Components Model

- Remember that we said that it was useful to think about the components representation

$$y_t = T_t + S_t + C_t$$

- Suppose that C_t is an AR(p) process
- What model does this imply for y_t?

Trend+Cycle Model

- For simplicity, we start with the trend-cycle model

$$y_t = T_t + C_t$$

- And specify the cycle as an AR(1)

$$C_t = \beta C_{t-1} + e_t$$

Intercept only

- Suppose the trend is just an intercept

$$T_t = \mu$$

- The model is

$$y_t = \mu + C_t$$

$$C_t = \beta C_{t-1} + e_t$$

Partial Differencing

- Lag the first equation, multiply by β and subtract

$$y_t = \mu + C_t$$

$$y_{t-1} = \mu + C_{t-1}$$

$$y_t - \beta y_{t-1} = (1-\beta)\mu + C_t - \beta C_{t-1}$$

Then use

$$C_t = \beta C_{t-1} + e_t$$

to find

$$y_t = (1-\beta)\mu + \beta y_{t-1} + e_t$$

Equivalence with AR(1)

- Thus

$$y_t = \mu + C_t$$

  implies

$$C_t = \beta C_{t-1} + e_t$$

$$y_t = \alpha + \beta y_{t-1} + e_t$$

  with

$$\alpha = (1 - \beta)\mu$$

- The model is just an AR(1) with intercept

Linear Trend

- Suppose the trend is a linear time trend

$$T_t = \mu_1 + \mu_2 t$$

- Then

$$y_t = \mu_1 + \mu_2 t + C_t$$

$$C_t = \beta C_{t-1} + e_t$$

Partial Differencing

- Lag the first equation, multiply by β and subtract, and use AR(1) equation

$$y_t = \mu_1 + \mu_2 t + C_t$$

$$y_{t-1} = \mu_1 + \mu_2 (t-1) + C_{t-1}$$

$$y_t = \beta y_{t-1} + (1-\beta)\mu_1 + \beta\mu_2 + (1-\beta)\mu_2 t + e_t$$

- We find

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 t + e_t$$

Summary : Trend+AR(1) Cycle

- When the trend is an intercept or a time trend
- The components model

$$y_t = T_t + C_t$$

$$C_t = \beta C_{t-1} + e_t$$

is equivalent with

$$y_t = T_t + \beta y_{t-1} + e_t$$

- The components model is equivalent with a regression on the trend variables and the lag

AR(p)+Trend

- A linear trend plus AR(p)

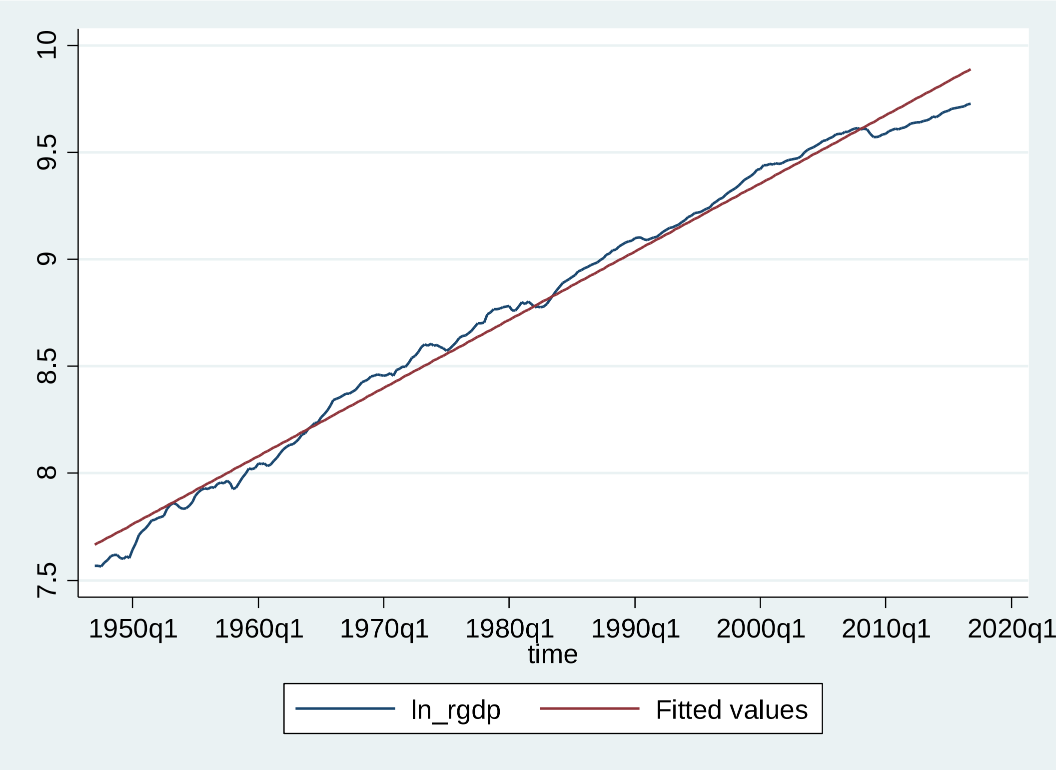
$$y_t = \mu_1 + \mu_2 t + C_t$$
$$C_t = \beta_1 C_{t-1} + \cdots + \beta_p C_{t-p} + e_t$$

is equivalent to

$$y_t = \alpha + \gamma t + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + e_t$$

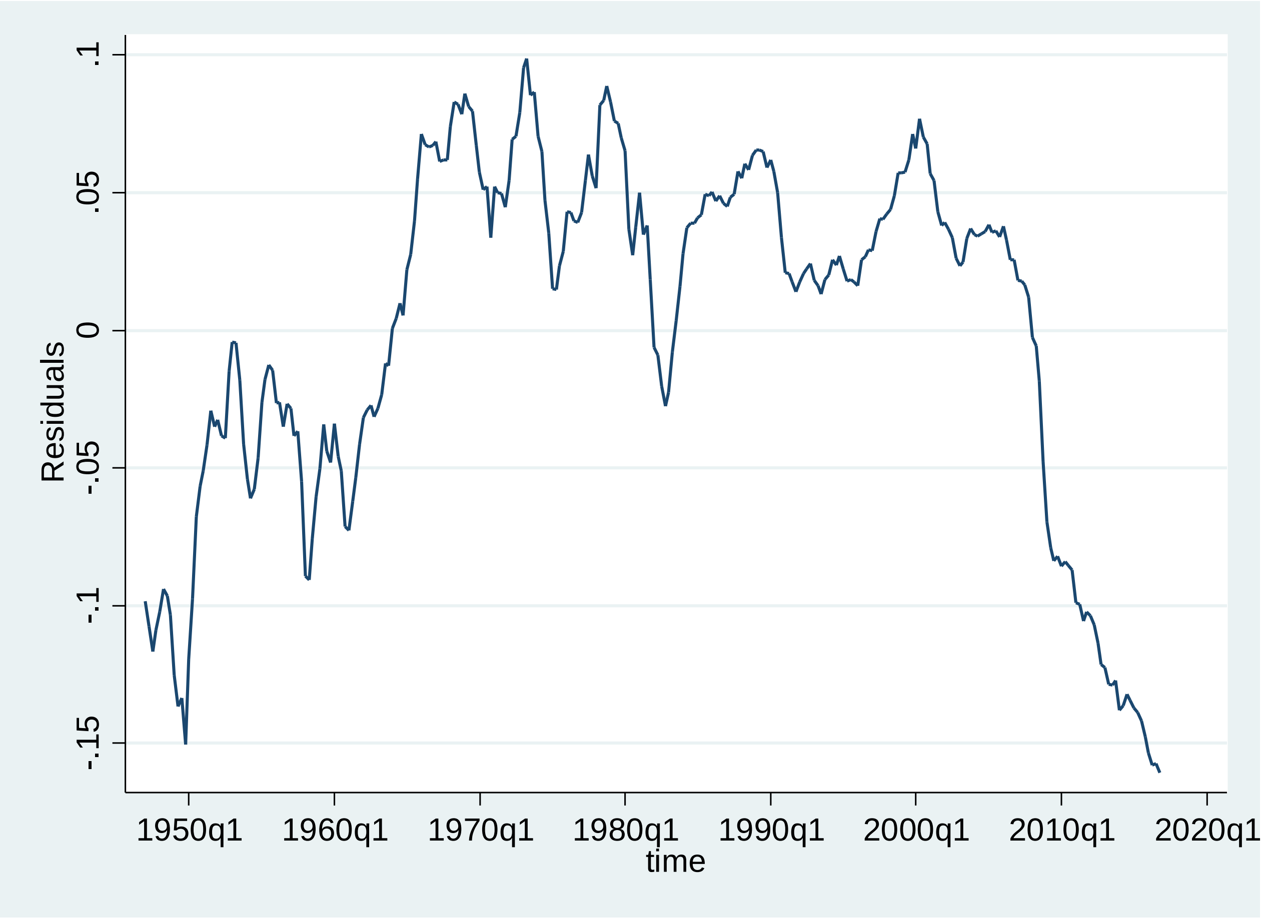
- A regression on a time trend plus *p* lags of *y*

Example: Real GDP



- Ln(rgdp) and linear trend

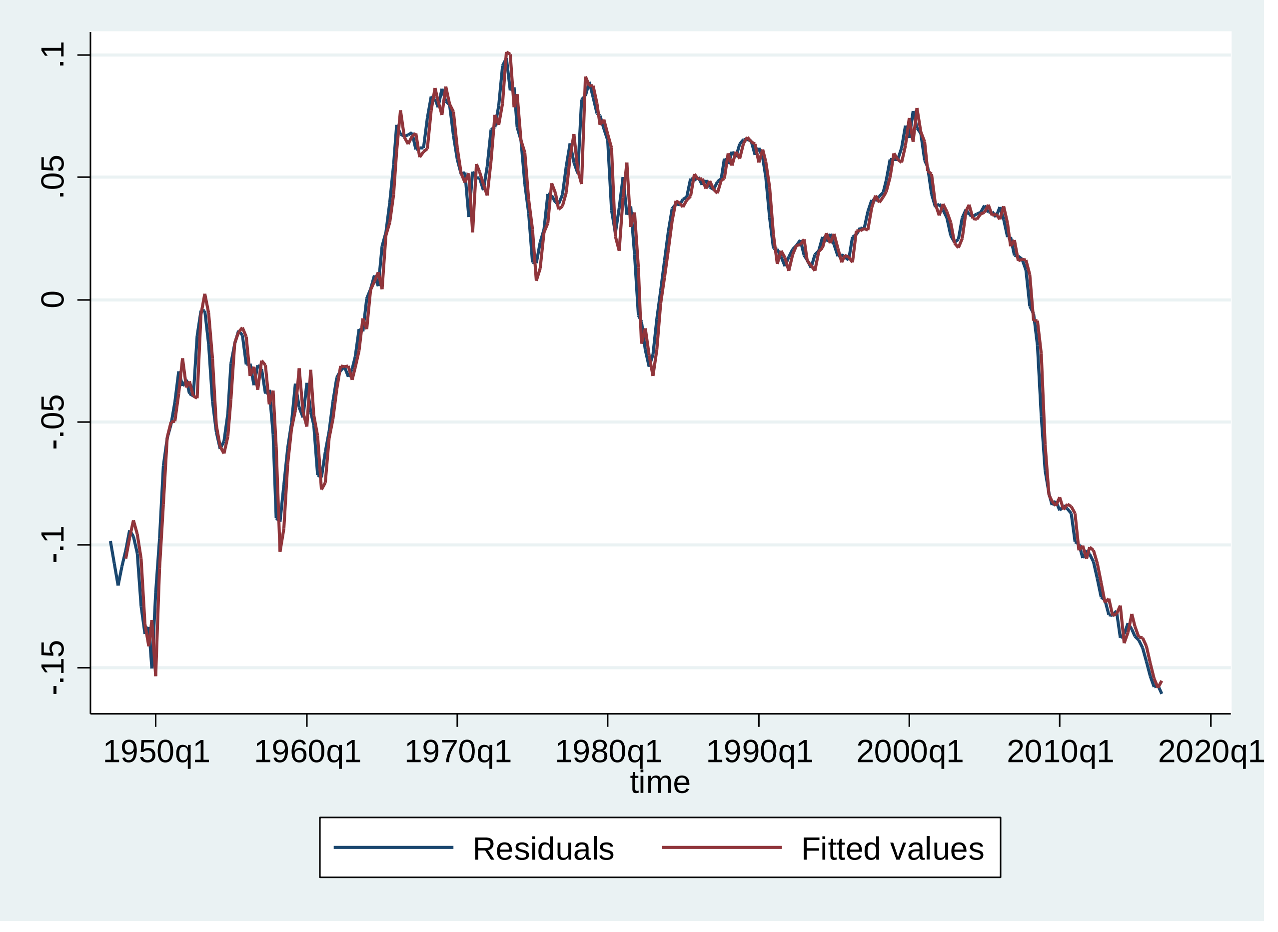
Residuals from Linear Trend



AR(4) on residuals

```
. reg u L(1/4).u

      Source |       SS           df       MS      Number of obs   =       276
-------------+----------------------------------   F(4, 271)       =   3643.88
       Model |  1.11484927         4  .278712318   Prob > F        =    0.0000
    Residual |  .020728205       271  .000076488   R-squared       =    0.9817
-------------+----------------------------------   Adj R-squared   =    0.9815
       Total |  1.13557748       275  .004129373   Root MSE        =    .00875

------------------------------------------------------------------------------
           u |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           u |
         L1. |   1.338388   .0604365    22.15   0.000     1.219403    1.457372
         L2. |  -.2142851   .1001542    -2.14   0.033    -.4114643   -.0171059
         L3. |   -.216372   .1001809    -2.16   0.032    -.4136037   -.0191403
         L4. |   .0809643   .0606865     1.33   0.183    -.0385128    .2004413
             |
       _cons |  -.0001075   .0005269    -0.20   0.838    -.0011449    .0009298
------------------------------------------------------------------------------
```

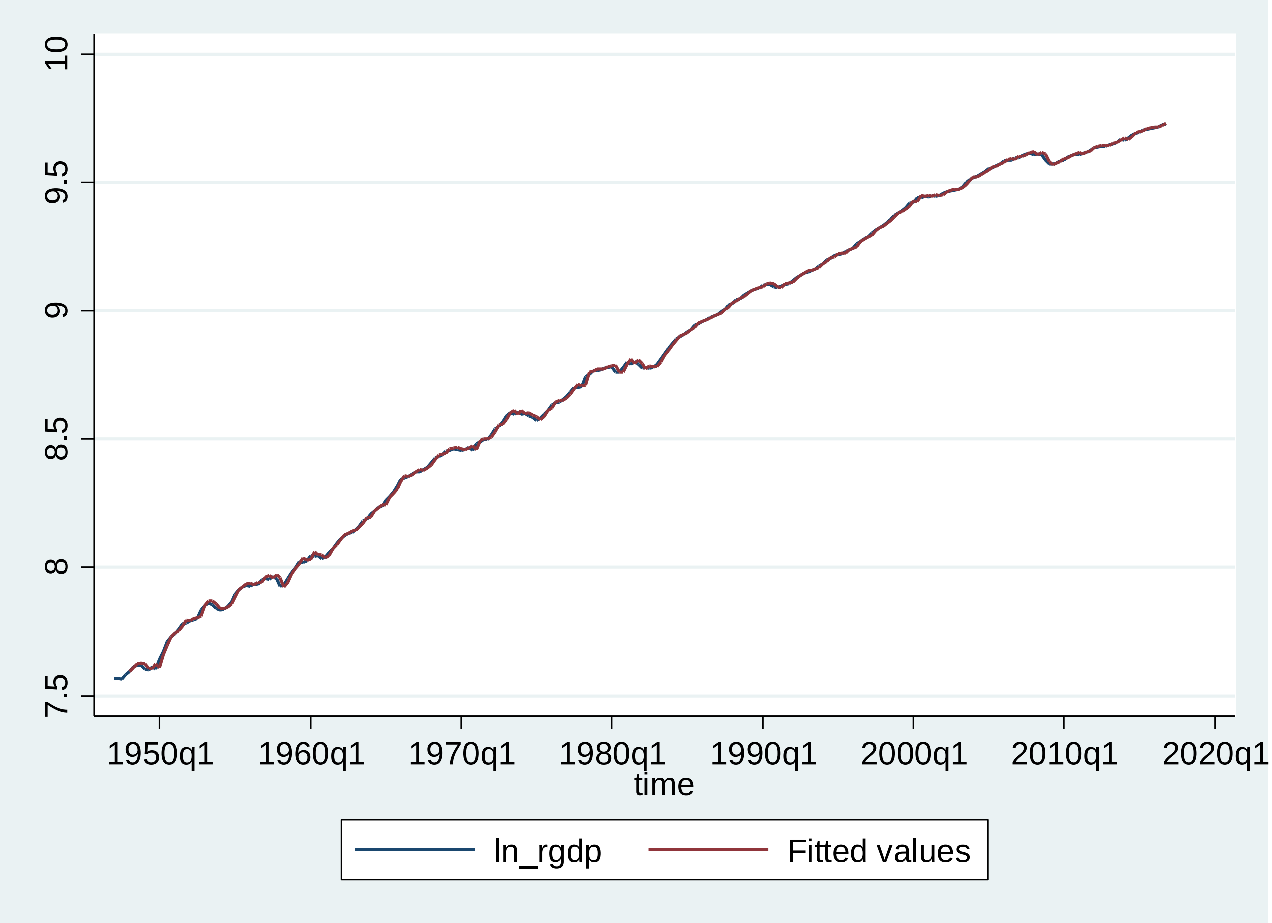
Fitted Values



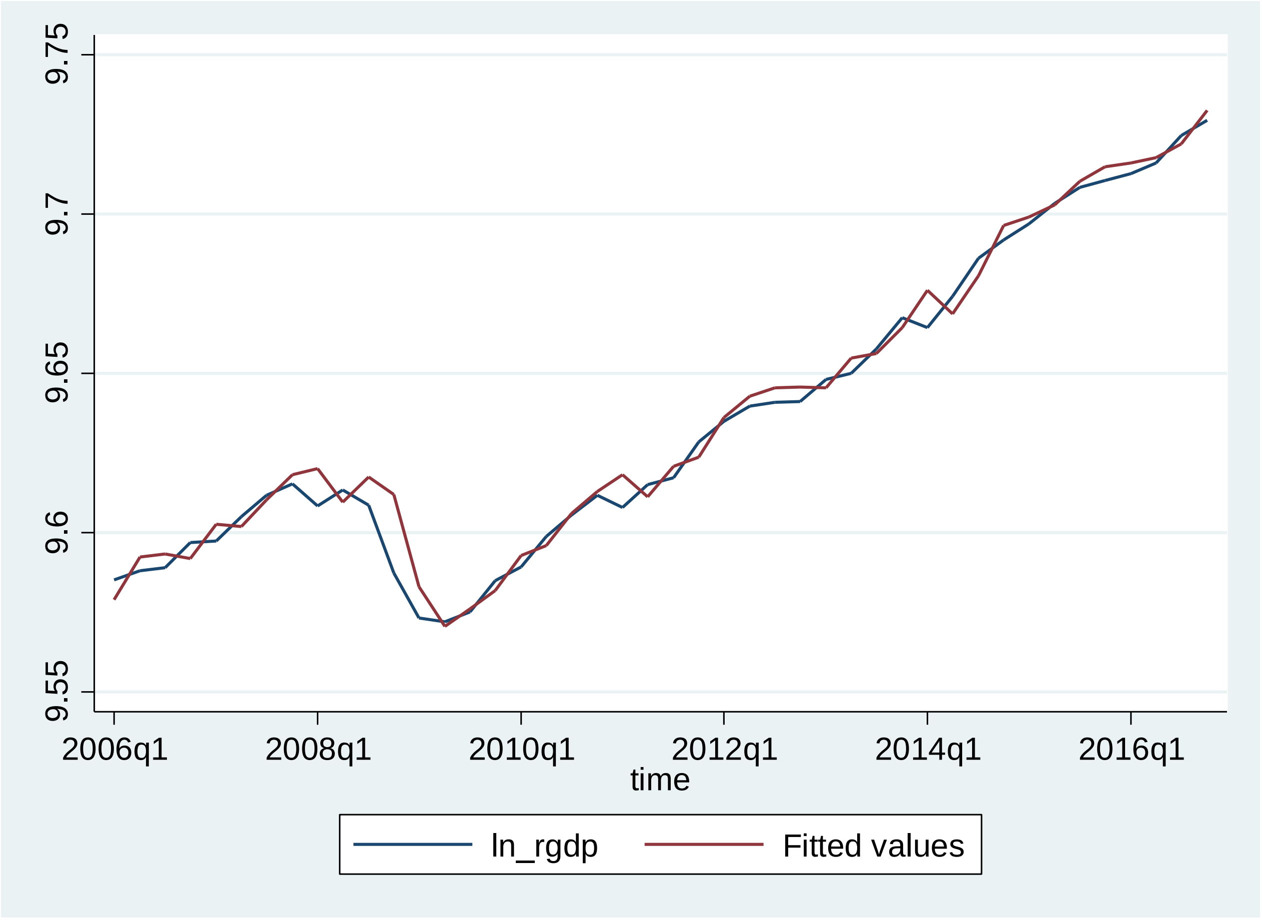
AR(4) with trend

```
. reg ln_rgdp time L(1/4).ln_rgdp

      Source |       SS           df       MS      Number of obs   =       276
-------------+----------------------------------   F(5, 270)       >  99999.00
       Model |  111.486043         5  22.2972086   Prob > F        =    0.0000
    Residual |  .020425309       270  .000075649   R-squared       =    0.9998
-------------+----------------------------------   Adj R-squared   =    0.9998
       Total |  111.506468       275  .405478066   Root MSE        =     .0087

------------------------------------------------------------------------------
     ln_rgdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |   .0000716   .0000679     1.05   0.293    -.0000622    .0002053
             |
     ln_rgdp |
         L1. |   1.324431   .0605076    21.89   0.000     1.205304    1.443558
         L2. |  -.2104888   .0996217    -2.11   0.036     -.406623   -.0143546
         L3. |  -.2185044   .0996359    -2.19   0.029    -.4146664   -.0223423
         L4. |   .0938848   .0606974     1.55   0.123    -.0256156    .2133852
             |
       _cons |   .0923474   .0680571     1.36   0.176    -.0416427    .2263375
------------------------------------------------------------------------------
```

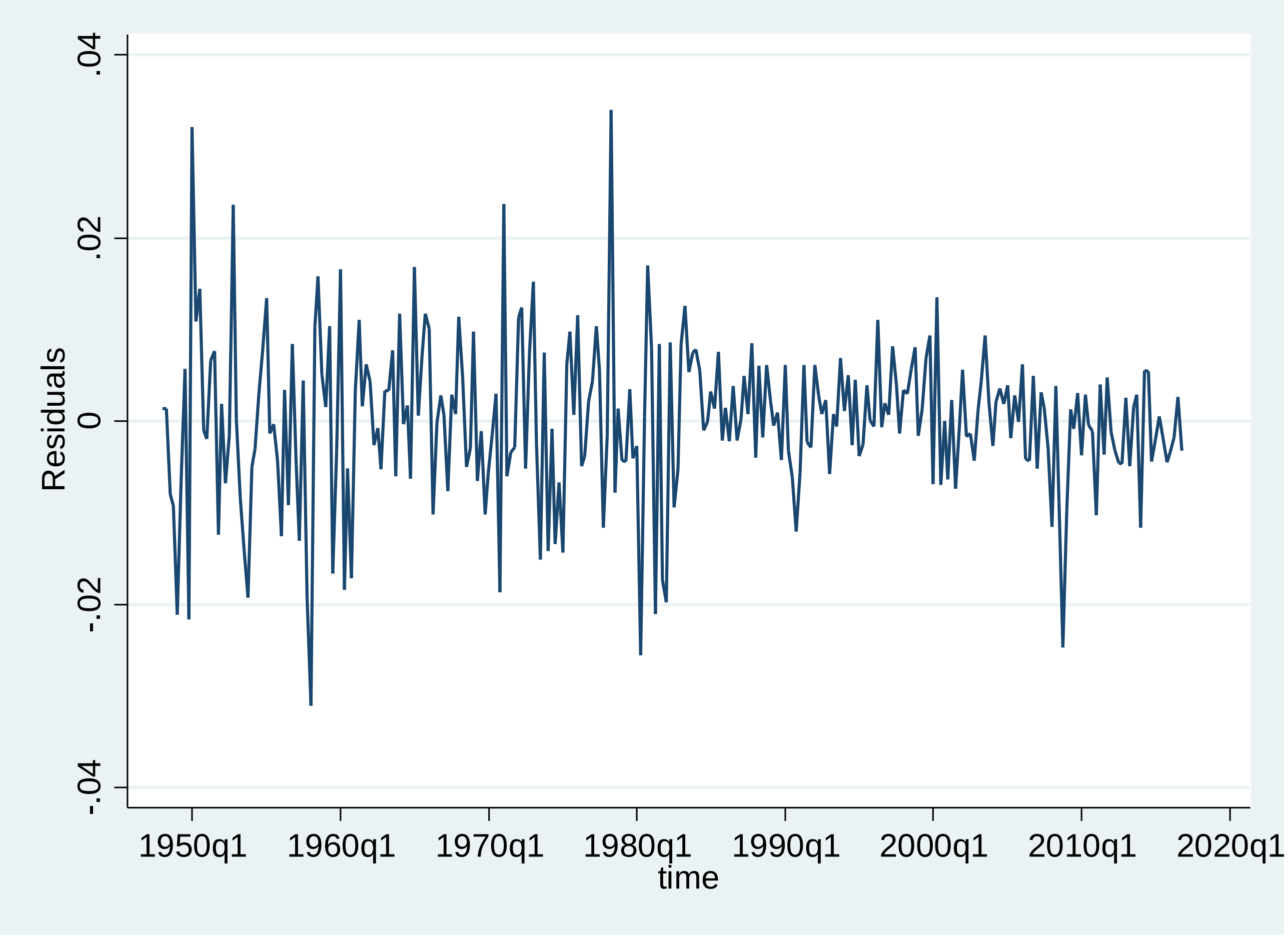
Fitted Values from AR(4) with Trend



Last 8 years


9.75
9.7
9.65
9.6
9.55
2006q1
2008q1
2010q1
2012q1
2014q1
2016q1
time
ln_rgdp
Fitted values

Residuals



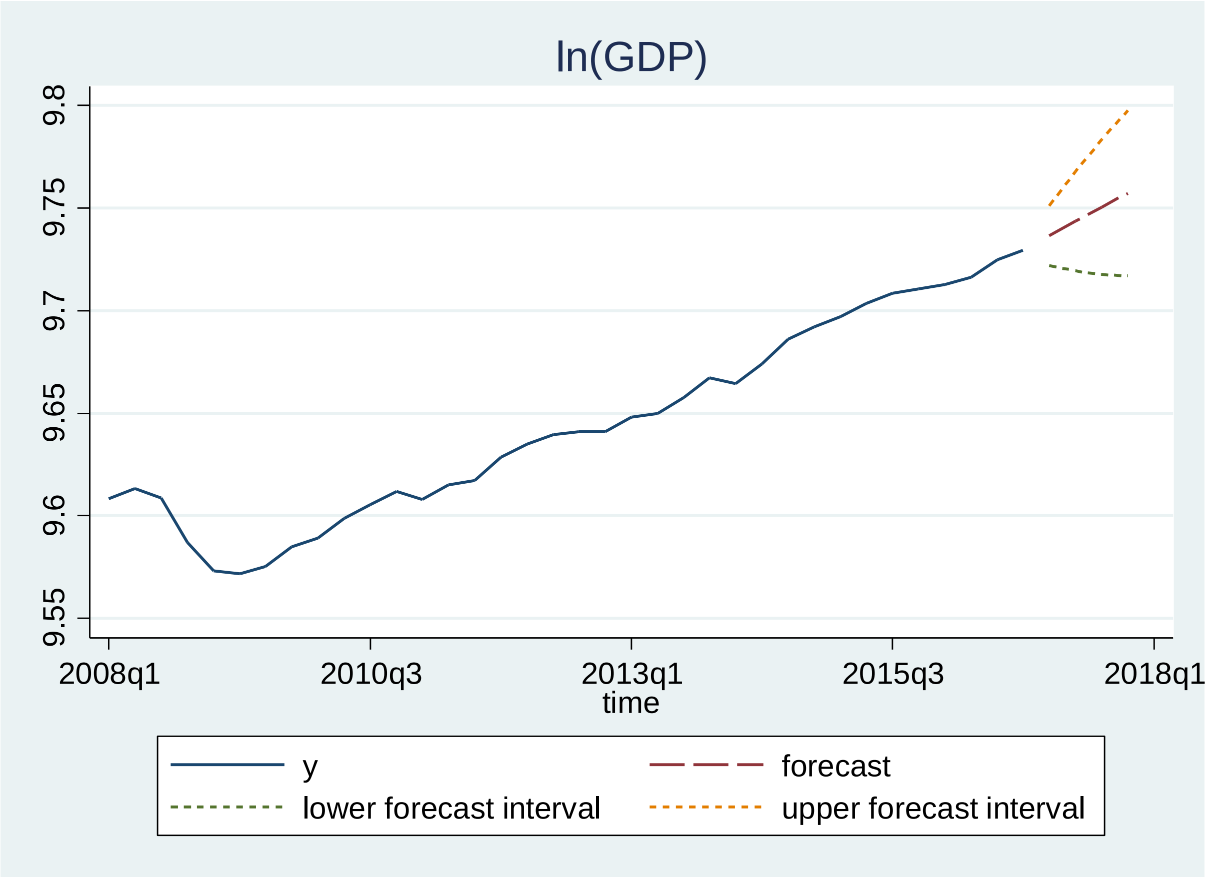
.04
.02
0
-.02
-.04
Residuals
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
time

Forecasts

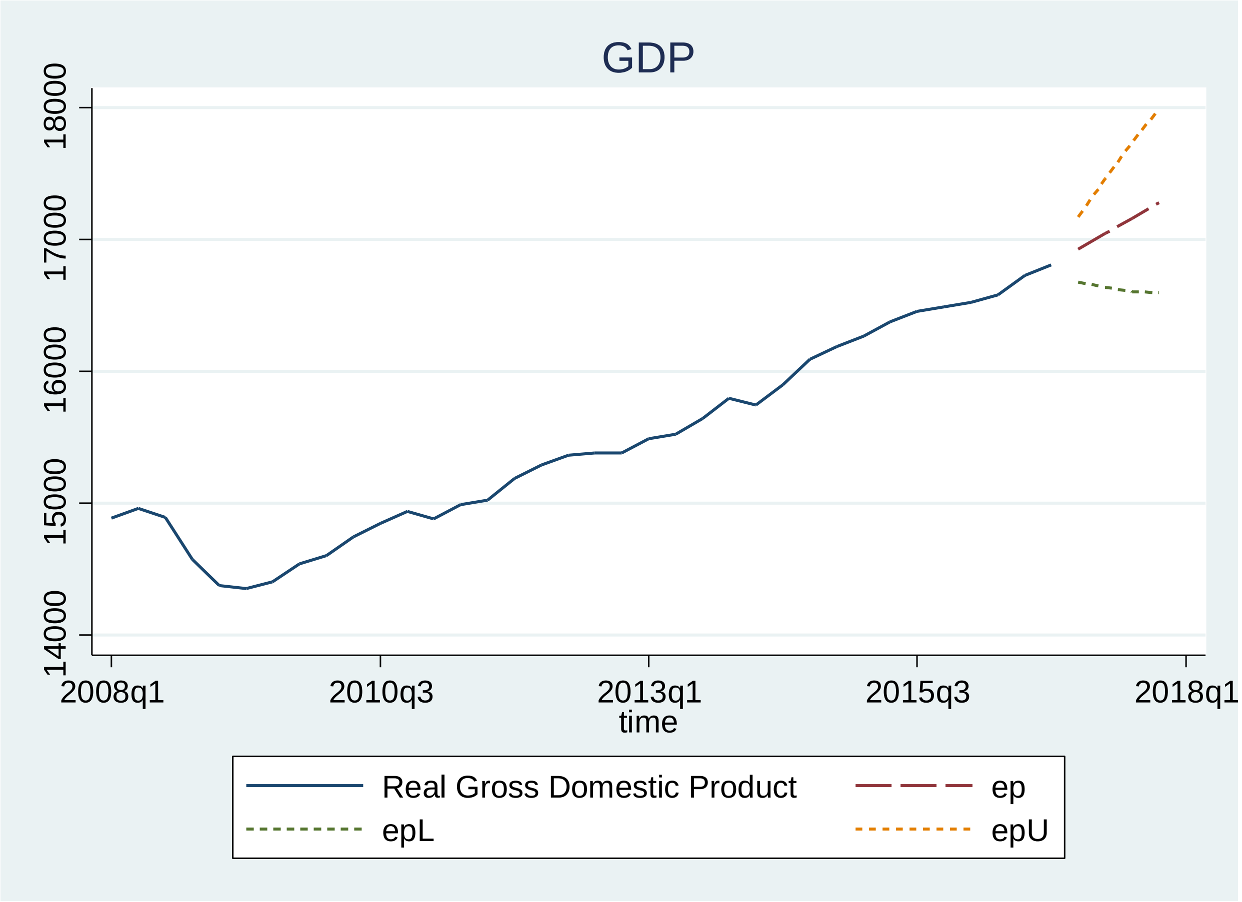
- Same as for AR(p) models, but include time trend as a regressor
- h-step forecast based on

$$y_t = \alpha + \gamma t + \beta_1 y_{t-h} + \cdots + \beta_p y_{t-h-p+1} + e_t$$

Forecast for ln(GDP) using AR(4)+trend



Forecast for GDP (using exponential)

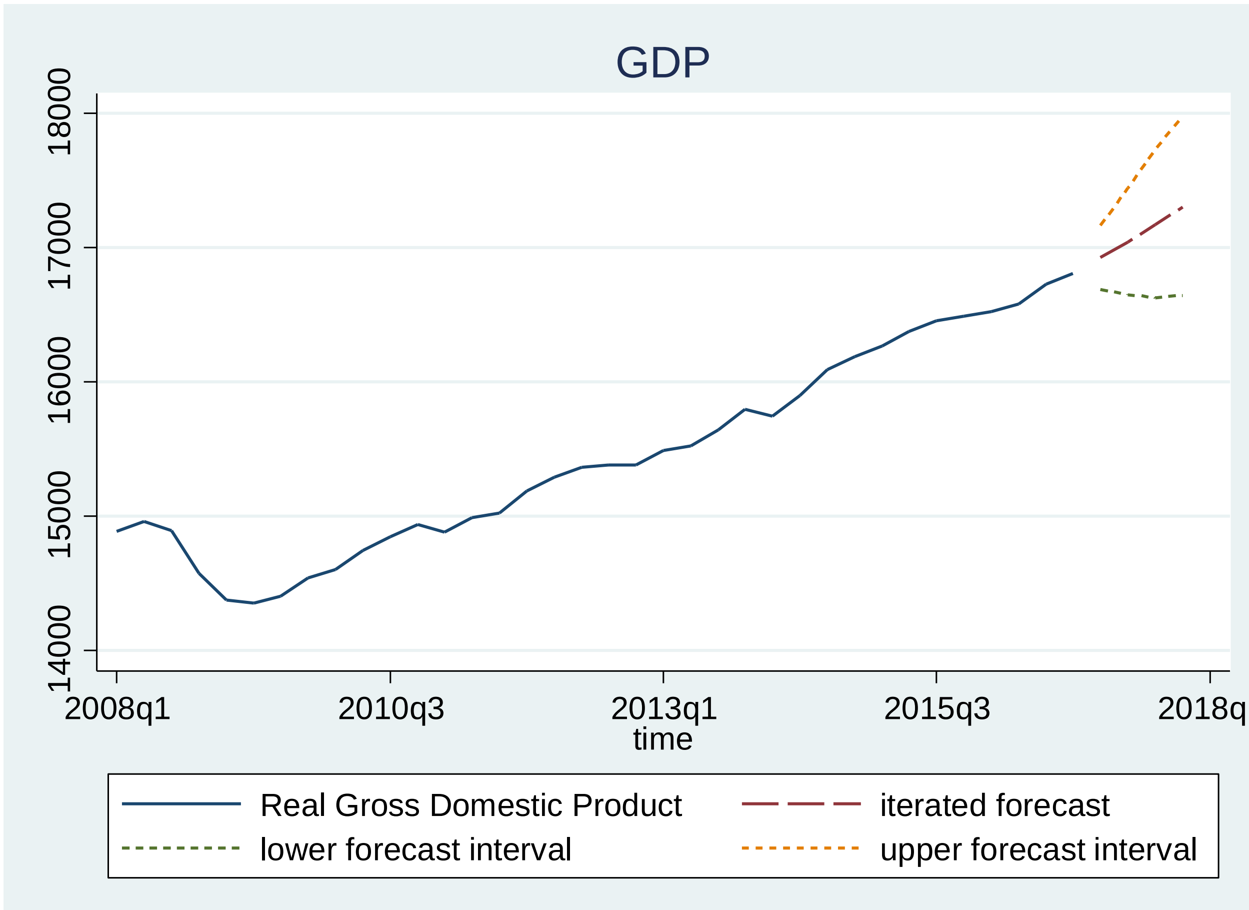


Direct Interval Forecasts

```
tsappend, add(4)
gen y=ln(gdp)
reg y time L(1/4).y
predict y1
predict sf1,stdf
gen y1L=y1-1.645*sf1
gen y1U=y1+1.645*sf1
reg y time L(2/5).y
predict y2
predict sf2,stdf
gen y2L=y2-1.645*sf2
gen y2U=y2+1.645*sf2
reg y time L(3/6).y
predict y3
predict sf3,stdf
gen y3L=y3-1.645*sf3
gen y3U=y3+1.645*sf3
```

```
reg y time L(4/7).y
predict y4
predict sf4,stdf
gen y4L=y4-1.645*sf4
gen y4U=y4+1.645*sf4
egen p=rowfirst(y1 y2 y3 y4) if t>=tq(2017q1)
egen pL=rowfirst(y1L y2L y3L y4L) if t>=tq(2017q1)
egen pU=rowfirst(y1U y2U y3U y4U) if t>=tq(2017q1)
label variable p "forecast"
label variable pL "lower forecast interval"
label variable pU "upper forecast interval"
tsline y p pL pU if time>=tq(2008q1), title(ln(GDP))
lpattern (solid longdash shortdash shortdash)
gen ep=exp(p)
gen epL=exp(pL)
gen epU=exp(pU)
tsline gdp ep epL epU if time>=tq(2008q1), title(GDP)
lpattern (solid longdash shortdash shortdash)
```

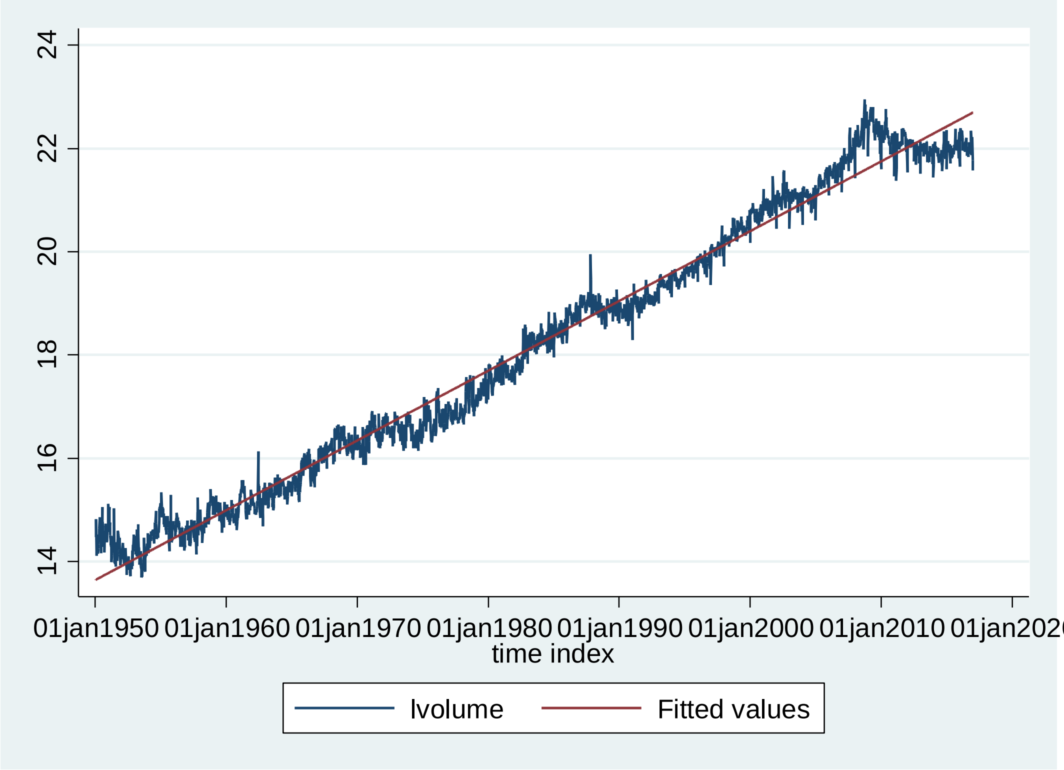
Iterated Forecast for GDP



Iterated Interval Forecasts

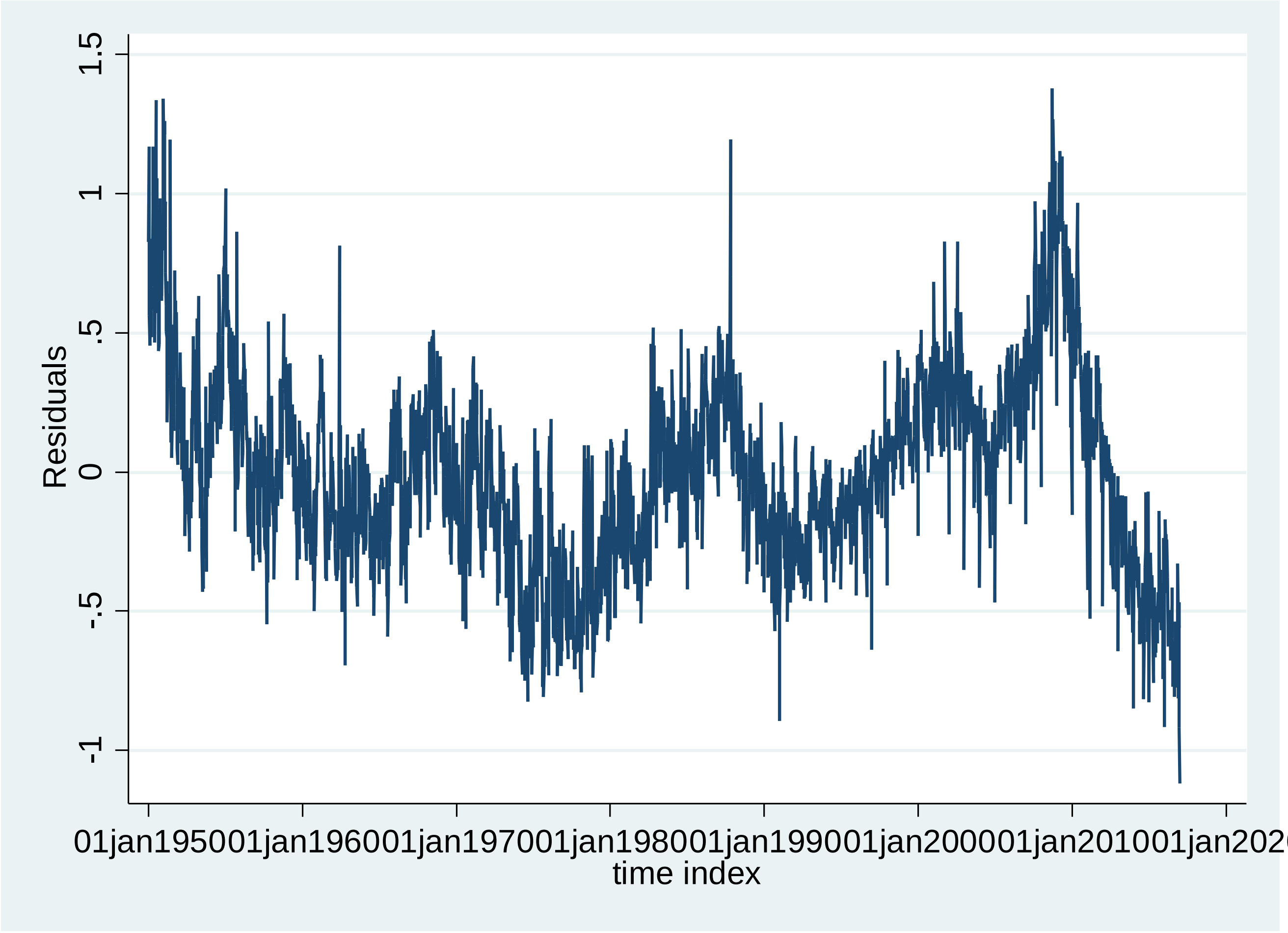
```
reg y time L(1/4).y
forecast create ar4
estimate store model1
forecast estimates model1
forecast solve, simulate(errors,statistic(stddev,prefix(sd_)) reps(1000) )
gen f=exp(f_y) if time>=tq(2017q1)
gen fL=exp(f_y-1.645*sd_y)
gen fU=exp(f_y+1.645*sd_y)
label variable f "iterated forecast"
label variable fL "lower forecast interval"
label variable fU "upper forecast interval"
tsline gdp f fL fU if time>=tq(2008q1), title(GDP) lpattern (solid longdash shortdash shortdash)
```

Example: Log Stock Volume



- Log Volume and Linear Trend

Residuals

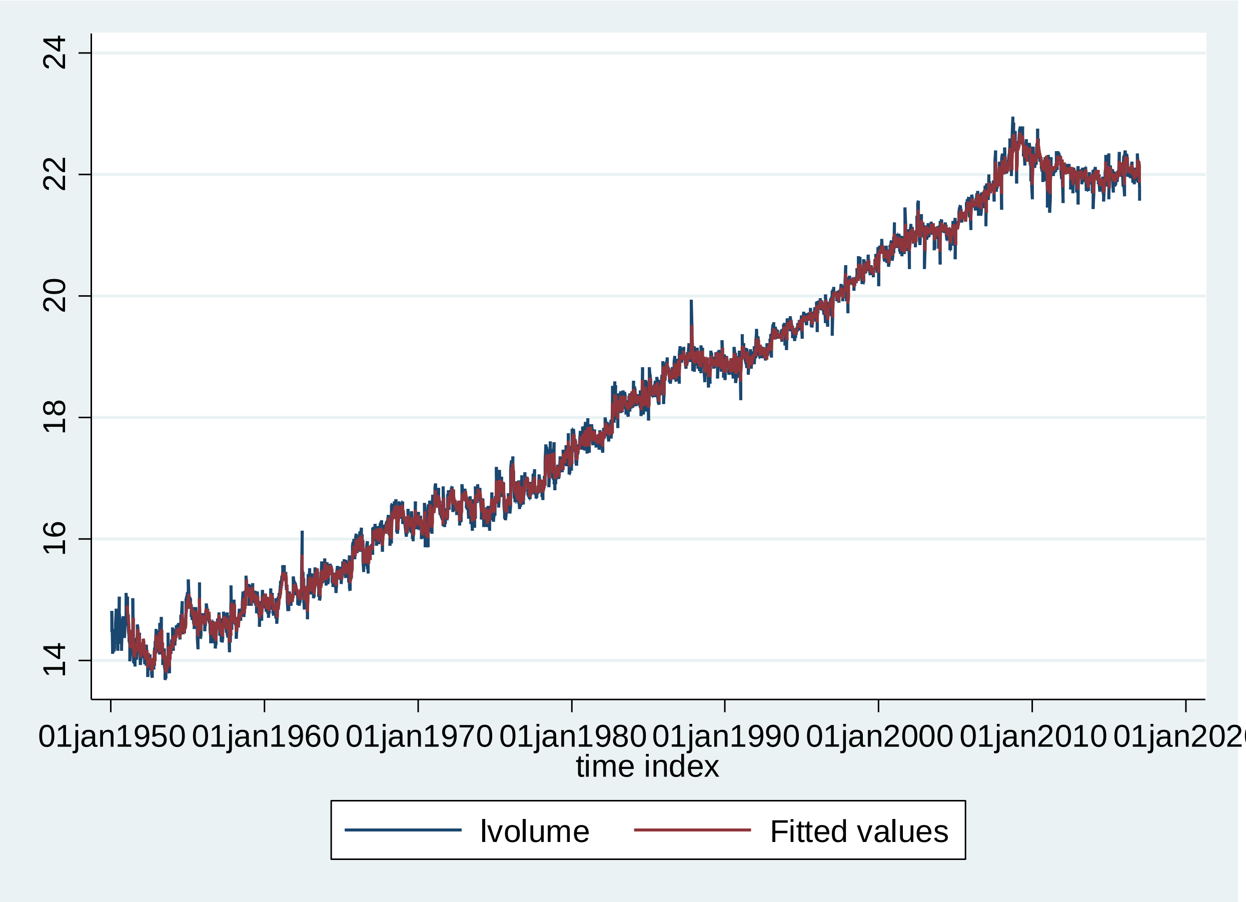


Residuals
1.5
1
.5
0
-.5
-1
01jan1950
01jan1960
01jan1970
01jan1980
01jan1990
01jan2000
01jan2010
01jan2020
time index

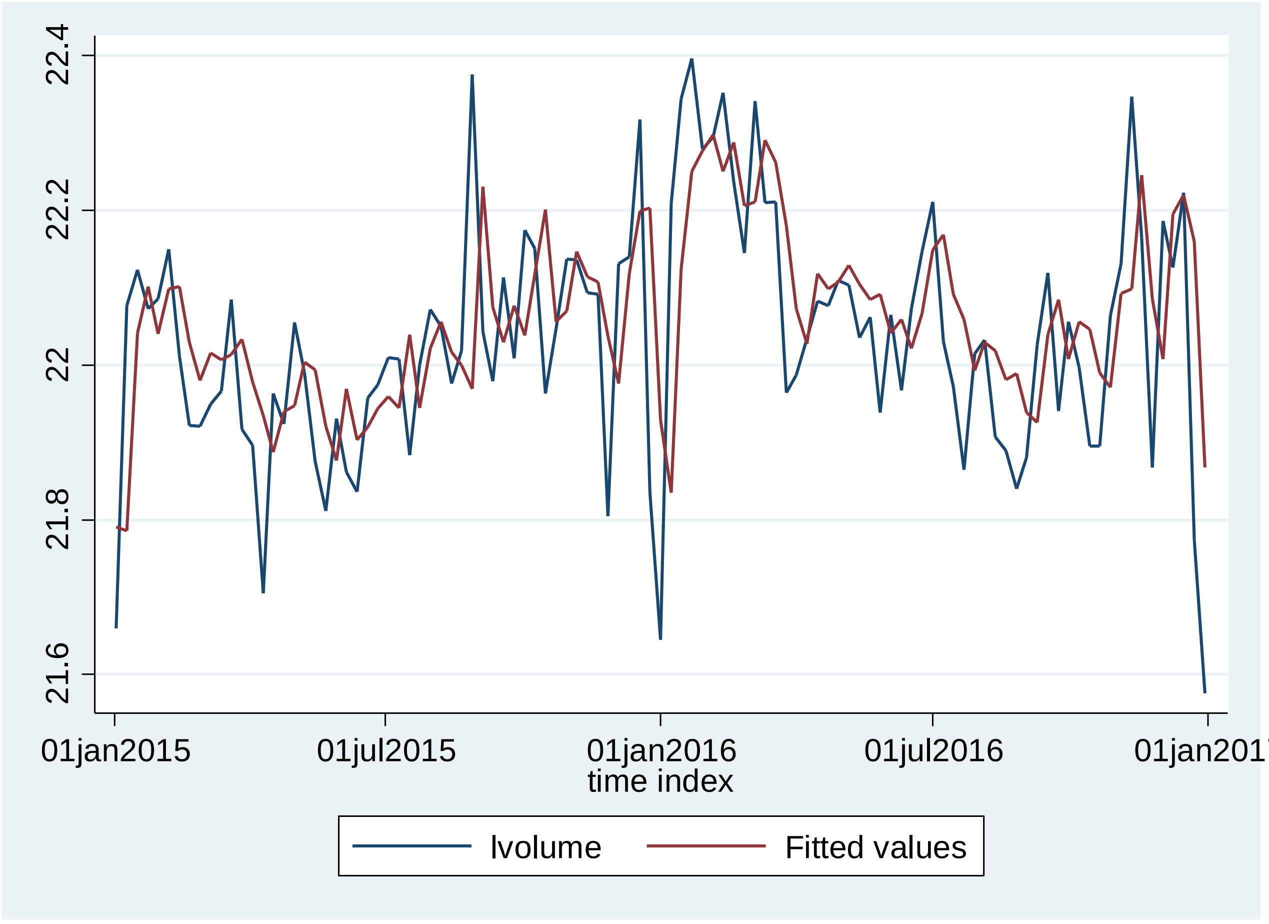
Model

- Weekly Data
- Fit AR(52)+trend

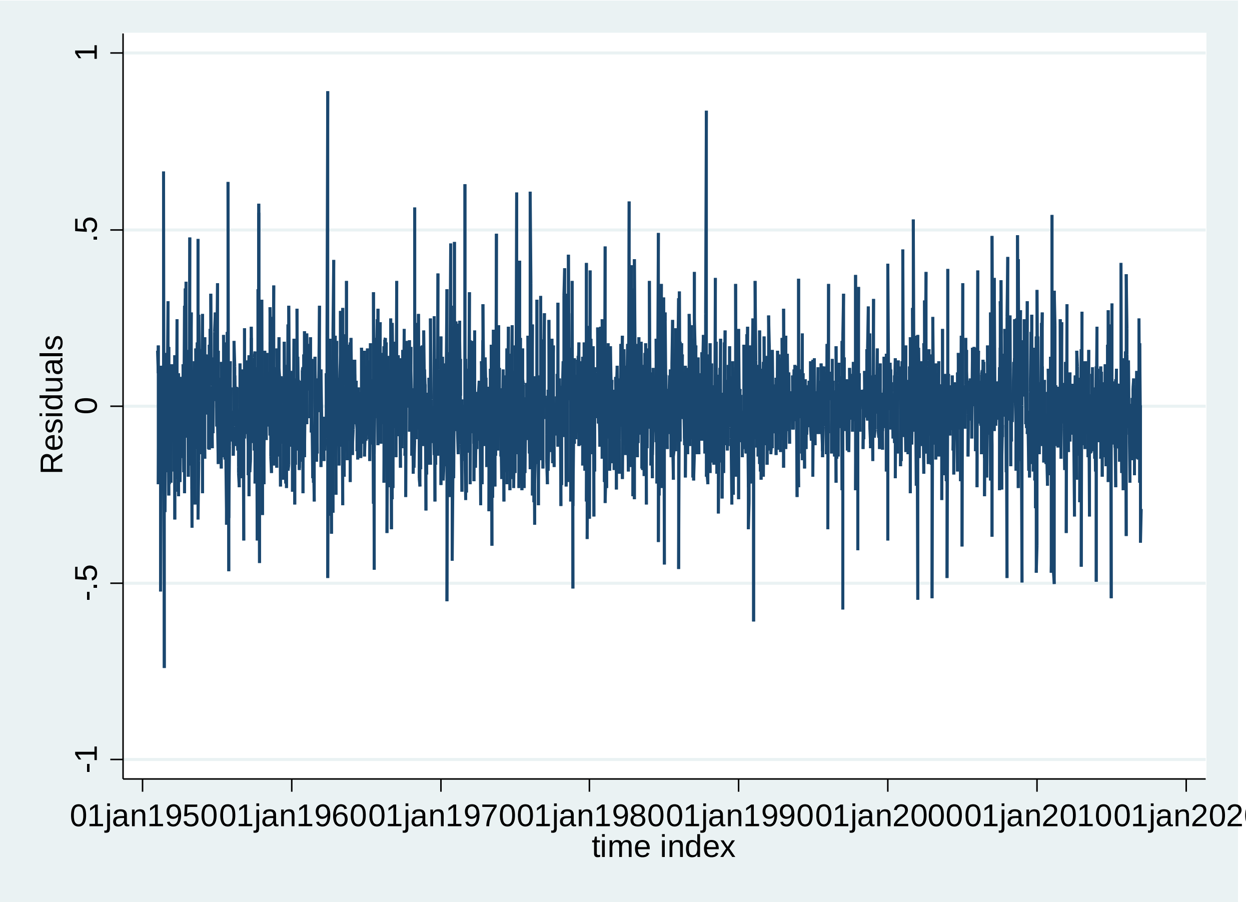
Data and Fitted



Last Two Years of Sample



Residuals



Residuals
1
.5
0
-.5
-1
01jan1950
01jan1960
01jan1970
01jan1980
01jan1990
01jan2000
01jan2010
01jan2020
time index

Trend Omission

- Suppose the truth is that the data have a trend, but you fit an AR model without a trend.
- What happens?
- Suppose

$$y_t = \mu_1 + \mu_2 t$$

- Then

$$y_t = y_{t-1} + \mu_2$$

Example

- Since

$$y_t = y_{t-1} + \mu_2$$

- If you estimate an AR(1), you obtain

$$y_t = \hat{\alpha} + \hat{\beta} y_{t-1} + \hat{e}_t$$

$$= \mu_2 + y_{t-1}$$

$$\hat{\alpha} = \mu_2$$

$$\hat{\beta} = 1$$

- You estimate a unit coefficient on the AR lag

General Effect of Trend Omission

- If the truth is

$$y_t = \mu_1 + \mu_2 t + \beta y_{t-1} + e_t$$

- But you estimate an AR(1) **without** a trend

$$y_t = \hat{\alpha} + \hat{\beta} y_{t-1} + \hat{e}_t$$

- Then you tend to find

$$\hat{\beta} \approx 1$$

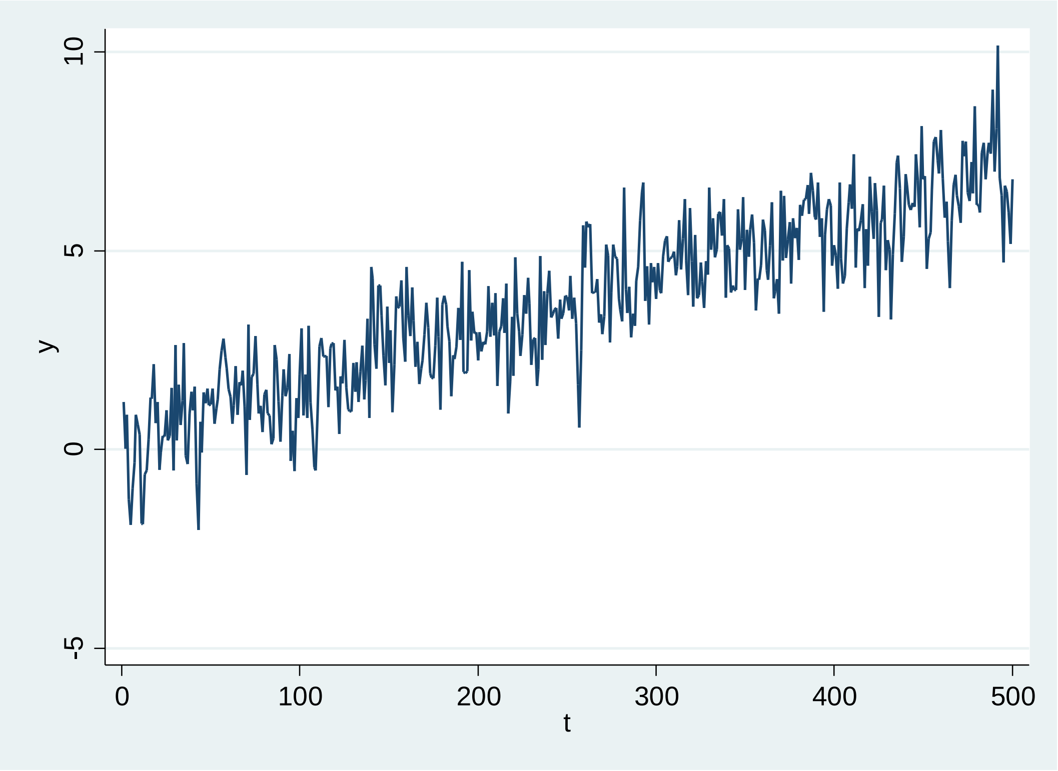
- This is due to model misspecification

Simulated Example

$$y_t = .01t + .3y_{t-1} + e_t$$

```
. gen e=rnormal(0)

. gen y=e

. replace y=.01*t+.3*L.y+e if t>1
(499 real changes made)
```

Simulated Process



Estimate AR(1) without Trend

```
. reg y L.y

      Source |       SS           df       MS      Number of obs   =       499
-------------+----------------------------------   F(1, 497)       =   1250.09
       Model |  1704.23834         1  1704.23834   Prob > F        =    0.0000
    Residual |  677.554342       497  1.36328841   R-squared       =    0.7155
-------------+----------------------------------   Adj R-squared   =    0.7150
       Total |  2381.79268       498  4.78271623   Root MSE        =    1.1676

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |    .846567   .0239436    35.36   0.000     .7995238    .8936102
             |
       _cons |   .5721782    .101954     5.61   0.000     .3718642    .7724922
------------------------------------------------------------------------------
```

- The estimated AR(1) coefficient is 0.85, much too large (true value was 0.3)

Estimate AR(1) with Trend

```
. reg y t L.y

      Source |       SS           df       MS      Number of obs   =       499
-------------+----------------------------------   F(2, 496)       =   1038.83
       Model |  1922.77107         2  961.385534   Prob > F        =    0.0000
    Residual |  459.021616       496  .925446806   R-squared       =    0.8073
-------------+----------------------------------   Adj R-squared   =    0.8065
       Total |  2381.79268       498  4.78271623   Root MSE        =      .962

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .0101097   .0006579    15.37   0.000     .0088171    .0114023
             |
           y |
         L1. |   .2523215   .0434121     5.81   0.000     .1670273    .3376157
             |
       _cons |   .2071672   .0872952     2.37   0.018     .0356533     .378681
------------------------------------------------------------------------------
```

- The estimated AR coef is 0.25, close to the true 0.3
- The estimated trend coef is 0.01, close to the true 0.10
- The root MSE decreases from 1.17 to 0.96

Midterm 1

- Tuesday March 7, during class
- Material:
  - Lectures through last week
  - Problem Sets 1-6
  - Diebold: Chapters 1-7
  - Not on midterm:
    - Chapter 9 & Today's lecture
- Review: During class Thursday
  - Bring Questions!!!!

Assignments

- Read Diebold Chapter 9
  - (material not on midterm)
- Read Chapter 6 from *The Signal and the Noise*
  - Reading Reflection
  - Due Thursday (3/2)
- Midterm 1
  - Tuesday (3/7)
- Problem Set # 7
  - Due Tuesday (3/14)