Econ 450-1: Problem Set 3

Gaston Illanes

Please turn in your answers in a typed-up pdf, and add your code as an appendix in a separate pdf.

Problem 1 - Identification of a Cobb-Douglas Production Function

Assume that the true production function is:

$$F(k_t, l_t, m_t) = K_t^{\alpha_k} L_t^{\alpha_l} M_t^{\alpha_m} \tag{1}$$

that productivity follows an AR(1) process:

$$\omega_t = \delta_0 + \delta_1 \omega_{t-1} + \eta_t \tag{2}$$

and that output and input prices are fixed (that is, ignore them).

1. Find the equivalent expression for equation (9) in Gandhi, Navarro and Rivers (2017) ("On the Identification of Gross Output Production Functions") in the context of this model. Leave $E[y_t|\Gamma_t]$ as a function of $E[m_t|\Gamma_t]$.
2. Use the firm's FOC to find the optimal quantity demanded of materials and replace into the previous expression.
3. Find the value of ϕ_{t-1} implied by this model, and replace into the previous expression.
4. Which coefficients of the production function are identified? Which are not? Why?
5. Find the share equation for this model. Are the parameters of the production function identified? Is the productivity process identified?

Problem 2 - Estimation of Production Functions

This question asks you to use several different methodologies to estimate production functions. I have two goals in mind. First, I'd like you to think carefully about different estimation techniques, including the standard panel data techniques that we did not cover in so much detail. Second, I want you to familiarize yourselves with programming GMM estimators. This will come in handy later in the course.

The problem set's companion data file is the replication dataset from Doraszelski and Jaumandreu (2013). It covers 10 years of data for Spanish firms in 18 different industries.

1. Load the dataset and report sample statistics (number of observations, mean, median, standard deviation, 25th and 75th percentiles) for output, investment, capital, total effective hours (our measure of labor), and intermediate consumption (our measure of materials expenditure). Also, report the number of firms per industry-year, and the count of firms with 0 investment/labor/materials use per industry-year.
2. Compare the aforementioned sample statistics to those obtained using only the balanced panel. How do they differ? What does this suggest?

3. Using only the balanced panel, assume that the production function is Cobb-Douglas in capital, labor and materials. Pick an industry of your choice, and estimate the following models: OLS, Fixed Effects, First differences, Long Differences (5 years) and Random Effects. Perform a Hausman test of Random Effects vs Fixed Effects. What do you learn from this exercise?

4. Repeat the previous exercise with the unbalanced panel (using the same industry). What do you learn from comparing the results?

5. Using the unbalanced panel for the industry of your choice, estimate the following models. Discuss the differences in results across models.

   - Arellano and Bond (1991). Hint: xtabond in Stata.
   - Blundell and Bond (1999). Hint: xtdpdsys in Stata.
   - Olley and Pakes (1996). Hint: opreg in Stata. Bonus: Can you code the one step estimator (and its SEs) in Matlab/Python/R?
   - Levinsohn and Petrin (2006). Hint: levpet in Stata. Bonus: Can you code the one step estimator (and its SEs) in Matlab/Python/R?
   - Ackerberg, Caves and Frazer (2015) (report point estimates and standard errors (bootstrap or GMM) for the relevant elasticities)
   - Gandhi, Navarro and Rivers (2017) (report point estimates and standard errors (bootstrap or GMM) for the relevant elasticities)

   You will probably need to use Matlab/Python/R for the last two estimators. If you are not familiar with any of them, I strongly suggest picking one and diving in with enthusiasm, as the next problem set will be more challenging than this one (and knowing how to program well is a critical skill for doing research in Empirical IO). Also, while these programs usually have their own optimization packages, I suggest learning how to use Knitro for this problem set. The marginal effort is very low, and the returns in the future will be large.

   Note: If you are having computational issues with the GMM minimization problems, consider programming the gradient and passing the function into the solver. You'll need the gradient to calculate GMM standard errors anyway.

   Note 2: I will set office hours for you to come and discuss any issues you are having.

   Note 3: Working in groups is encouraged, but all members of the group should program their own codes, write their own answers, and hand them in individually. Please present answers in tables. They don't have to be fancy, but handing in a bunch of Stata printouts will not do. Also, attach your code as an appendix.