Empirical Industrial Organization (ECO 310)
University of Toronto. Department of Economics
Fall 2019. Instructor: Victor Aguirregabiria

SOLUTION TO MIDTERM EXAM: Monday, October 28th, 2019

INSTRUCTIONS:
- **This is a closed-book exam.**
- **No study aids, including calculators, are allowed.**
- **Please, answer all the questions.**
- **Use the space provided between Questions to write your answers.**

TOTAL POINTS = 300

QUESTION 1 [150 points]. Consider the Cobb-Douglas production function $Y_{it} = A_{it}\ L_{it}^{\alpha_L}\ K_{it}^{\alpha_K}$.

Q1.1. [20 points] Write the equation for this production function as a linear regression (in logs).
ANSWER:

$$y_{it} = \alpha_L\ \ell_{it} + \alpha_K\ k_{it} + \omega_{it}$$

where: $y_{it} \equiv \ln(Y_{it})$; $\ell_{it} \equiv \ln(L_{it})$; $k_{it} \equiv \ln(K_{it})$; and $\omega_{it} \equiv \ln(A_{it})$.

Q1.2. [30 points] Why does the OLS method provide inconsistent estimates of parameters α_L and α_K?
ANSWER: Log-Total Factor Productivity (TFP), ω_{it}, is unobservable to the researcher. Therefore, it is the error term of this regression. A firm's profit maximization implies that – ceteris paribus – higher TFP is associated with a larger amount of labor and capital inputs. Therefore, we expect the regressors ℓ_{it} and k_{it} to be correlated with the error term ω_{it}. The OLS estimator is inconsistent when a regressor is correlated with the error term.

Q1.3. [30 points] Consider the method of instrumental variables (IV). Which are the conditions that a valid instrument should satisfy?
ANSWER: Let z_{1it} and z_{2it} be two observable variables that we want to use as instruments for ℓ_{it} and k_{it}. These instrumental variables should satisfy two conditions:

[No correlation with the error term] $Cov\left(z_{1it}, \omega_{it}\right) = Cov\left(z_{2it}, \omega_{it}\right) = 0$.

[Non zero correlation with the endogenous regressors] $Cov\left(z_{1it}, \ell_{it}\right) \neq 0$; $Cov\left(z_{2it}, \ell_{it}\right) \neq 0$; $Cov\left(z_{1it}, k_{it}\right) \neq 0$; $Cov\left(z_{2it}, k_{it}\right) \neq 0$.

Q1.4. [30 points] Describe a two step method to obtain the IV estimator.
ANSWER:

[Step 1]

- Run OLS regression of ℓ_{it} on z_{1it} and z_{2it}. Let $\widehat{\ell}_{it}$ be the fitted values from this regression.
- Run OLS regression of k_{it} on z_{1it} and z_{2it}. Let $\widehat{k}_{it}$ be the fitted values from this regression.

[Step 2]

- Run OLS regression of y_{it} on $\widehat{\ell}_{it}$ and $\widehat{k}_{it}$. This estimator is the IV estimator of α_L and α_K.

Q1.5. **[40 points]** You have a STATA datafile with the following variables: `id` (firm id number); `year` (year); `y` (firm's output); `n` (firm's labor); `k` (firm's capital). Write all the STATA command lines to implement the Fixed Effects estimator of the parameters α_L and α_K – with firm and time fixed effects.
ANSWER:

```
gen logy = ln(y)
gen logn = ln(n)
gen logk = ln(k)
xtset id year
xtreg logy logn logk i.year, fe
```

QUESTION 2 [150 points]. Consider a logit model for the demand of a differentiated product.

Q2.1. **[30 points]** Write the expression for the utility function of buying product (variety) j.
ANSWER: The utility for household h buying product j is:

$$U_{hj} = -\alpha\ p_j + X_{1j}\ \beta_1 + ... + X_{Kj}\ \beta_K + \xi_j + \varepsilon_{hj}$$

where: p_j = price; X_{1j} = observable characteristic 1; ...; X_{Kj} = observable characteristic K; ξ_j = unobservable characteristic; ε_{hj} = unobservable idiosyncratic taste of household h for variety j – logit error.

Q2.2. **[30 points]** For the logit model, write the expression for the market share of product j, as represented by s_j, as a function of prices and product characteristics (observable and unobservable)
ANSWER:

$$s_j = \frac{\exp\left\{-\alpha\ p_j + X_{1j}\ \beta_1 + ... + X_{Kj}\ \beta_K + \xi_j\right\}}{1 + \sum_{i=1}^{J} \exp\left\{-\alpha\ p_i + X_{1i}\ \beta_1 + ... + X_{Ki}\ \beta_K + \xi_i\right\}}$$

Q2.3. **[45 points]** Given the equation in Question 2.2, show that we can represent the logit model as a linear regression model. Write the expression for this regression model. What is the dependent variable? Which are the explanatory variables? What is the error term?
ANSWER: For the outside alternative, $j = 0$:

$$s_0 = \frac{1}{1 + \sum_{i=1}^{J} \exp\left\{-\alpha\ p_i + X_{1i}\ \beta_1 + ... + X_{Ki}\ \beta_K + \xi_i\right\}}$$

Therefore,

$$\ln\left(s_j/s_0\right) = \beta_p\ p_j + X_{1j}\ \beta_1 + ... + X_{Kj}\ \beta_K + \xi_j$$

This is a linear regression model with dependent variable $\ln\left(s_j/s_0\right)$, explanatory variables p_j, X_{1j}, ..., and X_{Kj}, and error term ξ_j.

Q2.4. **[45 points]** You have a STATA datafile for the J varieties of a product category – say automobiles – with the following variables: `id` (car id number); `q` (quantity sold); `p` (price); `index` (a quality index of the car from an expert report); `fuel` (fuel consumption). You know the value of H = Number of consumers in the market. Write all the STATA command lines to obtain the OLS estimator of the α and β parameters in the Logit model with product characteristics `p`, `index`, and `fuel`.

ANSWER: Suppose that $H = 100,000$.

```
gen H = 100000
gen s = ln(q/H)
egen sums = sum(s)
gen s0 = 1 - sums
gen y = ln(s/s0)
reg y p index fuel
```

END