ECON-C4200: Econometrics II exam 2.6.2023

There are three questions with subquestions. Answer all three. All questions have equal weight.

Question 1: Panel Data

1. Explain what type of unobservable effects one gets rid of with the within (fixed effects) estimator and how one achieves this.

2. What is the key assumption of the within estimator (that does not apply to (cross-section) OLS) regarding the error term and the explanatory variable(s) and what is its key economic implication?

3. What type of variation is needed in all explanatory variables when the within estimator is used?

Question 2: Machine Learning

1. Explain how a LASSO or a RIDGE regression model differs from OLS; for what type of an exercise are they potentially better suited than OLS; and why one may not care about individual coefficients or omitted variable bias when using either.

Question 3: Time series

1. What is a VAR model (you may use a model with two lags to illustrate) and what is it used for?

2. Explain in your own words what is **spurious correlation**.

3. You hire a research assistant to study an autoregressive time series. Your assistant comes back to you and says: "I did not bother with visualizing the data, but ran a first-order autoregressive model in levels. The coefficient of the lagged dependent variable is 0.987 and is highly significant. What's more, the R^2 is 0.99". What has or may have gone wrong?