Industrial Organization II (ECO 2901)
Winter 2018. Victor Aguirregabiria

Problem Set

Due on Sunday, March 18th, 2018

[**A. The Data**] The Stata datafile `mines_data_problemset_2018.dta` contains annual information on outputs, inputs, and input prices from 334 copper mines for the period 1992-2010 (19 years). The set of inputs includes labor (number of workers), capital (maximum capacity in physical units), electricity consumption (in physical units), fuel consumption (in physical units), and materials (in dollars). There is also information on the mine's annual ore grade (percentage of copper in mined ore), ore reserves, and input prices for labor, electricity, and fuel. The production of copper has two main stages: extraction of ore from the mine; and processing of the ore (in multiple sub-stages) until obtaining pure copper (99.9% of purity). At the extraction stage, it is important to distinguish between open pit and underground mines. At the processing stage, there are two main technologies: Pyrometallurgical (that here and in the dataset we denote as 'Conc'), and Hydrometallurgical (that we denote as 'SXEW'). Most mines use only one the two technologies, but a few mines use both. The datasets contains annual information on outputs, inputs, and input prices for each of the two technologies.

We assume that the production function is Cobb-Douglas in terms of labor, capital, fuel, electricity, materials, and ore grade. The parameters in the production function can be different for the 'Conc' and the 'SXEW' processing technologies.

Question 1. Obtain the Pooled OLS estimator and the Fixed Effects (Within-Groups) estimator for the following production functions: (a) with only labor and capital as inputs and restricting the two technologies to have the same parameters; (b) with labor, capital, electricity, fuel, materials, and grade as inputs and restricting the two technologies to have the same parameters; (c) with only labor and capital as inputs and allowing for different parameters in the two technologies; and (d) with labor, capital, electricity, fuel, materials, and grade as inputs and allowing for different parameters in the two technologies. Include time dummies in all the estimations. Comment the results.

[**B. First order conditions of optimality for variable inputs**] Suppose that these mines are price takers in the input markets. Consider that the variable inputs are labor, fuel, electricity, and materials. Ignoring investment and entry/exit decisions, the value of mine i at period t is

$$V_{it} = P_t \; q_{it} - \mathbf{w}_{it} \; \mathbf{x}_{it} + \beta \; EV_{i,t+1}(r_{it} - q_{it})$$

where P_t is the price of copper in the international market; q_{it} is mine i's output; $\mathbf{x}_{it} \equiv (n_{it}, f_{it}, e_{it}, m_{it})$ and $\mathbf{w}_{it} \equiv (w_{it}^n, w_{it}^f, w_{it}^e, w_{it}^m)$ are vectors with the amounts and the prices of variable inputs, respectively; β is the discount factor; $EV_{it}(.)$ is the continuation value function that depends on the remaining reserves of copper in the mine, $r_{i,t+1} = r_{it} - q_{it}$. The first order condition of optimality with respect to the variable inputs implies that:

$$\frac{\partial P_t}{\partial q_{it}} \frac{\partial q_{it}}{\partial \mathbf{x}_{it}} + P_t \; \frac{\partial q_{it}}{\partial \mathbf{x}_{it}} - \mathbf{w}_{it} - \beta \; \frac{\partial EV_{i,t+1}}{\partial r_{i,t+1}} \; \frac{\partial q_{it}}{\partial \mathbf{x}_{it}} = 0$$

or equivalently,

$$\frac{\partial q_{it}}{\partial \mathbf{x}_{it}} \left[\frac{\partial P_t}{\partial q_{it}} + P_t - \beta \; \frac{\partial EV_{i,t+1}}{\partial r_{i,t+1}} \right] = \mathbf{w}_{it}$$

The term $\frac{\partial P_t}{\partial q_{it}} + P_t$ is the marginal revenue. The term $-\beta \frac{\partial EV_{i,t+1}}{\partial r_{i,t+1}}$ represents the (negative) effect of current output on the future value of the mine through the depletion of reserves. Note that the term $\left[\frac{\partial P_t}{\partial q_{it}} + P_t - \beta \; \frac{\partial EV_{i,t+1}}{\partial r_{i,t+1}} \right]$ is the same for the marginal conditions of all the variable inputs. Therefore, though the model is dynamic, the following "static" conditions for cost minimization hold:

$$\frac{\partial q_{it} / \partial \mathbf{x}_{it}}{\partial q_{it} / \partial n_{it}} \quad = \quad \frac{\mathbf{w}_{it}}{w_{it}^n}$$

Given the Cobb-Douglas production function, we obtain:

$$\frac{\alpha_e}{\alpha_n} = \frac{w_{it}^e \; e_{it}}{w_{it}^n \; n_{it}}; \quad \frac{\alpha_f}{\alpha_n} = \frac{w_{it}^f \; f_{it}}{w_{it}^n \; n_{it}}; \quad \frac{\alpha_m}{\alpha_n} = \frac{w_{it}^m \; m_{it}}{w_{it}^n \; n_{it}}$$

Question 2. Given these expressions, obtain simple estimates for the parameters $\dfrac{\alpha_e}{\alpha_n}$, $\dfrac{\alpha_f}{\alpha_n}$, and $\dfrac{\alpha_m}{\alpha_n}$ using mean of the expenditure ratios. Obtain estimates pooling together the two technologies, and also allowing for different parameters in the two technologies. Comment the results.

[C. Acknowledging the existence of unobserved components in the cost of variable inputs] Note that the model in Question 2 implies that the expenditure shares (e.g., w_{it}^e $e_{it}/w_{it}^n \; n_{it}$) should be constant across mines and over time because the α parameters are constant. This restriction is obviously rejected by the data. A simple way to extend the model to explain this discrepancy is to consider that the is a component in the unit cost of a variable input that is not included in the expenditure that is observed by the researcher. More specifically, suppose that the unit cost of a variable input x_{it} (i.e., where x represents either labor, fuel, electricity, or materials) is $\exp\{\tau_{it}^{x}\}w_{it}^{x}$, where τ_{it}^{x} is a random variable that is unobservable to the researcher. Then, for any variable input x we have that $\dfrac{\alpha_x}{\alpha_n} = \dfrac{\exp\{\tau_{it}^{x}\}w_{it}^{x}\ x_{it}}{\exp\{\tau_{it}^{n}\}w_{it}^{n}\ n_{it}}$, or in logarithms:

$$\ln\left(w_{it}^{x}\ x_{it}/w_{it}^{n}\ n_{it}\right) = \theta_x + \left(\tau_{it}^{n} - \tau_{it}^{x}\right)$$

where $\theta_x \equiv \ln\left(\dfrac{\alpha_x}{\alpha_n}\right)$. We assume that the error terms $\tau_{it}^{n} - \tau_{it}^{x}$ have zero mean.

Question 3. Given this equation, obtain estimates of the parameters θ_e, θ_f, and θ_m, and of the standard deviation of the unobservables $(\tau_{it}^{n} - \tau_{it}^{e})$, $(\tau_{it}^{n} - \tau_{it}^{f})$, and $(\tau_{it}^{n} - \tau_{it}^{m})$. Obtain these estimates pooling together the two technologies, and also allowing for different parameters in the two technologies. Comment the results.

[D. 'Aggregating' variable inputs in the production function]. Given the results from Question 3, consider the following representation of the Cobb-Douglas production function:

$$\ln q_{it} = \alpha_n\ n_{it}^{*} + \alpha_k\ \ln k_{it} + \alpha_g\ \ln g_{it} + \omega_{it}$$

where $n_{it}^{*} \equiv \ln n_{it} + \exp\{\widehat{\theta}_e\}\ \ln e_{it} + \exp\{\widehat{\theta}_f\}\ \ln f_{it} + \exp\{\widehat{\theta}_m\}\ \ln m_{it}$, and $\widehat{\theta}_e$, $\widehat{\theta}_f$, and $\widehat{\theta}_m$ represents the estimates in Question 2.

Question 4. Using this specification, obtain the Fixed Effect (Within-Groups) estimator of the parameters α_n, α_k, and α_g: (a) restricting the two technologies to have the same parameters; and (b) allowing for different parameters in the two technologies. Include time dummies in all the estimations. Comment the results.

Question 5. Consider the specification of the production in Question 4. Obtain the following estimators of the parameters α_n, α_k, and α_g: (i) Blundell-Bond estimator with non-serially correlated transitory shock; (ii) Blundell-Bond estimator with AR(1) transitory shock; (iii) Olley-Pakes (using $\Delta \ln k_{it}$ as investment); and (iv) Levinshon-Petrin using materials. Obtain estimates both restricting the two technologies to have the same parameters, and allowing for different parameters. Comment the results. Based on these results, select your preferred estimates of the production function parameters. Explain your choice.

[E. Decomposition of Aggregate Industry TFP growth]. Following (among others) Foster, Haltiwanger, and Syverson (AER, 2008), define the industry-level productivity at period t as:

$$\ln A_t \equiv \sum_{i=1}^{N} s_{it} \ \omega_{it}$$

where $s_{it} = q_{it}/Q_t$ is the share of mine i in total industry output at year t. Consider the following decomposition of industry productivity growth.

$$\Delta \ln A_t \ = \ \underbrace{\sum_{i=1}^{N} s_{i,t-1} \ \Delta\omega_{it}}_{\text{Within}} \ + \ \underbrace{\sum_{i=1}^{N} \Delta s_{it} \ \omega_{i,t-1}}_{\text{Reallocation}}$$

Also, the reallocation component of productivity growth can be decomposed into three components: the contribution of continuing firms ($s_{i,t-1} > 0$ and $s_{it} > 0$), new entrants ($s_{i,t-1} = 0$ and $s_{it} > 0$), and exiting firms ($s_{i,t-1} > 0$ and $s_{it} = 0$).

$$\underbrace{\sum_{i=1}^{N} \Delta s_{it} \ \omega_{i,t-1}}_{\text{Reallocation}} =$$

$$= \underbrace{\sum_{s_{i,t-1}>0 \ \& \ s_{it}>0} \Delta s_{it} \ \omega_{i,t-1}}_{\text{Continuing}} + \underbrace{\sum_{s_{i,t-1}=0 \ \& \ s_{it}>0} \Delta s_{it} \ \omega_{i,t-1}}_{\text{New Entrants}} + \underbrace{\sum_{s_{i,t-1}>0 \ \& \ s_{it}=0} \Delta s_{it} \ \omega_{i,t-1}}_{\text{Exits}}$$

Question 6. Based on your estimates in Question 5, obtain the estimated Total Factor Productivity, ω_{it}, for every observation in the sample. Then, obtain the time-series of the industry productivity growth $\Delta \ln A_t$ and of its four components in the previous decomposition: Within, Reallocation-Continuing, Reallocation-Entrants, and Reallocation-Exits. Present figures with these time series. Comment the results.

Question 7. [Calculating Marginal Costs]. The variable cost function of a mine is the minimum cost of the variable inputs to produce an amount of output, taking as given input prices, the τ_{it}^{x} costs, capital stock, ore grade, and total factor productivity.

(a) Show that the marginal cost function of this model has the following form:

$$MC_{it} = c_{it} \ q_{it}^{(1-\alpha)/\alpha}$$

where c_{it} represents the exogenous component of the marginal cost function; and $\alpha \equiv \alpha_n + \alpha_e + \alpha_f + \alpha_m$.

(b) Obtain the expression of c_{it} in terms of variable input prices, $\tau's$, capital stock, ore grade, TFP, and production function parameters.

(c) Show that, given the estimation results in Questions 3 and 5, all the parameters and the variables that enter in the variable c_{it} are known to the researcher except for the labor cost variable τ_{it}^n. More specifically, show that:

$$\ln c_{it} = \ln \widetilde{c}_{it} + \alpha \ \tau_{it}^n$$

where $\ln \widetilde{c}_{it}$ is known given the estimation results in Questions 3 and 5.

Question 8. [Decomposition of the growth in industry Marginal Cost]. Assume (for the moment) that $\tau_{it}^n = 0$, and use $\ln \widetilde{c}_{it} + (1-\alpha)/\alpha \ \ln q_{it}$ to measure $\ln MC_{it}$. Repeat the decomposition exercise in Question 6 but now for the evolution of the industry average marginal cost. That is, define the industry marginal cost at period t as:

$$\ln MC_t \equiv \sum_{i=1}^{N} s_{it} \ \ln MC_{it}$$

and decompose the evolution of $\ln MC_t$ into four components: within, reallocation-continuing, reallocation-entrants, and reallocation-exits. Present figures with these time series. Comment the results.

[F. Equilibrium and Counterfactual Experiments]. We are interested in studying how different demand shocks (e.g., increasing demand from China) and supply shocks (e.g., changes in input prices, TFP, depletion of ore grade) have contributed to the observed evolution of the price of copper and the aggregate world production during the sample period . Question 10 below describes the specific counterfactual experiments that we want to obtain. To implement these counterfactual experiments, we need to complete the model with an specification of the demand function, and with an equilibrium model for firms' quantities.

We consider that the equation that describes the world demand for copper is isoelastic: $\ln Q_t = \gamma \ \ln P_t + \varepsilon_t$, where ε_t is an aggregate demand shock. Based on empirical evidence and estimates from other studies, we consider that the demand-price elasticity is $\gamma = -0.4$. It is well-known that the demand for copper is very inelastic because it basically does not have close substitutes in most of its applications (aluminium is the closer substitute).

For the equilibrium, we assume that firms compete ala Nash-Cournot. For the sake simplicity, in this problem set we make several (strong and not innocuous) simplifying assumptions: (1) we ignore that mines' output decisions are dynamic (i.e., we ignore here the depletion effect); and (2) we ignore the ownership structure of these mines (i.e., some companies own multiple mines). Under these conditions, the first order condition of optimality for the amount of output of a mine (i.e., marginal revenue equal to marginal cost) becomes:

$$P_t(Q_t)\left[1+\frac{1}{\gamma}\ \frac{q_{it}}{Q_t}\right]=c_{it}\ q_{it}^{(1-\alpha)/\alpha}$$

where $P_t(Q_t) = Q_t^{1/\gamma}\exp\{-\varepsilon_t/\gamma\}$ is the inverse demand function. We use these conditions to compute the equilibrium of the model under a counterfactual scenario.

Question 9. Given these Nash-Cournot equilibrium conditions, estimate the value of τ_{it}^n as a residual in this equation. Given τ_{it}^n, obtain the values of τ_{it}^e, τ_{it}^f, and τ_{it}^m (remeber that we have already estimated $\tau_{it}^n - \tau_{it}^x$ in Question 3). According to the model, the logarithm of the unit cost of variable input x is $\ln w_{it}^x + \tau_{it}^x$. For every variable input x, obtain the variance ratio $Var(\tau_{it}^x)/Var(\ln w_{it}^x + \tau_{it}^x)$. Comment the results. Are these variance ratios plausible? Why/why not?

[G. Fixed-point algorithm for the computation of a Nash-Cournot equilibrium]. The system of N_t equations (as many as activity firms at period t) that characterize a Nash-Cournot equilibrium in this model at period t can be represented as the following fixed-point problem:

$$\ln \mathbf{q}_t = \mathbf{F}_t\left(\ln \mathbf{q}_t\right)$$

where $\ln \mathbf{q}_t$ is the $N_t \times 1$ vector with elements $\ln q_{it}$ for every mine i active at year t; and $\mathbf{F}_t\left(\ln \mathbf{q}_t\right)$ is a vector of N_t functions, $\{F_{1t}, F_{2t}, ..., F_{N_t t}\}$ such that, for every mine i:

$$F_{it}\left(\ln \mathbf{q}_t\right) \equiv \delta_{0it} + \delta_1 \ln Q[\ln \mathbf{q}_t] + \delta_2\left[1+\frac{1}{\gamma}\ \frac{q_{it}}{Q[\ln \mathbf{q}_t]}\right]$$

with $\delta_{0it} \equiv \alpha/(1-\alpha)[-\varepsilon_t/\gamma - \ln c_{it}]$, $\delta_1 \equiv \alpha/\gamma(1-\alpha)$, $\delta_2 \equiv \alpha/(1-\alpha)$, and $Q[\ln \mathbf{q}_t]$ is the industry output function, $Q[\ln \mathbf{q}_t] = \sum_{i=1}^{N_t}\exp\{\ln q_{it}\}$. We can obtain the equilibrium value $\ln \mathbf{q}_t$ by applying the Fixed-Point algorithm to the mapping $\ln \mathbf{q} = \mathbf{F}_t\left(\ln \mathbf{q}\right)$. Given an initial value of $\ln \mathbf{q}$, say $\ln \mathbf{q}^0$, at each iteration k we update this vector using $\ln \mathbf{q}^k = \mathbf{F}_t\left(\ln \mathbf{q}^{k-1}\right)$. We iterate until $\left\|\ln \mathbf{q}^k - \ln \mathbf{q}^{k-1}\right\|$ is smaller than a small constant.

[H. Counterfactual experiments]. By construction, the vector of quantities $\ln \mathbf{q}_t$ that we observe in the data should satisfy the equilibrium condition $\ln \mathbf{q}_t = \mathbf{F}_t(\ln \mathbf{q}_t)$ when the vector of functions $\mathbf{F}_t(.)$ are constructed using the estimated parameters and residuals (remember that we have constructed τ_{it}^n as a residual to satisfy these equilibrium conditions in the data). We are interested in computing the equilibrium $\ln \mathbf{q}_t^*$ associated to a vector of functions $\mathbf{F}_t^*(.)$ that consists in some "counterfactual" variation of the original mapping $\mathbf{F}_t(.)$ in the DGP. We are interested in three counterfactual experiments.

> (Experiment 1). No aggregate demand shocks. The equilibrium mapping $\mathbf{F}_t^*(.)$ is constructed using the counterfactual demand shocks $\varepsilon_t^* = \overline{\varepsilon}$ for every period t, where $\overline{\varepsilon} = T^{-1} \sum_{t=1}^{T} \varepsilon_t$.
>
> (Experiment 2). No aggregate cost shocks. The equilibrium mapping $\mathbf{F}_t^*(.)$ is constructed using the counterfactual cost shocks $\ln c_{it}^* = \ln c_{it} - \ln \overline{c}_t + \ln \overline{c}$ for every (i,t), where $\ln \overline{c}_t = N_t^{-1} \sum_{i=1}^{N_t} \ln c_{it}$ and $\ln \overline{c} = T^{-1} \sum_{t=1}^{T} \ln \overline{c}_t$.
>
> (Experiment 3). No cross-sectional heterogeneity in (exogenous) marginal costs. The equilibrium mapping $\mathbf{F}_t^*(.)$ is constructed using the counterfactual cost shocks $\ln c_{it}^* = \ln \overline{c}_t$ for every (i,t), where $\ln \overline{c}_t = N_t^{-1} \sum_{i=1}^{N_t} \ln c_{it}$.

Question 10. For each of these three experiments, compute the Nash-Cournot equilibrium at every period t and present figures for the time-series of the predicted $\{P_t^*, Q_t^*\}$. Comment the results. Based on these experiments, what are the main factors that explain the sharp increase in copper prices between 2003 and 2006?