

- Write your name, ID number, course number, and homework number on page 1.
- Show your work in steps. Final answers alone will not receive full credit.
- Write clearly and avoid over-crowding.
- Upload your write-up, including STATA code in the appendix, on MTeams.
- Submit your solution on time. Late submissions will NOT be accepted.

General Tips

- Your write-up should be in the style of a report written in full sentences and paragraphs (not short answers)
- Read the questions carefully; many of them ask you to answer multiple questions together.
- Justify and explain your answers thoroughly in complete sentences and paragraphs
- Consult the STATA help files to learn about the options for commands

Families borrow money from banks (mortgage loan) to finance the purchase of their homes. An increase in the mortgage interest rate, thus, increases the cost of owning a house and lowers the demand for houses. In this assignment, you will analyze time series data on the relationship between the mortgage interest rate (R_t) and the number of houses (H_t) sold monthly in the market. The sample spans the period from 1992M1 to 2010M3. After you load the data onto STATA, *declare* it as time series data before you start the analysis.

1 Policy Analysis

Consider the ARDL(1,2) model (without the contemporaneous term, R_t):

$$H_t = \alpha + \theta_1 H_{t-1} + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \varepsilon_t \quad (1)$$

- Estimate model (1) and include a screenshot of the results table in the report. Interpret the 1-month and 2-month delay multipliers representing the effects of a temporary change in the mortgage rate on home sales and make an inferential statements about it. [8]
- Suppose, instead, the change in the mortgage rate is permanent. Compute the *total multiplier* and interpret it. Is it *economically significant*? Discuss. [4]

2 Stationarity in Levels

- What are the conditions that the two series H_t and R_t must satisfy in order for them to be considered stationary? [6]
- Plot the two series H_t and R_t , and include the figures in your write-up. By visual inspection, are they stationary? Describe the behaviors of the two series and indicate what types of underlying models do they resemble. [8]
- Perform the *standard* Dickey-Fuller unit root test (**dfuller**, without augmentation terms) on each of the two series. Justify your choice of test equation, report the relevant test results, and draw conclusions about the series stationarity at the 5% significance level. [8]
- Given the unit root test results, do you think the estimated multipliers in problem **1** can be interpreted as causal? Discuss. [3]

3 Stationarity in Differences

- (a) For any time series Y_t , the *first difference* ($\Delta Y_t = Y_t - Y_{t-1}$) can be generated using the command: **generate d1Y = D.Y**. Generate two new variables representing the *first differences* in home sales and mortgage rate: ΔH_t and ΔR_t . Include their plots in your write-up and discuss their stationarity and describe their behaviors by visual inspection. [4]
- (b) Perform the *standard* Dickey-Fuller test (without augmentation) on each of the first-differenced series. Justify your choice of test equation, report the relevant test results, and draw conclusions about the series stationarity at the 5% significance level. [6]
- (c) What do you conclude the *order of integration* for each of the two series to be? Justify your answer. [3]

4 Cointegration

Two I(1) series are said to be cointegrated if their residual ($\hat{\varepsilon}_t$) is stationary. Let's check if this is true. First, obtain the residuals using the command: **predict resid, res** (after running the regression). Check the *Data Browser* to verify that the residuals [*resid*] are added to the dataset.

- (a) Plot the residuals series and include the graph in your write-up. Comment on their behavior. [3]
- (b) Performing the *augmented* Dickey-Fuller test with two lags (**dfuller resid, lags(2)**) on the residuals. Report the test statistic and draw a conclusion about the cointegration between home sales and mortgage rate based on the 5% significance level. [4]
- (c) Does the cointegration test result change your conclusion about the interpretation of the multipliers in problem **1**? Explain. [3]

5 Forecasting Home Sales

Note: Answers to this problem are analytical (no STATA) but rely on data and results obtained previously. Type your answers neatly using a proper equation editor and use proper notation.

- (a) Write the *fitted equation* for $\hat{H}_T = \hat{H}_{2010M3}$ using the estimated coefficients from problem **1**. [2]
- (b) Iterate the fitted equation to write two equations for the 1-month and 2-months ahead forecasts: \hat{H}_{2010M4} and \hat{H}_{2010M5} . [4]
- (c) If the mortgage rate stayed the same in 2010M4 as it was in 2010M3 (e.g. because central bank did not change interest rates), will there be more or fewer homes sold in 2010M4 and 2010M5 compared to 2010M3? By how many? (Show your work). [4]
- (d) Suppose the central bank lowered the interest rate in 2010M4 such that the mortgage rate decreased by 0.5 percentage points relative to what it was in 2010M3. Will there be more or fewer homes sold in 2010M4 and 2010M5? By how many? (Show your work) [4]
- (e) Compare the forecasts under the two policy regimes and discuss the effect of the rate cut on the demand for houses in the months 2010M4 and 2010M5. [4]

6 Serial Correlation

- (a) Generate a graph of the correlogram (i.e. autocorrelation function (ACF)) for the residuals up to $s = 12$ lags using the following commands:

```
* Compute & save 95% confidence bands
scalar band = 1.96/sqrt(_N)
* Generate the ACF graph for s=12 lags
ac resid, lags(12) ciopts(color(none)) yline(`=scalar(band)' -`=scalar(band)')
```

Include the graph in your write-up and discuss it.

[4]

- (b) Perform the Breusch-Godfrey LM test assuming an AR(2) process for the errors (**estat bgodfrey, lags(2)**, after running the regression). Report the chi-squared test statistic and its p-value, and draw an inferential conclusion about the presence of serial correlation in the errors and the appropriateness of model (1) in terms of fully capturing the dynamic relationship between home sales and mortgage rate.

[3]

- (c) What are the implications of having serially correlated errors on the quality of the multiplier estimates and the forecasts? (Discuss thoroughly using economic and econometric reasoning; do not just cite assumption violations.)

[4]

7 Model Selection

- (a) Estimate four additional ARDL(p, q) models (other than model (1)) using different combinations of lags (p, q) of your choice (up to $p \leq 4$ and $q \leq 4$). Report the *Akaike Information Criterion* (AIC) estimate corresponding to each model. Which model do you choose as the most appropriate, and why?
- (b) Repeat parts (a) and (b) of problem 6 but using the chosen model. At the end, compare your findings with those corresponding to model (1). Is the newly chosen model an improvement over model (1)? Discuss why or why not.

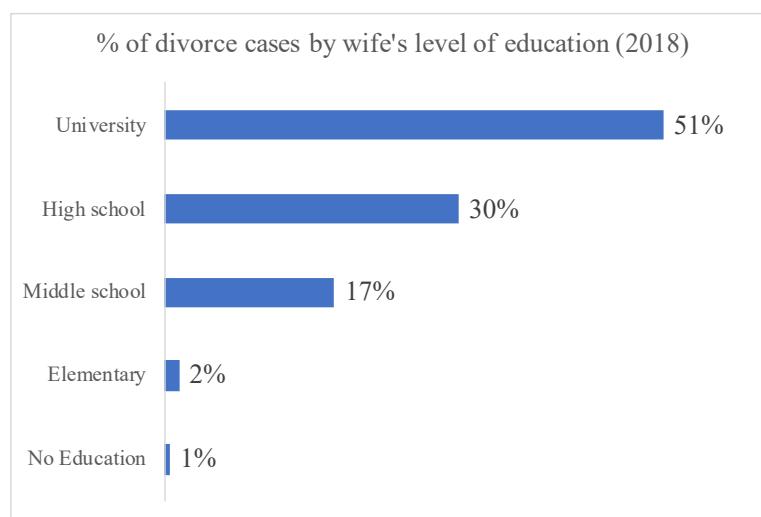
[7]

[7]

8 Essay

The following chart summarizes the findings of a 2018 study done by the Ministry of Justice in Kuwait:

[7]



The chart suggests that the share of divorce cases is positively correlated with the wife's level of education. Some analysts even suggest that women's educational attainment has had a *causal effect* on family instability by increasing the the probability of a divorce. Write a short essay (300–500 words) of what you think of this

problem and the claim of causality. Use the econometric concepts and terminology you have learned in this course to support your arguments. Some suggested points to address:

- What are the merits and/or flaws in this claim?
- How would you write a model for this relationship?
- What measurable or observable control variables should you include in it?
- What unobservable variables may still be omitted from the model? What is the sign of the bias that they may cause?
- How would you go about addressing any potential endogeneity problems?