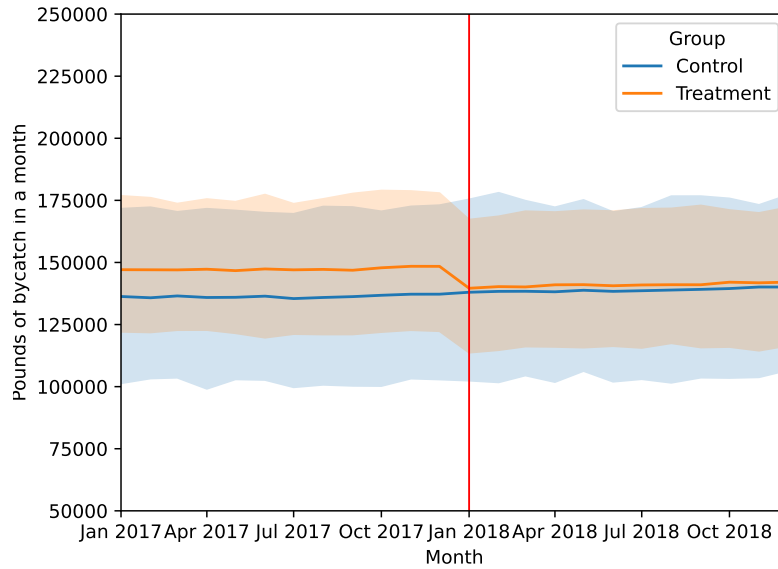


Figure 1: Bycatch by month plot

- DID estimates by using sample analog

Table 1: Parameter and average marginal effect estimates from Stata

	Sample analog value
$E[Y_{igt} g(i) = treatment, t = Pre] =$	148430.64
$E[Y_{igt} g(i) = treatment, t = Post] =$	139612.51
$E[Y_{igt} g(i) = control, t = Pre] =$	137228.60
$E[Y_{igt} g(i) = control, t = Post] =$	139612.51
DID=	-9591.35

3. Estimating DID using different specification in Python

Table 2: DID estimates from different specification

	(a)	(b)	(c)
DID estimates	-9591.35 (3198.64)	-8956.78 (3135.04)	-8436.28 (2795.47)
Group FE	✓	✓	✓
Month Indicator	✓	✓	✓
Controls	×	×	✓
Sample	Dec 2017 - Jan 2018	Jan 2017 - Dec 2018	Jan 2017 - Dec 2018

Standard errors are clustered at firm. I used pyfixest package for the estimation, the documentation is available at <https://github.com/s3alfisc/pyfixest>.

- (a) The following specification is used.

$$bycatch_{it} = \alpha + \lambda_{t=2017} + \gamma g(i) + \delta treat_{i,t} + \epsilon_{i,t} \quad (1)$$

By using sample from December 2017 and January 2018, the DID estimates is -9591.35 with a standard error of 3198.64. This means after the treatment started, the treated firms' bycatch yield is on average 9591.35 lbs less compared to the control group.

- (b) The following specification is used.

$$bycatch_{it} = \alpha + \lambda_t + \gamma g(i) + \delta treat_{i,t} + \epsilon_{i,t} \quad (2)$$

By using sample from January 2017 and December 2018, the DID estimates is -8956.78 with a standard error of 3135.04. This means after the treatment started, the treated firms' bycatch yield is on average 8956.78 lbs less compared to the control group. Using this specification we are now comparing the average of the entire after treatment period with the average of the entire before treatment period average, instead of only using 1 month of observations. This specification is more robust in capturing common time trends or seasonality of bycatch yields between both treated and untreated firms.

- (c) The following specification is used.

$$bycatch_{it} = \alpha + \lambda_t + \gamma g(i) + \delta treat_{i,t} + \beta X_{i,t} + \epsilon_{i,t} \quad (3)$$

By using sample from January 2017 and December 2018, the DID estimates is -8436.28 with a standard error of 2795.47. In this specification, we include firm size, salmon yields, and shrimp yields as control variables. This specification is more robust in capturing common time trends or seasonality of bycatch yields between both treated and untreated firms, and also controlling for other factors that might affect bycatch yields.

- (d) Table 2 show the DID estimates from the different specification shown in part (a), (b), and (c). The estimates from part (a) is similar with the result from computing the sample analog in question 2. The estimates from part (b) is more precise than (a) due to the larger sample size. The estimates from part (c) is more precise than (b) due to the inclusion of control variables.

Problem 2 Stata

1. Estimating DID using different specification in Stata

Table 3: DID estimates from different methods in Stata

	(a)	(b)
DID estimates	-8085.14 (2619.21)	-8149.06 (2489.02)
Method	Firm indicators	Within-transformation
Observations	1200	1200

Standard errors are clustered at firm.

- (a) The following specification is used.

$$bycatch_{it} = \alpha_i + \lambda_t + \gamma g(i) + \delta treat_{i,t} + \beta X_{i,t} + \epsilon_{i,t} \quad (4)$$

The estimates is shown in Table 3 column (a).

- (b) Let $\tilde{y}_{it} = y_{it} - \bar{y}_i$ denote the time-demeaned variable. The following specification which is equivalent with equation (4) is used.

$$\widetilde{bycatch}_{it} = \widetilde{\delta treat}_{i,t} + \beta \tilde{X}_{i,t} + \tilde{\epsilon}_{i,t} \quad (5)$$

Note that due to the time demeaning, firm size which is time-invariant control variable is dropped from the regression. The estimates is shown in Table 3 column (b).

- (c) Comparing estimates from (b) to (a), it looks like the estimates from (b) is more precise than (a) due to the smaller standard errors. However, the time demeaning removes the time-invariant control variables, i.e. firm size, from the regression, which could be a problem if we are to identify the coefficient for those variable.