Exam Solution Guide

Econometrics II

December 2018

Part 1
Forecasting GDP Growth

The Case The goal of this part of the exam is to estimate a univariate dynamic model and use the model to forecast the future growth rate of GDP.

The Data Graphs of the data and relevant transformations must be shown in the exam. It must be noted that the level of the GDP is clearly non-stationary, but the first-difference appears somewhat stationary.

Econometric Theory The econometric theory must include the following:

(1) A precise definition and interpretation of the model considered and its properties. Specifically, a univariate autoregressive (AR) or autoregressive moving-average (ARMA) model must be presented. Furthermore, a precise definition of the stationarity condition, the out-of-sample forecasts, and the forecast variances must be given.

(2) A precise description of the estimator used, in particular a precise account of the assumptions used to derive the estimator. Specifically, the method of moments (MM) or the maximum likelihood (ML) estimators can be used.

(3) A precise account of the necessary assumptions for consistent estimation and valid inference. This includes a precise definition of the null hypotheses, test statistics, and asymptotic distributions used to test relevant hypotheses.

(4) The theory must be presented precisely and in a logical order with a consistent and correct notation.

Empirical Results The empirical results must include the following:

(1) A presentation and discussion of the relevant empirical results, so that the reader is able to understand the steps carried out in the process as well as the conclusions made.

(2) A description of the model selection process based on a general-to-specific approach, information criteria, or both.

(3) A discussion of whether the assumptions for consistent estimation and valid inference are satisfied for the estimated models. Specifically, this includes misspecification testing, which must be presented and discussed before statistical testing is carried out.

(4) A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the conclusion. Specifically, the conclusion regarding the out-of-sample forecasts and the forecast variance must be presented and the limitations of the estimated models must be discussed in relation to the forecasts.

Part 2
The Effects of Housing and Financial Wealth on Consumption

The Case The goal of this part of the exam is to use cointegration techniques to estimate the short-run and long-run effect of income, housing wealth, and financial wealth on consumption for the United States.

The Data Graphs of the data and relevant transformations must be shown in the exam. It must be noted that the level of the variables are clearly non-stationary, but the first-differences appear somewhat stationary and they seem to move together over time indicating cointegration.

Econometric Theory The econometric theory must include the following:

(1) A precise definition and interpretation of the models considered and their properties. Specifically, an interpretation of cointegration must be presented along with a presentation of univariate autoregressive (AR) models used to test for unit roots and a single equation cointegration approach based on either the Engle-Granger two-step procedure or the autoregressive distributed lag (ADL) and error-correction models (ECM).

(2) A precise description of the estimator used, in particular a precise account of the assumptions used to derive an estimator.

(3) A precise account of the necessary assumptions for consistent estimation and valid inference. This includes a precise definition of the null hypotheses, test statistics, and asymptotic distributions used to test relevant hypotheses.

(4) The theory must be presented precisely and in a logical order with a consistent and correct notation.

Empirical Results The empirical results must include the following:

(1) A presentation and discussion of the relevant empirical results, so that the reader is able to understand the steps carried out in the process as well as the conclusions made.

(2) A description of the model selection process based on a general-to-specific approach, information criteria, or both.

(3) A discussion of whether the assumptions for consistent estimation and valid inference are satisfied for the estimated models. Specifically, this includes misspecification testing, which must be presented and discussed before statistical testing is carried out.

(4) A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the conclusion. Specifically, the conclusion regarding cointegration between consumption, income, housing wealth, and financial wealth must be presented and the limitations of the single-equation cointegration approach must be discussed in relation to the conclusion.

Part 3
Monetary Policy and Asset Price Volatility

The Case The goal of this part of the exam is to analyze if there is empirical evidence of the Federal Reserve setting its interest rate in accordance with a Taylor rule with interest rate smoothing extended with stock returns during the period where Alan Greenspan served as Chairman of the Federal Reserve.

The Data Graphs of the data and relevant transformations must be shown in the exam.

Econometric Theory The econometric theory must include the following:

(1) A precise definition and interpretation of the economic model considered and its properties.

(2) A precise description of the generalized method of moments (GMM) estimator used. In particular, this includes a precise account of how the set of moment conditions and valid instruments are derived from the theoretical model and how these moment conditions are used to derive the GMM estimator.

(3) A precise account of the necessary assumptions for consistent estimation and valid inference. This includes a precise definition of the null hypotheses, test statistics, and asymptotic distributions used to test relevant hypotheses.

(4) The theory must be presented in a logical order and with a consistent and correct notation.

Empirical Results The empirical results must include the following:

(1) A presentation and discussion of the relevant empirical results, so that the reader is able to understand the steps carried out in the process as well as the conclusions made.

(2) A discussion of whether the assumptions for consistent estimation and valid inference are satisfied for the estimated models.

(3) A robustness analysis of the estimated model.

(4) A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the conclusion. Specifically, if there is empirical evidence of the three different versions of the New Keynesian Phillips curve.

Part 4
Theoretical Problems

#4.1 Maximum Likelihood Estimation

Consider the model for z_t given by:

$$z_t = \rho_t \cdot z_{t-1} + \epsilon_t, \tag{4.1}$$
$$\rho_t = \rho + \sqrt{\alpha} \cdot \eta_{1t}, \tag{4.2}$$
$$\epsilon_t = \sqrt{\omega} \cdot \eta_{2t}, \tag{4.3}$$

for $t = 1, 2, \ldots, T$, where z_0 is given, $0 < \rho < 1$, $\alpha > 0$, and $\omega > 0$. The innovations $(\eta_{1t}, \eta_{2t})'$ follow the multivariate Gaussian process:

$$\begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \sim i.i.d.N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right). \tag{4.4}$$

Question 1

The conditional expectation $E(z_t|z_{t-1})$ is derived as:

$$\begin{aligned}
E\left(z_t \mid z_{t-1}\right) &= E\left(\rho_t z_{t-1} + \epsilon_t \mid z_{t-1}\right) \\
&= E\left((\rho + \sqrt{\alpha}\,\eta_{1t}) z_{t-1} + \sqrt{\omega}\,\eta_{2t} \mid z_{t-1}\right) \\
&= E\left(\rho z_{t-1} + \sqrt{\alpha}\,\eta_{1t} z_{t-1} + \sqrt{\omega}\,\eta_{2t} \mid z_{t-1}\right) \\
&= E\left(\rho z_{t-1} \mid z_{t-1}\right) + E\left(\sqrt{\alpha}\,\eta_{1t} z_{t-1} \mid z_{t-1}\right) + E\left(\sqrt{\omega}\,\eta_{2t} \mid z_{t-1}\right) \\
&= \rho z_{t-1} + \sqrt{\alpha}\,E\left(\eta_{1t} \mid z_{t-1}\right) z_{t-1} + \sqrt{\omega}\,E\left(\eta_{2t} \mid z_{t-1}\right) \\
&= \rho z_{t-1} + \sqrt{\alpha}\,E\left(\eta_{1t}\right) z_{t-1} + \sqrt{\omega}\,E\left(\eta_{2t}\right) \\
&= \rho z_{t-1} + \sqrt{\alpha} \cdot 0 \cdot z_{t-1} + \sqrt{\omega} \cdot 0 \\
&= \rho z_{t-1}.
\end{aligned} \tag{1}$$

Here, we have used that $E(z_{t-1}|z_{t-1}) = z_{t-1}$ and that $E(\eta_{1t}|z_{t-1}) = E(\eta_{1t}) = 0$ and $E(\eta_{2t}|z_{t-1}) = E(\eta_{2t}) = 0$ as η_{1t} and η_{2t} are independent of z_{t-1} and with mean zero.

Using the result in (1), the conditional variance $V(z_t|z_{t-1})$ is derived as:

$$\begin{aligned}
V\left(z_t \mid z_{t-1}\right) &= E\left(\left(z_t - E\left(z_t|z_{t-1}\right)\right)^2 \mid z_{t-1}\right) \\
&= E\left(\left(\rho_t z_{t-1} + \epsilon_t - \rho z_{t-1}\right)^2 \mid z_{t-1}\right) \\
&= E\left(\left((\rho + \sqrt{\alpha}\,\eta_{1t}) z_{t-1} + \sqrt{\omega}\,\eta_{2t} - \rho z_{t-1}\right)^2 \mid z_{t-1}\right) \\
&= E\left(\left(\rho z_{t-1} + \sqrt{\alpha}\,\eta_{1t} z_{t-1} + \sqrt{\omega}\,\eta_{2t} - \rho z_{t-1}\right)^2 \mid z_{t-1}\right) \\
&= E\left(\left(\sqrt{\alpha}\,\eta_{1t} z_{t-1} + \sqrt{\omega}\,\eta_{2t}\right)^2 \mid z_{t-1}\right) \\
&= E\left(\alpha \eta_{1t}^2 z_{t-1}^2 + \omega \eta_{2t}^2 + 2\sqrt{\alpha}\,\eta_{1t}\sqrt{\omega}\,\eta_{2t} \mid z_{t-1}\right) \\
&= \alpha E\left(\eta_{1t}^2 z_{t-1}^2 \mid z_{t-1}\right) + \omega E\left(\eta_{2t}^2 \mid z_{t-1}\right) + 2\sqrt{\alpha}\sqrt{\omega}\, E\left(\eta_{1t}\eta_{2t} \mid z_{t-1}\right) \\
&= \alpha E\left(\eta_{1t}^2\right) z_{t-1}^2 + \omega E\left(\eta_{2t}^2\right) + 2\sqrt{\alpha}\sqrt{\omega}\, E\left(\eta_{1t}\eta_{2t}\right) \\
&= \alpha \cdot 1 \cdot z_{t-1}^2 + \omega \cdot 1 + 2\sqrt{\alpha}\sqrt{\omega} \cdot 0 \\
&= \alpha z_{t-1}^2 + \omega. \qquad (2)
\end{aligned}$$

Again, we have used that $E(z_{t-1}|z_{t-1}) = z_{t-1}$ and that $E(\eta_{1t}|z_{t-1}) = E(\eta_{1t}) = 0$ and $E(\eta_{2t}|z_{t-1}) = E(\eta_{2t}) = 0$ as η_{1t} and η_{2t} are independent of z_{t-1}. Moreover, in the second last step, we have used that η_{1t} are η_{2t} are uncorrelated, $E(\eta_{1t}\eta_{2t}) = 0$, which follows from the specification in (4.4).

The result in (1) shows that the conditional mean of z_t depends on the level of z_{t-1}, while the result in (2) shows that the conditional variance of z_t depends on z_{t-1}^2. The parameter restrictions $\alpha > 0$ and $\omega > 0$ ensure that the conditional variance is strictly positive.

It can be noted that if $\alpha = 0$, the model simplifies to a standard (weakly) stationary first-order autoregressive model with i.i.d. normally distributed innovations ϵ_t with mean zero and variance ω.

Question 2

The results in (1) and (2) show that z_t given z_{t-1} is conditionally normally distributed with a conditional mean of ρz_{t-1} and a conditional variance of $\alpha z_{t-1}^2 + \omega$:

$$z_t \mid z_{t-1} \sim N(\rho z_{t-1},\, \alpha z_{t-1}^2 + \omega). \qquad (3)$$

It should be noted that the parameters of the model are given by $\theta = (\rho, \alpha, \omega)'$. These parameters can be estimated by maximum likelihood.

It can be noted that the joint density of $(z_0, z_1, z_2, \ldots, z_T)$ can be factorized as:

$$f(z_0, z_1, z_2, \ldots, z_T) = f(z_0) \cdot \prod_{t=1}^{T} f(z_t|z_{t-1}), \qquad (4)$$

such that the joint density of $(z_1, z_2, \ldots, z_T)$ conditional on the initial value z_0 can be written as:

$$f(z_1, z_2, \ldots, z_T \mid z_0) = \prod_{t=1}^{T} f(z_t|z_{t-1}). \qquad (5)$$

Given the conditional distribution of z_t given z_{t-1} in (3), the log-likelihood function for $(z_1, z_2, \ldots, z_T)$ conditional on the initial value z_0 is given by:

$$logL(\theta) = -\frac{T}{2} log(2\pi(\alpha z_{t-1}^2 + \omega)) - \sum_{t=1}^{T} \frac{(z_t - \rho z_{t-1})^2}{2(\alpha z_{t-1}^2 + \omega)}. \tag{6}$$

The maximum likelihood estimator $\widehat{\theta}_{ML}$ is found by maximizing the log-likelihood function with respect to θ. This can be done using numerical optimization.

#4.2 Mean-Squared Forecast Errors

Consider the model for y_t given by:

$$y_t = \rho \cdot y_{t-1} + \epsilon_t, \tag{4.5}$$
$$\epsilon_t = \eta_t + \alpha \cdot \eta_{t-1}, \qquad \eta_t \sim i.i.d.N(0, \sigma_\eta^2), \tag{4.6}$$

for $t = 1, 2, \ldots, T$, where $0 < \rho < 1$, $\alpha > 0$, and the initial values y_0 and η_0 are given.

Define the information set at time t as $\mathcal{I}_t = \{y_0, \eta_0, y_1, \eta_1, \ldots, y_t, \eta_t\}$.

Question 1

Using (4.5) and (4.6), we derive the forecast $\widehat{y}_{T+1|T} = E(y_{T+1}|\mathcal{I}_T)$ as:

$$\begin{aligned}
\widehat{y}_{T+1|T} &= E\left(y_{T+1} \mid \mathcal{I}_T\right) \\
&= E\left(\rho y_T + \epsilon_{T+1} \mid \mathcal{I}_T\right) \\
&= E\left(\rho y_T + \eta_{T+1} + \alpha\eta_T \mid \mathcal{I}_T\right) \\
&= \rho E\left(y_T \mid \mathcal{I}_T\right) + E\left(\eta_{T+1} \mid \mathcal{I}_T\right) + \alpha E\left(\eta_T \mid \mathcal{I}_T\right) \\
&= \rho y_T + 0 + \alpha\eta_T \\
&= \rho y_T + \alpha\eta_T.
\end{aligned} \tag{7}$$

It should be noted that η_T is included in the information set $\mathcal{I}_T$, so that $E(\eta_T|\mathcal{I}_T) = \eta_T$. Moreover, we have used that $E(\eta_{T+1}|\mathcal{I}_T) = 0$.

Using (4.5), (4.6), and (7), we derive the mean-squared error of the forecast $\widehat{y}_{T+1|T}$ as:

$$\begin{aligned}
MSE(\widehat{y}_{T+1|T}) &= E\left(\left(y_{T+1} - \widehat{y}_{T+1|T}\right)^2\right) \\
&= E\left(\left(\rho y_T + \epsilon_{T+1} - \rho y_T - \alpha\eta_T\right)^2\right) \\
&= E\left(\left(\rho y_T + \eta_{T+1} + \alpha\eta_T - \rho y_T - \alpha\eta_T\right)^2\right) \\
&= E\left(\eta_{T+1}^2\right) \\
&= \sigma_\eta^2.
\end{aligned} \tag{8}$$

It can be noted that the mean-squared error of the forecast $\widehat{y}_{T+1|T}$ equals the unconditional variance of the forecast error η_{T+1}.

Question 2

It can be noted that the model in (4.5)–(4.6) is a first-order autoregressive (AR) model with autocorrelated residuals ϵ_t, or, equivalently, a first-order autoregressive moving-average (ARMA) model.

To derive the unconditional mean of y_t, we first derive the moving average representation for y_t in terms of the innovations η_t by recursive substitution:

$$\begin{aligned}
y_t &= \rho y_{t-1} + \epsilon_t \\
&= \rho \left(\rho y_{t-2} + \eta_{t-1} + \alpha \eta_{t-2}\right) + \eta_t + \alpha \eta_{t-1} \\
&= \rho^2 y_{t-2} + \eta_t + (\alpha + \rho)\eta_{t-1} + \rho\alpha\eta_{t-2} \\
&= \rho^2 \left(\rho y_{t-3} + \eta_{t-2} + \alpha\eta_{t-3}\right) + \eta_t + (\alpha+\rho)\eta_{t-1} + \rho\alpha\eta_{t-2} \\
&= \rho^3 y_{t-3} + \eta_t + (\alpha+\rho)\eta_{t-1} + \rho(\alpha+\rho)\eta_{t-2} + \rho^2\alpha\eta_{t-3} \\
&= \rho^3 \left(\rho y_{t-4} + \eta_{t-3} + \alpha\eta_{t-4}\right) + \eta_t + (\alpha+\rho)\eta_{t-1} + \rho(\alpha+\rho)\eta_{t-2} + \rho^2\alpha\eta_{t-3} \\
&= \rho^4 y_{t-4} + \eta_t + (\alpha+\rho)\eta_{t-1} + \rho(\alpha+\rho)\eta_{t-2} + \rho^2(\alpha+\rho)\eta_{t-3} + \rho^3\alpha\eta_{t-4} \\
&\vdots \qquad (9) \\
&= \rho^t y_0 + \eta_t + (\alpha+\rho)\sum_{i=0}^{t-2}\rho^i\eta_{t-1-i} + (\alpha+\rho)\rho^{t-1}\eta_0. \qquad (10)
\end{aligned}$$

Continuing the recursive substitution infinitely back in time using that $0 < \rho < 1$, such that $\rho^i \to 0$ for $i \to \infty$, we get the infinite moving average representation:

$$y_t = \eta_t + \sum_{i=0}^{\infty}\rho^i(\alpha+\rho)\eta_{t-1-i} = \eta_t + (\alpha+\rho)\sum_{i=0}^{\infty}\rho^i\eta_{t-1-i} \qquad (11)$$

We can now find the unconditional mean of y_t using (11):

$$E\left(y_t\right) = E\left(\eta_t + (\alpha+\rho)\sum_{i=0}^{\infty}\rho^i\eta_{t-1-i}\right) = E\left(\eta_t\right) + (\alpha+\rho)\sum_{i=0}^{\infty}\rho^i E\left(\eta_{t-1-i}\right) = 0. \qquad (12)$$

Thus, we find an unconditional mean of zero. That implies that $E(y_t^2)$ equals the unconditional variance of y_t. Using that η_t is independent over time, such that $E(\eta_t\eta_s) = 0$ for all $t \neq s$, and that $0 < \rho < 1$ implies that $\sum_{i=0}^{\infty}\rho^{2i} = \frac{1}{1-\rho^2}$, we find the unconditional variance as:

$$\begin{aligned}
E\left(y_t^2\right) &= E\left(\left(\eta_t + (\alpha+\rho)\sum_{i=0}^{\infty}\rho^i\eta_{t-1-i}\right)^2\right) \\
&= E\left(\eta_t^2\right) + (\alpha+\rho)^2\sum_{i=0}^{\infty}\rho^{2i}E\left(\eta_{t-1-i}^2\right) \\
&= \sigma_\eta^2 + (\alpha+\rho)^2\sum_{i=0}^{\infty}\rho^{2i}\sigma_\eta^2 \\
&= \sigma_\eta^2 + \frac{(\alpha+\rho)^2}{1-\rho^2}\sigma_\eta^2. \qquad (13)
\end{aligned}$$

It can be noted that if $\alpha = 0$, the model simplifies to a stationary first-order autoregressive process with mean zero and an unconditional variance of $\frac{\sigma_\eta^2}{1-\rho^2}$.

QUESTION 3

The naive forecast is given by $\tilde{y}_{T+1|T} = y_T$. Inserting the naive forecast and the expression for y_{T+1} from (4.5) and (4.6), we derive the mean-squared error as:

$$\begin{aligned}
MSE(\tilde{y}_{T+1|T}) &= E\left(\left(y_{T+1} - \tilde{y}_{T+1|T}\right)^2\right) \\
&= E\left(\left(\rho y_T + \epsilon_{T+1} - y_T\right)^2\right) \\
&= E\left(\left((\rho-1)y_T + \epsilon_{T+1}\right)^2\right) \\
&= E\left((\rho-1)^2 y_T^2 + \epsilon_{T+1}^2 + 2(\rho-1)y_T\epsilon_{T+1}\right) \\
&= (\rho-1)^2 E\left(y_T^2\right) + E\left(\epsilon_{T+1}^2\right) + 2(\rho-1)E\left(y_T\epsilon_{T+1}\right). \qquad (14)
\end{aligned}$$

In Question 2, we found the unconditional variance $E(y_t^2) = V(y_t)$. As y_t is stationary, this unconditional variance is the same at all point in time, so $E(y_T^2) = E(y_t^2)$ as given in (13).

We use (4.6) to derive the unconditional expectation $E(\epsilon_{T+1}^2)$ as:

$$\begin{aligned}
E(\epsilon_{T+1}^2) &= E\left(\left(\eta_{T+1} + \alpha\eta_T\right)^2\right) \\
&= E\left(\eta_{T+1}^2 + \alpha^2\eta_T^2 + 2\eta_{T+1}\eta_T\right) \\
&= E\left(\eta_{T+1}^2\right) + \alpha^2 E\left(\eta_T^2\right) + 2E\left(\eta_{T+1}\eta_T\right) \\
&= \sigma_\eta^2 + \alpha^2\sigma_\eta^2 + 0 \\
&= (1+\alpha^2)\sigma_\eta^2, \qquad (15)
\end{aligned}$$

where we have used that η_t is independent over time, such that $E(\eta_{T+1}\eta_T) = 0$.

Finally, we derive the unconditional expectation $E(y_T\epsilon_{T+1})$ in (14) as:

$$\begin{aligned}
E\left(y_T\epsilon_{T+1}\right) &= E\left((\rho y_{T-1} + \eta_T + \alpha\eta_{T-1})(\eta_{T+1} + \alpha\eta_T)\right) \\
&= E\left(\rho y_{T-1}\eta_{T+1} + \alpha\rho y_{T-1}\eta_T + \eta_T\eta_{T+1} + \alpha\eta_T^2 + \alpha\eta_{T-1}\eta_{T+1} + \alpha^2\eta_{T-1}\eta_T\right) \\
&= \rho E\left(y_{T-1}\eta_{T+1}\right) + \alpha\rho E\left(y_{T-1}\eta_T\right) + E\left(\eta_T\eta_{T+1}\right) + \alpha E\left(\eta_T^2\right) \\
&\qquad + \alpha E\left(\eta_{T-1}\eta_{T+1}\right) + \alpha^2 E\left(\eta_{T-1}\eta_T\right) \\
&= \rho\cdot 0 + \alpha\rho\cdot 0 + 0 + \alpha\sigma_\eta^2 + \alpha\cdot 0 + \alpha^2\cdot 0 \\
&= \alpha\sigma_\eta^2. \qquad (16)
\end{aligned}$$

Here, we have used that $E(y_{T-1}\eta_{T+i}) = 0$ for all $i \geq 0$ as y_{T-1} only depends on η_{T-j} for $j = 1, 2, \ldots$. Moreover, we use have used that η_t is independent over time, such that $E(\eta_t\eta_s) = 0$ for all $t \neq s$.

We can now insert the expressions in (13), (15), and (16) into (14) to finde the mean-squared error of the naive forecast:

$$
\begin{aligned}
MSE(\tilde{y}_{T+1|T}) &= E\left(\left(y_{T+1} - \tilde{y}_{T+1|T}\right)^2\right) \\
&= (\rho-1)^2 E\left(y_T^2\right) + E\left(\epsilon_{T+1}^2\right) + 2(\rho-1)E\left(y_T\epsilon_{T+1}\right) \\
&= (\rho-1)^2\left(\sigma_\eta^2 + \frac{(\alpha+\rho)^2}{1-\rho^2}\sigma_\eta^2\right) + (1+\alpha^2)\sigma_\eta^2 + 2(\rho-1)\alpha\sigma_\eta^2 \\
&= \left((\rho-1)^2 + \frac{(\alpha+\rho)^2}{1-\rho^2} + 1 + \alpha^2 - 2\alpha + 2\alpha\rho\right)\sigma_\eta^2 \\
&= \left((\rho-1)^2 + \frac{(\alpha+\rho)^2}{1-\rho^2} + (1-\alpha)^2 + 2\alpha\rho\right)\sigma_\eta^2. \qquad (17)
\end{aligned}
$$

Note that the solution can be derived and expressed in different ways that are equivalently correct.

QUESTION 4

Inserting the parameter values $\rho = 0.5$ and $\alpha = 1$, we get the mean-squared errors of the two forecasts:

$$MSE(\widehat{y}_{T+1|T}) = \sigma_\eta^2. \qquad (18)$$

$$
\begin{aligned}
MSE(\tilde{y}_{T+1|T}) &= \left((0.5-1)^2 + \frac{(1+0.5)^2}{1-0.5^2} + (1-1)^2 + 2\cdot 1\cdot 0.5\right)\sigma_\eta^2 \\
&= (0.25 + 3 + 0 + 1)\,\sigma_\eta^2 = 4.25\cdot\sigma_\eta^2. \qquad (19)
\end{aligned}
$$

This illustrates that the mean-squared error is higher for the naive forecast $\tilde{y}_{T+1|T} = y_T$ compared to the forecast $\widehat{y}_{T+1|T} = E(y_{T+1}|\mathcal{I}_T) = \rho y_T + \alpha\eta_T$. That is, $MSE(\widehat{y}_{T+1|T}) < MSE(\tilde{y}_{T+1|T})$.

Though it not required to formally show this, it can be noted that this is a general result that holds for any parameter values.

In general, the forecast $\widehat{y}_{T+1|T} = E(y_{T+1}|\mathcal{I}_T)$ is the optimal forecast in the sense that, by construction, the mean-squared error will be equal to or larger for any other forecast: $MSE(\widehat{y}_{T+1|T}) \leq MSE(f(y_{T+1}|\mathcal{I}_T))$ for any forecast $f(y_{T+1}|\mathcal{I}_T)$. The intuition is that the optimal forecast given by the conditional expectation, by construction, minimizes the forecast error and thereby minimizes the mean-squared error of the forecast.