AR(1) Process

- The first-order autoregressive process, AR(1) is

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

where e_t is WN(0, σ^2)

Conditional Mean and Variance of AR(1)

- Conditional mean:

$$E(y_t \mid \Omega_{t-1}) = E(\beta y_{t-1} + e_t \mid \Omega_{t-1}) = \beta y_{t-1}$$

- Conditional variance:

$$\begin{aligned}\operatorname{var}(y_t \mid \Omega_{t-1}) &= \operatorname{var}(y_t - E(y_t \mid \Omega_{t-1}) \mid \Omega_{t-1}) \\ &= \operatorname{var}(e_t \mid \Omega_{t-1}) \\ &= \sigma^2\end{aligned}$$

Autocovariance of AR(1)

- Take the equation

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

- And then multiply both sides by y_{t-k}

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

- Then take expectations. Since e_t is white noise, it is uncorrelated with

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

or

$$\gamma(k) = \beta \gamma(k-1)$$

Autocorrelation of AR(1)

- Dividing by the variance, this implies

$$\rho(k) = \beta\rho(k-1)$$

- We know

$$\rho(0) = 1$$

- Then

$$\begin{aligned} \rho(1) &= \beta\rho(0) = \beta \\ \rho(2) &= \beta\rho(1) = \beta^2 \\ &\vdots \\ \rho(k) &= \beta^k \end{aligned}$$

Autocorrelation of AR(1)

- We have derived

$$\rho(k) = \beta^k$$

- The autocorrelation of the stationary AR(1) is a simple geometric decay ($|\beta|<1$)
- If β is small, the autocorrelations decay rapidly to zero with k
- If β is large (close to 1) then the autocorrelations decay moderately
- The AR(1) parameter describes the persistence in the time series

One-Step-Ahead Forecast

- As we showed earlier

$$E(y_t \mid \Omega_{t-1}) = \beta y_{t-1}$$

- Thus

$$E(y_{T+1} \mid \Omega_T) = \beta y_T$$

- The optimal one-step-ahead forecast is a linear function of the final observed value

2-step-ahead forecast

- By back-substitution

$$
\begin{aligned}
y_t &= \beta y_{t-1} + e_t \\
&= e_t + \beta\left(\beta y_{t-2} + e_{t-1}\right) \\
&= \beta^2 y_{t-2} + e_t + \beta e_{t-1}
\end{aligned}
$$

- Thus

$$
\begin{aligned}
E(y_t \mid \Omega_{t-2}) &= E(\beta^2 y_{t-2} + e_t + \beta e_{t-1} \mid \Omega_{t-2}) \\
&= \beta^2 y_{t-2}
\end{aligned}
$$

- and

$$
E(y_{T+2} \mid \Omega_T) = \beta^2 y_T
$$

2-step-ahead forecast

- This shows that the optimal 2-step-ahead forecast is also a linear function of the final observed value, but with the coefficient β^2.

$$E(y_{T+2} \mid \Omega_T) = \beta^2 y_T$$

h-step-ahead forecast

- Similarly

$$y_t = \beta^h y_{t-h} + e_t + \beta e_{t-1} + \cdots + \beta^{h-1} e_{t-h+1}$$

- So

$$E(y_t \mid \Omega_{t-h}) = E(\beta^h y_{t-h} + e_t + \beta e_{t-1} + \cdots + \beta^{h-1} e_{t-h+1} \mid \Omega_{t-2})$$
$$= \beta^h y_{t-h}$$

- Optimal forecast:

$$E(y_{T+h} \mid \Omega_T) = \beta^h y_T$$

Inversion of AR(1)

- By inverting the lag operator

$$\begin{aligned}(1-\beta L)y_t &= e_t \\ y_t &= (1-\beta L)^{-1} e_t \\ &= \left(\sum_{i=0}^{\infty} \beta^i L^i\right) e_t \\ &= \sum_{i=0}^{\infty} \beta^i e_{t-i}\end{aligned}$$

- Which is the same as found by back substitution

Condition for Invertibility

- The operator (1-βL) is invertible when |β|<1
- This is the same as for the MA(1) model
- β is the inverse of the root of the polynomial 1-βL
- The root of a function is the value where it crosses the x-axis
- The root of 1-βL is 1/β, the inverse of the root is β
- Invertibility requires that the inverse of the root be less than one

AR(1) with Intercept

- An AR(1) with intercept is

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

Taking expectations

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

- Thus

$$\mu = \alpha + \beta\mu$$

- and

$$\mu = \frac{\alpha}{1-\beta}$$

Best Linear Predictor

- A linear predictor of y_t given y_{t-1} is

$$\alpha + \beta y_{t-1}$$

- The forecast error is

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

- The linear predictor which minimizes the expected squared forecast error solves

$$\min_{\alpha,\beta} E(y_t - \alpha - \beta y_{t-1})^2$$

Least-Squares

- The estimate of the expected squared linear forecast error is the sum of squared errors
- The least squares estimate

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

minimizes the sum of squared errors, so is the estimate of the best linear predictor

- This is a linear regression, treating y_{t-1} as a regressor.

Unemployment Rate

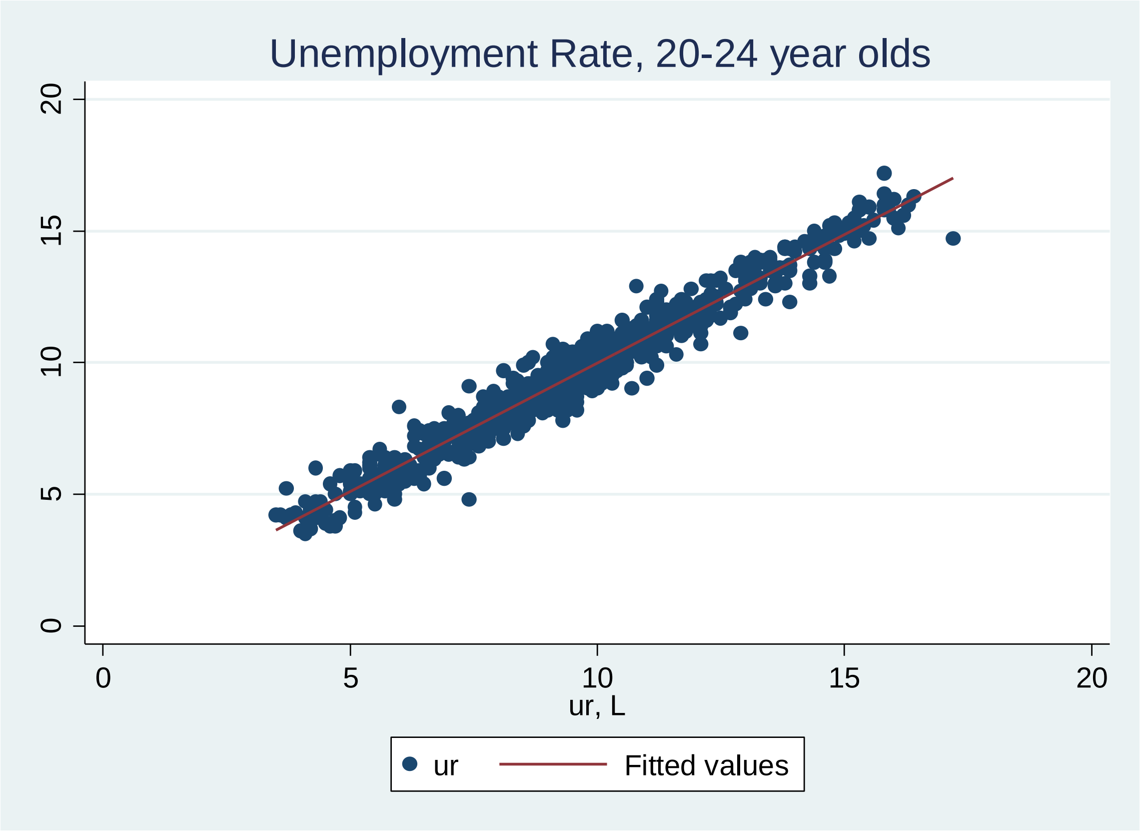
```
. regress ur L.ur

      Source |       SS           df       MS      Number of obs   =       828
-------------+----------------------------------   F(1, 826)       =  17943.11
       Model |  5337.15724         1  5337.15724   Prob > F        =    0.0000
    Residual |   245.69276       826  .297448861   R-squared       =    0.9560
-------------+----------------------------------   Adj R-squared   =    0.9559
       Total |     5582.85       827  6.75072551   Root MSE        =    .54539

------------------------------------------------------------------------------
          ur |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ur |
         L1. |   .9767288   .0072916   133.95   0.000     .9624165    .9910411
             |
       _cons |   .2213102   .0709754     3.12   0.002     .0819969    .3606236
------------------------------------------------------------------------------
```

Fitted AR(1)

- To plot a scatter and fitted regression,
  - `twoway scatter ur L.ur || lfit ur L.ur, title("Unemployment Rate, 20-24 year olds")`

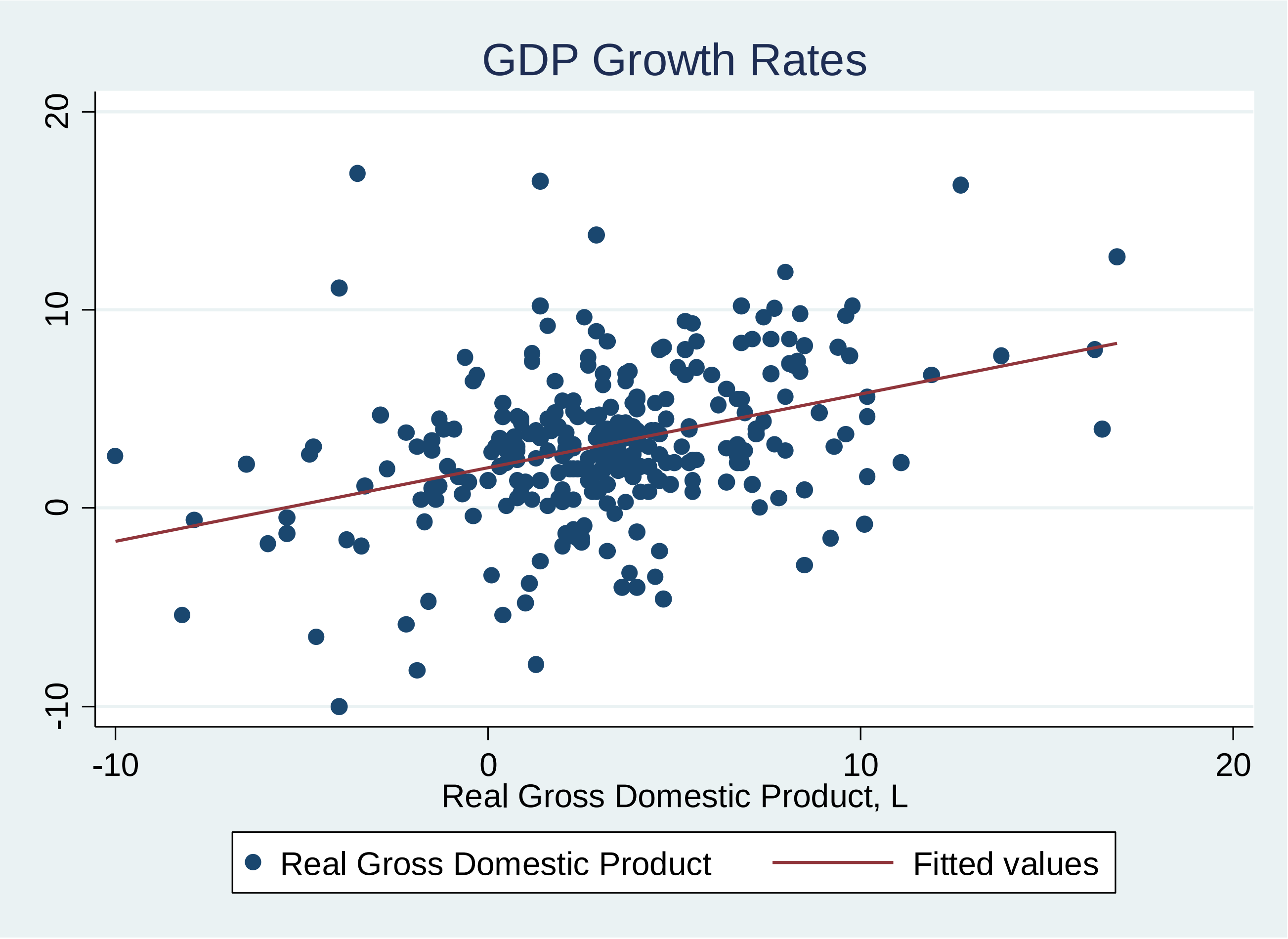


GDP Growth Rates

```
. regress gdp L.gdp

      Source |       SS           df       MS      Number of obs   =       278
-------------+----------------------------------   F(1, 276)       =     44.32
       Model |  585.172068         1  585.172068   Prob > F        =    0.0000
    Residual |  3643.98246       276   13.202835   R-squared       =    0.1384
-------------+----------------------------------   Adj R-squared   =    0.1352
       Total |  4229.15453       277  15.2677059   Root MSE        =    3.6336

------------------------------------------------------------------------------
         gdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         gdp |
         L1. |   .3714752   .0557984     6.66   0.000     .2616307    .4813197
             |
       _cons |   2.036509   .2826904     7.20   0.000     1.480006    2.593013
------------------------------------------------------------------------------
```



GDP Growth Rates
20
10
0
-10
-10
0
10
20
Real Gross Domestic Product, L
Real Gross Domestic Product
Fitted values

One-Step-Ahead Forecast

- The optimal forecast for T+1 given T is

$$\hat{y}_{T+1|T} = \alpha + \beta y_T$$

- The forecast using the estimates is

$$\hat{y}_{T+1|T} = \hat{\alpha} + \hat{\beta} y_T$$

Unemployment Rate, 20-24 year olds

- The estimates were

ur	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ur						
L1.	.9767288	.0072916	133.95	0.000	.9624165	.9910411
_cons	.2213102	.0709754	3.12	0.002	.0819969	.3606236

$$y_t = 0.22 + 0.977 y_{t-1} + \hat{e}_t$$

- The value for Jan 2016 is 8.3%, so

$$\hat{y}_{2017:2} = 0.22 + 0.977 \times 8.3 = 8.3$$

- Point forecast is 8.3% (unchanged from current)

- list time ur L.ur p if time>=tm(2016m1)

	time	ur	L.ur	p
817.	2016m1	8.2	9.4	9.402561
818.	2016m2	8.6	8.2	8.230487
819.	2016m3	8.4	8.6	8.621178
820.	2016m4	8.8	8.4	8.425833
821.	2016m5	8.3	8.8	8.816524
822.	2016m6	8.6	8.3	8.328159
823.	2016m7	8.9	8.6	8.621178
824.	2016m8	8.1	8.9	8.914197
825.	2016m9	8.2	8.1	8.132813
826.	2016m10	8.4	8.2	8.230487
827.	2016m11	8.1	8.4	8.425833
828.	2016m12	8.2	8.1	8.132813
829.	2017m1	8.3	8.2	8.230487
830.	2017m2	.	8.3	8.328159

Example – GDP Growth

- The estimates were

gdp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
gdp						
L1.	.3714752	.0557984	6.66	0.000	.2616307	.4813197
_cons	2.036509	.2826904	7.20	0.000	1.480006	2.593013

$$y_t = 2.04 + 0.371 y_{t-1} + \hat{e}_t$$

- The value for 4th quarter 2016 is 1.9%, so

$$\hat{y}_{2017:1} = 2.04 + 0.371 \times 1.9 = 2.7\%$$

GDP Growth

	time	gdp	L. gdp	p
272.	2015q1	2	2.3	2.890902
273.	2015q2	2.6	2	2.77946
274.	2015q3	2	2.6	3.002345
275.	2015q4	.9	2	2.77946
276.	2016q1	.8	.9	2.370837
277.	2016q2	1.4	.8	2.333689
278.	2016q3	3.5	1.4	2.556575
279.	2016q4	1.9	3.5	3.336673
280.	2017q1	.	1.9	2.742312

One-Step-Ahead Forecast Error

- The forecast error is

$$
\begin{aligned}
y_{T+1} - \hat{y}_{T+1|T} &= \alpha + \beta y_T + e_{T+1} - \left(\alpha + \beta y_T\right) \\
&= e_{T+1}
\end{aligned}
$$

- The forecast variance is

$$\operatorname{var}(y_{T+1} - \hat{y}_{T+1|T}) = \operatorname{var}(e_{T+1}) = \sigma^2$$

Forecast variance estimation

- Average of squared residuals

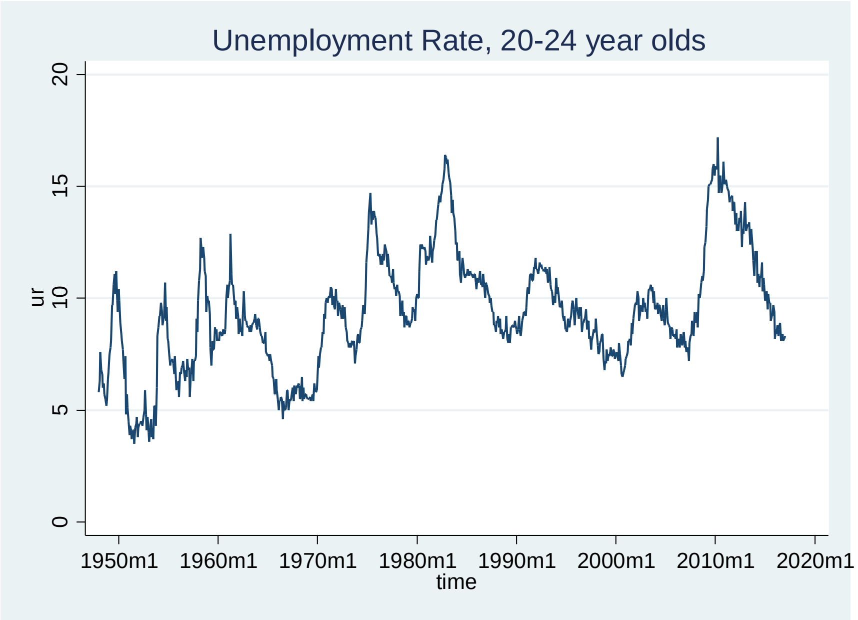
$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{e}_t^2$$

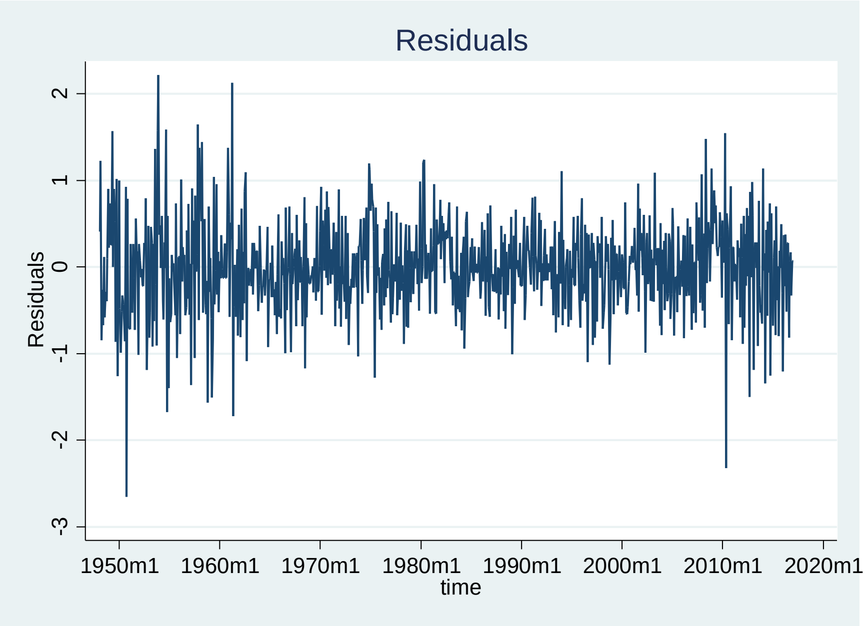
$$\hat{\sigma} = \sqrt{\hat{\sigma}^2}$$

where the least-squares residuals are

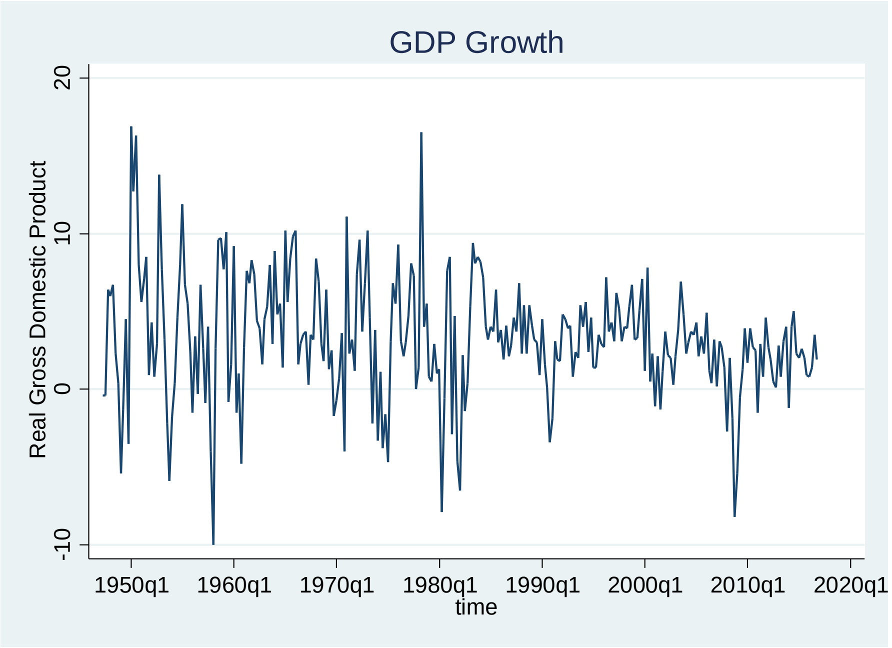
$$\hat{e}_t = y_t - \hat{\alpha} - \hat{\beta} y_{t-1}$$

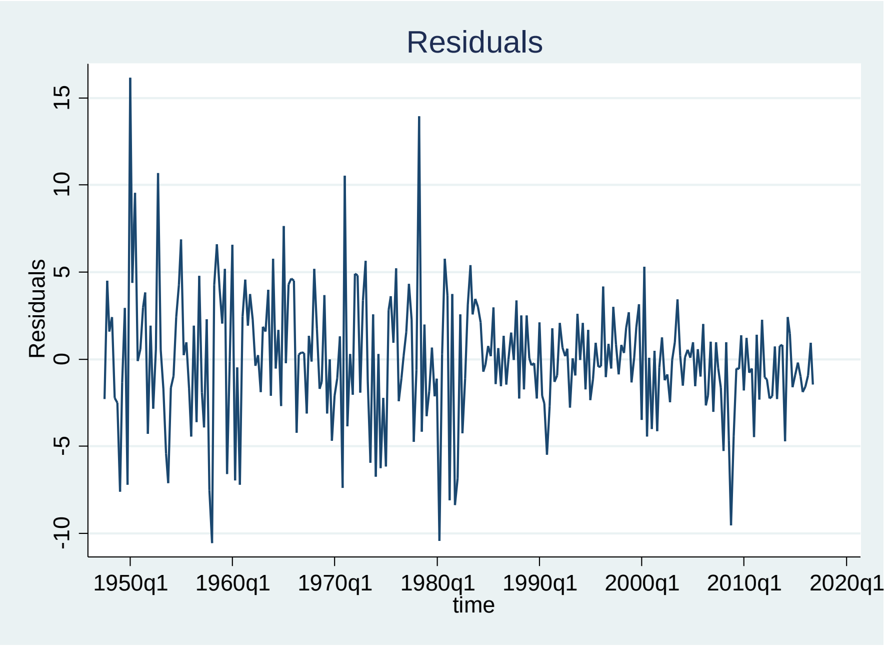
Unemployment Rate, 20-24 year olds





GDP Growth


GDP Growth
Real Gross Domestic Product
20
10
0
-10
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
time


Residuals
Residuals
15
10
5
0
-5
-10
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
time

One-Step-Ahead Intervals

- Normal Method
  - Assume forecast error is normally distributed
  - Forecast interval is point estimate, plus and minus the *standard deviation of forecast* multiplied by a normal percentile
  - For a 95% interval:

$$\hat{y}_{T+1|T} \pm \hat{\sigma} \cdot z_{.025} = \hat{y}_{T+1|T} \pm \hat{\sigma} \cdot 1.96$$

  - For a 90% interval

$$\hat{y}_{T+1|T} \pm \hat{\sigma} \cdot z_{.05} = \hat{y}_{T+1|T} \pm \hat{\sigma} \cdot 1.645$$

Standard Dev of Forecast

- The forecast variance is the variance of the error, plus the variance of the estimate
- The standard deviation of the forecast is its square root
- In most cases the major component of the forecast variance is the variance of the error
- Thus a simple method to estimate the standard deviation of the forecast is to use σ , "root mean squared error", from the regression estimate
- In STATA, for stdf you can use the command

  predict s, stdf
- The difference between the rmse and stdf grows when the number of regressors is large relative to the sample size.

Example: GDP

- The "Root MSE" is 3.63
- The stdf is 3.64, only slightly higher!

```
. regress gdp L.gdp

      Source |       SS           df       MS      Number of obs   =       278
-------------+----------------------------------   F(1, 276)       =     44.32
       Model |  585.172068         1  585.172068   Prob > F        =    0.0000
    Residual |  3643.98246       276   13.202835   R-squared       =    0.1384
-------------+----------------------------------   Adj R-squared   =    0.1352
       Total |  4229.15453       277  15.2677059   Root MSE        =    3.6336

------------------------------------------------------------------------------
         gdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         gdp |
         L1. |   .3714752   .0557984     6.66   0.000     .2616307    .4813197
             |
       _cons |   2.036509   .2826904     7.20   0.000     1.480006    2.593013
------------------------------------------------------------------------------
```

Forecast Interval Construction

- `tsappend, add(1)`
- `predict p`
- `predict s, stdf`
- `gen p1 = p – 1.645*s`
- `gen p2 = p + 1.645*s`
- `list time p1 p p2 s if time>tq(2016q4)`

	time	p1	p	p2	s
280.	2017q1	-3.246891	2.742312	8.731515	3.640853

- Point estimate = 2.7%
- Stdf = 3.64
- 90% Interval = [-3.2%, 8.7%]

Assignments

- Read Diebold through Chapter 7
- Problem Set # 5
  - Due Tuesday (2/21)
- Read Chapter 5 from *The Signal and the Noise*
  - Reading Reflection
  - Due Thursday (2/23)