Time Series and Dynamic Econometrics
Assignment 3

Last updated: September 29, 2025

Instructions:

1. The assignment is made in groups of **exactly two students**. The same group is maintained for Assignments 1, 2, and 3.
2. The assignment is made using a programming language of choice, such as `python`/`R`/`matlab`/`Ox`. Software packages such as `Eviews` or `STATA` can **not** be used.
3. Within your chosen program language, you may use packages for computing statistical estimates (such as least squares estimates). However, simulations should be done manually (using for loops).[1]
4. The assignment must be **uploaded in PDF format** on Canvas before the **deadline**:

   Monday, October 13, 2025, at 23:59.

   The code must be uploaded as separate file(s) from the PDF.
5. The assignment is graded as either "complete" or "incomplete"; in the latter case the assignment must be revised. To pass the TSDE course, all three assignments must be graded as "complete."
6. On the first page of the assignment state the names and student numbers of the group members.
7. Do **not** copy the questions in your assignment (this invokes the plagiarism detection software). Simply write:
   1. <YOUR ANSWER TO SUBQUESTION 1>.
   2. <YOUR ANSWER TO SUBQUESTION 2>.
   3. ...

Remark (Plagiarism). *The VU has a **zero tolerance** policy for plagiarism, which can lead to exclusion from the Minor "Applied Econometrics: A Big Data Experience for All" without graduating. All assignments are checked for plagiarism using dedicated software, and cases that raise suspicion are automatically sent to the examination board for review.*

[1] You may of course use a package for drawing normal random variables.

Part I: Spurious Regression and Unit Roots

This part of the assignment starts with analysis of the spurious regression problem using a Monte Carlo simulation study. The assignment then continues with an empirical application.

Monte Carlo simulation study

1. Simulate two independent random walks

$$X_t = X_{t-1} + v_t, \qquad v_t \sim \mathrm{N}(0,1),$$
$$Y_t = Y_{t-1} + w_t, \qquad w_t \sim \mathrm{N}(0,1),$$

   for $t = 1, \ldots, T$ with $X_0 = Y_0 = 0$, and perform the regression $Y_t = \alpha + \beta X_t + \varepsilon_t$ a large number of times for different sample sizes T to demonstrate that

   - $\hat{\beta} \xrightarrow{d}$ Random variable
   - $\hat{\beta}/\mathrm{SE}(\hat{\beta}) \xrightarrow{p} \pm\infty$ (the t-statistic diverges)
   - $R^2 \xrightarrow{d}$ Random variable.

See also the corresponding plots in the slides of Week 5 for examples of how to visualize this.

Empirical application

Suppose you are asked to analyze investment opportunities in the stock market by studying the dynamic behavior of 10 major stock prices. In particular, you are asked to study the stocks of Apple, Intel, Microsoft, Ford, General Electrics, Netflix, Nokia, Exxon Mobil, and Yahoo, as well as the S&P500 stock market index. You can find the entire sample of daily data at your disposal in the csv file named `data_tsde_assignment_3_part_1.csv`. This data set contains the daily stock prices of all the companies mentioned above and ranges from the 14th of February 2007 to the 28th of January of 2013.

2. Provide plots for two stock market time series at your choice and report 12-period ACF functions for those two time series. What does the sample ACF tell you about the dynamic properties of these stocks?

3. Perform an ADF unit root test for each of the 10 time series. Use the Bayesian Information Criterion (BIC) to select the order p, and report the values of the ADF test statistics. Is the unit root hypothesis rejected for any time series at the 10% significance level? Did you expect to reject the unit root hypothesis for some time series? Justify your answer.

4. Investigate the following claim:

   > *"Financial analysts have found that changes in the price of Microsoft stocks can be largely explained by fluctuations in the market value of Exxon Mobile. According to these analysts, this shows the extent to which the Microsoft Corporation is currently exposed to the market performance of the oil and gas industry."*

   Do you find a statistically significant contemporaneous relation between Microsoft and Exxon Mobile stock prices? Do you agree that changes in Microsoft stock prices are largely explained by fluctuations in the stock price of Exxon Mobile? Justify your answer.

Part II: Cointegration and Error Correction Models

This part of the assignment starts by an analysis of cointegration using a Monte Carlo simulation study. The assignment then continues with an empirical application.

Monte Carlo simulation study

1. Show that the spurious regression problem does not occur when two $I(1)$ time series are *cointegrated* using a Monte Carlo simulation study. Furthermore, analyze the claim of super-consistency of the estimator of the static cointegrating regression. You can achieve the above by simulating two cointegrated $I(1)$ time series using the following triangular system,

$$\begin{aligned} X_t &= X_{t-1} + v_t, \qquad v_t \sim \mathrm{N}(0,1), \\ Y_t &= \gamma X_t + w_t, \qquad w_t \sim \mathrm{N}(0,1), \end{aligned}$$

for some $\gamma \neq 0$ (for example, $\gamma = 1$) with $X_0 = 0$, and performing the regression $Y_t = \alpha + \beta X_t + \varepsilon_t$ a large number of times for different sample sizes T. Super-consistency can be demonstrated by comparing the results with the outcomes of the same regression but with X_t generated as

$$X_t = \phi X_{t-1} + v_t, \qquad v_t \sim \mathrm{N}(0,1),$$

for some $|\phi| < 1$. See also the corresponding plots in the slides of Week 6 for examples of how to visualize this.

Empirical application

A large Dutch retailer of consumer goods is interested in predicting the effects of a potential increase in VAT and other consumption taxes over its sales. In particular, this retailer would like you to explore the relation between the total consumption of non-durable goods and the fluctuations in the total disposable income of families in the Netherlands. The sample of quarterly aggregate consumption in the Netherlands and the sample of aggregate household income that you have at your disposal ranges from the first quarter of 1988 to the first quarter of 2012. You can find this data set in the csv file named `data_tsde_assignment_3_part_2.csv`.

2. Plot both the aggregate consumption and aggregate income time series (these series are called *cons* and *inc* respectively in the csv file). Compute and report 12-period ACF functions for each series and comment on their shape.

3. Suppose that you have concluded that both time series are $I(1)$. Test for cointegration between *consumption* and *income* by regressing *consumption* on *income* and performing a unit root test on the residuals. Report the estimated regression coefficients. Plot the regression residuals. Use the Akaike Information Criterion (AIC) to determine the number of ADF lags in your unit root residual test. Report the cointegration test statistic. What do you conclude regarding cointegration at the 10% significance level?

4. Estimate an error correction model for consumption using the estimated residuals from the cointegration regression above. Use the AIC to select the model specification. Report the estimated model and the long-run equilibrium. Interpret the value of the error correction coefficient; is there correction to the long-run equilibrium?