Ch13_2e_ChapterExamples

```
require(knitr)
require(haven)
#library(car)  #install.packages("car")
library(AER)
library(tseries)
library(apsrtable);
library(urca)  # for ur.df function

opts_chunk$set(echo = TRUE)
options(digits = 3)
opts_knit$set(root.dir =
"C:/Users/baileyma/Dropbox/Bailey_RealStats_RealEconometrics_2e_Supplements_March2021/ChapterExamples_R

# Remove objects from the previous session
rm(list = ls(all = TRUE))
```

Preparation

Global warming

We have updated data to 2019. The ".old" versions show the data through 2012 in earlier editions. The ".new" versions use data through 2019.

Figure 13.2

```
load("Ch7_ChapterExample_GlobalWarming.RData")
dta$YearSq = dta$Year^2

# Through 2018
Warming.ols.new= lm(observedtemp_new ~ Year, data = dta)
options(scipen = 999, digits = 4)       ## Turn scientific notation off
summary(Warming.ols.new)
```

```
##
## Call:
## lm(formula = observedtemp_new ~ Year, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4989 -0.1157 -0.0036  0.1240  0.6466
##
## Coefficients:
##               Estimate Std. Error t value            Pr(>|t|)
## (Intercept) -17.819376   0.826989   -21.6 <0.0000000000000002 ***
## Year          0.009174   0.000424    21.6 <0.0000000000000002 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.201 on 137 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.773,  Adjusted R-squared:  0.772
## F-statistic:  468 on 1 and 137 DF,  p-value: <0.0000000000000002
```

```
options(scipen = 0)     ## Turn scientific notation back on
```

Figure 13.2 (see also Chapter 7 (Figure: Figure13dot2NEW.eps)

```
par(mfrow=c(2, 1), mar=c(4, 4,  2.,  2.5), oma=c(1.2, 1.5, 1.5,  0.8))   # mar(south, west, north, east
plot (dta$Year, dta$observedtemp_new , main="", col= "black", cex.axis =0.8,
      type="n", ylim = c(-0.6, 1.4), xlim = c(1880, 2020), xlab = "", ylab = "", xaxt='n', yaxt='n')
mtext("Year",           side = 1, line = 1.4, cex = 0.8)
mtext("(a)",            side = 2, line = 3.8, , las = 1, cex = 0.8)
mtext("Temperature\ndeviation",     las = 1, side = 3, at = 1852, line =-0.6, cex = 0.8)
axis(1, at = seq(1880, 2020,by=10), labels =  seq(1880, 2020,by=10), tick = T, cex.axis = .8, mgp = c(2
axis(2, las = 1, at = seq(-0.6, 1.8, by=0.4), labels =  round(seq(-0.6, 1.8, by=0.4), 2), tick = T, cex
points(dta$Year, dta$observedtemp_new, pch=16, col="#3c78c3", cex=0.75)
abline(Warming.ols.new, col="#3c78c3", lwd = 1.6)

plot(dta$Year[2:140], Warming.ols.new$residuals, main="", type="p", xlab = "", ylab = "", xaxt='n', yax
#   axis(2, las = 1, tick = T, cex.axis = .8, mgp = c(2,.7,0))
  axis(2, las = 1, at = seq(-0.6, 0.6, by=0.3), labels =  round(seq(-0.6, 0.6, by=0.3), 2), tick = T, c
    axis(1, tick = T, at= seq(1880, 2020, by=10), labels =seq(1880, 2020, by=10),cex.axis = .8, mgp = c
mtext("(b)",            side = 2, line = 3.8, , las = 1, cex = 0.8)
    mtext("Residuals",  las = 1, side = 3, at = 1856, line = -0.25, cex = 0.8)
    mtext("Year",           side = 1, line = 1.4, cex = 0.8)
abline(h=0, col="#3c78c3", lwd = 1.6)
```

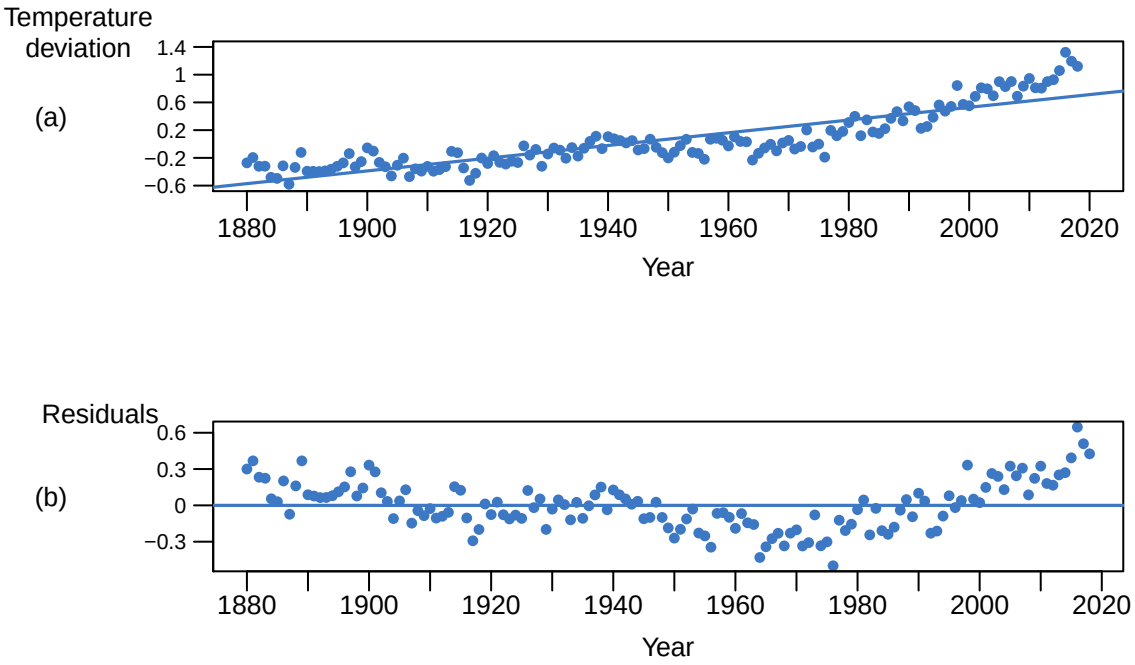


Table 13.1 Lagged residuals

```
dta$Residuals= NA
dta$Residuals[1:length(Warming.ols.new$residuals)] = Warming.ols.new$residuals

dta$Lag.Residuals = rep(NA, length(dta[,1]))
dta$Lag.Residuals[1:length(Warming.ols.new$residuals)] = c(NA, Warming.ols.new$residuals[1:(length(Warm

Warming.ols.LagResid = lm(Residuals ~ Lag.Residuals, data = dta)

summary(Warming.ols.LagResid)
```

```
##
## Call:
## lm(formula = Residuals ~ Lag.Residuals, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3115 -0.0908  0.0026  0.0845  0.3505
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.000152   0.011292    0.01     0.99
## Lag.Residuals 0.754152   0.057425   13.13   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.133 on 136 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.559,  Adjusted R-squared:  0.556
## F-statistic:  172 on 1 and 136 DF,  p-value: <2e-16
```

```
Rho = summary(Warming.ols.LagResid)$coefficients[2]
Rho
```

```
## [1] 0.7542
```

Chapter 13 Case Study: Quadratic models for global climate data

Figure 13.3 ()

```
par(mfrow=c(1, 1), mar=c(4, 4,  2.,  2.5), oma=c(1.2, 1.5, 1.5,  0.8))   # mar(south, west, north, east
plot (dta$Year, dta$observedtemp_new , main="", col= "black", cex.axis =0.8,
      type="n", ylim = c(-0.55, 1.28), xlim = c(1880, 2020), xlab = "", ylab = "", xaxt='n', yaxt='n')
mtext("Year",           side = 1, line = 1.4, cex = 0.8)
mtext("Temperature\ndeviation",     las = 1, side = 3, at = 1852, line =-0.6, cex = 0.8)
axis(1, at = seq(1880, 2020,by=10), labels =  seq(1880, 2020,by=10), tick = T, cex.axis = .8, mgp = c(2
axis(2, las = 1, at = seq(-0.75, 1.75, by=0.25), labels =  round(seq(-0.75, 1.75, by=0.25), 2), tick =
points(dta$Year, dta$observedtemp_new, pch=16, col="#3c78c3", cex=0.75, type = "l", lwd = 2)
```

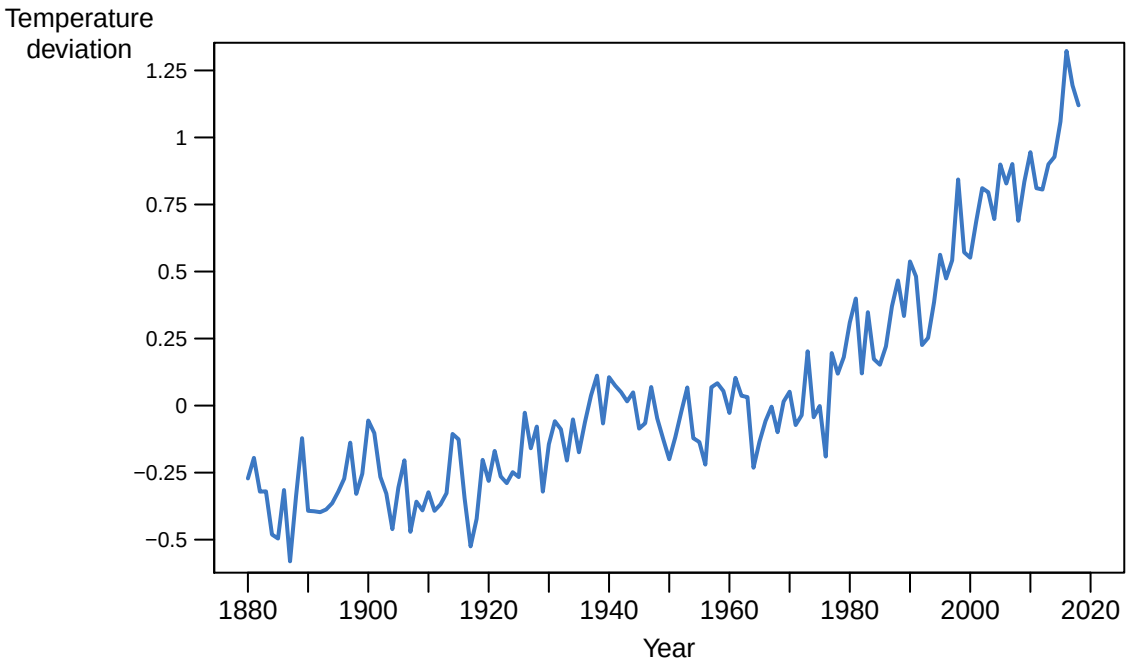


Table 13.3 - rho calculations

```
# Through 2018
Warming.quad.new        = lm(observedtemp_new ~ Year + YearSq, data = dta)
options(scipen = 999)        ## Turn scientific notation off
summary(Warming.quad.new) #confint(Warming.quad.new)
```

```
##
## Call:
## lm(formula = observedtemp_new ~ Year + YearSq, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4121 -0.0878 -0.0058  0.0977  0.3628
##
## Coefficients:
##                 Estimate   Std. Error t value            Pr(>|t|)
## (Intercept) 356.45808818  31.62089992    11.3 <0.0000000000000002 ***
## Year         -0.37505990   0.03245796   -11.6 <0.0000000000000002 ***
## YearSq        0.00009857   0.00000833    11.8 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.141 on 136 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.888,  Adjusted R-squared:  0.887
## F-statistic:  541 on 2 and 136 DF,  p-value: <0.0000000000000002
```

```
options(scipen = 0)        ## Turn scientific notation back on

## Get t-stats with greater precision
#summary(Warming.quad.new)$coefficients[,1]/summary(Warming.quad.new)$coefficients[,2]

length(dta$observedtemp_new)^(0.25)
```

```
## [1] 3.446
```

```
dta$Residuals2= NA
dta$Residuals2[1:length(Warming.quad.new$residuals)] = Warming.quad.new$residuals

dta$Lag.Residuals2 = rep(NA, length(dta[,1]))
dta$Lag.Residuals2[1:length(Warming.quad.new$residuals)] = c(NA,  Warming.quad.new$residuals[1: (length


Warming.quad.LagResid = lm(Residuals2 ~ Lag.Residuals2, data = dta)

summary(Warming.quad.LagResid)
```

```
##
## Call:
## lm(formula = Residuals2 ~ Lag.Residuals2, data = dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30737 -0.06445 -0.00597  0.09130  0.30114
```

```
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.000495   0.010412    0.05     0.96
## Lag.Residuals2 0.502511   0.074379    6.76  3.8e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.122 on 136 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.251,  Adjusted R-squared:  0.246
## F-statistic: 45.6 on 1 and 136 DF,  p-value: 3.79e-10
```

```
Rho2 = summary(Warming.quad.LagResid)$coefficients[2]
```

Table 13.3 - Newey-West

```
library(sandwich)
NeweyWest(Warming.quad.new, lag = 3, prewhite = FALSE)
```

```
##             (Intercept)       Year     YearSq
## (Intercept)   2.451e+03 -2.518e+00  6.462e-04
## Year         -2.518e+00  2.586e-03 -6.638e-07
## YearSq        6.462e-04 -6.638e-07  1.704e-10
```

```
sqrt(diag(NeweyWest(Warming.quad.new, lag = 3, prewhite = FALSE, adjust = TRUE)))
```

```
## (Intercept)        Year      YearSq
##   5.005e+01   5.141e-02   1.320e-05
```

```
options(scipen = 999)        ## Turn scientific notation off
cbind(Warming.quad.new$coefficients, sqrt(diag(NeweyWest(Warming.quad.new, lag = 3, prewhite = FALSE, ad
      Warming.quad.new$coefficients/sqrt(diag(NeweyWest(Warming.quad.new, lag = 3, prewhite = FALSE, adj
```

```
##                      [,1]       [,2]   [,3]
## (Intercept) 356.45808818 50.0501230  7.122
## Year         -0.37505990  0.0514119 -7.295
## YearSq        0.00009857  0.0000132  7.469
```

Table 13.3 - Rho-transform using packages

```
library(orcutt)
summary(cochrane.orcutt(Warming.quad.new))
```

```
## Call:
## lm(formula = observedtemp_new ~ Year + YearSq, data = dta)
##
##                Estimate  Std. Error t value      Pr(>|t|)
## (Intercept) 362.4974594  56.3077829    6.44 0.00000000196 ***
## Year         -0.3812844   0.0577533   -6.60 0.00000000085 ***
## YearSq        0.0001002   0.0000148    6.77 0.00000000037 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.1228 on 135 degrees of freedom
## Multiple R-squared:  0.7329 ,  Adjusted R-squared:  0.7289
## F-statistic: 185.2 on 2 and 135 DF,  p-value: < 2.013e-39
##
## Durbin-Watson statistic
## (original):    0.99500 , p-value: 4e-10
## (transformed): 2.09032 , p-value: 6.404e-01
```

Table 13.3 - Rho-transform manually (to get $\hat{\beta}_0$ and $\hat{\rho}$)

```
dta$LagTemp         = c(NA, dta$observedtemp_new[1:(length(dta$observedtemp_new)-1)])

dta$LagYear         = c(NA, dta$Year[1:(length(dta$observedtemp_new)-1)])

dta$LagYearSq       = c(NA, dta$YearSq[1:(length(dta$observedtemp_new)-1)])

AvgTemp.rho     = dta$observedtemp_new  - Rho2*dta$LagTemp
Year.rho          = dta$Year                - Rho2*dta$LagYear
YearSq.rho    = dta$YearSq          - Rho2*dta$LagYearSq
Warming.rho     = lm(AvgTemp.rho ~ Year.rho + YearSq.rho )
summary(Warming.rho)
```

```
##
## Call:
## lm(formula = AvgTemp.rho ~ Year.rho + YearSq.rho)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30706 -0.06569 -0.00752  0.08963  0.29788
##
## Coefficients:
##                Estimate  Std. Error t value      Pr(>|t|)
## (Intercept) 180.3368583  28.0037846    6.44 0.00000000194 ***
## Year.rho     -0.3812808   0.0577353   -6.60 0.00000000084 ***
## YearSq.rho    0.0001002   0.0000148    6.77 0.00000000036 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.123 on 135 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.733,  Adjusted R-squared:  0.729
## F-statistic:  185 on 2 and 135 DF,  p-value: <0.0000000000000002
```

```
dta$Residuals.rho= NA
dta$Residuals.rho[1:length(Warming.rho$residuals)] = Warming.rho$residuals

dta$Lag.Residuals.rho = rep(NA, length(dta[,1]))
dta$Lag.Residuals.rho[1:length(Warming.rho$residuals)] = c(NA,  Warming.rho$residuals[1: (length(Warming

Warming.rho.LagResid = lm(Residuals.rho ~ Lag.Residuals.rho, data = dta)

summary(Warming.rho.LagResid)
```

```
##
```

```
## Call:
## lm(formula = Residuals.rho ~ Lag.Residuals.rho, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3107 -0.0667 -0.0109  0.0912  0.3036
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -0.000541   0.010462   -0.05     0.96
## Lag.Residuals.rho -0.046340   0.085858   -0.54     0.59
##
## Residual standard error: 0.122 on 135 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.00215,    Adjusted R-squared:  -0.00524
## F-statistic: 0.291 on 1 and 135 DF,  p-value: 0.59
```

```
Rho.rho = summary(Warming.rho.LagResid)$coefficients[2]
Rho.rho
```

```
## [1] -0.04634
```

```
# Estimate of \hat{\beta}_0
180.34/(1-0.503)
```

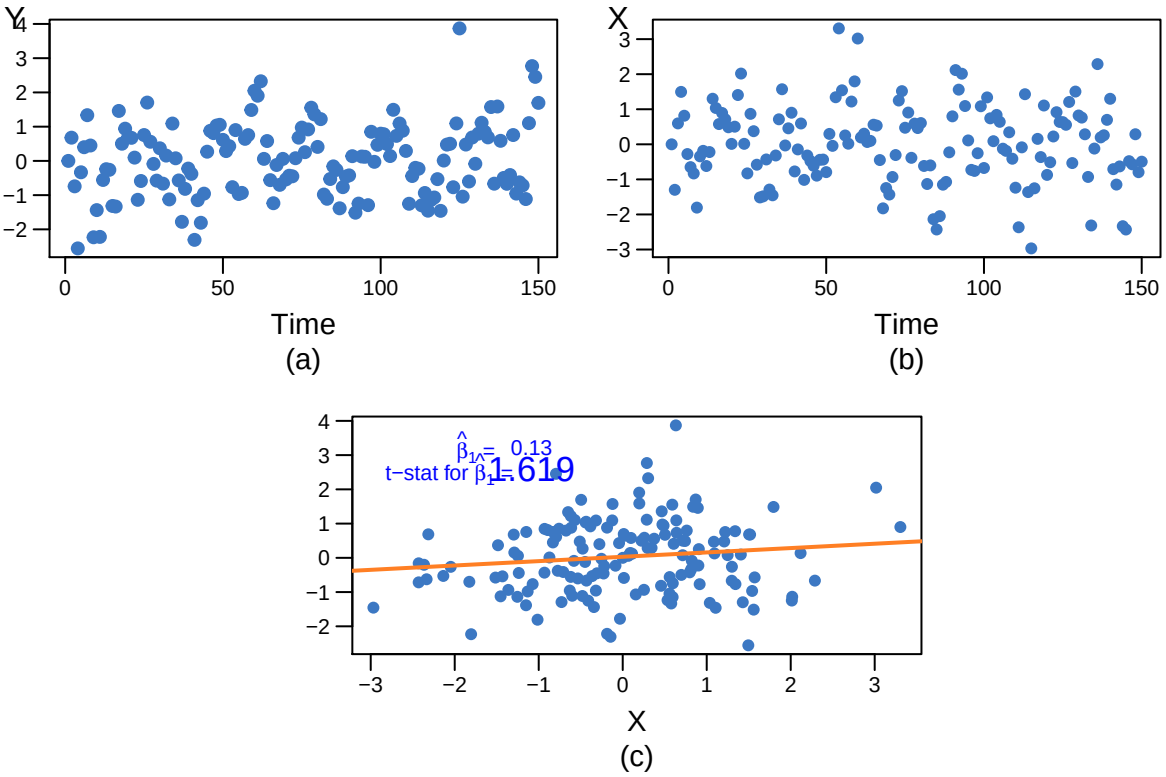
```
## [1] 362.9
```

Figure 13.4 Unit roots and spurious regression (simulation)

Figure 13.5 Non-unit root regression (simulation)

```
# Set coefficients on lagged values
Nsim        = 150       ## Number of observations in each simulation
YGamma = 0.5        ## Set to 1 for unit root
XGamma = 0.5        ## Set to 1 for unit root

# Create X and Y series
    Y = 0
        X = 0
        for(ii in 1:Nsim){
            Y = c(Y, YGamma*Y[ii-1] + rnorm(1))         ## Generate Y as dynamic process independent of
            X = c(X, XGamma*X[ii-1] + rnorm(1)) }       ## Generate X as dynamic process independent of
        Unit.root   = summary(lm(Y~X))$coef[2, 3]

## Figure: NotUnitRootSimulation.eps
# Note 1: need to set seed if want same result every time
# Note 2: if control for lagY in this model (much less likely to be statistically significant - relevan
par(mfrow=c(1, 1), mar=c(4.0,2.5,1.,0.5), oma=c(0.2,1,2.,0.3))
layout(matrix(c(1,2,3,3), 2, 2, byrow = TRUE))  ## this says fig 1 in upper left, fig 2 in upper right,
    plot(1:Nsim, Y, main="", col= "black", cex.axis =0.8, type="n", xlab = "", ylab = "", xaxt='n', yax
    axis(2, las = 1, tick = T, cex.axis = .8, mgp = c(2,.7,0))
    axis(1, tick = T, cex.axis = .8, mgp = c(2,.4,0))   #at = seq(-32, 44, by=8), labels =  seq(-32, 44
    mtext("Y",  las = 1, side = 3, at = -16, line = -0.5, cex = 1.)
    mtext("Time",           side = 1, line = 1.6, cex = 0.9)
    mtext("(a)",            side = 1, line = 2.7, cex = 0.9)
    points(1:Nsim, Y, cex=1, col="#3c78c3", pch = 19)
```

```r
  plot(1:Nsim, X, main="", col= "black", cex.axis =0.8, type="n", xlab = "", ylab = "", xaxt='n', yax
  axis(2, las = 1, tick = T, cex.axis = .8, mgp = c(2,.7,0))
  axis(1, tick = T, cex.axis = .8, mgp = c(2,.4,0))   #at = seq(-32, 44, by=8), labels =  seq(-32, 44
  mtext("X",  las = 1, side = 3, at = -16, line = -0.5, cex = 1.)
  mtext("Time",           side = 1, line = 1.6, cex = 0.9)
  mtext("(b)",            side = 1, line = 2.7, cex = 0.9)
  points(1:Nsim, X, cex=1, col="#3c78c3", pch = 16)

  par(mar=c(4.0, 12, 1., 8)) # mar(south, west, north, east)
  plot(X, Y, main="", col= "black", cex.axis =0.8, type="n", xlab = "", ylab = "", xaxt='n', yaxt='n'
  axis(2, las = 1, tick = T, cex.axis = .8, mgp = c(2,.7,0))
  axis(1, tick = T, cex.axis = .8, mgp = c(2,.4,0))   #at = seq(-32, 44, by=8), labels =  seq(-32, 44
  mtext("Y",  las = 1, side = 3, at = min(X)-5, line = -0.5, cex = 1.)
  mtext("X",          side = 1, line = 1.6, cex = 0.9)
  mtext("(c)",            side = 1, line = 2.7, cex = 0.9)
  points(X, Y, cex=1, col="#3c78c3", pch = 16)
      XYsim = lm(Y~X)
      abline(XYsim, col= "#FF7F24", lwd=2)
      B1temp  = round(summary(XYsim)$coefficients[2, 1], 2)
      B1tempt = round(summary(XYsim)$coefficients[2, 3], 3)

## Locate results on scale of actual data
Yloc  = (B1temp< 0) *(min(Y) + 0.2*(max(Y)- min(Y))) +(B1temp> 0) *(max(Y) - 0.1*(max(Y)- min(Y)))
Xloc  = min(X) + 0.2*(max(X)- min(X))
Yloc2  = (B1temp< 0) *(min(Y) + 0.1*(max(Y)- min(Y))) +(B1temp> 0) *(max(Y) - 0.2*(max(Y)- min(Y)))
Xloc2 = min(X) + 0.15*(max(X)- min(X))
Xloc3 = min(X) + 0.38*(max(X)- min(X))
  text(Xloc, Yloc, expression(paste(hat(beta)[1], " = ")),    cex=0.85, col="blue")
  text(Xloc - 0.5 + abs(0.18*(max(X)- min(X)) ), Yloc, B1temp,                        cex=0.85, col="
  text(Xloc2, Yloc2, expression(paste("t-stat for ", hat(beta)[1], " = ")) , cex=0.85, col="blue")
  text(Xloc3 - 0.5, Yloc2, B1tempt , cex=1.3, col="blue")
```



Case Study: Gloabl warming data and Dickey Fuller unit root tests

Temperature

```
dta$ChangeTemp     = dta$observedtemp_new - dta$LagTemp
dta$ChangeTempLag     = c(NA, dta$ChangeTemp[1:(length(dta$ChangeTemp)-1)])
dta$ChangeTempLag     = c(NA, dta$ChangeTemp[1:(length(dta$ChangeTemp)-1)])

AugDickeyFuller.1    = lm(ChangeTemp ~ Year + LagTemp + ChangeTempLag, data = dta)
summary(AugDickeyFuller.1)
```

```
##
## Call:
## lm(formula = ChangeTemp ~ Year + LagTemp + ChangeTempLag, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3102 -0.0922  0.0017  0.0870  0.3369
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -4.032004   1.207511   -3.34   0.0011 **
## Year           0.002079   0.000621    3.35   0.0011 **
## LagTemp       -0.193212   0.061496   -3.14   0.0021 **
## ChangeTempLag -0.219749   0.085797   -2.56   0.0115 *
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.13 on 133 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.171,  Adjusted R-squared:  0.152
## F-statistic: 9.12 on 3 and 133 DF,  p-value: 0.0000157
```

```
## Packages for Augmented Dickey-Fuller Test Unit Root Test
options(digits = 5)        ## Turn scientific notation off
library(tseries)
adf.test(na.omit(dta$observedtemp_new), k = 1)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  na.omit(dta$observedtemp_new)
## Dickey-Fuller = -3.14, Lag order = 1, p-value = 0.1
## alternative hypothesis: stationary
```

```
summary(ur.df(na.omit(dta$observedtemp_new), type="trend"))
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3102 -0.0922  0.0017  0.0870  0.3369
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.122804   0.042186   -2.91   0.0042 **
## z.lag.1     -0.193212   0.061496   -3.14   0.0021 **
## tt           0.002079   0.000621    3.35   0.0011 **
## z.diff.lag  -0.219749   0.085797   -2.56   0.0115 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.13 on 133 degrees of freedom
## Multiple R-squared:  0.171,  Adjusted R-squared:  0.152
## F-statistic: 9.12 on 3 and 133 DF,  p-value: 0.0000157
##
##
## Value of test-statistic is: -3.1418 4.2338 5.6719
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
```

```
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

Carbon dioxide

```
dta$CO2ppm_new<- dta$CO2ppm
  # Old variable has years prior to 1959 subsetting that data in for the     # Fill in missing variable
  dta$CO2ppm_new[dta$Year>1959]<-dta$CO2NOAA[dta$Year>1959]

dta$LagCO2ppm    = c(NA, dta$CO2ppm_new[1:(length(dta$CO2ppm_new)-1)])
dta$ChangeCO2ppm     = dta$CO2ppm_new    - dta$LagCO2ppm
dta$ChangeCO2ppmLag= c(NA, dta$ChangeCO2ppm [1:(length(dta$ChangeCO2ppm )-1)])

AugDickeyFuller.2    = lm(ChangeCO2ppm ~ Year + LagCO2ppm + ChangeCO2ppmLag, data = dta)
summary(AugDickeyFuller.2)
```

```
##
## Call:
## lm(formula = ChangeCO2ppm ~ Year + LagCO2ppm + ChangeCO2ppmLag,
##     data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8786 -0.1968 -0.0160  0.0815  1.4436
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -5.199243   3.064898   -1.70  0.09214 .
## Year             0.000617   0.001940    0.32  0.75109
## LagCO2ppm        0.014137   0.003180    4.45 0.000018 ***
## ChangeCO2ppmLag  0.313921   0.085064    3.69  0.00032 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.348 on 134 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.797,  Adjusted R-squared:  0.793
## F-statistic:  176 on 3 and 134 DF,  p-value: <0.0000000000000002
```

```
adf.test(na.omit(dta$CO2ppm_new), k = 1)
```

```
## Warning in adf.test(na.omit(dta$CO2ppm_new), k = 1): p-value greater than
## printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  na.omit(dta$CO2ppm_new)
## Dickey-Fuller = 4.45, Lag order = 1, p-value = 0.99
## alternative hypothesis: stationary
```

```
#adf.test(na.omit(dta$CO2ppm_new), k = 1, alternative = "explosive")

library(urca)
summary(ur.df(na.omit(dta$CO2ppm_new), type="trend"))
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8786 -0.1968 -0.0160  0.0815  1.4436
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.040764   0.885374   -4.56 0.000011 ***
## z.lag.1      0.014137   0.003180    4.45 0.000018 ***
## tt           0.000617   0.001940    0.32  0.75109
## z.diff.lag   0.313921   0.085064    3.69  0.00032 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.348 on 134 degrees of freedom
## Multiple R-squared:  0.797,  Adjusted R-squared:  0.793
## F-statistic:  176 on 3 and 134 DF,  p-value: <0.0000000000000002
##
##
## Value of test-statistic is: 4.4461 20.303 25.9
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

Differenced models: ChangeTemp and ChangeCO2ppm

```
# Temperature
summary(ur.df(na.omit(dta$ChangeTemp), type="trend"))
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -0.30832 -0.08620  0.00431  0.08831  0.29017
##
## Coefficients:
##              Estimate Std. Error t value             Pr(>|t|)
## (Intercept) -0.011572   0.022622   -0.51              0.60984
## z.lag.1     -1.698956   0.136898  -12.41 < 0.0000000000000002 ***
## tt           0.000428   0.000286    1.50              0.13691
## z.diff.lag   0.284680   0.084039    3.39              0.00093 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.13 on 132 degrees of freedom
## Multiple R-squared:  0.688,  Adjusted R-squared:  0.681
## F-statistic: 96.9 on 3 and 132 DF,  p-value: <0.0000000000000002
##
##
## Value of test-statistic is: -12.41 51.37 77.054
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

```
adf.test(na.omit(dta$ChangeTemp), k = 1)
```

```
## Warning in adf.test(na.omit(dta$ChangeTemp), k = 1): p-value smaller than
## printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  na.omit(dta$ChangeTemp)
## Dickey-Fuller = -12.4, Lag order = 1, p-value = 0.01
## alternative hypothesis: stationary
```

```
# CO2
summary(ur.df(na.omit(dta$ChangeCO2ppm), type="trend"))
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1548 -0.1978 -0.0173  0.1121  1.5105
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.09960    0.06629   -1.50    0.135
## z.lag.1     -0.36998    0.07962   -4.65 0.000008 ***
## tt           0.00608    0.00146    4.17 0.000055 ***
## z.diff.lag  -0.20221    0.08422   -2.40    0.018 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.365 on 133 degrees of freedom
## Multiple R-squared:  0.267,  Adjusted R-squared:  0.25
## F-statistic: 16.1 on 3 and 133 DF,  p-value: 0.00000000533
##
##
## Value of test-statistic is: -4.6468 7.3614 10.913
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

```
adf.test(na.omit(dta$ChangeCO2ppm), k = 1)
```

```
## Warning in adf.test(na.omit(dta$ChangeCO2ppm), k = 1): p-value smaller than
## printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  na.omit(dta$ChangeCO2ppm)
## Dickey-Fuller = -4.65, Lag order = 1, p-value = 0.01
## alternative hypothesis: stationary
```

Figure 13.6 Temperature and CO2

```
    # First get both temp and CO2 on same scale
    c(mean(dta$CO2ppm_new, na.rm= T), sqrt(var(dta$CO2ppm_new, na.rm= T)))
```

```
## [1] 324.78  31.85
```

```
    c(min(dta$CO2ppm_new, na.rm= T), max(dta$CO2ppm_new, na.rm= T))
```

```
## [1] 290.23 408.52
```

```
    c(max(dta$observedtemp_new, na.rm= T), min(dta$observedtemp_new, na.rm= T))
```

```
## [1]  1.32250 -0.58083
```

```
    mm = 1.9/118    #OLD mm = 1.1/91
    bb = -0.58 - 290*mm
    CO2plot = bb + mm*dta$CO2ppm_new
    c(min(CO2plot, na.rm= T), max(CO2plot, na.rm= T))
```

```
## [1] -0.57638  1.32837
```

```
par(mfrow=c(1, 1), mar=c(2.5, 4.2, 0.5, 4.2), oma=c(0.2, 1.0, 1.0, 1.0))     # mar(south, west, north,
plot(dta$Year, dta$observedtemp_new, type="n", ylim = c(-0.6, 1.35), xlab= "", ylab= "", xaxt='n',  yax
    points(dta$Year, dta$observedtemp_new, type="l", col = "#3c78c3")
```

```
    mtext("Year", side = 1, line = 1.5, cex = 0.9)
    mtext("Temperature",    las = 1, side = 3, at = 1853, line = -0.1, cex = 0.9)
#   mtext("(deviation\nfrom\npre-industrial\naverage,\nin Fahrenheit)",     las = 1, side = 3, at = 185
    axis(1, at=seq(1880, 2020, by=20),  labels=seq(1880, 2020, by=20), cex.axis=0.85, padj=-1.2)
    axis(2, las = 1, at = seq(-0.5, 1.75 , by=0.25), labels =  seq(-0.5, 1.75 , by=0.25), tick = T,
        cex.axis = .8, mgp = c(2,.7,0))

    ## Add CO2
    mtext("Carbon\ndioxide",    las = 1, side = 3, at = 2035, line = -0.7, cex = 0.90)
#   mtext("(parts per\nmillion)", las = 1, side = 3, at = 2035, line = -2.1, cex = 0.65)
    axis(4, las= 1, at = seq(-0.5, 1.25 , by=0.25), labels =  seq(295, 410, by=15), tick = T,
        cex.axis = .8, mgp = c(2,.7,0))
    points(dta$Year, CO2plot, type="l", lty=2 , lwd=2.0, col = "#FF7F24")
    lines(c(1880, 1900), c(1.2, 1.2), col="blue")
    lines(c(1880, 1900), c(1.1, 1.1), lty=2 , lwd=2.0, col="#FF7F24")
    text(1933, 1.2, "Temperature (left-hand scale)",    cex= 0.75)
    text(1936.5, 1.1, "Carbon dioxide (right-hand scale)", cex= 0.75)
```

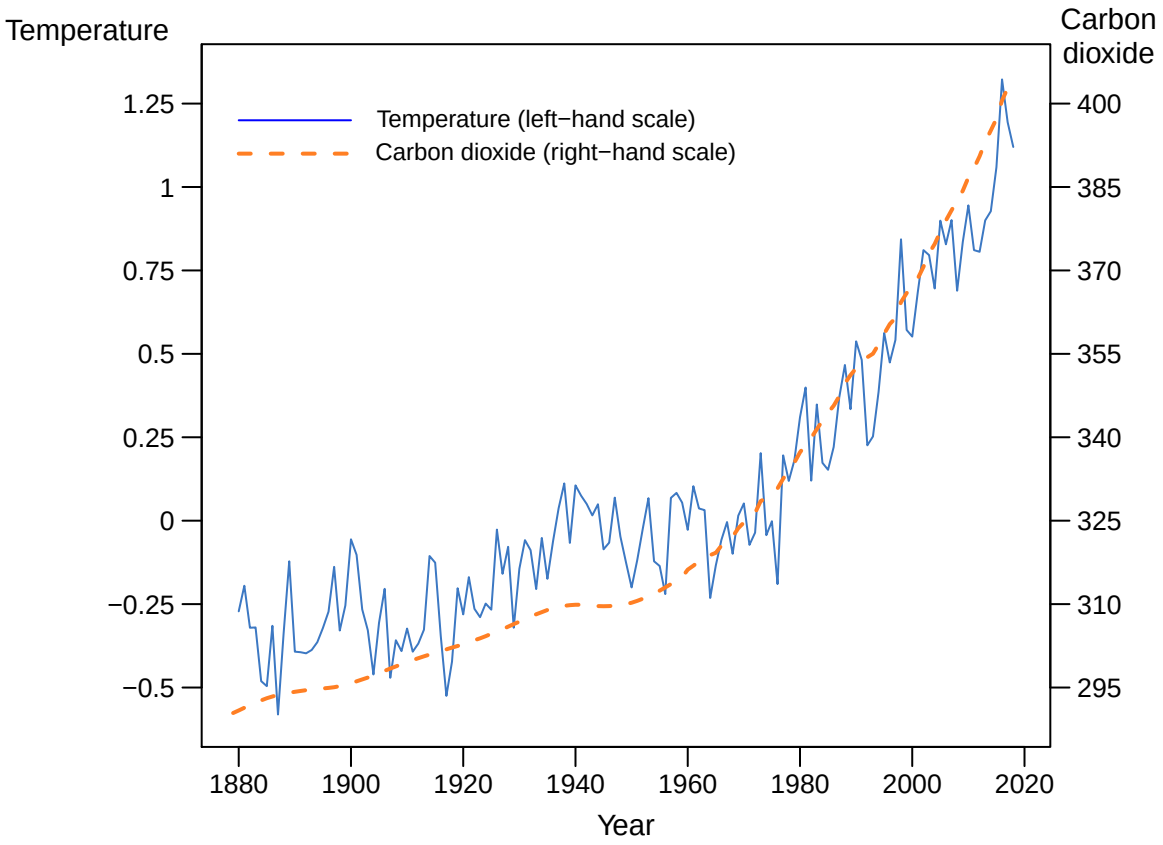


Table Regression model with CO2 as independent variable

```
ols.co2 = lm(observedtemp_new ~ LagTemp + CO2ppm_new + Year + YearSq, data = dta)
summary(ols.co2)
```

```
##
## Call:
```

```
## lm(formula = observedtemp_new ~ LagTemp + CO2ppm_new + Year +
##     YearSq, data = dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27699 -0.07744 -0.00012  0.07695  0.26721
##
## Coefficients:
##                 Estimate   Std. Error t value Pr(>|t|)
## (Intercept) -106.6011669   75.3208071   -1.42     0.16
## LagTemp        0.3346643    0.0803417    4.17 0.000056 ***
## CO2ppm_new     0.0117057    0.0026973    4.34 0.000028 ***
## Year           0.1079021    0.0781564    1.38     0.17
## YearSq        -0.0000283    0.0000205   -1.38     0.17
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.116 on 133 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.926,  Adjusted R-squared:  0.924
## F-statistic:  419 on 4 and 133 DF,  p-value: <0.0000000000000002
```

Table 13.5 Regression models using differenced variables

```
dta$lagYearSq       = c(NA, dta$YearSq[1:(length(dta$YearSq)-1)])
dta$ChangeYearSq        = dta$YearSq                    - dta$lagYearSq
ols.co2.diff = lm(ChangeTemp ~ ChangeTempLag + ChangeCO2ppm + ChangeYearSq, data = dta)
summary(ols.co2.diff)
```

```
##
## Call:
## lm(formula = ChangeTemp ~ ChangeTempLag + ChangeCO2ppm + ChangeYearSq,
##     data = dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26912 -0.09772  0.00349  0.08996  0.28338
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.195893   0.924157    2.38  0.01892 *
## ChangeTempLag -0.366286   0.078806   -4.65 0.000008 ***
## ChangeCO2ppm   0.095655   0.025166    3.80  0.00022 ***
## ChangeYearSq  -0.000581   0.000241   -2.40  0.01756 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.128 on 133 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.196,  Adjusted R-squared:  0.178
## F-statistic: 10.8 on 3 and 133 DF,  p-value: 0.00000206
```