In-Class Lab 11

ECON 4223 (Prof. Tyler Ransom, U of Oklahoma)

February 26, 2019

The purpose of this in-class lab is to use R to practice estimating time series regression models and to test for serial correlation. The lab should be completed in your group. To get credit, upload your .R script to the appropriate place on Canvas.

For starters

First, install the `pdfetch` and `tsibble` packages. `pdfetch` stands for "Public Data Fetch" and is a slick way of downloading statistics on stock prices, GDP, inflation, unemployment, etc. `tsibble` is a package useful for working with time series data. It is the "tibble" for time series data.

Open up a new R script (named `ICL11_XYZ.R`, where `XYZ` are your initials) and add the usual "preamble" to the top:

```
# Add names of group members HERE
library(tidyverse)
library(wooldridge)
library(broom)
library(car)
library(pdfetch)
library(magrittr)
library(tsibble)
```

Load the data

We're going to use data on US macroeconomic indicators. The `wooldridge` data set is called `intdef`.

```
df.ts <- as_tsibble(intdef, key=id(), index=year)
```

Now it will be easy to include lags of various variables into our regression models.

Plot time series data

Let's have a look at the inflation rate for the US over the period 1948–2003:

```
ggplot(df.ts, aes(year, inf)) + geom_line()
```

Determinants of the interest rate

Now let's estimate the following regression model:

$$i3_t = \beta_0 + \beta_1 inf_t + \beta_2 inf_{t-1} + \beta_3 inf_{t-2} + \beta_4 def_t + u_t$$

where $i3$ is the 3-month Treasury Bill interest rate, inf is the inflation rate (as measured by the CPI), and def is the budget deficit as a percentage of GDP.

```
est <- lm(i3 ~ inf + lag(inf,1) + lag(inf,2) + def, data=df.ts)
```

1. Are any of these variables significant determinants of the interest rate? If so, which ones?

Testing for Serial Correlation

Now let's test for serial correlation in our model. Serial correlation is defined as

$$u_t = \rho u_{t-1}$$

with $|\rho| > 0$.

We want to test

$$H_0 : \rho = 0$$

To do so, we need to run a regression of residuals (from `est`) on lagged residuals and look at the t-stat.

```
residtemp <- c(NA,NA,resid(est)) #first two undefined because of t-1 and t-2 lags
df.ts %<>% mutate(resids = residtemp)
est.resid <- lm(resids ~ lag(resids), data=df.ts)
tidy(est.resid)
```

2. What is the outcome of the hypothesis test?

When the x's aren't strictly exogenous

When x is correlated with lags of u, we need to modify the above test to include our x's from our original regression:

```
est.resid.endog <- lm(resids ~ lag(resids) + inf + lag(inf,1) + lag(inf,2) + def, data = df.ts)
```

3. What do you conclude about serial correlation in this more general case?