<u>Part A</u>

1) Write the model $y_t = \alpha + \delta t + \rho y_{t-1} + \beta x_{t-1} + r_t$ in first differences.

$$y_t - y_{t-1} = \left(\alpha + \delta t + \rho y_{t-1} + \beta x_{t-1} + r_t\right) - \left(\alpha + \delta\left(t-1\right) + \rho y_{t-2} + \beta x_{t-2} + r_{t-1}\right)$$
$$\Delta y_t = \left(\alpha + \delta t + \rho y_{t-1} + \beta x_{t-1} + r_t\right) - \left(\alpha + \delta t - \delta + \rho y_{t-2} + \beta x_{t-2} + r_{t-1}\right)$$
$$\Delta y_t = \delta + \rho \Delta y_{t-1} + \beta \Delta x_{t-1} + \Delta r_t$$

2) Suppose after first differencing a model is $\Delta y_t = \delta - \phi + 2\phi t + \rho \Delta y_{t-1} + \beta \Delta x_{t-1} + \Delta r_t$. What was it before the first difference was taken? (Hint: both *t* and $t^2$ are in it.)

Probably the best way to approach this is to make an educated guess and verify your guess is correct. If not, see what is different between the difference of what you guessed and what you wanted, then try a different guess. The key to guess well is to see that since there is a time trend in the difference, there is an accelerating time trend in the undifferenced model, and the accelerator term must have been multiplied by $\phi$. In polynomials, accelerating trends come from squared terms. So it is reasonable to conjecture $\phi t^2$ must appear in the undifferenced equation, in addition to everything in the model in #1. To verify:

$$y_t = \alpha + \delta t + \phi t^2 + \rho y_{t-1} + \beta x_{t-1} + r_t$$
$$y_t - y_{t-1} = \left(\alpha + \delta t + \phi t^2 + \rho y_{t-1} + \beta x_{t-1} + r_t\right) - \left(\alpha + \delta\left(t-1\right) + \phi\left(t-1\right)^2 + \rho y_{t-2} + \beta x_{t-2} + r_{t-1}\right)$$
$$\Delta y_t = \left(\alpha + \delta t + \phi t^2 + \rho y_{t-1} + \beta x_{t-1} + r_t\right) - \left(\alpha + \delta t - \delta + \phi t^2 - 2\phi t + \phi + \rho y_{t-2} + \beta x_{t-2} + r_{t-1}\right)$$
$$\Delta y_t = \delta - \phi + 2\phi t + \rho \Delta y_{t-1} + \beta \Delta x_{t-1} + \Delta r_t$$

3) Suppose you are originally interested in the model $y_t = \alpha + \delta t + \rho y_{t-1} + \beta x_{t-1} + r_t$, where $r_t = \gamma r_{t-1} + \varepsilon_t$ and $\varepsilon_t$ is an independent random disturbance. Write the dynamically complete model in first differences. Hint: first substitute to make the model dynamically complete, and then take the first difference.

First, note that:

$$r_t = y_t - \alpha - \delta t - \rho y_{t-1} - \beta x_{t-1}$$
$$r_{t-1} = y_{t-1} - \alpha - \delta\left(t-1\right) - \rho y_{t-2} - \beta x_{t-2}$$

From there:

$$y_t = \alpha + \delta t + \rho y_{t-1} + \beta x_{t-1} + \gamma\left(y_{t-1} - \alpha - \delta\left(t-1\right) - \rho y_{t-2} - \beta x_{t-2}\right) + \varepsilon_t$$
$$y_t = \left(1-\gamma\right)\alpha + \gamma\delta + \delta\left(1-\gamma\right)t + \left(\rho+\gamma\right)y_{t-1} + \beta x_{t-1} - \gamma\rho y_{t-2} - \gamma\beta x_{t-2} + \varepsilon_t$$
$$\Delta y_t = \delta\left(1-\gamma\right) + \left(\rho+\gamma\right)\Delta y_{t-1} - \gamma\rho \Delta y_{t-2} + \beta \Delta x_{t-1} - \gamma\beta \Delta x_{t-2} + \varepsilon_t - \varepsilon_{t-1}$$

<u>Part B</u>

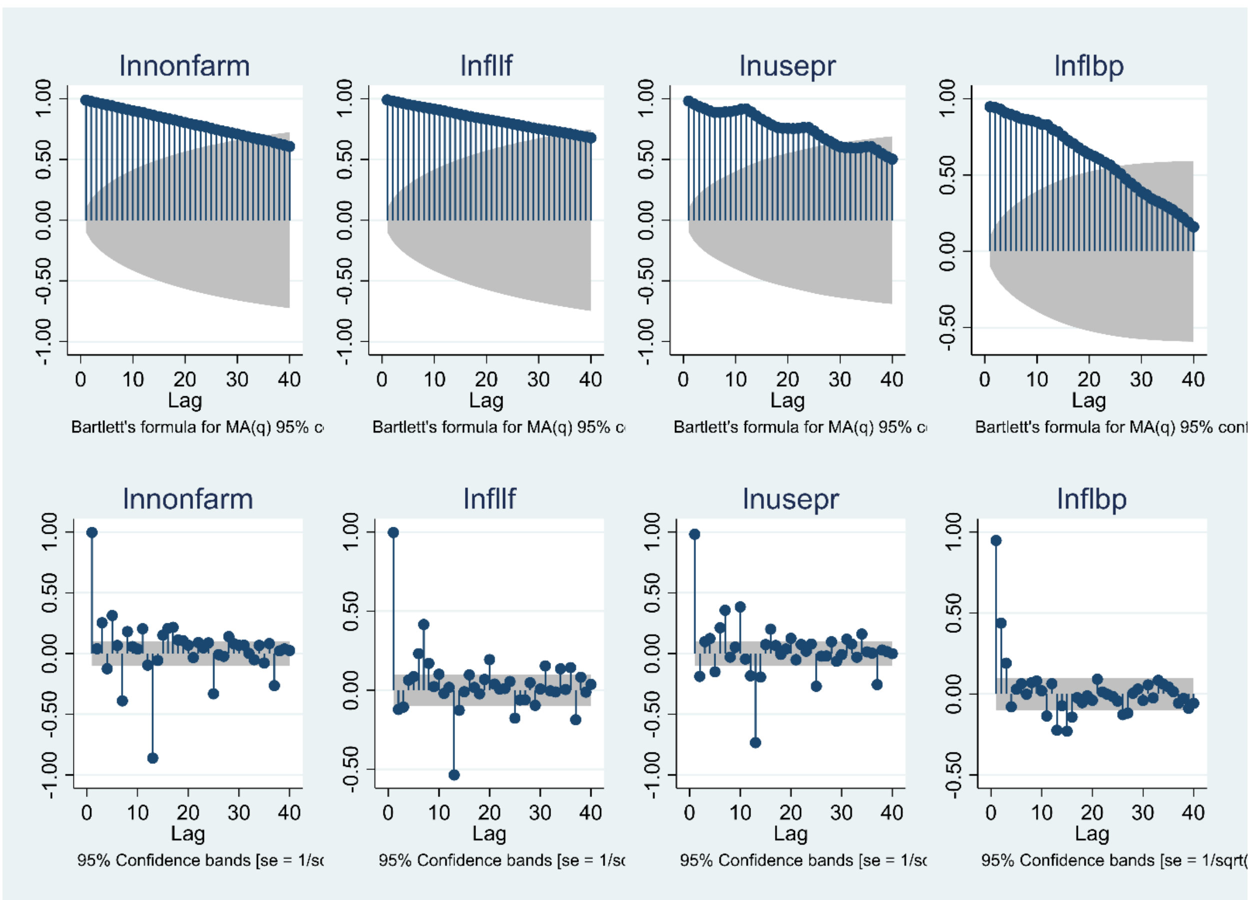
2) Prepare to analyze the data:
See the do file, Appendix A

3) Evaluate Autocorrelation and Weak Dependence

a) Obtain the correlation of each variable with its one period lag.

| Variable | Correlation with lag |
|---|---|
| lnflnonfarm | 0.9981 |
| lnfllf | 0.9994 |
| lnusepr | 0.9821 |
| lnflbp | 0.9477 |

b) Obtain the autocorrelogram and partial autocorrelagram for each variable.



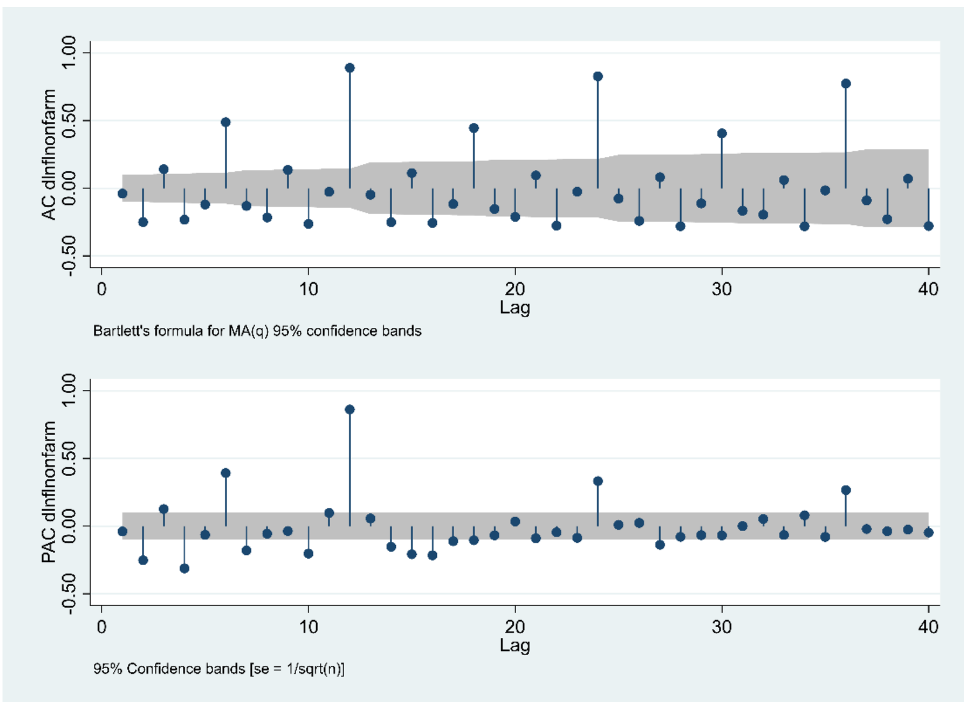
c) Conduct the Dickey-Fuller unit root rest for each variable.

See the log file for the full results.

| Variable | Dickey-Fuller p-value |
|---|---|
| lnflnonfarm | 0.0328 |
| lnfllf | 0.6285 |
| lnusepr | 0.2246 |
| lnflbp | 0.7774 |

d) Interpret these results.

Looking at the AC and PAC, all four show strong enough first order autoregressive relationships to merit differencing. We can reject the null of an I(1) process for the log of non-farm employment. But, the partial autocorrelation coefficient is so close to one that we should difference anyway. The AC and PAC for the log difference of non-farm employment are bwlow, illustrating the differences are clearly not I(1).

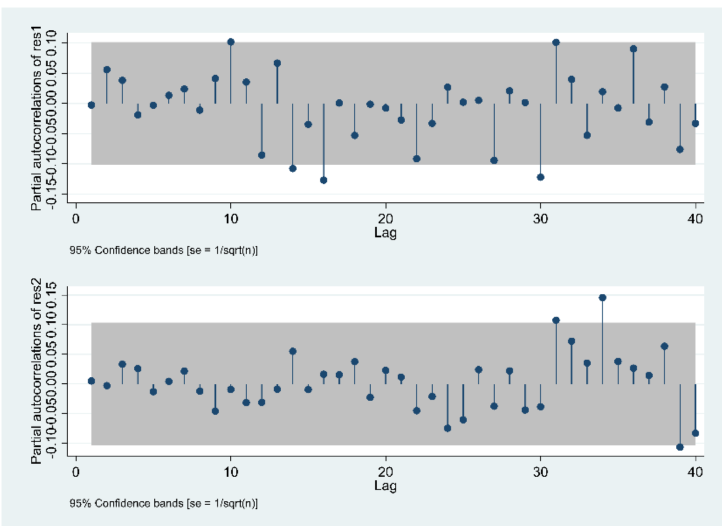


4) Given the results of the previous question, transform the data as needed and estimate a dynamically complete ARDL model for non-farm employment. Include at least one lag of the relevant dependent variable. How many additional lags of the dependent variable, and how many lags of which independent variables you include, are up to you. Looking back at what you did for Problem Set 1 might be informative, but don't be limited by it. Produce and interpret the AC and PAC for the residuals and the results of a Breusch-Godfrey test. In your write up, justify your specification and interpret the results.

I estimated two models, one with all lags back 12 months and one going back 24 months. Breush-Godfrey test results are in the table below. In the first case, the null of no serial correlation is rejected. For the second, the null can't be rejected at 24 lags, but it neither is it convincingly rejected ($p$=0.16). However, examining the PACs for the residuals in the figure below gives a bit more confidence in the second model. It also suggests some lags from year 3 and 4 may be worth including.

**Breush-Godfrey tests for question 4**

| | $p$ > chi | |
|---|---|---|
| Lags | Model 1 | Model 2 |
| 1 | 0.8812 | 0.4861 |
| 2 | 0.0332 | 0.778 |
| 3 | 0.0074 | 0.0585 |
| 4 | 0.0129 | 0.0386 |
| 5 | 0.0266 | 0.0709 |
| 6 | 0.0475 | 0.0774 |
| 7 | 0.0787 | 0.1049 |
| 8 | 0.0453 | 0.1426 |
| 9 | 0.068 | 0.1042 |
| 10 | 0.0467 | 0.1464 |
| 11 | 0.0688 | 0.1728 |
| 12 | 0.005 | 0.2121 |
| 13 | 0.0035 | 0.2321 |
| 14 | 0.0047 | 0.1206 |
| 15 | 0.0064 | 0.1304 |
| 16 | 0.0007 | 0.1483 |
| 17 | 0.0012 | 0.1816 |
| 18 | 0.0019 | 0.2039 |
| 19 | 0.0028 | 0.2119 |
| 20 | 0.0037 | 0.2138 |
| 21 | 0.0056 | 0.255 |
| 22 | 0.006 | 0.2065 |
| 23 | 0.0066 | 0.2485 |
| 24 | 0.0079 | 0.1618 |




A more parsimonious model, possibly with selected lags out further, might be a good idea. However, with some careful thought and exploration, I still have not come up with one that passed a Breusch-Godfrey test. Perhaps you did… Really, we will need more model selection tools to help us choose if we want to forecast. If we need to estimate parameters, we need to choose the appropriate model for the purpose, even if not dynamically complete, and then use appropriately adjusted standard errors. That is the point of the next problem.

5) Suppose you are interested in the relationship between the first difference in non-farm employment and the lags 0 to 4 of the differences of Florida building permits, controlling for seasonal impacts, but not controlling for any other variables or lags, including lags of employment. That is, you explicitly do not want to a dynamically complete model. (Don't worry about why, for this purpose.) Estimate the model both with and without Newey-West standard errors and discuss the difference that makes.

The results of interest are in the table at right. Note that the Newey-West standard errors are larger for the first three coefficients and smaller for the last. The regular standard errors are misleading regarding the precision of the estimates.

**Models for question 5**

| Std Err | Regular | Newey-West |
|---|---|---|
| D.lnflbp | 0.00820$^{***}$ | 0.00820$^{**}$ |
| | (0.00203) | (0.00250) |
| LD.lnflbp | 0.00793$^{***}$ | 0.00793$^{**}$ |
| | (0.00236) | (0.00294) |
| L2D.lnflbp | 0.00627$^{*}$ | 0.00627 |
| | (0.00244) | (0.00348) |
| L3D.lnflbp | 0.00730$^{**}$ | 0.00730$^{*}$ |
| | (0.00237) | (0.00300) |
| L4D.lnflbp | 0.00430$^{*}$ | 0.00430$^{*}$ |
| | (0.00204) | (0.00199) |
| $N$ | 379 | 379 |
| $R^2$ | 0.764 | |

Standard errors in parentheses
$^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$
Constant, trend, and month coefficients omitted for space

6) Provide a neat report writing up your answers to 2-4. Do not simply post screenshots of the Stata results window. Rather, provide neat professional looking tables of the results you obtain. Make sure any results you refer to in written answers to questions appear near the written answers so that your overall submission is easy to make sense of.

Your report should not look like this solution. This solution just describes what should be in your report. Your answer should resemble a professional report.

7) As Appendix A, include the clean do file to replicate your analysis.

See Appendix A.

8) As Appendix B, include the log file of a run of your clean do file to your write up.

See Appendix B

## Appendix A: Do File

```
*Time Series - Problem Set 2 Solution
*Spring 2020
clear
set more off
cd "C:\Users\jdewey\Documents\A S20 Time Series\Problem Sets\"
log using "Problem Set 2 Work", replace
import delimited using "us and florida economic time series.txt"

**2 - data prep

rename observation_date datestring
gen dateday=date(datestring,"YMD")
gen date=mofd(dateday)
format date %tm
tsset date
generate month=month(dateday)

keep if tin(1988m1,2019m12)

rename flbppriv fl_bp
rename fllfn fl_lf
rename flnan fl_nonfarm
rename lnu02300000_20200110 us_epr

gen lnflnonfarm=ln( fl_nonfarm)
gen lnfllf=ln( fl_lf)
gen lnusepr = ln(us_epr)
gen lnflbp=ln( fl_bp)


**#3 Evaluate Autocorrelation and Weak Dependence

corr lnflnonfarm l.lnflnonfarm
corr lnfllf l.lnfllf
corr lnusepr l.lnusepr
corr lnflbp l.lnflbp

ac lnflnonfarm, ytitle("") title("lnnonfarm") saving("AC FL Nonfarm Employemt", replace)
pac lnflnonfarm, ytitle("") title("lnnonfarm") saving("PAC FL Nonfarm Employment",
replace)

ac lnfllf, ytitle("") title("lnfllf") saving("AC FL Labor Force", replace)
pac lnfllf, ytitle("") title("lnfllf") saving("PAC FL Labor Force", replace)

ac lnusepr, ytitle("") title("lnusepr") saving("AC US Employment-Population Ratio",
replace)
```

```
pac lnusepr, ytitle("") title("lnusepr") saving("PAC US Employment-Population Ratio",
replace)
ac lnflbp, ytitle("") title("lnflbp")saving("AC FL Building Permits", replace)
pac lnflbp, ytitle("") title("lnflbp")saving("PAC FL Building Permits", replace)

graph combine "AC FL Nonfarm Employemt" "AC FL Labor Force" ///
      "AC US Employment-Population Ratio" "AC FL Building Permits" ///
      "PAC FL Nonfarm Employment" "PAC FL Labor Force" ///
      "PAC US Employment-Population Ratio" "PAC FL Building Permits", rows(2)
graph export "ac pac.emf" , replace

dfuller lnflnonfarm, trend  lag(12)
dfuller lnfllf, trend lag(12)
dfuller lnusepr, trend lag(12)
dfuller lnflbp, trend lag(12)

ac d.lnflnonfarm, ytitle("AC dlnflnonfarm") saving("AC Dif FL Nonfarm Employemt",
replace)
pac d.lnflnonfarm, ytitle("PAC dlnflnonfarm") saving("PAC Dif FL Nonfarm Employment",
replace)
graph combine "AC Dif FL Nonfarm Employemt" ///
      "PAC Dif FL Nonfarm Employment" , rows(2)
graph export "dif ac pac.emf", replace

**#4 Estimate a dynamically complete model for nonfarm employment

quietly reg d.lnflnonfarm l(1/12)d.lnflnonfarm l(0/12)d.lnfllf ///
  l(0/12)d.lnusepr l(0/12)d.lnflbp i.month
estimates store m41
bgodfrey , lag(1/24)
predict res1, residual
pac res1, saving(pacres1, replace)

quietly reg d.lnflnonfarm l(1/24)d.lnflnonfarm l(0/24)d.lnfllf ///
  l(0/24)d.lnusepr l(0/24)d.lnflbp i.month
estimates store m42
bgodfrey , lag(1/24)
predict res2, residual
pac res2, saving(pacres2, replace)

graph combine "pacres1" "pacres2", rows(2)
graph export "res4142.emf", replace
*5
quietly reg d.lnflnonfarm l(0/4)d.lnflbp i.month date
estimates store m51
quietly newey d.lnflnonfarm l(0/4)d.lnflbp i.month date , lag(24)
estimates store m52
esttab m51 m52 using "models5", se r2 onecell compress replace

log close
```

**Appendix B: Log File**

```
-------------------------------------------------------------------------------------------
      name:  <unnamed>
       log:  C:\Users\jdewey\Documents\A S20 Time Series\Problem Sets\Problem Set 2
Work.smcl
  log type:  smcl
 opened on:   3 Feb 2020, 20:17:39

. import delimited using "us and florida economic time series.txt"
(5 vars, 972 obs)

.
. **2 - data prep
.
. rename observation_date datestring

. gen dateday=date(datestring,"YMD")

. gen date=mofd(dateday)

. format date %tm

. tsset date
        time variable:  date, 1939m1 to 2019m12
                delta:  1 month

. generate month=month(dateday)

.
. keep if tin(1988m1,2019m12)
(588 observations deleted)

.
. rename flbppriv fl_bp

. rename fllfn fl_lf

. rename flnan fl_nonfarm

. rename lnu02300000_20200110 us_epr

.
. gen lnflnonfarm=ln( fl_nonfarm)

. gen lnfllf=ln( fl_lf)

. gen lnusepr = ln(us_epr)
```

```
. gen lnflbp=ln( fl_bp)

.
. **#3 Evaluate Autocorrelation and Weak Dependence
.
. corr lnflnonfarm l.lnflnonfarm
(obs=383)

             |                 L.
             | lnflno~m lnflno~m
-------------+------------------
 lnflnonfarm |
         --. |   1.0000
         L1. |   0.9981   1.0000

. corr lnfllf l.lnfllf
(obs=383)

             |                 L.
             |   lnfllf   lnfllf
-------------+------------------
      lnfllf |
         --. |   1.0000
         L1. |   0.9994   1.0000

. corr lnusepr l.lnusepr
(obs=383)

             |                 L.
             |  lnusepr  lnusepr
-------------+------------------
     lnusepr |
         --. |   1.0000
         L1. |   0.9821   1.0000

. corr lnflbp l.lnflbp
(obs=383)

             |                 L.
             |   lnflbp   lnflbp
-------------+------------------
      lnflbp |
         --. |   1.0000
         L1. |   0.9477   1.0000


.
. ac lnflnonfarm, ytitle("") title("lnnonfarm") saving("AC FL Nonfarm Employemt",
replace)
```

```
(file AC FL Nonfarm Employemt.gph saved)

. pac lnflnonfarm, ytitle("") title("lnnonfarm") saving("PAC FL Nonfarm Employment",
replace)
(file PAC FL Nonfarm Employment.gph saved)

.
. ac lnfllf, ytitle("") title("lnfllf") saving("AC FL Labor Force", replace)
(file AC FL Labor Force.gph saved)

. pac lnfllf, ytitle("") title("lnfllf") saving("PAC FL Labor Force", replace)
(file PAC FL Labor Force.gph saved)

.
. ac lnusepr, ytitle("") title("lnusepr") saving("AC US Employment-Population Ratio",
replace)
(file AC US Employment-Population Ratio.gph saved)

. pac lnusepr, ytitle("") title("lnusepr") saving("PAC US Employment-Population Ratio",
replace)
(file PAC US Employment-Population Ratio.gph saved)

.
. ac lnflbp, ytitle("") title("lnflbp")saving("AC FL Building Permits", replace)
(file AC FL Building Permits.gph saved)

. pac lnflbp, ytitle("") title("lnflbp")saving("PAC FL Building Permits", replace)
(file PAC FL Building Permits.gph saved)

.
. graph combine "AC FL Nonfarm Employemt" "AC FL Labor Force" ///
>         "AC US Employment-Population Ratio" "AC FL Building Permits" ///
>         "PAC FL Nonfarm Employment" "PAC FL Labor Force" ///
>         "PAC US Employment-Population Ratio" "PAC FL Building Permits", rows(2)

. graph export "ac pac.emf" , replace
(file C:\Users\jdewey\Documents\A S20 Time Series\Problem Sets\ac pac.emf written in
Enhanced Me
> tafile format)

.
. dfuller lnflnonfarm, trend  lag(12)

Augmented Dickey-Fuller test for unit root         Number of obs   =       371

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -3.567            -3.985            -3.425            -3.130
```

```
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0328

. dfuller lnfllf, trend lag(12)

Augmented Dickey-Fuller test for unit root         Number of obs   =       371

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.950            -3.985            -3.425            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.6285

. dfuller lnusepr, trend lag(12)

Augmented Dickey-Fuller test for unit root         Number of obs   =       371

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.728            -3.985            -3.425            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.2246

. dfuller lnflbp, trend lag(12)

Augmented Dickey-Fuller test for unit root         Number of obs   =       371

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.637            -3.985            -3.425            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.7774


.
. ac d.lnflnonfarm, ytitle("AC dlnflnonfarm") saving("AC Dif FL Nonfarm Employemt",
replace)
(file AC Dif FL Nonfarm Employemt.gph saved)

. pac d.lnflnonfarm, ytitle("PAC dlnflnonfarm") saving("PAC Dif FL Nonfarm Employment",
replace)
(file PAC Dif FL Nonfarm Employment.gph saved)
```

```
. graph combine "AC Dif FL Nonfarm Employemt" ///
>         "PAC Dif FL Nonfarm Employment" , rows(2)

.
.
. **#4 Estimate a dynamically complete model for nonfarm employment
.
. quietly reg d.lnflnonfarm l(1/12)d.lnflnonfarm l(0/12)d.lnfllf ///
>   l(0/12)d.lnusepr l(0/12)d.lnflbp i.month

. estimates store m41

. bgodfrey , lag(1/24)

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          0.022               1                   0.8812
       2     |          6.813               2                   0.0332
       3     |         11.994               3                   0.0074
       4     |         12.680               4                   0.0129
       5     |         12.681               5                   0.0266
       6     |         12.734               6                   0.0475
       7     |         12.740               7                   0.0787
       8     |         15.802               8                   0.0453
       9     |         15.948               9                   0.0680
      10     |         18.528              10                   0.0467
      11     |         18.592              11                   0.0688
      12     |         28.311              12                   0.0050
      13     |         30.906              13                   0.0035
      14     |         31.525              14                   0.0047
      15     |         32.034              15                   0.0064
      16     |         40.189              16                   0.0007
      17     |         40.226              17                   0.0012
      18     |         40.226              18                   0.0019
      19     |         40.474              19                   0.0028
      20     |         40.991              20                   0.0037
      21     |         41.006              21                   0.0056
      22     |         42.124              22                   0.0060
      23     |         43.149              23                   0.0066
      24     |         43.866              24                   0.0079
---------------------------------------------------------------------------
                        H0: no serial correlation

. predict res1, residual
(13 missing values generated)

. pac res1, saving(pacres1, replace)
```

```
(file pacres1.gph saved)

.
. quietly reg d.lnflnonfarm l(1/24)d.lnflnonfarm l(0/24)d.lnfllf ///
>   l(0/24)d.lnusepr l(0/24)d.lnflbp i.month

. estimates store m42

. bgodfrey , lag(1/24)

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          0.485               1                   0.4861
       2     |          0.502               2                   0.7780
       3     |          7.463               3                   0.0585
       4     |         10.111               4                   0.0386
       5     |         10.156               5                   0.0709
       6     |         11.378               6                   0.0774
       7     |         11.870               7                   0.1049
       8     |         12.198               8                   0.1426
       9     |         14.546               9                   0.1042
      10     |         14.624              10                   0.1464
      11     |         15.216              11                   0.1728
      12     |         15.563              12                   0.2121
      13     |         16.324              13                   0.2321
      14     |         20.313              14                   0.1206
      15     |         21.205              15                   0.1304
      16     |         21.845              16                   0.1483
      17     |         22.081              17                   0.1816
      18     |         22.663              18                   0.2039
      19     |         23.601              19                   0.2119
      20     |         24.685              20                   0.2138
      21     |         24.821              21                   0.2550
      22     |         27.126              22                   0.2065
      23     |         27.176              23                   0.2485
      24     |         30.726              24                   0.1618
---------------------------------------------------------------------------
                        H0: no serial correlation

. predict res2, residual
(25 missing values generated)

. pac res2, saving(pacres2, replace)
(file pacres2.gph saved)

.
. graph combine "pacres1" "pacres2", rows(2)
```

```
. graph export "res4142.emf", replace
(note: file res4142.emf not found)
(file C:\Users\jdewey\Documents\A S20 Time Series\Problem Sets\res4142.emf written in
Enhanced M
> etafile format)

.
. *5
. quietly reg d.lnflnonfarm l(0/4)d.lnflbp i.month date

. estimates store m51

. quietly newey d.lnflnonfarm l(0/4)d.lnflbp i.month date , lag(24)

. estimates store m52

.
. esttab m51 m52 using "models5", se r2 onecell compress replace
(output written to models5.txt)

.
. log close
      name:  <unnamed>
       log:  C:\Users\jdewey\Documents\A S20 Time Series\Problem Sets\Problem Set 2
Work.smcl
  log type:  smcl
 closed on:   3 Feb 2020, 20:17:57
-----------------------------------------------------------------------------
```