Problem Set 2: Product Function Estimation
Econ 760: Empirical Industrial Organization

Matthew Grennan*

Due Date: Mar 5, 2025

The problem set is based on the dataset GMdata.csv that was the basis of the Griliches and Mairesse (1995) paper. Refer to this paper for more detail on the construction of the data and the variables, but in brief, the variables are:

1. index: firm ID
2. sic3: 3 digit SIC code
3. yr : year (73, 78, 83, 88)
4. ldsal : log of deflated sales
5. lemp : log of employment
6. ldnpt : log of deflated capital
7. ldrst : log of deflated R&D capital
8. ldrnd : log of deflated R&D
9. ldinv: log of deflated investment

For the entire problem set, you should consider the following production function

$$y_{jt} = \beta_0 + \beta_l l_{jt} + \beta_k k_{jt} + \omega_{jt} + u_{jt} \tag{1}$$

where j is firm ID, t is year (73, 78, 83, 88), y is log of deflated sales, l is log of employment , k is log of deflated capital.

1. Summarize the data by describing the variables, and in particular by focusing on the balanced and unbalanced nature of the panel.

*This Problem Set is based off work by Allan Collard-Wexler and Jan De Loecker. Any errors are my own.

2. Estimate the two coefficients of the production function using OLS, for both the balanced and unbalanced panel. Discuss your results and provide a reason for the different estimators. Now re-estimate your model, adding both year and industry dummies. Provide an interpretation on the coefficients on the dummies.

   [From now on consider the unbalanced panel only!]

3. Estimate the coefficients of the model, including firm and year fixed effects. Discuss your results and compare them with 2.

4. Estimate the probit model of survival as a function of investment and capital, including time and industry dummies. Discuss the signs on investment and capital. How is this regression of interest in the context of the production function?

5. Estimate the coefficients relying on the Olley and Pakes (1996) method with and without the selection control.

6. Compute the OP decomposition of aggregate productivity (across all sectors, that is market share is firm sales divided by total sales per year). Do the same analysis for sic3 357 and discuss the difference.