Studienzentrum Gerzensee Doctoral Program in Economics
Econometrics 2019 Week 3 Take-Home Empirical Exercise

The level of per-capita GDP currently stands at approximately \$40,000 (in US \$2011) in the world's industrialized economies. What will its level be 50 years from now? 100 years from now? 200 years from now? And, how certain can we be about these values? Answers to these questions are important for evaluating long-term economic policies such as policies aimed at mitigating climate change. In this exercise you will use historical data on GDP to construct predictive distributions for per-capita values of GDP at long horizons.

The file **Gerz_OECD_w3.xlsx** contains annual observations on per-capita values of GDP for the OECD from 1915-1914. Call this series Q_t. Let $Y_t = \ln(Q_t/Q_{t-1})$ denote the annual growth rate of Q_t.

1. Summary statistics:

(a) Plot Q_t. Plot Y_t.

(b) Calculate the sample mean and standard deviation of Y_t over the full sample period.

(c) Repeat (b) after dividing the sample into 25-year sub-samples.

(d) Comment on the evolution of Q_t and Y_t over the past century.

(e) (e.i) Compute the first 3 sample autocorrelations of Y_t.
(e.ii) Estimate an AR(2) model for Y_t. (Be sure to include a constant.)
(e.iii) Comment on the serial correlation properties of Y_t.

2. Suppose Y_t is a stationary stochastic process that is well-described by an AR(1) model: $Y_t = \mu + u_t$, where $u_t = \phi u_{t-1} + \varepsilon_t$. Suppose you have data $Y_1, \ldots, Y_T$. Let $\bar{Y}_{1:T} = \frac{1}{T}\sum_{t=1}^{T} Y_t$

(a) Let μ_Y denote the mean of Y. Show that $\sqrt{T}(\bar{Y}_{1:T} - \mu_Y) \xrightarrow{d} N(0,\Omega)$, and derive an expression for Ω.

(b) Suppose you want to forecast the average value of Y from T+1 to period T+h; that is, you are interested in $\bar{Y}_{T+1:T+h} = \frac{1}{h}\sum_{j=1}^{h} Y_{T+j}$. Suppose that h is large.

(b.i) Show that $\sqrt{h}(\bar{Y}_{T+1:T+h} - \mu_Y) \xrightarrow{d} N(0,V)$, and derive an expression for V.

(b.ii) Suppose you knew the value of μ_Y. A researcher constructs a 95% "prediction interval" for $\bar{Y}_{T+1:T+h}$ as $\mu_Y \pm 1.96 \times \sqrt{h^{-1}V}$. Will this interval contain the realized value of $\bar{Y}_{T+1:T+h}$ with 95% probability? Explain.

(b.iii) Suppose that $h = T$ (so you are interested in forecasting the average value of Y for T periods in the future). Show that

$$\sqrt{T}\begin{bmatrix} (\bar{Y}_{1:T} - \mu_Y) \\ (\bar{Y}_{T+1:2T} - \mu_Y) \end{bmatrix} \xrightarrow{d} N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega & 0 \\ 0 & \Omega \end{pmatrix}\right].$$

(b.iv) Consider a 95% "prediction interval" for $\bar{Y}_{T+1:2T}$ as $\bar{Y}_{1:T} \pm 1.96 \times \sqrt{T^{-1}(\Omega + \Omega)}$. Will this interval contain the realized value of $\bar{Y}_{T+1:T+h}$ with 95% probability? Explain.

(b.v) Discuss how the results in (b.i)-(b.iv) change if $h = \rho T$ where ρ is a fixed number that is greater than zero (for example, $\rho = 0.6$ or $\rho = 2$).

(c) Using the data for Y_t = growth rate of OECD per-capita GDP.

(c.i) Estimate an AR(1) model for Y_t (i.e., $Y_t = \mu + u_t$ with $u_t = \phi u_{t-1} + \varepsilon_t$. Equivalently, you can estimate $Y_t = \alpha + \phi Y_{t-1} + \varepsilon_t$).

(c.ii) Use the estimates to construct an estimate of Ω.

(c.iii) Use your analysis in parts (a)-(b) to form a 95% "prediction interval" for the average value of Y_t over the 2015-2054. Over 2015-2114.

(c.iv) Per capita OECD GDP in 2014 was $39,124. What is a 95% confidence interval for the value of per capita GDP in 2054? In 2114?

(d) Your answer in (c) assumed that Y_t followed an AR(1) model. Suppose you were unsure of this, and instead used a Newey-West estimator for Ω.

(d.i) Repeat part (c.ii) and (c.iv) using a Newey-West estimator with $k = 4$ lags.

(d.ii) How sensitive is your answer to using $k = 4$ lags? What about $k = 3$ or $k = 5$? How many lags do you think you should use? Why?

3. In question 2(a)-(c), you used the model $Y_t = \mu + u_t$, where $u_t = \phi u_{t-1} + \varepsilon_t$. In this question I want you to use the model:

$$Y_t = \mu_t + u_t$$
$$u_t = \phi u_{t-1} + \varepsilon_t$$
$$\mu_t = \mu_{t-1} + e_t$$

where $$\begin{bmatrix} \varepsilon_t \\ e_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_e^2 \end{bmatrix} \right)$$

(a) Suppose $\sigma_e = 0$. Show that this is same model as in parts 2(a)-(c).

(b) I want you to think about reasonable values for σ_e.

(b.i) Show the standard deviation of $(\mu_{t+k} - \mu_t) = k^{1/2}\sigma_e$.

(b.ii) Using $\sigma_e = 0.002$, show that the standard deviation of $(\mu_{t+25} - \mu_t) = 0.01$, and the standard deviation of $(\mu_{t+50} - \mu_t) = 0.014$.

(b.iii) Using $\sigma_e = 0.003$, show that the standard deviation of $(\mu_{t+25} - \mu_t) = 0.015$, and the standard deviation of $(\mu_{t+50} - \mu_t) = 0.021$.

(b.iv) Using $\sigma_e = 0.001$, show that the standard deviation of $(\mu_{t+25} - \mu_t) = 0.005$, and the standard deviation of $(\mu_{t+50} - \mu_t) = 0.007$.

(b.v) Discuss the 'reasonableness' of these values of σ_e for the growth rate of per-capita GDP for the OECD.

(c) Using $\sigma_e = 0.002$ and the values of ϕ and σ_ε you computed in (2.c.i), use the Kalman filter to compute estimates of μ_t for $t = 1915$-2014. Initialize the Kalman filter using

$$\begin{bmatrix} u_0 \\ \mu_0 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{\sigma_\varepsilon^2}{1-\phi^2} & 0 \\ 0 & 1 \end{bmatrix} \right).$$

(c.i) Why did I suggest this initialization for u_0? Is it reasonable?

(c.ii) How 'informative' is the initialization about the value of μ_0? Does it restrict μ_0 in an important way?

(c.iii) After running the Kalman filter, plot the values of Y_t and $\mu_{t|t}$. Discuss the plot.

(c.iv) Consider the interval $\mu_{t|t} \pm 1.96\sqrt{P_{t|t}(\mu)}$ where $P_{t|t}(\mu)$ is "2,2" element of the Kalman filter $P_{t|t}$ matrix. Does this interval contain μ_t with probability 0.95. Explain.

(c.v) Plot $\mu_{t|t}$ along with these 95% error bands. Discuss the plot.

(d) Repeat your calculations using $\sigma_e = 0.001$ and $\sigma_e = 0.003$. Discuss how these values of σ_e change the estimates of μ_t and the error bands.

(e) Suppose you have a prior that $\sigma_e = 0.001$, 0.002 and 0.003 each with probability 1/3. Using the Y_t data, compute the posterior distribution of σ_e.

(f) (Somewhat harder) Using $\sigma_e = 0.002$,

(f.i) Construct a 95% "prediction interval" for the average value of Y_t over the 2015 – 2054. Over 2015-2014.

(f.ii) Construct a 95% confidence interval for the value of per capita GDP in 2054? In 2114?

(g) (Harder still) Using the posteriors for σ_e in (e), repeat (f).