Problem Set 1

Empirical Asset Pricing

January 12, 2025

Instructions

- This problem set is due on February 26, 2025.
- Submit your solution and all relevant files as a zip folder to juan.imbet@dauphine.psl.eu with the Subject PROBLEM SET 1 - LASTNAME - FIRST NAME
- Show all your work to receive full credit.

Problem Set

In this problem set you will study the properties of the bias correction presented in Stambaugh (1999) Predictive Regressions, Journal of Financial Economics.

Setup

: Consider the following system of predictive regressions

$$y_{t+1} = \alpha + \beta x_t + u_{t+1}$$
$$x_{t+1} = \theta + \rho x_t + \nu_{t+1}$$

Where u and ν follow a bivariate normal distribution $\mathcal{N}(0, \Sigma)$ where

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{u\nu} \\ \sigma_{u\nu} & \sigma_\nu^2 \end{bmatrix}$$

The following are the **true** values of the parameters:

Parameter	Value
α	0.01
β	0.05
θ	0.01
ρ	0.3
σ_u^2	0.6
σ_ν^2	0.5
$\sigma_{u\nu}$	-0.5

Question 1, 4 points

Create a function (routine) that given a set of parameters, and a sample size T simulates the dynamics of the system above. The function should return both time series x_t and y_t. Assume that both processes begin at their unconditional mean, e.g.

$$x_0 = \frac{\theta}{1-\rho}$$
$$y_0 = \alpha + \beta x_0$$

Plot the dynamics of x_t and y_t for $T = 100$. Do not forget that the residuals u_t and ν_t are correlated. Investigate how to generate correlated normal random variables in the programming language of your choice.

Question 2, 4 points

Create a function that given simulated data, estimates all of the parameters of the system above using OLS.

Question 3, 4 points

Fix a sample size of $T = 100$ and perform $N = 10,000$ simulations of the system above. For every simulation estimate β. Plot the distribution of $\hat{\beta}$ and show graphically how the estimator is biased.

Question 4, 4 points

Now you are going to fix $N = 100$ and compute the bias of $\hat{\beta}$ for different sample sizes (T). Plot the bias for 500 different points in the interval $[50, 1000]$.

Question 5, 4 points

Using the results from the question above, fit the following regression using OLS

$$\text{Bias}_i = \gamma_0 + \gamma_1 \frac{1}{T_i} + \gamma_2 \frac{1}{T_i^2} + \epsilon_i$$

Compute the t-statistics of all coefficients and comment, more specifically compare γ_1 with the equivalent term in the Stambaugh's bias definition.