

Industrial Organization II (ECO 3901)

Winter 2022. Victor Aguirregabiria

Problem Set

Due on Thursday, March 17th, 2022 [before 11:59pm]

INSTRUCTIONS.

1. Write your answers electronically in a word processor.
2. For the answers that involve coding, include in the document the code that you have used to obtain your empirical results, indicating the language of the code (e.g., R, Matlab, Julia, Python) as well as the version of the software/compiler you have used.
3. Convert the document with your answers and code to PDF format and submit the PDF online via Quercus.
4. Please, submit your completed problem set before 11:59pm on Thursday, March 17, 2022.
5. Problem sets should be written individually.

The total number of marks is 500.

This problem set deals with **solution, estimation, and simulation, of a (single-agent) dynamic structural model of firms' investment and entry and exit decisions.**

A. DATA

The Stata datafile `mines_data_problemset_2022.dta` contains annual information from 279 copper mines for the period 1992-2010 (19 years). It is apparently a balanced panel, but not all the mines are active every period. The following is a description of the variables.

Variable name	Description
<code>id</code>	: Mine identification number
<code>mine</code>	: Name of the mine
<code>year</code>	: Year [from 1992 to 2010]
<code>country</code>	: Country where the mine is located
<code>active</code>	: Binary indicator of the event "mine is active during the year "
<code>prod_tot</code>	: Annual production of pure copper of the mine [in thousands of tonnes]
<code>labor_n_tot</code>	: Total number of workers per year (annual equivalent)
<code>cap_tot</code>	: Measure of capital [maximum production capacity of the mine]
<code>invest</code>	: Investment in capacity
<code>lme_2010</code>	: Price of copper at London Metal Exchange market (deflated 2010)

Zero values for variables `prod_tot`, `labor_n_tot`, and `cap_tot` means that the mine was idle during that year – not actively extracting ore and producing copper. There are not missing values in this panel.

B. MAIN FEATURES OF THE MODEL

We consider a pretty standard dynamic structural model of investment under uncertainty. The following is a description of the model.

[1] Mines are indexed by i and time (years) by t . Variable k_{it} represents the installed capacity (or capital) of mine i at year t , and it corresponds to variable `cap_tot` in the Stata datafile. Binary variable x_{it} represents if the mine is active (producing) at year t ($x_{it} = 1$) or is inactive / not producing ($x_{it} = 0$). It corresponds to variable `active` in the Stata datafile. For an inactive mine, current capacity k_{it} is zero.

[2] Every year t , the manager of a mine makes two decisions. She decides the amount of output to produce and sell in the market, that we represent as q_{it} and corresponds to variable `prod_tot` in the Stata datafile. She also chooses investment in capacity, closing (exit), and reopening (entry). Variable I_{it} represents the mine's investment in capacity at year t , and it corresponds to variable `invest` in the Stata datafile. Variable I_{it} also implicitly provides information about the mine's closing and reopening decisions. If an incumbent mine at year t (i.e., a mine with $x_{it} = 1$, or equivalently, with $k_{it} > 0$) decides to continue active at $t + 1$, then $x_{i,t+1} = 1$, and this decision is equivalent to $I_{it} > -k_{it}$. If this incumbent decides to close, then $x_{i,t+1} = 0$, and this is equivalent to $I_{it} = -k_{it}$. If a mine that is temporarily closed at period t (i.e., a mine with $x_{it} = 0$, or equivalently, with $k_{it} = 0$) decides to remain closed, then $x_{i,t+1} = 0$, and this is equivalent to $I_{it} = 0$. If this mine decides to reopen, then $x_{i,t+1} = 1$, which is equivalent to $I_{it} > 0$. Accordingly, we can use I_{it} to represent investment, closing, and opening decisions.

[3] There is a one year time-to-build in investment, entry, and exit decisions. That is, these decisions at year t are not effective until year $t + 1$. The transition rule for capacity is:

$$k_{i,t+1} = k_{it} + I_{it}, \tag{1}$$

where I_{it} is restricted to $I_{it} \geq -k_{it}$, and we have that $I_{it} = -k_{it}$ is equivalent to $x_{i,t+1} = 0$.

[4] We make four main simplifying assumptions in the model for this problem set. First,

we assume that a firm's choice of output is a static decision.¹ Second, though copper is an homogenous product (a commodity), there is spatial differentiation due to transportation costs. We abstract from this spatial aspect. Third, though there is empirical evidence that some copper companies have market power, in this problem set we abstract from this relevant aspect and we assume that all the mines are price takers in the international metal market (e.g., in the London Metal Exchange). The price in the metal exchange is p_t , and it corresponds to variable `lme_2010` in the Stata datafile. From the point of view of a mine, this price is an exogenous state variable that follows a first order Markov process. Finally, for simplicity, we ignore the ownership structure of the mines and treat each mine as if it were an independent firm.

[5] Investment in capacity, closing a mine, and reopening a mine are dynamic and forward-looking decisions. A mine makes these decisions to maximize its expected and discounted stream of profits $\mathbb{E}_t(\sum_{s=0}^{\infty} \beta^s \Pi_{i,t+s})$, where $\beta \in (0, 1)$ is the discount factor.

[6] The profit function of a mine is:

$$\Pi_{it} = p_t q_{it} - VC_{it} - IC_{it}(I_{it}, k_{it}), \quad (2)$$

where VC_{it} is the variable cost function, IC_{it} is the investment cost function. This investment cost includes also as particular cases the cost of closing (when $I_{it} = -k_{it}$) and the cost of reopening (when $I_{it} > 0$ and $k_{it} = 0$). We specify these functions below.

[7] The state variables in the dynamic programming (DP) problem of a mine are $(p_t, k_{it}, \omega_{it})$ where p_t is the price of copper in the international market, k_{it} is the mine's capacity or capital stock at year t , and ω_{it} is a total factor productivity (TFP) shock.

[8] It is convenient to represent capacity and investment decisions using variables in logarithms. Define $lk_{it} \equiv \ln(1 + k_{it})$ and $a_{it} \equiv \ln(1 + I_{it}/(1 + k_{it}))$. Using these definitions, it is straightforward to show that the transition rule $k_{i,t+1} = k_{it} + I_{it}$ is equivalent to the transition rule:

$$lk_{i,t+1} = lk_{it} + a_{it} \quad (3)$$

[9] A mine's production technology can be described by a Cobb-Douglas production

¹In this model, we abstract from the firm's concern for the depletion of the mine's reserves. Note that the amount of reserves of a mine is a non-renewable asset. The choice of output implies a depletion in this amount of reserves. Therefore, there is a relevant dynamic aspect in the output choice of a mine. For simplicity, we ignore this depletion effect. A possible interpretation of this assumption is that all these mines have very large reserves relative to their production capacities.

function in labor, capital, and TFP: $q_{it} = \ell_{it}^{\alpha_\ell} k_{it}^{\alpha_k} \omega_{it}$, where α_ℓ and α_k are parameters.² Capital and TFP are fixed inputs, and labor is the only variable input. Let r_ℓ be the price of labor.³ Then, it is straightforward to show that minimization of labor costs for a given amount of output implies the following variable cost function:

$$VC_{it} = r_\ell \left[\frac{q_{it}}{k_{it}^{\alpha_k} \omega_{it}} \right]^{1/\alpha_\ell} \quad (4)$$

[10] Specification of investment, reopening, and closing costs.

$$IC_{it} = \begin{cases} \theta^{XC} - \theta^{IC-} a_{it} + \varepsilon_{it}(a_{it}) & \text{if } lk_{it} > 0 \text{ \& } a_{it} = -lk_{it} \\ -\theta^{IC-} a_{it} + \varepsilon_{it}(a_{it}) & \text{if } lk_{it} > 0 \text{ \& } -lk_{it} < a_{it} \leq 0 \\ \theta^{IC+} a_{it} + \varepsilon_{it}(a_{it}) & \text{if } lk_{it} > 0 \text{ \& } a_{it} \geq 0 \\ 0 + \varepsilon_{it}(a_{it}) & \text{if } lk_{it} = 0 \text{ \& } a_{it} = 0 \\ \theta^{EC} + \theta^{IC+} a_{it} + \varepsilon_{it}(a_{it}) & \text{if } lk_{it} = 0 \text{ \& } a_{it} > 0 \end{cases} \quad (5)$$

where θ^{IC+} , θ^{IC-} , θ^{EC} , and θ^{XC} are parameters that represent, respectively, unit costs of positive investments, unit costs of negative investments, fixed entry cost, and fixed exit cost. Variables $\varepsilon_{it}^{IC}(a)$ are unobservable state variables that are *i.i.d.* over mines, time, and choice alternatives a , and they have a extreme value type I distribution.

C. ROAD MAP.

The final goal of this problem set is to obtain estimates of the structural parameters in the Investment Cost function. With this goal in mind, we proceed as follows. In **Question 1**, we derive a simple expression for a mine's variable profit function that comes from the combination of a Cobb-Douglas production function and an assumption of perfect competition in the copper international market. In **Question 2**, we construct the variables that we will use in our empirical analysis and present some figures and descriptive statistics. In **Question 3**, we obtain reduced form estimates of the transition probability functions of the exogenous state variables (tfp and price) and of the conditional choice probabilities

²In this problem set, we abstract from other inputs which are observable in the original dataset and important in this industry, such as electricity, fuel, materials, reserves, and ore grade. For this problem set, we want to keep the dynamic programming problem as simple and parsimonious as possible.

³Here we also abstract from heterogeneity across firms or over time in the price of labor.

(CCPs) for investment conditional on the state variables. In **Questions 4 and 5**, we use these reduced form estimates to construct a matrix of transition probabilities for the vector of state variables and a matrix of choice probabilities for the decision variables conditional on the state variables. In **Question 6**, we use these matrices and a value for the time discount factor parameter to construct vectors of present values for each component of the profit function. In **Question 7**, we use these estimated present values to obtain estimates of the structural parameters in the investment cost function. In **Question 8**, we implement a simple (but imperfect) counterfactual experiment that tries to measure the effect of price uncertainty on mines' investment decisions.

QUESTION 1 [40 points]. Derive the following equations under the condition that a mine is a price taker in the international copper market (perfect competition), labor is the only variable input, and capital/capacity is a fixed input.

Question 1(A) [20 points] Obtain the expression for the optimal amount of output q_{it} as a function of the state variables $(p_t, lk_{it}, \omega_{it})$.

Question 1(B) [20 points] Suppose that $\alpha_\ell = \alpha_k = 0.5$. Particularize the result in Question 1(A) to these values of α_ℓ and α_k . Plug this expression of the equilibrium value of q_{it} into the variable profit $p_t q_{it} - VC_{it}(q_{it})$ to obtain the following expression for the **indirect profit function**:

$$VP(p_t, k_{it}, \omega_{it}) = \frac{1}{4 r_\ell} k_{it} (p_t \omega_{it})^2 \quad (6)$$

QUESTION 2 [50 points]. In the dataset, variable `prod_tot` is q_{it} ; `labor_n_tot` is ℓ_{it} ; `cap_tot` is k_{it} ; and `lme_2010` is p_t .

Question 2(A) [25 points] Construct the following variables.

- (i) `a_invest`. It represents the variable $a_{it} \equiv \ln(1 + I_{it}/(1 + k_{it}))$.
- (ii) `active_next`. It represents next year value of the binary variable `active`. Therefore, (`active`=1 & `active_next`=0) means the closing of a mine, and (`active`=0 & `active_next`=1) means the reopening of a mine.
- (iii) `tfp`. It represents the TFP variable ω_{it} . It is equal to $\omega_{it} = q_{it}/\sqrt{\ell_{it} k_{it}}$. This variable contains missing values at every observation where the mine is not active (such that $\ell_{it} = k_{it} = 0$).
- (iv) `lkdisc`. It is a discretized version of variable $\ln(1 + k_{it})$. We consider the following discretization: `lkdisc` = `round(ln(1+cap_tot),0.05)`.

(v) `adisc`. It is a discretization of variable `a_invest` (or a_{it}). We construct this variable as follows: $\text{adisc}_{it} = \text{lkdisc}_{i,t+1} - \text{lkdisc}_{it}$, or using Stata syntax, `adisc = lkdisc[_n+1] - lkdisc`.

(vi) `tfpdisc`. It is a discretized version of TFP. We consider the following discretization: `tfpdisc = round(tfp, 0.05)`.

(viii) `lpdisc`. It is a discretized version of the logarithm of price, $\ln(\text{lme})$. We consider the following discretization: `lpdisc = round(ln(lme_2010), 0.1)`.

Question 2(B) [25 points]

(i) Present a table with means and the standard deviations of the variables `lkdisc`, `adisc`, `tfpdisc`, `lpdisc`, `active`, and `active_next`.

(ii) Present a table with the transition probabilities from `active` to `active_next`.

(iii) Present figures with the histograms of the variables `lkdisc` and `tfpdisc` for the subsample of observations where the `active=1`.

(iv) Present a figure with the histogram of variable `adisc` for the subsample of observations where the `active_next=1`.

(v) Present a figure with the time series of `lpdisc`.

*** For the rest of this problem set, we use only the discrete variables `lkdisc`, `adisc`, `tfpdisc`, `lpdisc`, `active`, and `active_next`.

QUESTION 3 [50 points]. Obtain estimates for the following reduced form equations. Interpret the results.

Question 3(A) [10 points] Logit model for the stay/closing decision: dependent variable `active_next`; explanatory variables `lkdisc`, `tfpdisc`, and `lpdisc`; using the subsample of observations with `active=1`.

Question 3(B) [10 points] Logit model for the reopening decision: dependent variable `active_next`; explanatory variable `lpdisc`; using the subsample of observations with `active=0`.

Question 3(C) [10 points] A linear regression model for the `adisc` investment decision using the subsample of observations with `active=1` and `active_next=1` and with explanatory variables `lkdisc`, `tfpdisc`, and `lpdisc`.

Instead of this regression model, you can estimate a Multinomial Logit model where the dependent variable is `adisc` and the explanatory variables are `lkdisc`, `tfpdisc`, and `lpdisc`. Using this Multinomial Logit (instead of a regression model)

will facilitate your answer to Question 4(C), as you will not need using Tauchen's method.

And a different linear regression model for the `adisc` investment decision using the subsample of observations with `active=0` and `active_next=1` and with explanatory variable `lpdisc`.

Similarly as above, instead of this regression model, you can estimate a Multinomial Logit model where the dependent variable is `adisc` and the explanatory variable is `lpdisc`. Using this Multinomial Logit (instead of a regression model) will facilitate your answer to Question 4(C), as you will not need using Tauchen's method.

Question 3(D) [10 points] Linear AR(1) for the stochastic process of `tfp`: dependent variable `tfpdisc`; explanatory variable the lagged value of `tfpdisc`; using the subsample of observations with `active=1` and `active[t-1]=1`.

Question 3(E) [10 points] Linear AR(1) for the stochastic process of log-price: dependent variable `lpdisc`; explanatory variable the lagged value of `lpdisc`.

QUESTION 4 [90 points]. Given the reduced form estimates in Question 3, obtain the following matrices of transition probabilities and conditional choice probabilities. For Question 4(A), I provide here a detailed explanation for the construction of the transition probability. The construction of the other probabilities should apply the same procedure.

Question 4(A) [20 points] Using the estimated reduced form model in Question 3(D), obtain the transition probability matrix for `tfpdisc` (\mathbf{F}_{tfp}).

- Suppose that `tfpdisc` can take any value within the uniform grid between 0.0 and 9.0 with step 0.05, that is the grid of points $\{0.0, 0.05, 0.10, \dots, 0.85, 0.90\}$. Let $K_{\text{tfp}} = 19$ be the number of points in this grid. And let `tfpvals` be the $K_{\text{tfp}} \times 1$ vector with the values in this grid.

- Let `thres` be a $(K_{\text{tfp}} - 1) \times 1$ vector with the following "threshold" values:

```
thres[1]=(tfpvals[1]+tfpvals[2])/2;
thres[2]=(tfpvals[2]+tfpvals[3])/2;
...
thres[K_tfp-1]=(tfpvals[K_tfp-1]+tfpvals[K_tfp])/2;
```

- The matrix of transition probabilities \mathbf{F}_{tfp} has dimension $K_{\text{tfp}} \times K_{\text{tfp}}$. Element $F_{\text{tfp}}[\text{row}, \text{col}]$ in this matrix is equal to $\Pr(\text{tfpdisc}_{t+1} = \text{tfpvals}[\text{col}] \mid \text{tfpdisc}_t = \text{tfpvals}[\text{row}])$.

• How do we obtain these probabilities? We apply a popular method proposed by George Tauchen (*Economics Letters*, 1986). Let $\hat{\beta}_0$ and $\hat{\beta}_1$ be the estimated intercept and slope in the reduced-form regression in Question 3(D). And let $\hat{\sigma}$ be the estimate of the standard deviation of the residuals in that regression. We assume logistic distribution for the shock in the AR(1) process. Then:

$$F_{\text{tfp}}[\text{row}, \text{col}] = \begin{cases} \Lambda\left(\frac{\text{thres}[1] - \hat{\beta}_0 - \hat{\beta}_1 \text{tfpvals}[\text{row}]}{\hat{\sigma}}\right) & \text{if } \text{col} = 1 \\ \Lambda\left(\frac{\text{thres}[\text{col}] - \hat{\beta}_0 - \hat{\beta}_1 \text{tfpvals}[\text{row}]}{\hat{\sigma}}\right) & \text{if } 2 \leq \text{col} \leq K_{\text{tfp}} - 1 \\ -\Lambda\left(\frac{\text{thres}[\text{col}-1] - \hat{\beta}_0 - \hat{\beta}_1 \text{tfpvals}[\text{row}]}{\hat{\sigma}}\right) & \\ 1 - \Lambda\left(\frac{\text{thres}[K_{\text{tfp}}-1] - \hat{\beta}_0 - \hat{\beta}_1 \text{tfpvals}[\text{row}]}{\hat{\sigma}}\right) & \text{if } \text{col} = K_{\text{tfp}} \end{cases} \quad (7)$$

where $\Lambda(\cdot)$ is the CDF of a logistic random variable, i.e., the logistic function $\Lambda(u) = e^u / (1 + e^u)$.

• If the implementation is correct, the sum of elements in a row should be equal to one. Check this in your code.

Question 4(B) [20 points] Using the estimated reduced form model in Question 3(E), obtain the transition probability matrix for `lpdisc` ($\mathbf{F}_{\text{lpdisc}}$).

The support set for `lpdisc` is a uniform grid of points between 4.5 and 5.9 with step equal to 0.1, such that $K_{\text{price}} = 15$. Obtain the transition probability matrix for `lpdisc` ($\mathbf{F}_{\text{price}}$). This matrix has dimension $K_{\text{price}} \times K_{\text{price}}$.

Question 4(C) [20 points] Using the estimated reduced form models in Questions 3(A), 3(B), and 3(C), obtain the conditional choice probability (CCP) matrix for investment in capacity `adisc` conditional on all the state variables (`lkdisc`, `lpdisc`, `tfpdisc`), $\mathbf{P}_{\text{adisc}}$.

Suppose that the state variable `lkdisc` can take any value within the uniform grid between 0.0 and 7.3 with step 0.1, that is the grid of points $\{0.0, 0.1, 0.2, \dots, 7.2, 7.3\}$. Let $K_{\text{lkdisc}} = 74$ be the number of points in this grid. This means that the whole vector of state variables (`lkdisc`, `lpdisc`, `tfpdisc`) has a support set with $K_{\mathbf{x}} = K_{\text{lk}} * K_{\text{price}} * K_{\text{tfp}} = 74 * 15 * 19 = 21,090$ points. Suppose that the decision variable `adisc` can take any value within the uniform grid between -6.0 and $+6.0$ with step 0.1, that is the grid of points $\{-6.0, -5.9, \dots, -0.1, 0.0, 0.1, \dots, 5.9, 6.9\}$. Let $K_{\text{adisc}} = 121$ be the number of points in this grid. Therefore, CCP matrix $\mathbf{P}_{\text{adisc}}$ has dimension $K_{\mathbf{x}} \times K_{\text{adisc}}$, that is, $21,090 \times 121$.

Note that matrix $\mathbf{P}_{\text{adisc}}$ should incorporate different restrictions that imply that many elements in this matrix are equal to zero. For instance: it is not possible to make a disinvestment larger than current capital stock; it is not possible to make a positive investment

that generates a capital stock next period larger than the upper limit in the support of this variable.

Question 4(D) [10 points] Taking into account the transition rule $lk_{i,t+1} = lk_{it} + a_{it}$, construct the transition matrices for capital lk for each possible value of investment a . That is, $\mathbf{F}_{lk|a}$ is the $K_{lk} \times K_{lk}$ transition matrix from `lkdisc` given a particular value of `adisc`. There are as many of these matrices as values in the support set of `adisc`. Since the transition rule is deterministic, these transition matrices are also deterministic, i.e., each row contains an element equal to 1 and the rest of the elements are zeroes. Also, note that given the bounded support for `lkdisc`, there is a truncation at the maximum and minimum values of `lkdisc`.

QUESTION 5 [60 points].

Question 5(A) [20 points] Using the matrices \mathbf{F}_{tfp} , $\mathbf{F}_{\text{price}}$, $\mathbf{F}_{lk|a}$ that you have obtained in Question 4, construct the conditional transition matrix for the vector of the three state variables, $\mathbf{F}_{x|a}$. Note that there is one of these transition matrices for each possible value in the support set of `adisc`.

Hint. $\mathbf{F}_{x|a}$ is a matrix with dimension $K_x \times K_x$, with $K_x = K_{lk} * K_{\text{price}} * K_{\text{tfp}}$. Element $\mathbf{F}_{x|a}[\text{row}, \text{col}]$ in this matrix is the probability:

$$\mathbf{F}_{x|a}[\text{row}, \text{col}] = \Pr \left(\begin{array}{l} (\text{lkdisc}_{t+1}, \text{lpdisc}_{t+1}, \text{tfpdisc}_{t+1}) = \\ (\text{lkval}[\text{col}], \text{lpval}[\text{col}], \text{tfpval}[\text{col}]) \\ | \\ (\text{lkdisc}_t, \text{lpdisc}_t, \text{tfpdisc}_t) = \\ | (\text{lkval}[\text{row}], \text{lpval}[\text{row}], \text{tfpval}[\text{row}]), \\ \text{adisc}_t = a \end{array} \right) \quad (8)$$

where $(\text{lkval}[\text{col}], \text{lpval}[\text{col}], \text{tfpval}[\text{col}])$ is the value in the support of the state variables that corresponds to column "col", and similarly $(\text{lkval}[\text{row}], \text{lpval}[\text{row}], \text{tfpval}[\text{row}])$ is the value in the support of the state variables that corresponds to row "row".

You can use loops to fill one by one the elements in matrix $\mathbf{F}_{x|a}$ given the matrices you have constructed in Question 4. Alternatively, it is possible to use matrix operations (Kronecker products combined with element-by-element products) to obtain directly $\mathbf{F}_{x|a}$ from \mathbf{F}_{tfp} , $\mathbf{F}_{\text{price}}$, and $\mathbf{F}_{lk|a}$. Using loops can be safer and easier to code, though the computation time will be longer.

Question 5(B) [10 points] Given the conditional choice transition matrices $\mathbf{F}_{x|a}$ and the

CCP matrix $\mathbf{P}_{\text{adisc}}$, obtain the unconditional transition matrix \mathbf{F}_x^P which is defined as follows:

$$\mathbf{F}_x^P = \sum_a \mathbf{F}_{x|a} * \mathbf{P}_{\text{adisc}}(a) \quad (9)$$

where $*$ is the element-by-element or Hadamard product, and $\mathbf{P}_{\text{adisc}}(a)$ represents the a -th row of the CCP matrix $\mathbf{P}_{\text{adisc}}$.

NOTE: For the Questions below, it is convenient to define the following $K_x \times 1$ vector of per-period profits, Π^P :

$$\Pi^P = \sum_a \mathbf{P}_{\text{adisc}}(a) * \Pi(a) \quad (10)$$

where $*$ is the element-by-element or Hadamard product, $\mathbf{P}_{\text{adisc}}(a)$ represents the a -th row of the CCP matrix $\mathbf{P}_{\text{adisc}}$, and $\Pi(a)$ is the $K_x \times 1$ vector of per-period profits when choice alternative is a .

Question 5(C) [30 points] Valuation matrix. Define the following matrix:

$$\Lambda_{PV} = (I - \beta \mathbf{F}_x^P)^{-1} \quad (11)$$

where I is the identity matrix, β is the time discount factor, and \mathbf{F}_x^P is the unconditional transition matrix defined in Question 5(B) above. Calculate matrix Λ_{PV} using $\beta = 0.95$.

Matrix Λ_{PV} is a Present Value operator. **And Π^P is the $K_x \times 1$ vector of per-period profits for each possible value of the state variables, that we have defined above.** Then, by pre-multiplying Π^P with matrix Λ_{PV} we get a vector of present values:

$$\mathbf{PV} = \Lambda_{PV} \Pi^P \quad (12)$$

Element $\mathbf{PV}[\text{row}]$ in vector \mathbf{PV} give us the present value of the stream of current and future profits in Π^P given that the current state is the one associated with the value of the state variables in row "row".

To provide some intuition, consider first the scalar case. Let π be a constant per period profit. Then, $PV = \sum_{t=0}^{\infty} \beta^t \pi$ is the present value of this permanent per-period profit over an infinite horizon. It should be clear that $PV = (1 - \beta)^{-1} \pi$, and $(1 - \beta)^{-1}$ is a particular case of matrix Λ_{PV} when there is only one possible state. Note, that we can also define the present value PV using the recursive expression $PV = \pi + \beta PV$. Now, consider the general

case with vector $\mathbf{\Pi}^P$ and matrix \mathbf{F}_x^P . The recursive expression for the vector of present values is:

$$\mathbf{PV} = \mathbf{\Pi}^P + \beta \mathbf{F}_x^P \mathbf{PV} \quad (13)$$

Solving for \mathbf{PV} in this recursive equation, we get:

$$\mathbf{PV} = (I - \beta \mathbf{F}_x^P)^{-1} \mathbf{\Pi}^P \quad (14)$$

Similarly, we can obtain the present value $\mathbf{\Pi}^P$ conditional on the current choice a as:

$$\mathbf{PV}(a) = \mathbf{\Pi}(a) + \beta \mathbf{F}_{x|a} \mathbf{\Lambda}_{PV} \mathbf{\Pi}^P \quad (15)$$

QUESTION 6 [40 points]. Vector with Present values of Variable profits. The profit function (without including the unobservable state variables $\varepsilon_{it}(a_{it})$) can be represented as:

$$\Pi_{it} = h(a_{it}, p_t, k_{it}, \omega_{it})' \boldsymbol{\theta} \quad (16)$$

where:

$$\boldsymbol{\theta} = (\theta_{vp}, \theta^{XC}, \theta^{IC-}, \theta^{IC+}, \theta^{EC})' \quad (17)$$

with $\theta_{vp} = 1/4r_\ell$, and

$$\begin{aligned} h(a_{it}, p_t, k_{it}, \omega_{it})' &= \\ &= (k_{it} \ (p_t \ \omega_{it})^2, \ -1\{k_{it} > 0 \ \& \ a_{it} = -k_{it}\}, \ 1\{k_{it} > 0 \ \& \ a_{it} < 0\}a_{it}, \\ &\quad -1\{a_{it} > 0\}a_{it}, \ -1\{k_{it} = 0 \ \& \ a_{it} > 0\}) \end{aligned} \quad (18)$$

For some descriptions below, it will be convenient to represent the 5 elements of the vector $h()$ as $h_{it}^1, h_{it}^2, \dots, h_{it}^5$.

Question 6(A) [20 points] For each of the 5 elements in the vector $h()$, construct a matrix with dimension $K_x \times K_a$ with the value of that component for every value of the state variables and the decision variables. That is, let \mathbf{H}^3 be the matrix for the third element, $1\{k_{it} > 0 \ \& \ a_{it} < 0\}a_{it}$. Each row in this matrix corresponds to a particular value in the support of the state variables, and each column corresponds to a particular value in the support of the decision variables. Then, element $\mathbf{H}^3[row, col]$ contains the value of $1\{k_{it} > 0 \ \& \ a_{it} < 0\}a_{it}$ when the value of the state variables corresponds to the value of these variables in row "row", and the value of the decision variable corresponds to the value of this variable in column "col".

Question 6(B) [20 points] Given the Present Value operator matrix Λ_{PV} that you have calculated in Question 5, and applying equation (15, **obtain the conditional choice present values for each of the 5 components in vector $h()$: that is, for $j = 1, 2, 3, 4, 5$:**

$$\mathbf{PV}^j(a) = \mathbf{H}^j(a) + \beta \mathbf{F}_{x|a} \Lambda_{PV} \mathbf{H}^{j,P} \quad (19)$$

with

$$\mathbf{H}^{j,P} = \sum_a \mathbf{P}_{\text{adisc}}(a) * \mathbf{H}^j(a) \quad (20)$$

QUESTION 7 [100 points]. Estimation.

Question 7(A) [20 points] Given the data and the vectors of present values in Question 6, explain how to obtain an estimator of the structural parameters in θ .

Question 7(B) [40 points] Present the code to implement that estimator.

Question 7(C) [20 points] Interpret the results: parameter estimates; economic interpretation; potential biases.

Question 7(D) [20 points] Use the estimated choice probabilities from the structural model to provide measures of goodness of fit. More specifically, compare statistics on investment, entry, and exit decisions from the raw data with those implied by the estimated structural model: e.g., frequency of zeroes in investment, entry rate, exit rate, time series of aggregate capacity and aggregate investment.

QUESTION 8 [70 points]. Counterfactuals.

Commodity markets are characterized by large volatility in prices. This is partly due to a very inelastic aggregate supply in the short run that does not respond to transitory demand shocks. Studying the contribution of these supply side frictions (adjustment costs in capacity) to the volatility of prices would require a dynamic equilibrium model of supply and demand. We leave these counterfactuals for the final exam. Instead, here we focus on the effect of the volatility of prices on mines' investment decisions.

Question 8(A) [20 points] Suppose that the price were constant over time and equal to its median value. We can represent this counterfactual as a new transition probability for price, say $\mathbf{F}_{\text{price}}^*$, such that all the columns contain zeroes except the column for the median price

that is a column of ones. Given this counterfactual matrix, repeat the exercise in Questions 5 and 6 to obtain the valuation operator matrix and present values, and use the estimated structural parameter to obtain new counterfactual choice probabilities for investment, entry, and exit.

Question 8(B) [30 points] Use the counterfactual choice probabilities to generate the same statistics than for the goodness-of-fit analysis in Question 7(D). Comment the results by comparing factual and counterfactual statistics.

Question 8(C) [10 points] Explain why the counterfactual exercise in Questions 8(A)-(B) has some limitations (Hint: in the valuation exercise we are using the same choice probabilities as in the factual scenario). Explain how to implement a correct counterfactual experiment.

Question 8(D) [10 points] Suppose that you know (have estimated) the aggregate demand curve of copper in the international market. Explain how you would implement a counterfactual experiment that tries to evaluate how supply side frictions in investment affect the volatility of prices.