Escola de Economia de São Paulo - Fundação Getulio Vargas

STUDENT NAME:

Course: Econometria 2

Instructor: Vitor Possebom

Final Exam 2023 - Tuesday, June 27th, 9:00-12:50 — Total = 470 points

Instructions:

This exam lasts 3 hours and 50 minutes. No extra time will be provided and you must hand me your exam before 12:50 pm. I will display a clock using the projector and this clock will be our official time.

While taking the exam, you can eat your own food, drink your own water or any other type of liquid, and go to the restroom. However, you cannot leave the floor to buy food or drinks.

You cannot check any notes, books or electronic devices. You cannot talk to your colleagues.

The exam is worth 470 points distributed in the following way:

- Question 1 - Local Projection (Theory): 240 points
  - Item a: 60 points
  - Item b: 20 points
  - Item c: 10 points
  - Item d: 40 points
  - Item e: 20 points
  - Item f: 60 points
  - Item g: 20 points
  - Item h: 10 points
- Question 2 - Local Projection (Monte Carlo): 40 points
  - Item a: 30 points
  - Item b: 10 points
- Question 3 - Local Projection (Empirical Application): 90 points

- Item a: 20 points
- Item b: 30 points
- Item c: 40 points

- Question 4 - Estimating an ARMA(1,2) using GMM: 40 points
  - Item a: 10 points
  - Item b: 10 points
  - Item c: 20 points
- Question 5 - Dynamic Multipliers of Nonstationary Processes: 60 points
  - Item a: 30 points
  - Item b: 20 points
  - Item c: 10 points

Each question-item pair has an assigned space for its answer. This assigned space starts immediately below the question-item pair and ends at the bottom of the page before the next question-item pair. I will not grade anything that is not properly located in its correct place.

You can write your answers using any writing material, including quills, pieces of charcoal or a lipstick. I recommend using a pencil, a mechanic pencil or a standard pen. Have in mind that I can only grade what I am actually able to read and understand.

You can write answers in any language that I can fluently understand. Since I am dumb, you can only write in Portuguese or English. If you write in Spanish, I can try to understand, but I cannot promise you anything.

At the end of the exam, there are 10 blank pages. You can use them as you please. You can even write a poem or draw a cartoon. I recommend using them to sketch your own work. Importantly, I will not read nor grade anything in those sketching pages.

At the top of the first page, you must write your FULL NAME in the assigned space. You must write your FULL NAME even if you are "Pedro de Alcântara Francisco Antônio João Carlos Xavier de Paula Miguel Rafael Joaquim José Gonzaga Pascoal Cipriano Serafim de Bragança e Bourbon", also known as Pedro I.

You cannot detach any sheet of paper from this exam. All sheets must be kept attached as they were handed to you.

Questions 4 and 5 are based on Problem Set Exercises. Probably, it is a good idea to start your exam with Questions 4 and 5, guaranteeing their points.

Question 1 is the hardest question in this exam. Although Questions 2 and 3 are connected to Question 1, you do not need to solve Question 1 to answer Questions 2 and 3. Simply reading Question 1 will help you to answer Questions 2 and 3.

Question 1 (Local Projection (Theory) - 240 points)

Questions 1-3 are based on a article written by Jorda (2005).

Researchers in macroeconomics often compute dynamic multipliers of interest, such as impulse responses, by specifying a VAR even though the VAR per se is often of no particular interest. In practice, VARs may be a significantly misspecified representation of the true data generating process (DGP).

This misspecification issue may not be relevant if our goal is to predict our variable of interest one-period ahead. Since a VAR is a linear global approximation of the DGP and is optimally designed for one-period ahead forecasting, it may still produce reasonable one-period ahead forecasts.

However, an impulse response is a function of forecasts at increasingly distant horizons. Consequently, misspecification errors are compounded with the forecast horizon, implying that VAR models may perform poorly when analyzing long impulse response functions.

To circumvent this problem, Jorda (2005) proposes to use a collection of projections local to each forecast horizon. By doing so, he matches design and evaluation.

Jorda (2005) defines an impulse response function as the difference between two forecasts:

$$\mathrm{IR}\left(t, s, \mathbf{d}_i\right)=E\left(\mathbf{y}_{t+s} \mid \mathbf{v}_t=\mathbf{d}_i ; \mathbf{X}_t\right)-E\left(\mathbf{y}_{t+s} \mid \mathbf{v}_t=\underline{\mathbf{0}} ; \mathbf{X}_t\right) \quad s=0,1,2, \ldots \tag{1}$$

where $\mathbf{y}_t$ *is an* $n \times 1$ *random vector;* $\mathbf{X}_t \equiv \left(\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots\right)'$; $\underline{\mathbf{0}}$ *is of dimension* $n \times 1$*;* $\mathbf{v}_t$ *is the* $n \times 1$ *vector of reduced-form disturbances; and* $\mathbf{D}$ *is an* $n \times n$ *matrix, whose columns* $\mathbf{d}_i$ *contain the relevant experimental shocks.*[1]

Expression (1) shows that the statistical objective in calculating impulse responses is to obtain the best, mean-squared, multi-step predictions. One natural way to obtain these predictions is to directly use forecasting models that are re-estimated for each forecast horizon. Therefore, consider <u>*projecting*</u> $\mathbf{y}_{t+s}$ *onto the linear space generated by* $\left(\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}\right)'$*, specifically*

$$\mathbf{y}_{t+s}=\boldsymbol{\alpha}^s+\mathbf{B}_1^{s+1} \mathbf{y}_{t-1}+\mathbf{B}_2^{s+1} \mathbf{y}_{t-2}+\cdots+\mathbf{B}_p^{s+1} \mathbf{y}_{t-p}+\mathbf{u}_{t+s}^s \quad s=0,1,2, \ldots, h \tag{2}$$

where $\boldsymbol{\alpha}^s$ *is an* $n \times 1$ *vector of constants, and the* $\mathbf{B}_i^{s+1}$ *are matrices of coefficients for each lag* i *and*

[1] Intuitively, our experimental variable of interest is given by $\tilde{\mathbf{y}}_t \coloneqq \mathbf{y}_t+\mathbf{v}_t$.

horizon $s+1$. We denote the collection of h regressions in Equation (2) as local-linear projections.

1.a (60 points): *Derive the Impulse-Response Functions for any t and $s = 0, 1, 2, \dots, h$ based on local-linear projections.* ***Hint:*** *A good starting point is to write $\mathbf{y}_{t+s}$ as a function of $\mathbf{y}_t$ and combine Footnote 1 with Equation (1).*

1.b (20 points): *Propose a reasonable estimator for the local-linear projection in Equation (2) and use it to propose an estimator for the impulse response functions you found in Question 1.a.*

1.c (10 points): *Describe one way to determine the maximum lag p in each linear regression defined in Equation (2).*

1.d (40 points):

In this question, we want to derive the relationship between the local-linear projection approach and the VAR approach. To do so, we will impose that the true DGP follows a VAR(p) model.

A VAR(p) specifies that the $n \times 1$ vector $\mathbf{y}_t$ depends linearly on $\mathbf{X}_t \equiv (\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots,\ \mathbf{y}_{t-p})'$, through the expression

$$\mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\Pi}'\mathbf{X}_t + \mathbf{v}_t$$

where $\mathrm{v_t}$ is an i.i.d. vector of disturbances and $\Pi' \equiv [\Pi_1 \quad \Pi_2 \quad \dots \quad \Pi_p]$. The VAR(1) companion form to this VAR can be expressed by defining

$$\mathbf{W}_t \equiv \begin{bmatrix} \mathbf{y}_t - \boldsymbol{\mu} \\ \mathbf{y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{y}_{t-p+1} - \boldsymbol{\mu} \end{bmatrix}; \mathbf{F} \quad \equiv \begin{bmatrix} \boldsymbol{\Pi}_1 & \boldsymbol{\Pi}_2 & \cdots & \boldsymbol{\Pi}_{p-1} & \boldsymbol{\Pi}_p \\ \mathbf{I} & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{I} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{I} & 0 \end{bmatrix}; \boldsymbol{\nu}_t \equiv \begin{bmatrix} \mathbf{v}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and then realizing that the last two equations imply that

$$\mathbf{W}_t = \mathbf{F}\mathbf{W}_{t-1} + \boldsymbol{\nu}_t$$

from which s-step ahead forecasts can be easily computed since

$$\mathbf{W}_{t+s} = \boldsymbol{\nu}_{t+s} + \mathbf{F}_{t+s-1}\boldsymbol{\nu}_{t+s-1} + \cdots + \mathbf{F}^{s}\boldsymbol{v}_{t} + \mathbf{F}^{s+1}\mathbf{W}_{t-1}$$

and therefore

$$\mathbf{y}_{t+s} - \boldsymbol{\mu} = \mathbf{v}_{t+s} + \mathbf{F}_{1}^{1}\mathbf{v}_{t+s-1} + \cdots + \mathbf{F}_{1}^{s}\mathbf{v}_{t} + \mathbf{F}_{1}^{s+1}\left(\mathbf{y}_{t-1} - \boldsymbol{\mu}\right) + \cdots + \mathbf{F}_{p}^{s+1}\left(\mathbf{y}_{t-p} - \boldsymbol{\mu}\right) \tag{3}$$

where $\mathbf{F}_{i}^{s}$ is the i^{th} upper, $n \times n$ block of the matrix $\mathbf{F}^{s}$ (i.e., $\mathbf{F}$ raised to the power s).

Write $\boldsymbol{\alpha}^{s}$, $\mathbf{B}_{1}^{s+1}$ and $\mathbf{u}_{t+s}^{s}$ in Equation (2) as functions of $\boldsymbol{\mu}$, $\mathbf{F}_{i}^{l}$ and $\mathbf{v}_{t+l}$ for $l = 0, 1, 2, \ldots, s, s+1$ and $i = 1, \ldots, p$.

1.e (20 points): *Now, we want to conduct inference about the coefficients of the local-linear projection in Equation (2). Propose one reasonable way to compute standard errors for your estimator of B_{1}^{s}. Justify your answer.*

1.f (60 points): *Linear models, such as VARs, bestow four restrictive properties on their implied impulse response functions: (1) symmetry, where responses to positive and negative shocks are mirror images of each other; (2) shape invariance, where responses to shocks of different magnitudes are scaled versions of one another; (3) history independence, where the shape of the responses is independent of the local conditional history (i.e., impulse response functions do not depend on the level of the variables of interest); and (4) multidimensionality, where responses are high-dimensional nonlinear functions of parameter estimates which complicate the calculation of standard errors and quickly compound misspecification errors.*

Although local-linear projection methods (Equation (2)) dispense with the fourth of these problems, they are indeed linear and thus constrained by properties (1)–(3). To circumvent these three problems, Jorda (2005) proposes to use local-cubic projections:

$$\mathbf{y}_{t+s} = \boldsymbol{\alpha}^{s} + \mathbf{B}_{1}^{s+1}\mathbf{y}_{t-1} + \mathbf{Q}_{1}^{s+1}\mathbf{y}_{t-1}^{2} + \mathbf{C}_{1}^{s+1}\mathbf{y}_{t-1}^{3} + \mathbf{B}_{2}^{s+1}\mathbf{y}_{t-2} + \cdots + \mathbf{B}_{p}^{s+1}\mathbf{y}_{t-p} + \mathbf{u}_{t+s}^{s} \quad s = 0, 1, 2, \ldots, h \tag{4}$$

where Jorda (2005) do not allow for cross-product terms so that $\mathbf{y}_{t-1}^{k} = \left(y_{1,t-1}^{k}, y_{2,t-1}^{k}, \ldots, y_{n,t-1}^{k}\right)'$ for $k = 1, 2$, as a matter of choice and parsimony.

Derive the Impulse-Response Functions for any t and $s = 0, 1, 2, \ldots, h$ based on local-cubic projections. ***Hint:*** *A good starting point is to write $\mathbf{y}_{t+s}$ as a function of $\mathbf{y}_t$ and combine Footnote 1 with Equation (1).*

1.g (20 points): *Propose a reasonable estimator for the local-cubic projection in Equation (4) and use it to propose an estimator for the impulse response functions you found in Question 1.f.*

1.h (10 points): *Consider the impulse-response functions derived in Question 1.f. In which conditions will they be asymmetric, shape variant and history dependent?*

Answer 1.a:

The definition of impulse response functions is given by Equation (1). This expression suffices to derive the impulse response function for any t and $s = 0$ because $\text{IR}(t, 0, \mathbf{d}_i)$ is simply our experimental shock of interest:

$$\text{IR}(t, 0, \mathbf{d}_i) = \mathbf{d}_i. \quad \text{(10 points)}$$

To derive $\text{IR}(t, s, \mathbf{d}_i)$ for any t and $s = 1, 2, \ldots, h$ based on local-linear projections, we need to combine Equations (1) and (2). Fix t and $s = 1, 2, \ldots, h$ arbitrarily. From Equation (2) and $\tau \coloneqq t - 1$, we have that

$$\mathbf{y}_{\tau+1+s} = \boldsymbol{\alpha}^s + \mathbf{B}_1^{s+1}\mathbf{y}_\tau + \mathbf{B}_2^{s+1}\mathbf{y}_{\tau-1} + \cdots + \mathbf{B}_p^{s+1}\mathbf{y}_{\tau-p+1} + \mathbf{u}_{\tau+1+s}^s. \quad \text{(10 points)}$$

From the last equation and $\tilde{s} \coloneqq s + 1$, we have that

$$\mathbf{y}_{\tau+\tilde{s}} = \boldsymbol{\alpha}^{\tilde{s}-1} + \mathbf{B}_1^{\tilde{s}}\mathbf{y}_\tau + \mathbf{B}_2^{\tilde{s}}\mathbf{y}_{\tau-1} + \cdots + \mathbf{B}_p^{\tilde{s}}\mathbf{y}_{\tau-p+1} + \mathbf{u}_{\tau+\tilde{s}}^{\tilde{s}-1}. \quad \text{(10 points)}$$

Now, we can combine the last equation, Equation (1) and Footnote 1 to find

$$\begin{aligned}
\text{IR}(t, s, \mathbf{d}_i) &\coloneqq E(\mathbf{y}_{t+s} \mid \mathbf{v}_t = \mathbf{d}_i; \mathbf{X}_t) - E(\mathbf{y}_{t+s} \mid \mathbf{v}_t = \underline{\mathbf{0}}; \mathbf{X}_t) \\
&= \boldsymbol{\alpha}^{s-1} + \mathbf{B}_1^s\tilde{\mathbf{y}}_t + \mathbf{B}_2^s\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p^s\mathbf{y}_{t-p+1} + \mathbf{u}_{ts}^{s-1} \\
&\quad - \left(\boldsymbol{\alpha}^{s-1} + \mathbf{B}_1^s\mathbf{y}_t + \mathbf{B}_2^s\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p^s\mathbf{y}_{t-p+1} + \mathbf{u}_{ts}^{s-1}\right) \quad \text{(10 points)} \\
&= \mathbf{B}_1^s \cdot (\mathbf{y}_t + \mathbf{d}_i - \mathbf{y}_t) \quad \text{(10 points)} \\
&= \mathbf{B}_1^s\mathbf{d}_i. \quad \text{(10 points)}
\end{aligned}$$

Answer 1.b: To estimate each element in $\boldsymbol{\alpha}^s$ and each row j in matrices B_i^{s+1} for $i = 1, \ldots, p$ and $s = 0, 1, \ldots, h$, we can regress $y_{j,t}$ on $\mathbf{X}_t \equiv (\mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p})'$ and an intercept using OLS. (10 points)

Afterwards, we multiply our estimators of B_1^s by $\mathbf{d}_i$ to find an estimator for $\text{IR}(t, s, \mathbf{d}_i)$. (10 points)

Answer 1.c:

One way to determine the maximum lag p in each linear regression defined in Equation (2) is to run these regressions for different values of p and choose the specification with the smallest value of AIC or

BIC statistics for each regression. (10 points)

Answer 1.d:

We start by rearranging Equation (3): (10 points)

$$\mathbf{y}_{t+s} = \left(\mathbf{I} - \mathbf{F}_1^{s+1} - \cdots - \mathbf{F}_p^{s+1}\right)\boldsymbol{\mu} + \mathbf{F}_1^{s+1}\mathbf{y}_{t-1} + \cdots + \mathbf{F}_p^{s+1}\mathbf{y}_{t-p} + \left(\mathbf{v}_{t+s} + \mathbf{F}_1^{1}\mathbf{v}_{t+s-1} + \cdots + \mathbf{F}_1^{s}\mathbf{v}_t\right).$$

Now, we match the coefficients and the error term in the last equation with the coefficients and error term in Equation (2) to find:

$$\boldsymbol{\alpha}^s = \left(\mathbf{I} - \mathbf{F}_1^{s+1} - \cdots - \mathbf{F}_p^{s+1}\right)\boldsymbol{\mu} \quad \text{(10 points)}$$

$$\mathbf{B}_1^{s+1} = \mathbf{F}_1^{s+1} \quad \text{(10 points)}$$

$$\mathbf{u}_{t+s}^s = \left(\mathbf{v}_{t+s} + \mathbf{F}_1^{1}\mathbf{v}_{t+s-1} + \cdots + \mathbf{F}_1^{s}\mathbf{v}_t\right). \quad \text{(10 points)}$$

Answer 1.e:

In Question 1.d, we found that $\mathbf{u}_{t+s}^s = \left(\mathbf{v}_{t+s} + \mathbf{F}_1^{1}\mathbf{v}_{t+s-1} + \cdots + \mathbf{F}_1^{s}\mathbf{v}_t\right)$. Consequently, the error term in Equation (2) is serially correlated. (10 points)

To account for the autocorrelation of the error term, we can use the Newey-West Standard Error Estimator to compute the standard error of the OLS estimator proposed in Question 1.c. (10 points)

Answer 1.f:

The definition of impulse response functions is given by Equation (1). This expression suffices to derive the impulse response function for any t and $s = 0$ because $\mathrm{IR}\left(t, 0, \mathbf{d}_i\right)$ is simply our experimental shock of interest:

$$\mathrm{IR}\left(t, 0, \mathbf{d}_i\right) = \mathbf{d}_i \quad \text{(10 points)}$$

To derive $\mathrm{IR}\left(t, s, \mathbf{d}_i\right)$ for any t and $s = 1, 2, \ldots, h$ based on local-cubic projections, we need to combine Equations (1) and (4). Fix t and $s = 1, 2, \ldots, h$ arbitrarily. From Equation (4) and $\tau := t - 1$, we have that

$$\mathbf{y}_{\tau+1+s} = \boldsymbol{\alpha}^s + \mathbf{B}_1^{s+1}\mathbf{y}_\tau + \mathbf{Q}_1^{s+1}\mathbf{y}_\tau^2 + \mathbf{C}_1^{s+1}\mathbf{y}_\tau^3 + \mathbf{B}_2^{s+1}\mathbf{y}_{\tau-1} + \cdots + \mathbf{B}_p^{s+1}\mathbf{y}_{\tau-p+1} + \mathbf{u}_{\tau+1+s}^s. \quad \text{(10 points)}$$

From the last equation and $\tilde{s} \coloneqq s+1$, we have that

$$\mathbf{y}_{\tau+\tilde{s}} = \boldsymbol{\alpha}^{\tilde{s}-1} + \mathbf{B}_1^{\tilde{s}}\mathbf{y}_\tau + \mathbf{Q}_1^{\tilde{s}}\mathbf{y}_\tau^2 + \mathbf{C}_1^{\tilde{s}}\mathbf{y}_\tau^3 + \mathbf{B}_2^{\tilde{s}}\mathbf{y}_{\tau-1} + \cdots + \mathbf{B}_p^{\tilde{s}}\mathbf{y}_{\tau-p+1} + \mathbf{u}_{\tau+\tilde{s}}^{\tilde{s}-1}. \quad \text{(10 points)}$$

Now, we can combine the last equation, Equation (1) and Footnote 1 to find

$$\begin{aligned}
\mathrm{IR}\left(t,s,\mathbf{d}_i\right) &\coloneqq E\left(\mathbf{y}_{t+s} \mid \mathbf{v}_t = \mathbf{d}_i; \mathbf{X}_t\right) - E\left(\mathbf{y}_{t+s} \mid \mathbf{v}_t = \underline{\mathbf{0}}; \mathbf{X}_t\right) \\
&= \boldsymbol{\alpha}^{s-1} + \mathbf{B}_1^s\tilde{\mathbf{y}}_t + \mathbf{Q}_1^s\tilde{\mathbf{y}}_t^2 + \mathbf{C}_1^s\tilde{\mathbf{y}}_t^3 + \mathbf{B}_2^s\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p^s\mathbf{y}_{t-p+1} + \mathbf{u}_{t+s}^{s-1} \\
&\quad - \left(\boldsymbol{\alpha}^{s-1} + \mathbf{B}_1^s\mathbf{y}_t + \mathbf{Q}_1^s\mathbf{y}_t^2 + \mathbf{C}_1^s\mathbf{y}_t^3 + \mathbf{B}_2^s\mathbf{y}_{t-1} + \cdots + \mathbf{B}_p^s\mathbf{y}_{t-p+1} + \mathbf{u}_{t+s}^{s-1}\right) \quad \text{(10 points)} \\
&= \mathbf{B}_1^s\left(\mathbf{y}_t + \mathbf{d}_i\right) + \mathbf{Q}_1^s\left(\mathbf{y}_t + \mathbf{d}_i\right)^2 + \mathbf{C}_1^s\left(\mathbf{y}_t + \mathbf{d}_i\right)^3 - \left(\mathbf{B}_1^s\mathbf{y}_t + \mathbf{Q}_1^s\mathbf{y}_t^2 + \mathbf{C}_1^s\mathbf{y}_t^3\right) \quad \text{(10 points)} \\
&= \mathbf{B}_1^s\left(\mathbf{y}_t + \mathbf{d}_i\right) + \mathbf{Q}_1^s\left(\mathbf{y}_t^2 + 2\cdot\mathbf{y}_t\cdot\mathbf{d}_i + \mathbf{d}_i^2\right) + \mathbf{C}_1^s\left(\mathbf{y}_t^3 + 3\cdot\mathbf{y}_t^2\cdot\mathbf{d}_i + 3\cdot\mathbf{y}_t\cdot\mathbf{d}_i^2 + \mathbf{d}_i^3\right) \\
&\quad - \left(\mathbf{B}_1^s\mathbf{y}_t + \mathbf{Q}_1^s\mathbf{y}_t^2 + \mathbf{C}_1^s\mathbf{y}_t^3\right) \\
&= \mathbf{B}_1^s\mathbf{d}_i + \mathbf{Q}_1^s\left(2\cdot\mathbf{y}_t\cdot\mathbf{d}_i + \mathbf{d}_i^2\right) + \mathbf{C}_1^s\left(3\cdot\mathbf{y}_t^2\cdot\mathbf{d}_i + 3\cdot\mathbf{y}_t\cdot\mathbf{d}_i^2 + \mathbf{d}_i^3\right). \quad \text{(10 points)}
\end{aligned}$$

Answer 1.g: To estimate each element in $\boldsymbol{\alpha}^s$, each row j in matrices B_i^{s+1} for $i = 1,\ldots,p$ and $s = 0,1,\ldots,h$, each row j in matrices Q_1^{s+1} for $s = 0,1,\ldots,h$, and each row j in matrices C_1^{s+1} for $s = 0,1,\ldots,h$, we can regress $y_{j,t}$ on $\mathbf{X}_t \equiv \left(\mathbf{y}_{t-1}, \mathbf{y}_{t-1}^2, \mathbf{y}_{t-1}^3, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}\right)'$ and an intercept using OLS. (10 points)

To estimate the impulse response functions derived in the last question, we denote our OLS estimators using the conventional hat notation and estimate

$$\hat{\mathrm{IR}}\left(t,s,\mathbf{d}_i\right) = \hat{\mathbf{B}}_1^s\mathbf{d}_i + \hat{\mathbf{Q}}_1^s\left(2\cdot\mathbf{y}_t\cdot\mathbf{d}_i + \mathbf{d}_i^2\right) + \hat{\mathbf{C}}_1^s\left(3\cdot\mathbf{y}_t^2\cdot\mathbf{d}_i + 3\cdot\mathbf{y}_t\cdot\mathbf{d}_i^2 + \mathbf{d}_i^3\right). \quad \text{(10 points)}$$

Answer 1.h:

Recall the impulse-response functions derived in Question 1.f.:

$$\mathrm{IR}\left(t,s,\mathbf{d}_i\right) = \mathbf{B}_1^s\mathbf{d}_i + \mathbf{Q}_1^s\left(2\cdot\mathbf{y}_t\cdot\mathbf{d}_i + \mathbf{d}_i^2\right) + \mathbf{C}_1^s\left(3\cdot\mathbf{y}_t^2\cdot\mathbf{d}_i + 3\cdot\mathbf{y}_t\cdot\mathbf{d}_i^2 + \mathbf{d}_i^3\right).$$

They will be asymmetric, shape variant and history dependent if some of the terms in $\mathbf{Q}_1^s$ and $\mathbf{C}_1^s$ are non-zero. (10 points)

Question 2 (Local Projection (Monte Carlo) - 40 points)

Questions 1-3 are based on a article written by Jorda (2005).

Jorda (2005) uses three Monte Carlo simulations to evaluate the performance of local projections for impulse response estimation and inference. In this exam, we will focus on his first two Monte Carlo Simulation.

In his first two simulation experiments, Jorda (2005) uses a standard monetary VAR as the true DGP. He collects American monthly data from January 1960 to February 2001 (494 observations). First, he estimates a VAR of order 12 on the following variables:

1. *EM, log of non-agricultural payroll employment;*
2. *P, log of personal consumption expenditures deflator (1996 = 100);*
3. *PCOM, annual growth rate of the index of sensitive materials prices issued by the Conference Board;*
4. *FF, federal funds rate;*
5. *NBRX, ratio of nonborrowed reserves plus extended credit to total reserves; and*
6. *$\Delta M2$, annual growth rate of M2 stock.*

He, then, saves the coefficient estimates from this VAR and simulate 500 series of 494 observations using multivariate normal residuals and the variance-covariance matrix from the estimation stage, and uses the first 12 observations from the data to initialize all 500 runs.

His first Monte Carlo experiment compares (i) the impulse responses that would result from fitting a VAR of order two against (ii) local-linear and local-cubic projections of order two as well. Since the true DGP contains 12 lags, all three models are misspecified.

Each panel in Figure 1 displays the impulse response of a variable in our system due to a shock in the variable FF. The meaning of each line is explained below the figure.

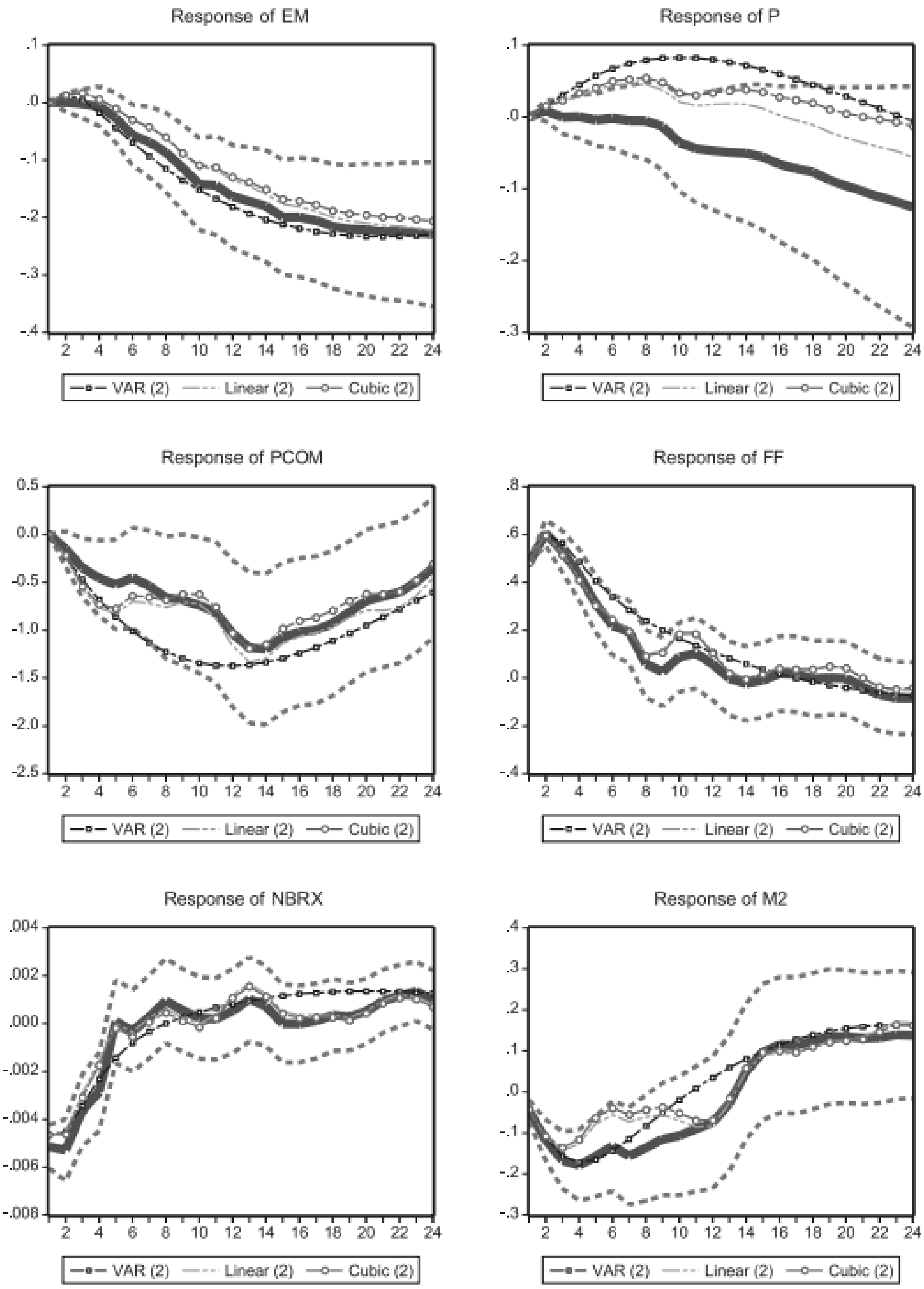

FIGURE 1. IMPULSE RESPONSES TO A SHOCK IN *FF*. LAG LENGTH: 2

Notes: Evans and Marshall (1998) VAR(12) Monte Carlo Experiment. The thick line is the true impulse response based on a VAR(12). The thick dashed lines are Monte Carlo two standard error bands. Three additional impulse responses are compared, based on estimates involving two lags only: (1) the response calculated by fitting a VAR(2) instead, depicted by the line with squares; (2) the response calculated with a local-linear projection, depicted by the dashed line; and (3) the response calculated with a local-cubic projection, depicted by the line with circles. 500 replications.

Figure 1: Figure 1 by Jorda (2005)

2.a (30 points): *Interpret the performance of all three estimators when estimating the true impulse response function of P to a shock in FF.*

2.b (10 points): *His second Monte Carlo experiment compares (i) the impulse responses that would result from fitting a VAR of order 12 against (ii) local-linear and local-cubic projections of order 12 as well. Since the true DGP contains 12 lags, all three models are correctly specified.*

Each panel in Figure 2 displays the impulse response of a variable in our system due to a shock in the variable FF. The meaning of each line is explained below the figure.

Interpret the performance of all three estimators when estimating the true impulse responses to a shock in FF.

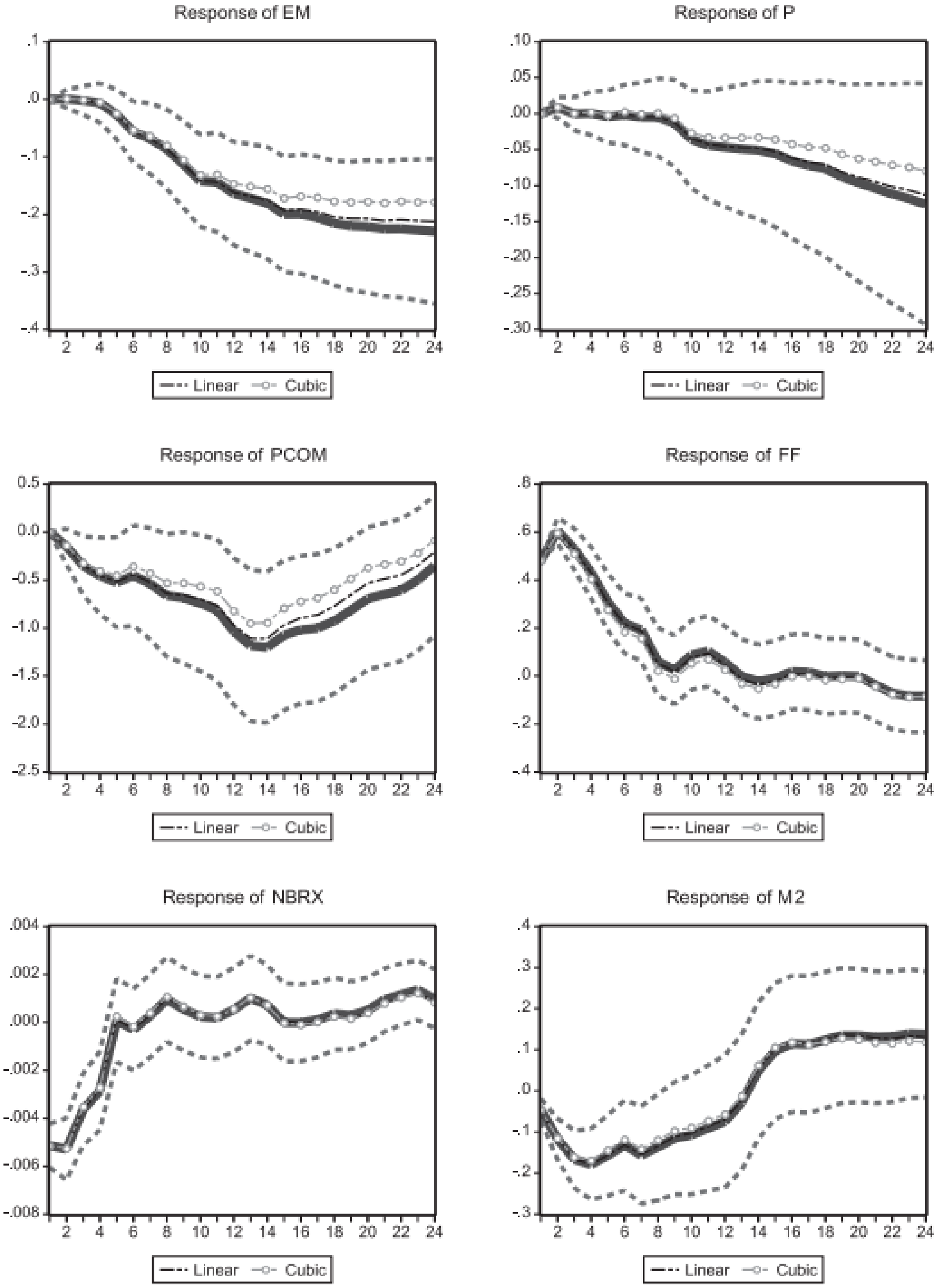

FIGURE 2. IMPULSE RESPONSES TO A SHOCK IN *FF*. LAG LENGTH: 12

Notes: Evans and Marshall (1998) VAR(12) Monte Carlo Experiment. The thick line is the true impulse response based on a VAR(12). The thick dashed lines are Monte Carlo two standard error bands. Two additional impulse responses are compared: (1) the response calculated with a local-linear projection with 12 lags, depicted by the dashed line; and (2) the response calculated with a local-cubic projection, depicted by line with circles. 500 replications.

Figure 2: Figure 2 by Jorda (2005)

Answer 2.a:

The response based on the VAR(2) model is statistically different from the true response for the first 17 periods and suggests that prices increase in response to an increase in the federal funds rate over 23 out of the 24 periods displayed. (10 points)

The local-linear projection is virtually within the true two standard error bands throughout the 24 periods depicted, and is strictly negative for the last 7 periods. (10 points) This result suggest that local-linear projections are robust to misspecification.

The local-cubic projection is virtually within the true two standard error bands throughout the 24 periods depicted. (10 points)

Answer 2.b:

For all three estimators, all estimated impulse response functions literally lie on top of the true response with occasional minor differences. (10 points)

Question 3 (Local Projection (Empirical Application) - 90 points)

Questions 1-3 are based on a article written by Jorda (2005).

One research area in Macroeconomics focus on the efficacy, optimality, credibility and robustness of interest rates rules for monetary policy. To implement this analysis, we frequently use a simple, new-Keynesian, closed-economy model which, at a minimum, can be summarized by three fundamental expressions: an IS equation, a Phillips relation, and the candidate policy rule itself.

To estimate these expressions, Jorda (2005) uses the following variables:

- y_t *is the percentage gap between real GDP and potential GDP;*
- π_t *is quarterly inflation in the GDP, chain-weighted price index in percent at annual rate; and*
- i_t *is the quarterly average of the federal funds rate in percent at an annual rate.*

The sample of quarterly data runs over the period 1955:I–2003:I.

Figure 3 displays the impulse responses based on a VAR(4), a local-linear projection with four lags and local-cubic projection with four lags, all identified with a standard Cholesky decomposition and the Wold-causal order y_t, π_t *and* i_t.

Each panel in Figure 3 displays the impulse response of the column variable to a shock in the row variable. The meaning of each line is explained below the figure.

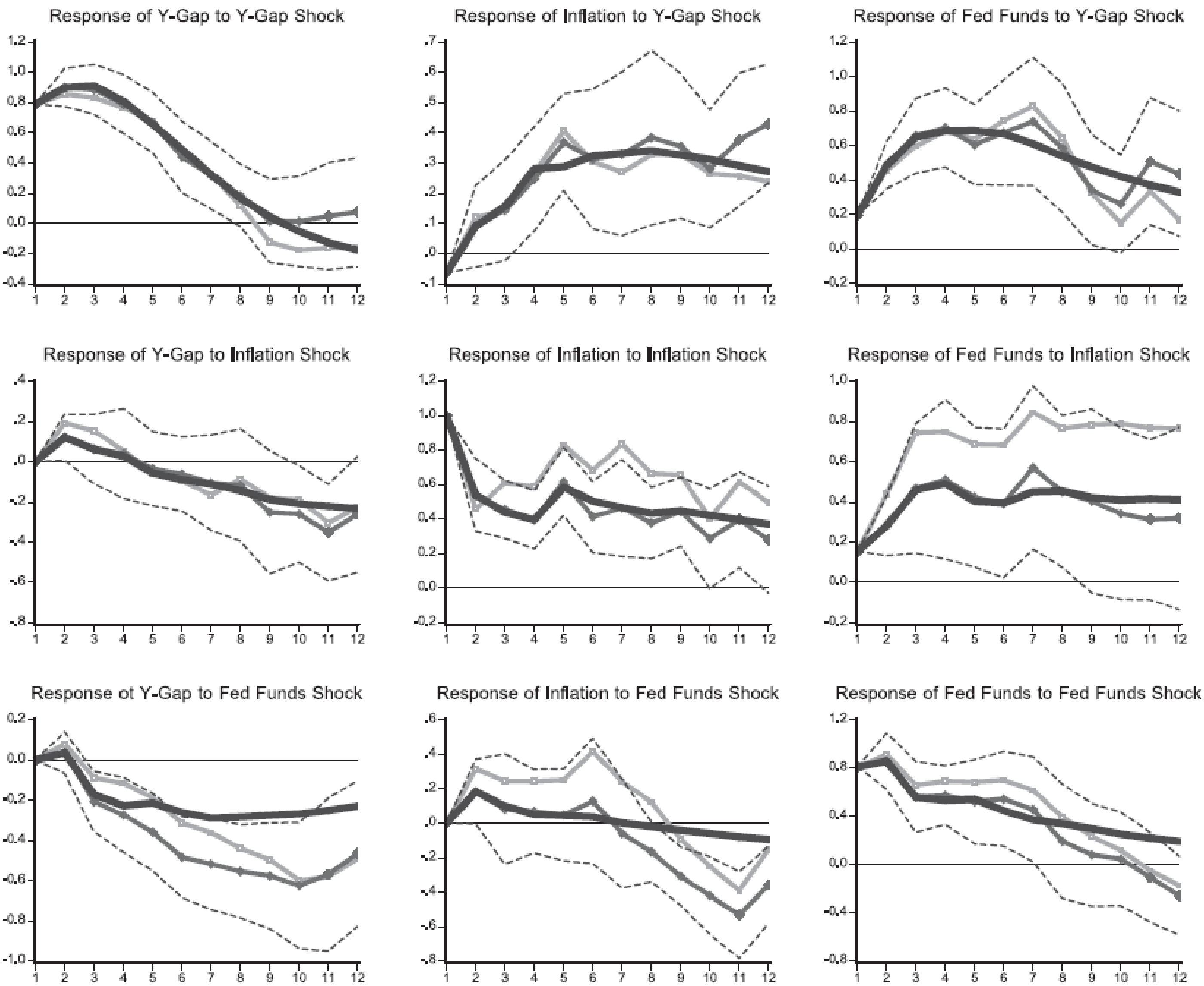

FIGURE 5. IMPULSE RESPONSES FOR THE NEW KEYNESIAN MODEL BASED ON A VAR, AND LINEAR AND CUBIC PROJECTIONS

Notes: The thick line is the response calculated from a VAR; the solid line with crosses is the response calculated by linear projection; the two dashed lines are 95-percent confidence level error bands for the individual coefficients of the linear projection response; and the solid line with circles is the response calculated by cubic projection evaluated at the sample mean. All responses calculated with four lags.

Figure 3: Figure 5 by Jorda (2005)

3.a (20 points): *Focus on the impulse response of inflation to inflation shocks. Compare the results achieved by our three estimation methods.*

3.b (30 points): *Focus on the impulse response of the output gap to interest rate shocks. Compare the results achieved by our three estimation methods.*

3.c (40 points): *Focus on the impulse response of inflation to interest rate shocks. Interpret (separately) the results achieved by each one of your three estimation methods.*

Answer 3.a:

The results achieved by our VAR estimator and our local-linear projection are similar to each other. (10 points) However, the local-cubic projection response shows that inflation is considerably more persistent to its own shocks than what is reflected by the responses calculated by either linear method. (10 points)

Answer 3.b:

In the short run, all estimators produce similar results. (10 points) However, in the long run, there are interesting differences in the point-estimates. While both local-projection methods produce similar estimates (10 points), our VAR model estimates a much weaker response of output to interest rate shocks. (10 points)

Answer 3.c:

The VAR estimates are very close to zero. (10 points)

The local-cubic estimates are initially positive, but become negative after the eighth quarter. (10 points)

The local-linear estimates are close to zero up to the seventh quarter. (10 points) After the eighth quarter, they become negative and statistically different from zero. (10 points)

Question 4 (Estimating an ARMA(1,2) using GMM - 40 points)

ARMA(p,q) models are usually estimated by Maximum Likelihood. However, we can also estimate them by GMM.

Let $\{Y_t\}$ *be a stationary stochastic process such that*

$$Y_t = \phi_1 \cdot Y_{t-1} + U_t \tag{5}$$

and

$$U_t = \epsilon_t + \theta_1 \cdot \epsilon_{t-1} + \theta_2 \cdot \epsilon_{t-2}, \tag{6}$$

where $\phi_1, \theta_1, \theta_2 > 0$ *and* $\epsilon_t \sim i.i.d.N\,(0,1)$.

4.a (10 points): *Show that you cannot consistently estimate Equation (5) using OLS.*

4.b (10 points): *Show that* Y_{t-2} *is not a valid instrument for* Y_{t-1} *in Equation (5).*

4.c (20 points): *Show that* Y_{t-3} *is a valid instrument for* Y_{t-1} *in Equation (5).*

Answer 4.a:

We must show that the assumption behind OLS estimation do not hold. In particular, we will show that $\mathbb{E}\left[Y_{t-1} \cdot U_t\right] \neq 0$. (10 points including the proof)

$$
\begin{aligned}
\mathbb{E}\left[Y_{t-1} \cdot U_t\right] &= \mathbb{E}\left[\left(\phi_1 \cdot Y_{t-2} + U_{t-1}\right) \cdot U_t\right] \\
&\qquad \text{according to Equation (5)} \\
&= \mathbb{E}\left[\left(\phi_1^2 \cdot Y_{t-3} + \phi_1 \cdot U_{t-2} + U_{t-1}\right) \cdot U_t\right] \\
&\qquad \text{according to Equation (5)} \\
&= \mathbb{E}\left[\begin{array}{l} \left(\phi_1^2 \cdot Y_{t-3} + \phi_1 \cdot \{\epsilon_{t-2} + \theta_1 \cdot \epsilon_{t-3} + \theta_2 \cdot \epsilon_{t-4}\} + \{\epsilon_{t-1} + \theta_1 \cdot \epsilon_{t-2} + \theta_2 \cdot \epsilon_{t-3}\}\right) \\ \qquad \cdot \left(\epsilon_t + \theta_1 \cdot \epsilon_{t-1} + \theta_2 \cdot \epsilon_{t-2}\right) \end{array}\right] \\
&\qquad \text{according to Equation (6)} \\
&= \theta_2 \cdot (\phi_1 + \theta_1) \cdot \mathbb{E}\left[\epsilon_{t-2}^2\right] + \theta_1 \cdot \mathbb{E}\left[\epsilon_{t-1}^2\right] \\
&\qquad \text{because } \epsilon_t \text{ is independently distributed} \\
&= \theta_2 \cdot (\phi_1 + \theta_1) + \theta_1 \\
&\qquad \text{because } \epsilon_t \text{ is identically distributed distributed with variance 1} \\
&\neq 0
\end{aligned}
$$

because $\phi_1, \theta_1, \theta_2 > 0$.

Answer 4.b:

We must show that the assumption behind IV estimation do not hold when we use Y_{t-2} to instrument for Y_{t-1} in Equation (5). In particular, we will show that $\mathbb{E}\left[Y_{t-2} \cdot U_t\right] \neq 0$. (10 points including the proof)

$$
\begin{aligned}
\mathbb{E}\left[Y_{t-2} \cdot U_t\right] &= \mathbb{E}\left[\left(\phi_1 \cdot Y_{t-3} + \epsilon_{t-2} + \theta_1 \cdot \epsilon_{t-3} + \theta_2 \cdot \epsilon_{t-4}\right) \cdot \left(\epsilon_t + \theta_1 \cdot \epsilon_{t-1} + \theta_2 \cdot \epsilon_{t-2}\right)\right] \\
&= \theta_2 \neq 0.
\end{aligned}
$$

Answer 4.c:

We must show that $\mathbb{E}\left[Y_{t-1} \cdot Y_{t-3}\right] \neq 0$ and $\mathbb{E}\left[Y_{t-3} \cdot U_t\right] = 0$.

Note that $\mathbb{E}[Y_{t-3} \cdot U_t] = 0$ because Y_{t-3} depends only on terms $\{\epsilon_{t-j}\}_{j=3}^{+\infty}$ while U_t depends only on terms ϵ_t, ϵ_{t-1} and ϵ_{t-2}. (10 points including the proof)

Now, we will show that $\mathbb{E}[Y_{t-1} \cdot Y_{t-3}] \neq 0$. (10 points including the proof)

$$
\begin{aligned}
\mathbb{E}[Y_{t-1} \cdot Y_{t-3}] &= \mathbb{E}[(\phi_1 \cdot Y_{t-2} + U_{t-1}) \cdot Y_{t-3}] \\
&= \mathbb{E}\left[\left(\phi_1^2 \cdot Y_{t-3} + \phi_1 \cdot U_{t-2} + U_{t-1}\right) \cdot Y_{t-3}\right] \\
&= \phi_1^2 \cdot \mathbb{E}\left[Y_{t-3}^2\right] + \mathbb{E}[(\phi_1 \cdot U_{t-2} + U_{t-1}) \cdot Y_{t-3}] \\
&= \phi_1^2 \cdot \sigma_Y^2 + \mathbb{E}[(\phi_1 \cdot U_{t-2} + U_{t-1}) \cdot Y_{t-3}] \\
&\qquad \text{because } \{Y_t\} \text{ is a stationary process} \\
&= \phi_1^2 \cdot \sigma_Y^2 + \mathbb{E}[(\phi_1 \cdot U_{t-2} + U_{t-1}) \cdot (\phi_1 \cdot Y_{t-4} + U_{t-3})] \\
&= \phi_1^2 \cdot \sigma_Y^2 + \mathbb{E}\left[(\phi_1 \cdot U_{t-2} + U_{t-1}) \cdot \left(\phi_1^2 \cdot Y_{t-5} + \phi_1 \cdot U_{t-4} + U_{t-3}\right)\right] \\
&= \phi_1^2 \cdot \sigma_Y^2 + \mathbb{E}[(\phi_1 \cdot U_{t-2} + U_{t-1}) \cdot (\phi_1 \cdot U_{t-4} + U_{t-3})] \\
&= \phi_1^2 \cdot \sigma_Y^2 + \mathbb{E}\left[\begin{array}{l} (\phi_1 \cdot \{\epsilon_{t-2} + \theta_1 \cdot \epsilon_{t-3} + \theta_2 \cdot \epsilon_{t-4}\} + \{\epsilon_{t-1} + \theta_1 \cdot \epsilon_{t-2} + \theta_2 \cdot \epsilon_{t-3}\}) \\ \qquad \cdot (\phi_1 \cdot \{\epsilon_{t-4} + \theta_1 \cdot \epsilon_{t-5} + \theta_2 \cdot \epsilon_{t-6}\} + \{\epsilon_{t-3} + \theta_1 \cdot \epsilon_{t-4} + \theta_2 \cdot \epsilon_{t-5}\}) \end{array}\right] \\
&= \phi_1^2 \cdot \sigma_Y^2 + \phi_1 \cdot \theta_1 + \theta_2 + \phi_1 \cdot \theta_2 \cdot (\phi_1 + \theta_1) \\
&> 0.
\end{aligned}
$$

Question 5 (Dynamic Multipliers of Nonstationary Processes - 60 points)

For any $s \in \{0\} \cup \mathbb{N}$, the dynamic multiplier s-periods ahead for a stochastic process $\{Y_t\}$ is given by

$$\frac{\partial Y_{t+s}}{\partial \epsilon_t}$$

and it captures the consequences for Y_{t+s} if ϵ_t were to increase by one unit with ϵ's for all other dates unaffected.

5.a (30 points): *Let $\{Y_t\}$ be a MA(1) process with a deterministic time trend, i.e.,*

$$Y_t = \alpha + \delta \cdot t + \epsilon_t + \theta \cdot \epsilon_{t-1} \tag{7}$$

where $\{\epsilon_t\}$ is a white noise process.

What is the dynamic multiplier s-periods ahead for this stochastic process and any $s \in \{0\} \cup \mathbb{N}$?

5.b (20 points): *Let $\{Y_t\}$ be a I(1) process with a drift, i.e.,*

$$Y_t = \delta + Y_{t-1} + \epsilon_t \tag{8}$$

where $\{\epsilon_t\}$ is a white noise process.

What is the dynamic multiplier s-periods ahead for this stochastic process and any $s \in \{0\} \cup \mathbb{N}$?

5.c (10 points): *Given the results above, explain why unit root processes are considered to have infinite memory.*

Answer 5.a:

For $s = 0$, Equation (7) implies that $\dfrac{\partial Y_t}{\partial \epsilon_t} = 1$. (10 points including the proof)

For $s = 1$, Equation (7) implies that

$$Y_{t+1} = \alpha + \delta \cdot (t+1) + \epsilon_{t+1} + \theta \cdot \epsilon_t,$$

implying that $\dfrac{\partial Y_{t+1}}{\partial \epsilon_t} = \theta$. (10 points including the proof)

For $s \geq 2$, Equation (7) implies that

$$Y_{t+s} = \alpha + \delta \cdot (t+s) + \epsilon_{t+s} + \theta \cdot \epsilon_{t+s-1},$$

implying that $\dfrac{\partial Y_{t+s}}{\partial \epsilon_t} = 0$ because $t + s - 1 > t$. (10 points including the proof)

Answer 5.b:

Fix $s \in \{0\} \cup \mathbb{N}$ arbitrarily. Note that Equation (8) implies that

$$\begin{aligned}
Y_{t+s} &= \delta + Y_{t+s-1} + \epsilon_{t+s} \\
&= \delta + (\delta + Y_{t+s-2} + \epsilon_{t+s-1}) + \epsilon_{t+s} \\
&= 2 \cdot \delta + Y_{t+s-2} + \epsilon_{t+s} + \epsilon_{t+s-1} \\
&\vdots \\
&= (s+1) \cdot \delta + Y_{t-1} + \epsilon_t + \sum_{\tau=t+1}^{t+s} \epsilon_\tau, \qquad \text{(10 points)}
\end{aligned}$$

implying that $\dfrac{\partial Y_{t+1}}{\partial \epsilon_t} = 1$. (10 points)

Answer 5.c:

Based on the results of Question 5.b, we observe that the effect of a shock in period t never dissipates in an I(1) process with drift. For this reason, we say that unit root processes have infinite memory. (10 points)

In contrast, we note that the effect of a shock in period t quickly dissipates in a MA(1) process with a deterministic time trend.

References

Jorda, O. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review 95*(1), pp. 161–182. (Cited on pages 4, 6, 12, 13, 15, 17, and 18.)

Sketching

The next pages are here for you to sketch your work. They will not be read nor graded. The instructor will most definitely not read them.