Industrial Organization II (ECO 2901)
Winter 2021. Victor Aguirregabiria

Problem Set + Final Exam
Due on Monday, April 12, 2021 [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. You should submit your completed problem set before 11:59pm on Monday, April 12, 2021.
5. Problem sets should be written individually.

The total number of marks is 500.

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

A. DATA

The Stata datafile `data_mines_eco2901_2021.dta` contains annual information on output and inputs from 333 copper mines for the period 1992-2010 (19 years). It is a balanced panel, though not all the mines are active every period. The following is a description of the variables.

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

Zero values for variables `prod_tot`, `labor_n_tot`, and `cap_tot` mean 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 relatively simple dynamic structural model of investment under uncertainty for these mines. The following is a description of the model.

(*i*) 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. Binary variable a_{it}^{active} represents if the mine is active (producing) at year t ($a_{it}^{active} = 1$) or is inactive / not producing ($a_{it}^{active} = 0$). For an inactive mine, current capacity k_{it} is zero.

(*ii*) 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}. She also chooses investment in capacity, closing (exit), and reopening (entry). Incumbent mines at year t (those with positive capacity) decide whether to continue active ($a_{i,t+1}^{active} = 1$) or to temporarily close ($a_{i,t+1}^{active} = 0$). If they choose to stay active, then they decide how much to invest or disinvest in capacity (variable I_{it}). Mines that are temporarily closed decide whether to remain closed ($a_{i,t+1}^{active} = 0$) or to reopen ($a_{i,t+1}^{active} = 0$). If they decide to reopen, then they choose the amount of investment I_{it}.

(*iii*) 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:

$$k_{i,t+1} = a_{i,t+1}^{active} \ (k_{it} + I_{it}) \tag{1}$$

(iv) The choice of output is a static decision.[1] Copper is an homogenous product (a commodity) and we assume that all the mines are price takers in the international metal market (e.g., in the London Metal Exchange)[2]. The price in the metal exchange is p_t. From the point of view of a mine, this price is an exogenous state variable that follows a first order Markov process.

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

(vi) The profit function of a mine is $\Pi_{it} = p_t\ q_{it} - VC_{it}(q_{it}) - IC_{it}(I_{it}, k_{it}) - EC_{it} - XC_{it}$, where $VC_{it}(q_{it})$ is the variable cost function, IC_{it} is the investment cost function, EC_{it} is the reopening cost, and XC_{it} is the closing cost. We specify these functions below.

(vii) The state variables in the dynamic programming (DP) decision 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 the mine's capacity or capital stock at year t, and ω_{it} is a total factor productivity (TFP) shock.

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

$$lk_{i,t+1} = a_{i,t+1}^{active}\ \left(lk_{it} + a_{it}^{invest}\right) \tag{2}$$

(ix) A mine's production technology can be described by a Cobb-Douglas production 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.[3] Capital and TFP are fixed inputs, and labor is the only variable input. Let w_ℓ be the price of labor.[4] Then, it is straightforward to show that the variable cost function is:

$$VC_{it}(q_{it}) = w_\ell\ \left[\frac{q_{it}}{k_{it}^{\alpha_k}\ \omega_{it}}\right]^{1/\alpha_\ell} \tag{3}$$

[1] 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. Here, 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.

[2] There is empirical evidence that some copper companies have market power. In this problem set we abstract from this relevant aspect. We also ignore the ownership structure of the mines and treat them as if each mine were an independent firm.

[3] In 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.

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

(x) Specification of investment, reopening, and closing costs.

$$\left\{ \begin{array}{lcl} IC_{it} & = & \theta^{IC+} 1\{a_{it}^{invest} > 0\} a_{it}^{invest} + \theta^{IC-} 1\{a_{it}^{invest} < 0\}(-a_{it}^{invest}) + \varepsilon_{it}^{IC}(a_{it}^{invest}) \\ \\ EC_{it} & = & \theta^{EC} + \varepsilon_{it}^{EC} \\ \\ XC_{it} & = & \theta_0^{XC} + \theta_1^{XC} \; lk_{it} + \varepsilon_{it}^{XC} \end{array} \right. \tag{4}$$

where θ^{IC+}, θ^{IC-}, θ^{EC}, θ_0^{XC}, and θ_1^{XC} are parameters, and $\varepsilon_{it}^{IC}(a)$, ε_{it}^{EC}, ε_{it}^{XC} are unobservable state variables that are *i.i.d.* over mines, time, and choice alternatives, 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, Reopening Cost, and Closing Cost functions. 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 (CCPs) for the investment, entry, and exit decisions 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.

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

(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**:

$$VP(p_t, k_{it}, \omega_{it}) = \frac{1}{4w_\ell}\ k_{it}\ \left(p_t\ \omega_{it}\right)^2 \tag{5}$$

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) `invest.` It represents the variable I_{it}, and is equal to $I_{it} = k_{i,t+1} - k_{it}$. Note that this variable takes missing values at the last year of the sample (year 2010) for every mine.

(ii) `a_invest.` It represents the variable $a_{it}^{invest} \equiv \ln(1 + I_{it}/(1 + k_{it}))$.

(iii) `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.

(iv) `tfp.` It represents the 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 $\ell_{it} = k_{it} = 0$).

(v) `lkdisc.` It is a discretized version of variable $\ln(1 + k_{it})$. We consider the following discretization: `lkdisc = round(ln(1+cap_tot),0.1)`. This implies that 0.1 units increase in variable `lkdisc` is equivalent to 10% increase in capacity.

(vi) `adisc.` It is a discretization of variable `a_invest` (or a_{it}^{invest}). We construct this variable as follows: $\texttt{adisc}_{it} = \texttt{lkdisc}_{i,t+1} - \texttt{lkdisc}_{it}$.

(vii) `tfpdisc.` It is a discretized version of TFP. We consider the following discretization: `tfpdisc = round(tfp,0.1)`.

(viii) `lpdisc.` It is a discretized version of the logarithm of price, `ln(lme)`. We consider the following discretization: `lpdisc = round(ln(lme_2010),0.1)`. This implies that one 0.1 units increase in variable `lpdisc` is equivalent to 10% increase in price.

Question 2(B) [25 points]

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

(ii) Present table with 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.

(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`.

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

(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`.

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`.

(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`.

(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 a detailed explanation for the construction of the transition probability. The construction of the other probabilities should apply the same procedure.

(A) [**20 points**] Transition probability matrix for `tfpdisc` ($\mathbf{F}_{\texttt{tfp}}$).

• Let $K_{\texttt{tfp}}$ be the number of possible values that `tfpdisc` can take: according to my results, $K_{\texttt{tfp}} = 10$, with values in the uniform grid $\{0.0, 0.1, 0.2, ..., 0.9\}$. Let `tfpvals` be the $K_{\texttt{tfp}} \times 1$ vector with the values in the support of variable `tfpdisc`.

• Let `thres` be a $(K_{\texttt{tfp}} - 1) \times 1$ 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 $\mathbf{F}_{\texttt{tfp}}$ has dimension $K_{\texttt{tfp}} \times K_{\texttt{tfp}}$. Element $F_{\texttt{tfp}}[\text{row}, \text{col}]$ in this matrix is equal to $\Pr(\texttt{tfpdisc}_{t+1} = \texttt{tfpvals[col]} \mid \texttt{tfpdisc}_t = \texttt{tfpvals[row]})$.

• How do we obtain these probabilities? We apply a popular method proposed by George Tauchen (*Economics Letters*, 1986). Let $\widehat{\beta}_0$ and $\widehat{\beta}_1$ be the estimated intercept and slope in the reduced-form regression in Question 3(D). And let $\widehat{\sigma}$ 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:

$$F_{\texttt{tfp}}[\text{row}, \text{col}] = \begin{cases} \Lambda\left(\frac{\texttt{thres[1]-}\widehat{\beta}_0 - \widehat{\beta}_1 \texttt{ tfpvals[row]}}{\widehat{\sigma}}\right) & \text{if} \quad \text{col} = 1 \\ \\ \Lambda\left(\frac{\texttt{thres[col]-}\widehat{\beta}_0 - \widehat{\beta}_1 \texttt{ tfpvals[row]}}{\widehat{\sigma}}\right) \\ -\Lambda\left(\frac{\texttt{thres[col-1]-}\widehat{\beta}_0 - \widehat{\beta}_1 \texttt{ tfpvals[row]}}{\widehat{\sigma}}\right) & \text{if} \quad 2 \leq \text{col} \leq K_{\texttt{tfp}} - 1 \\ \\ 1 - \Lambda\left(\frac{\texttt{thres[}K_{\texttt{tfp}}-1\texttt{]-}\widehat{\beta}_0 - \widehat{\beta}_1 \texttt{ tfpvals[row]}}{\widehat{\sigma}}\right) & \text{if} \quad \text{col} = K_{\texttt{tfp}} \end{cases} \tag{6}$$

where $\Lambda(.)$ 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.

(B) [**20 points**] Transition probability matrix for `lpdisc` ($\mathbf{F}_{\texttt{price}}$). This matrix has dimension $K_{\texttt{price}} \times K_{\texttt{price}}$. According to my results 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 $K_{\texttt{price}} = 15$.

(C) [**30 points**] Conditional choice probability matrix for investment in capacity `adisc` ($\mathbf{P}_{\texttt{adisc}}$).conditional on all the state variables (`lkdisc, lpdisc, tfpdisc`). This matrix has dimension $K_{\texttt{x}} \times K_{\texttt{adisc}}$, with $K_{\texttt{x}} = K_{\texttt{lk}} * K_{\texttt{price}} * K_{\texttt{tfp}}$. According to my results, `lkdisc` has 74 points of support (i.e., a uniform grid between 0.0 and 7.3 with step 0.1). This means that the whole vector of state variables (`lkdisc, lpdisc, tfpdisc`) has a support set with

$K_{\mathtt{x}} = 74 * 15 * 10 = 11,100$ points. The investment decision variable `adisc` has $K_{\mathtt{adisc}} = 121$ points of support (i.e., a uniform grid between -6.0 and $+6.0$ with step 0.1). Therefore, matrix $\mathbf{P}_{\mathtt{adisc}}$ has dimension 11100×121.

(D) [**10 points**] Conditional choice probability vector for the closing decision, `active_next=0` given that `active=1` ($\mathbf{P}_{\mathtt{exit}}$). This vector has dimension $K_{\mathtt{x}} \times 1$. Note that for this vector of CCPs you do not have to use Tauchen's method. The Logit model provides the exit probabilities for every possible value of the state variables.

(E) [**10 points**] Conditional choice probability vector for the reopening decision, `active_next=1` given that `active=0` ($\mathbf{P}_{\mathtt{entry}}$). This vector has dimension $K_{\mathtt{x}} \times 1$. The Logit model provides the reopening probabilities for every possible value of the state variables. Since the only explanatory variable for the reopening position is `lpdisc`, this vector could have dimension $K_{\mathtt{price}} \times 1$. However, for later calculations, it is convenient to expand this vector for every value of all the state variables.

QUESTION 5 [60 points].

(A) [**40 points**] Using the matrices $\mathbf{F}_{\mathtt{tfp}}$, $\mathbf{F}_{\mathtt{price}}$, and $\mathbf{P}_{\mathtt{adisc}}$, and vectors $\mathbf{P}_{\mathtt{entry}}$ and $\mathbf{P}_{\mathtt{exit}}$ that you have obtained in Question 4, obtain the transition matrix of the vector of state variables that is induced by the choice probabilities in $\mathbf{P}_{\mathtt{adisc}}$, $\mathbf{P}_{\mathtt{entry}}$, and $\mathbf{P}_{\mathtt{exit}}$.

Hint. Let me provide some details on this matrix and its calculation. Let $\mathbf{F}_{\mathtt{x}}$ be this matrix that has dimension $K_{\mathtt{x}} \times K_{\mathtt{x}}$, with $K_{\mathtt{x}} = K_{\mathtt{lk}} * K_{\mathtt{price}} * K_{\mathtt{tfp}}$. Element $\mathbf{F}_{\mathtt{x}}[\text{row}, \text{col}]$ in this matrix is the probability:

$$\mathbf{F}_{\mathtt{x}}[\text{row}, \text{col}] = \Pr\left(\begin{array}{c} (\mathtt{lkdisc}_{t+1}\texttt{, } \mathtt{lpdisc}_{t+1}\texttt{, } \mathtt{tfpdisc}_{t+1}) = \\ (\mathtt{lkval}[\text{col}]\texttt{, } \mathtt{lpval}[\text{col}]\texttt{, } \mathtt{tfpval}[\text{col}]) \\ \mid \quad (\mathtt{lkdisc}_{t}\texttt{, } \mathtt{lpdisc}_{t}\texttt{, } \mathtt{tfpdisc}_{t}) = \\ (\mathtt{lkval}[\text{row}]\texttt{, } \mathtt{lpval}[\text{row}]\texttt{, } \mathtt{tfpval}[\text{row}]) \end{array}\right) \tag{7}$$

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

$$\begin{aligned} \mathbf{F}_{\mathtt{x}}[\text{row}, \text{col}] = \; & \Pr\left(\mathtt{tfpdisc}_{t+1} = \mathtt{tfpval}[\text{col}] \mid \mathtt{tfpdisc}_{t} = \mathtt{tfpval}[\text{row}]\right) \\ & \Pr\left(\mathtt{lpdisc}_{t+1} = \mathtt{lpval}[\text{col}] \mid \mathtt{lpdisc}_{t} = \mathtt{lpval}[\text{row}]\right) \\ & \Pr\left(\begin{array}{c} \mathtt{lkdisc}_{t+1} = \mathtt{lkval}[\text{col}] \\ \mid \quad (\mathtt{lkdisc}_{t}\texttt{, } \mathtt{lpdisc}_{t}\texttt{, } \mathtt{tfpdisc}_{t}) = \\ (\mathtt{lkval}[\text{row}]\texttt{, } \mathtt{lpval}[\text{row}]\texttt{, } \mathtt{tfpval}[\text{row}]) \end{array}\right) \end{aligned} \tag{8}$$

and these probabilities can be obtained from $\mathbf{F}_{\texttt{tfp}}$, $\mathbf{F}_{\texttt{price}}$, $\mathbf{P}_{\texttt{adisc}}$, $\mathbf{P}_{\texttt{entry}}$ and $\mathbf{P}_{\texttt{exit}}$.

You can use loops to fill one by one the elements in matrix $\mathbf{F}_{\mathtt{x}}$. Alternatively, it is possible to use matrix operations (Kronecker products combined with element-by-element products and a couple of 'tricks') to obtain directly $\mathbf{F}_{\mathtt{x}}$ from $\mathbf{F}_{\texttt{tfp}}$, $\mathbf{F}_{\texttt{price}}$, $\mathbf{P}_{\texttt{adisc}}$, $\mathbf{P}_{\texttt{entry}}$ and $\mathbf{P}_{\texttt{exit}}$. Using loops can be safer and easier to code, though the computation time will be longer.

(B) **[20 points] Valuation matrix.** Define the following matrix:

$$\mathbf{\Lambda}_{PV} = (I - \beta\ \mathbf{F}_{\mathtt{x}})^{-1} \tag{9}$$

where I is the identity matrix, β is the time discount factor, and $\mathbf{F}_{\mathtt{x}}$ is the transition matrix that you have obtained above. Calculate matrix $\mathbf{\Lambda}_{PV}$ using $\beta = 0.95$.

Matrix $\mathbf{\Lambda}_{PV}$ is Present Value operator. More specifically, let $\mathbf{\Pi}$ be a $K_{\mathtt{x}} \times 1$ vector of per-period profits for each possible value of the state variables. Then, by pre-multiplying $\mathbf{\Pi}$ with matrix $\mathbf{\Lambda}_{PV}$ we get a vector of present values:

$$\mathbf{PV} = \mathbf{\Lambda}_{PV}\ \mathbf{\Pi} \tag{10}$$

Element $\mathbf{PV}[\,\text{row}\,]$ in vector $\mathbf{PV}$ give us the present value of the stream of current and future profits in $\mathbf{\Pi}$ 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 $\mathbf{\Lambda}_{PV}$ 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 $\mathbf{\Pi}$ and matrix $\mathbf{F}_{\mathtt{x}}$. The recursive expression for the vector of present values is:

$$\mathbf{PV} = \mathbf{\Pi} + \beta\ \mathbf{F}_{\mathtt{x}}\ \mathbf{PV} \tag{11}$$

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

$$\mathbf{PV} = (I - \beta\ \mathbf{F}_{\mathtt{x}})^{-1}\ \mathbf{\Pi} \tag{12}$$

QUESTION 6 [40 points]. Vector with Present values of Variable profits. Let $(a_{i,t+1}^{active}, a_{it}^{invest})$ be the decision of mine i at year t. The profit function (without including the unobservable state variables) can be represented as:

$$\Pi_{it} = h(a_{i,t+1}^{active}, a_{it}^{invest}; p_t, k_{it}, \omega_{it})'\ \boldsymbol{\theta} \tag{13}$$

where:

$$\boldsymbol{\theta} = (\theta_{vp},\ \theta^{IC+},\ \theta^{IC-},\ \theta^{EC},\ \theta_0^{XC},\ \theta_1^{XC})' \tag{14}$$

with $\theta_{vp} = 1/4w_\ell$; and

$$h(a_{i,t+1}^{active}, a_{it}^{invest}; p_t, k_{it}, \omega_{it})' =$$

$$= \left(k_{it}\ \left(p_t\ \omega_{it}\right)^2,\ 1\{a_{it}^{invest} > 0\}a_{it}^{invest},\ ^{IC-}1\{a_{it}^{invest} < 0\}(-a_{it}^{invest}), \right. \tag{15}$$

$$\left. 1\{a_{i,t+1}^{active} = 1\ \&\ k_{it} = 0\},\ 1\{a_{i,t+1}^{active} = 0\ \&\ k_{it} > 0\},\ 1\{a_{i,t+1}^{active} = 0\ \&\ k_{it} > 0\} * lk_{it}\right)$$

For some descriptions below, it will be convenient to represent the 6 elements of the vector $h()$ as h_{it}^1, h_{it}^2, ..., h_{it}^6.

(A) **[20 points]** For each of the 6 elements in the vector $h()$, construct a matrix with dimension $K_{\mathtt{x}} \times K_{\mathtt{a}}$ with the value of that component for every value of the state variables and the decision variables. That is, let $\mathbf{H}^1$ be the matrix for the first element, $k_{it}\ (p_t\ \omega_{it})^2$. 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 $\mathbf{H}^1[\text{row}, \text{col}]$ contains the value of $k_{it}\ (p_t\ \omega_{it})^2$ when the value of the state variables corresponds to the value of these variables in row "row", and the value of the decision variables corresponds to the value of these variables in column "col". In this particular case, because $k_{it}\ (p_t\ \omega_{it})^2$ does not depend on the decision variables, all columns in $\mathbf{H}^1$ will be the same.

(B) **[20 points]** Given the Present Value operator matrix $\mathbf{\Lambda}_{PV}$ obtained in Question 5, obtained the vector of present values for each of the 6 components in vector $h()$: that is, $\mathbf{PV}^1 = \mathbf{\Lambda}_{PV}\ \mathbf{H}^1$, $\mathbf{PV}^2 = \mathbf{\Lambda}_{PV}\ \mathbf{H}^2$, ..., $\mathbf{PV}^6 = \mathbf{\Lambda}_{PV}\ \mathbf{H}^6$.

QUESTION 7 [100 points]. Estimation.

(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 $\boldsymbol{\theta}$.

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

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

(D) **[20 points]** Used 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.

(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 $\mathbf{F}^{*}_{\texttt{price}}$, 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.

(B) [**30 points**] Used 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.

(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.

(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.