Wisconsin Unemployment Rate Forecast Revisited

- Forecast in Lecture 1
- Wisconsin unemployment
- November 2016 was 4.1%
- Forecasts

	Point Forecast	**50% Interval Forecast**	**80% Interval Forecast**
December 2016	4.0%	(4.0%, 4.0%)	(3.95%, 4.05%)
January 2017	4.0%	(4.0%, 4.1%)	(3.9%, 4.1%)
February 2017	4.0%	(3.9%, 4.1%)	(3.8%, 4.2%)

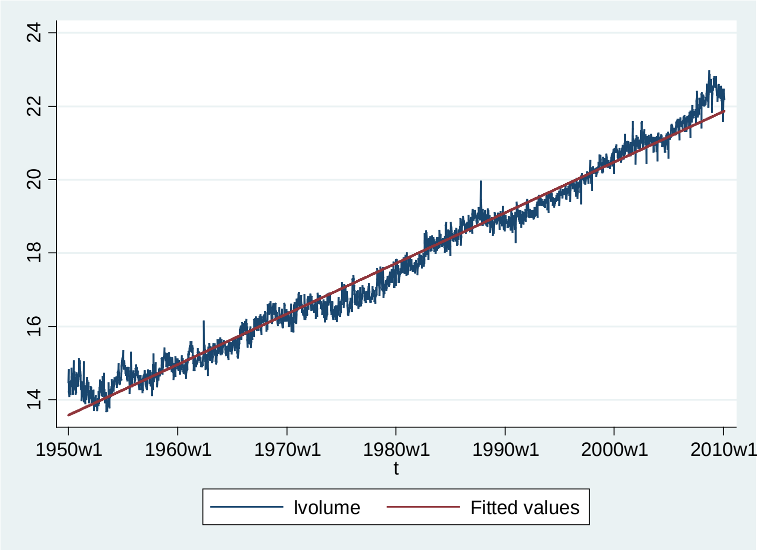
- Realizations
  - December 2016: 4.1%
  - January 2017: 3.9%

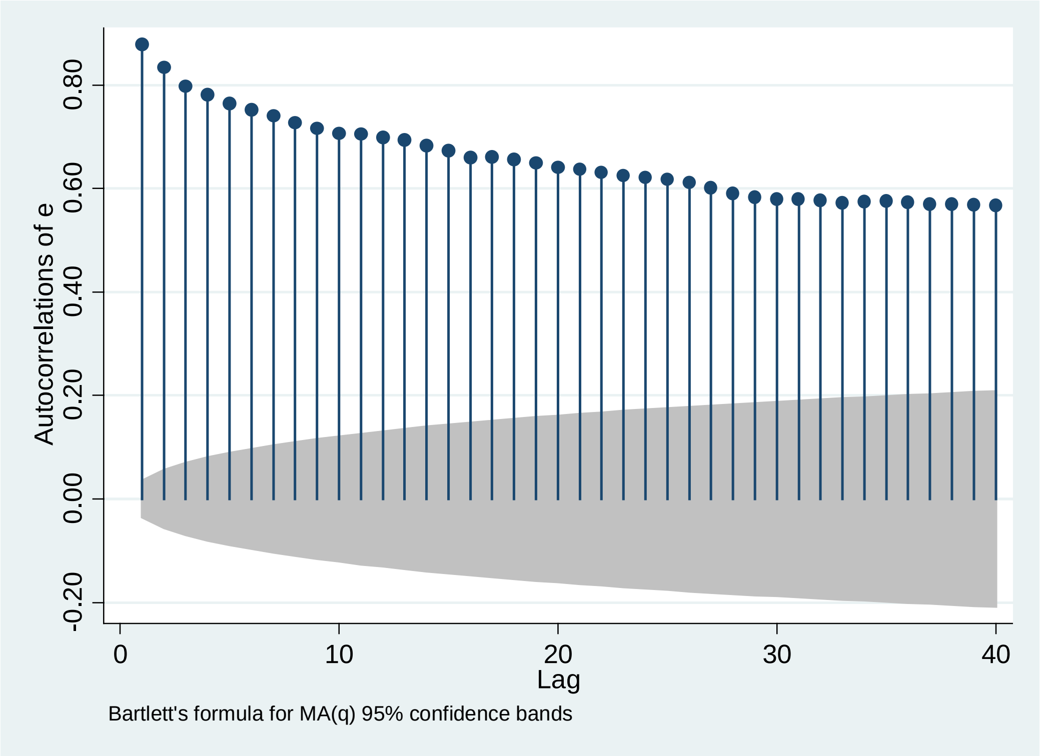
Regression with Correlated Errors

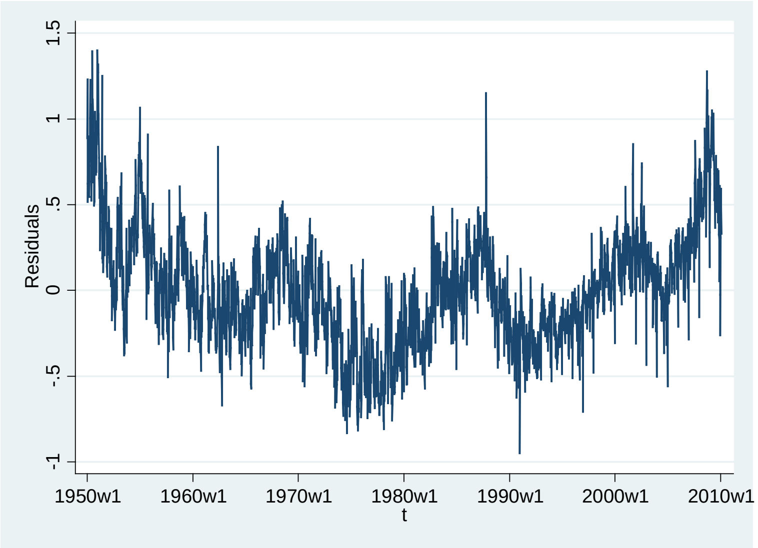
$$y_t = \alpha + \beta x_t + e_t$$

- In some regression models, the errors are correlated
  - Pure Trend Models
  - Pure Seasonality Models
- In these models the errors can be correlated
- Classical and robust standard errors are not appropriate

Example: Stock Volume







Least-Squares Variance Formula

Recall for $v_t = x_t e_t$

$$\operatorname{var}\left(\hat{\beta}\right) \overset{a}{\sim} \frac{\operatorname{var}\left(\sum_{t=1}^{T} v_t\right)}{\left[T \operatorname{var}(x_t)\right]^2}$$

When the v are uncorrelated

$$\operatorname{var}\left(\sum_{t=1}^{T} v_t\right) = \sum_{t=1}^{T} \operatorname{var}(v_t) = T \operatorname{var}(v_t)$$

$$\operatorname{var}\left(\hat{\beta}\right) \overset{a}{\sim} \frac{\operatorname{var}(v_t)}{T\left[\operatorname{var}(x_t)\right]^2}$$

General Formula

Define

$$f_T = \frac{\operatorname{var}\left(\sum_{t=1}^{T} v_t\right)}{T \operatorname{var}(v_t)}$$

When the v are uncorrelated f_T=1, otherwise not.

Then

$$\operatorname{var}\left(\hat{\beta}\right) \overset{a}{\sim} \frac{\operatorname{var}(x_t e_t)}{T\left[\operatorname{var}(x_t)\right]^2} f_T$$

Adjustment Factor

- The asymptotic variance of least-squares is the conventional, multiplied by an adjustment factor for the serial correlation

$$\text{var}\left(\hat{\beta}\right) \overset{a}{\sim} \frac{\text{var}\left(x_t e_t\right)}{T\left[\text{var}\left(x_t\right)\right]^2} f_T$$

Autocovariance of v

- We want a useful formula for

$$f_T = \frac{\operatorname{var}\left(\sum_{t=1}^{T} v_t\right)}{T \operatorname{var}(v_t)}$$

- Since $E(v_t)=0$, then

$$E\left(v_t^2\right) = \operatorname{var}(v_t)$$

$$E\left(v_t v_j\right) = \operatorname{cov}\left(v_t v_j\right) = \gamma(t-j)$$

the autocovariance of v_t

Variance of sum of correlated v

$$\operatorname{var}\left(\sum_{t=1}^{T} v_t\right) = E\left(\sum_{t=1}^{T} v_t\right)^2$$

$$= E\left(\sum_{t=1}^{T} v_t \sum_{j=1}^{T} v_j\right)$$

$$= \sum_{t=1}^{T} \sum_{j=1}^{T} E\left(v_t v_j\right)$$

$$= \sum_{t=1}^{T} \sum_{j=1}^{T} \gamma\left(t - j\right)$$

Adjustment Factor

$$f_T = \frac{\mathrm{var}\left(\sum_{t=1}^{T} v_t\right)}{T\,\mathrm{var}(v_t)} = \frac{1}{T}\sum_{t=1}^{T}\ \sum_{j=1}^{T}\rho(t-j)$$

- Where the ρ(t-j) are the autocorrelations of v_t

- This double sum is the sum of all the elements in the matrix

$$\begin{bmatrix} \rho(0) & \rho(1) & \rho(2) & \cdots & \rho(T-1) \\ \rho(1) & \rho(0) & \rho(1) & \cdots & \rho(T-2) \\ \rho(2) & \rho(1) & \rho(0) & \cdots & \rho(T-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(T-1) & \rho(T-2) & \rho(T-3) & \cdots & \rho(0) \end{bmatrix}$$

- There are
  - *T* of the ρ(0)
  - 2*(T-1)* of the ρ(1)
  - 2*(T-2)* of the ρ(2)
  - …

$$T + \sum_{j=1}^{T-1} 2(T-j)\rho(j)$$

Adjustment Factor

- Dividing by T

$$f_T = \frac{1}{T}\sum_{t=1}^{T}\ \sum_{j=1}^{T}\rho(t-j)$$

$$= 1 + \sum_{j=1}^{T-1} 2\left(\frac{T-j}{T}\right)\rho(j)$$

- If T is large

$$f_T \rightarrow 1 + 2\sum_{j=1}^{\infty}\rho(j) = f$$

Summary: Least-Squares Variance

- When the errors are correlated

$$\operatorname{var}\left(\hat{\beta}\right) \overset{a}{\sim} \frac{\operatorname{var}\left(x_t e_t\right)}{T\left[\operatorname{var}\left(x_t\right)\right]^2} f$$

$$f = 1 + 2\sum_{j=1}^{\infty} \rho\left(j\right)$$

- The conventional formula is multiplied by an adjustment for autocorrelation

HAC Estimation

- Estimation of f
  - For variances and standard errors under autocorrelation
- Called heteroskedasticity and autocorrelation consistent (HAC) variance estimation
- Multiply conventional variance estimates by estimates of f

HAC Estimation

- The adjustment is

$$f = 1 + 2\sum_{j=1}^{\infty} \rho(j)$$

where ρ(j) are the autocorrelations of $v_t = x_t e_t$

- Estimate ρ(j) by sample autocorrelations using least-squares residuals
- But in a sample of length T we cannot estimate all autocorrelations well

Unweighted HAC Estimator

- For some **truncation parameter** m,

$$\hat{f} = 1 + 2\sum_{j=1}^{m} \hat{\rho}(j)$$

- Original proposal
  - L. Hansen, Hodrick (1978)
  - Hal White (1982)

Lars Hansen

- Professor Lars Hansen, U Chicago
- Invented Generalized Method of Moments, the leading estimation method for applied econometrics
- Introduced unweighted HAC estimator for multi-step regression models
- 2013 Nobel Prize in economics

Deficiencies of Unweighted Estimator

$$\hat{f} = 1 + 2\sum_{j=1}^{m} \hat{\rho}(j)$$

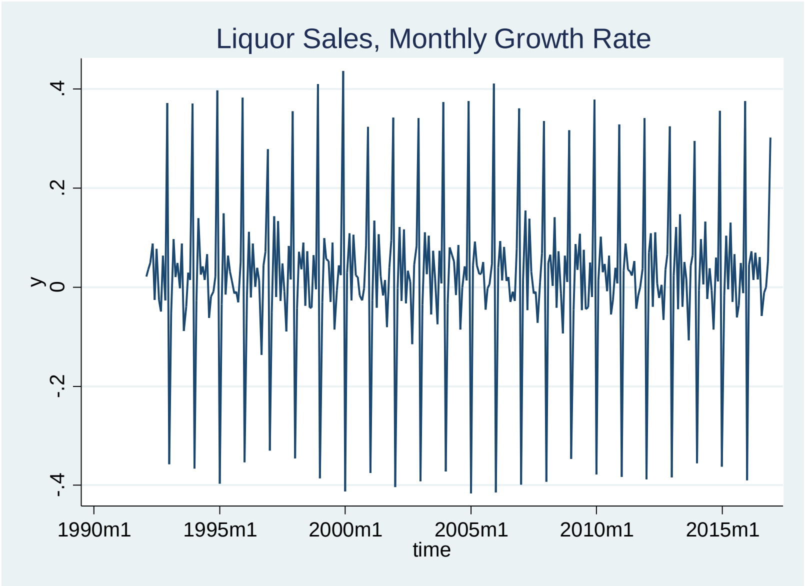
- The estimator is not smooth in the truncation parameter
- The sample estimate can be negative
- Example of negative estimate when m=1

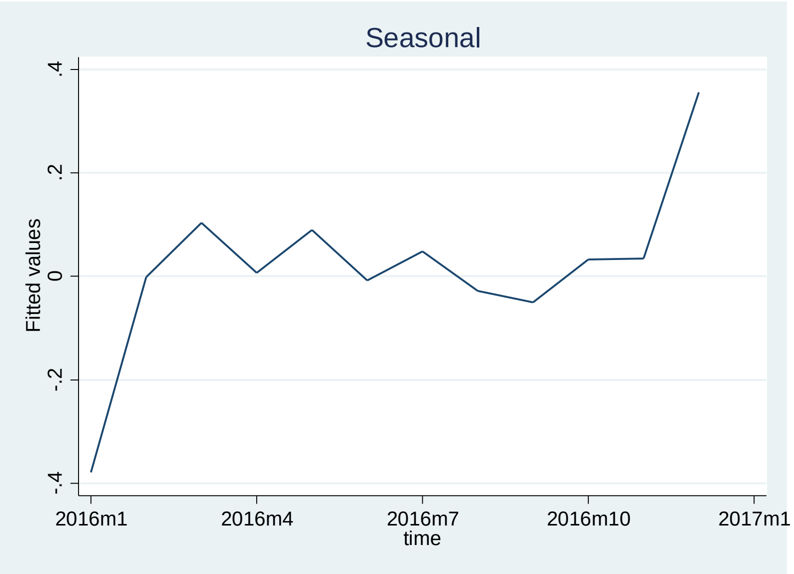
$$\hat{f} = 1 + 2\hat{\rho}(1)$$

  - Negative if $\rho(1)<-1/2$

Example: Liquor Sales

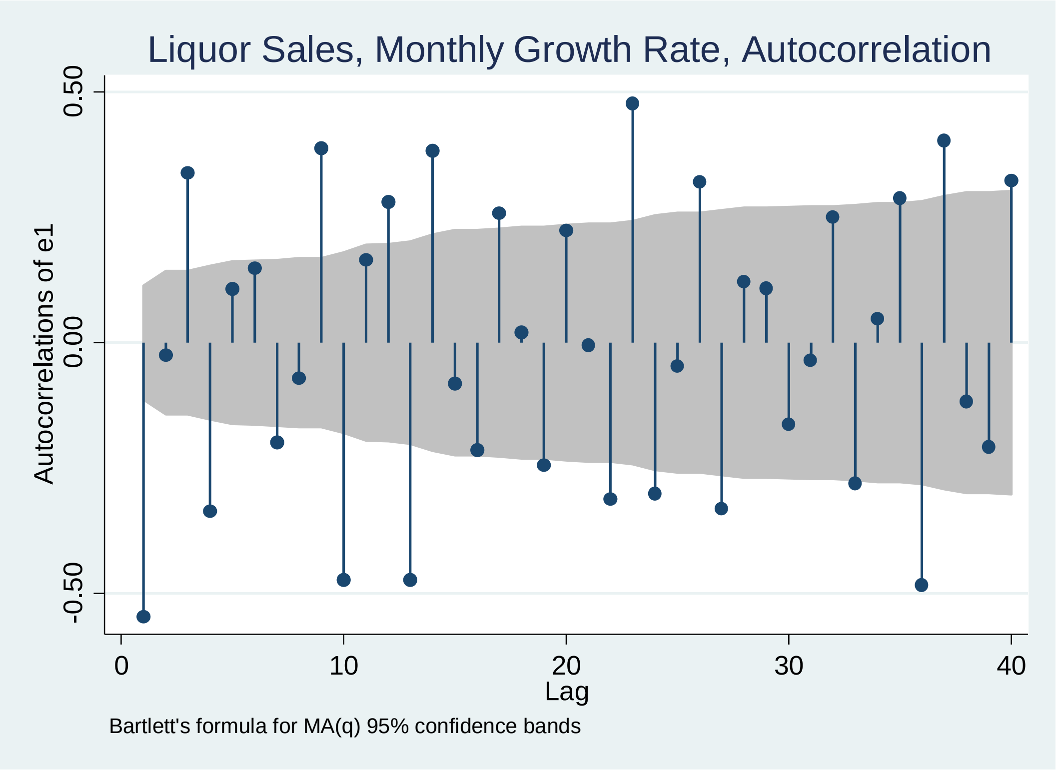
- Take monthly growth rates
- Regress on Seasonal Dummies to obtain seasonal pattern





Autocorrelation of Residual

- The first autocorrelation is less than -1/2



Weighted HAC Estimator

$$\hat{f} = 1 + 2\sum_{j=1}^{m}\left(\frac{m-j}{m}\right)\hat{\rho}(j)$$

- Called Newey-West variance estimator
  - Whitney Newey, Ken West (1987)
- This weighted estimator is always positive
- Smoothly changes in truncation parameter *m*

Whitney Newey and Ken West

- Professor Whitney Newey, MIT
  - Leading econometric theorist
- Professor Ken West, Wisconsin
  - Macroeconomist & econometrician
  - Forecast evaluation and comparison
- Joint paper in 1987
  - Weighted HAC estimator
  - One of the most referenced papers in econometrics

Computation

- In STATA, replace **regress** command with **newey** command

  .newey y x, lag(m)

- You supply the truncation parameter "m"
- Similar to regression with robust standard errors
- These are identical

  .newey y x, lag(0)

  .reg y x, r

Example: Liquor Sales

```
. regress y b12.m, r
```

Linear regression

Number of obs	=	299
F(11, 287)	=	839.72
Prob > F	=	0.0000
R-squared	=	0.9575
Root MSE	=	.0333

y	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-.7345193	.0089466	-82.10	0.000	-.7521286	-.7169101
2	-.3569828	.0105029	-33.99	0.000	-.3776554	-.3363103
3	-.2518665	.009429	-26.71	0.000	-.2704252	-.2333078
4	-.3484662	.0101598	-34.30	0.000	-.3684634	-.328469
5	-.2652584	.010626	-24.96	0.000	-.2861732	-.2443437
6	-.363618	.009398	-38.69	0.000	-.3821157	-.3451202
7	-.3071406	.0099995	-30.72	0.000	-.3268222	-.287459
8	-.3839307	.0101597	-37.79	0.000	-.4039276	-.3639338
9	-.4055328	.010923	-37.13	0.000	-.4270322	-.3840334
10	-.3228669	.0098322	-32.84	0.000	-.3422193	-.3035145
11	-.3209942	.0103776	-30.93	0.000	-.3414201	-.3005683
_cons	.3554102	.0076051	46.73	0.000	.3404414	.370379

With Newey-West standard errors

```
. newey y b12.m, lag(12)

Regression with Newey-West standard errors      Number of obs     =        299
maximum lag: 12                                 F( 11,       287) =    1405.25
                                                Prob > F          =     0.0000
```

y	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-.7345193	.0107666	-68.22	0.000	-.7557109	-.7133278
2	-.3569828	.0105326	-33.89	0.000	-.3777138	-.3362518
3	-.2518665	.0098502	-25.57	0.000	-.2712543	-.2324787
4	-.3484662	.0083028	-41.97	0.000	-.3648082	-.3321241
5	-.2652584	.0131356	-20.19	0.000	-.2911129	-.239404
6	-.363618	.0087986	-41.33	0.000	-.3809359	-.3463
7	-.3071406	.0088429	-34.73	0.000	-.3245457	-.2897354
8	-.3839307	.0125861	-30.50	0.000	-.4087035	-.3591579
9	-.4055328	.0071964	-56.35	0.000	-.4196972	-.3913684
10	-.3228669	.0106235	-30.39	0.000	-.3437767	-.3019571
11	-.3209942	.0124128	-25.86	0.000	-.3454259	-.2965626
_cons	.3554102	.0076822	46.26	0.000	.3402897	.3705307

Truncation Parameter

- *m* should be large when autocorrelation is large
- Sophistical data-dependent methods to pick m have been developed, but are not in STATA
- Stock-Watson default (explanatory x's)

$$m = 0.75T^{1/3}$$

- Trend/Seasonal default

$$m = 1.4T^{1/3}$$

Derivation of Defaults

- Due to Donald Andrews (1991)
- The optimal m minimizes the mean-squared error of the estimate of f
- When v_t is an AR(1) with coefficient ρ, Andrews found the optimal m is

$$m = CT^{1/3}$$

$$C = \left(\frac{6\rho^2}{(1-\rho^2)^2} \right)^{1/3}$$

Donald Andrews

- Leading econometric theorist
- Contributions to time-series
  - Optimal selection of truncation parameter
  - Tests for structural change

Default Values

$$m = CT^{1/3}$$

$$C = \left(\frac{6\rho^2}{\left(1-\rho^2\right)^2} \right)^{1/3}$$

- Stock-Watson
  - If both x_t and e_t are AR(1) with coef ½, then $v_t = x_t e_t$ has AR(1) coefficient ρ=.25. Plug this in, and C=.75
- Trend-Seasonal
  - If x_t is trend and/or seasonal and e_t are AR(1) with coef ½, then $v_t = x_t e_t$ has AR(1) coefficient ρ=.5. Plug this in, and C=1.4

Liquor Sales again

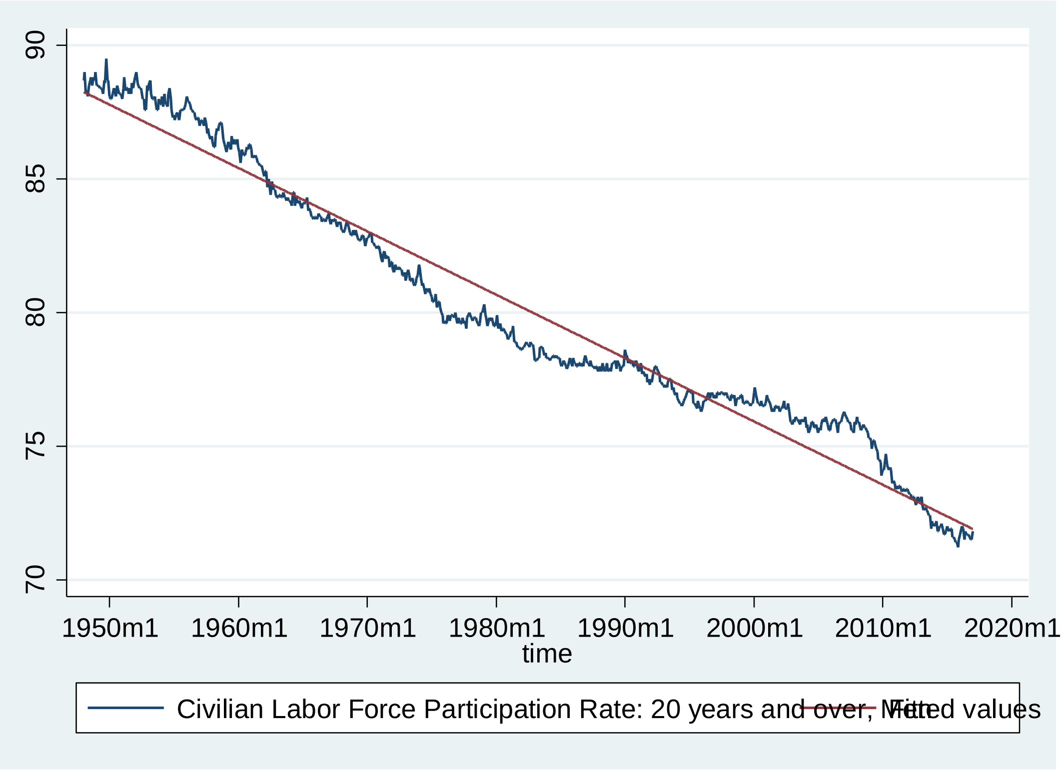
```
. dis 1.4*e(N)^(1/3)
9.3616363

. newey y b12.m, lag(9)

Regression with Newey-West standard errors      Number of obs     =        299
maximum lag: 9                                  F( 11,       287) =    1228.72
                                                Prob > F          =     0.0000
```

y	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-.7345193	.0106633	-68.88	0.000	-.7555075	-.7135312
2	-.3569828	.0098266	-36.33	0.000	-.3763242	-.3376415
3	-.2518665	.0096794	-26.02	0.000	-.2709181	-.2328149
4	-.3484662	.0088933	-39.18	0.000	-.3659705	-.3309619
5	-.2652584	.0124058	-21.38	0.000	-.2896763	-.2408406
6	-.363618	.0088542	-41.07	0.000	-.3810455	-.3461905
7	-.3071406	.0089587	-34.28	0.000	-.3247736	-.2895076
8	-.3839307	.0119737	-32.06	0.000	-.4074981	-.3603633
9	-.4055328	.0080752	-50.22	0.000	-.4214269	-.3896387
10	-.3228669	.0098372	-32.82	0.000	-.3422291	-.3035048
11	-.3209942	.0123806	-25.93	0.000	-.3453624	-.296626
_cons	.3554102	.0076051	46.73	0.000	.3404414	.370379

Example: Men's Labor Force Participation Rate, Trend Model



```
. reg men  time, r

Linear regression                                      Number of obs     =        829
                                                       F(1, 827)         =   24055.36
                                                       Prob > F          =     0.0000
                                                       R-squared         =     0.9634
                                                       Root MSE          =     .92307

------------------------------------------------------------------------------
             |               Robust
         men |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |  -.0197667   .0001274  -155.10   0.000    -.0200168   -.0195165
       _cons |   85.41493   .0467774  1825.99   0.000     85.32311    85.50675
------------------------------------------------------------------------------


. dis 1.4*e(N)^(1/3)
13.151629

. newey men time, lag(13)

Regression with Newey-West standard errors      Number of obs     =        829
maximum lag: 13                                 F(  1,       827) =    1849.84
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
         men |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        time |  -.0197667   .0004596   -43.01   0.000    -.0206688   -.0188646
       _cons |   85.41493   .1684105   507.18   0.000     85.08437    85.74549
------------------------------------------------------------------------------
```

Summary

- In one-step-ahead forecast regressions
- If the errors are serially uncorrelated
  - Use Robust standard errors
    - reg with r option
- If the errors are correlated
  - Use Newey-West standard errors
    - newey y x, lag(m)
  - In pure trend or seasonality models
    - Set $m=1.4T^{1/3}$
  - In dynamic regression
    - Set $m=.75T^{1/3}$

h-step-ahead forecasts

- In the AR(1) Model

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

- The optimal h-step forecasting regression takes the form

$$y_t = \alpha + \beta^h y_{t-h} + u_t$$

$$u_t = e_t + \beta e_{t-1} + \beta^2 e_{t-2} + \cdots + \beta^{h-1} e_{t-h+1}$$

- The error u_t is a correlated MA(h-1)
  - Unless β=0

h-step-ahead models

- In any h-step model

$$y_t = \alpha + \beta y_{t-h} + u_t$$

  the variable $v_t = y_{t-h} e_t$ is generally serially correlated

- Generally MA(h-1)
- Correct adjustment term

$$f = 1 + 2\sum_{j=1}^{h-1} \rho(j)$$

Newey-West Standard Errors

- Standard errors can be estimated using the Newey-West method
- Truncation parameter set to forecast horizon
  - m=h

$$\hat{f} = 1 + 2\sum_{j=1}^{h-1}\left(\frac{h-j}{h}\right)\hat{\rho}(j)$$

Example: Unemployment Rate

- 12-month-ahead forecast with 4 AR lags
  - Robust standard errors:

```
. reg ur L(12/15).ur, r

Linear regression                               Number of obs     =        814
                                                F(4, 809)         =     210.69
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5282
                                                Root MSE          =     1.7825
```

ur	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
ur						
L12.	.9525096	.1385007	6.88	0.000	.6806465	1.224373
L13.	.2404673	.17578	1.37	0.172	-.1045713	.5855059
L14.	-.1563478	.1646489	-0.95	0.343	-.4795374	.1668417
L15.	-.3427953	.1337901	-2.56	0.011	-.605412	-.0801785
_cons	2.906192	.2436478	11.93	0.000	2.427935	3.384448

Example: Unemployment Rate

- Newey-West standard errors:
- Standard errors on lag 13 and 14 decrease by half
- Standard error on constant more than doubles

```
. newey ur L(12/15).ur, lag(12)

Regression with Newey-West standard errors      Number of obs     =        814
maximum lag: 12                                 F(  4,       809) =      32.83
                                                Prob > F          =     0.0000
```

ur	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
ur						
L12.	.9525096	.1633674	5.83	0.000	.6318357	1.273183
L13.	.2404673	.0826344	2.91	0.004	.0782642	.4026705
L14.	-.1563478	.0659848	-2.37	0.018	-.2858694	-.0268263
L15.	-.3427953	.158594	-2.16	0.031	-.6540996	-.0314909
_cons	2.906192	.6652979	4.37	0.000	1.600278	4.212105

newey and forecasting

- **predict** works after **newey** command, but not with **stdf** option
- **newey** not appropriate for iterated forecasts
- Use **newey** to assess model and examine coefficients
- Use **reg** to compute out-of-sample forecast intervals

Summary

- In one-step-ahead forecast regressions
  - If the errors are serially uncorrelated, use r option
  - If the errors are correlated
    - Use **newey** for standard errors
      - In pure trend or seasonality models set $m=1.4T^{1/3}$
      - In dynamic regression set $m=.75T^{1/3}n$
    - Use **reg** and **predict sf, stdf** for forecast intervals, or iterated forecasts with **forecast**
- In h-step-ahead forecast regressions
  - Use **newey** with m=h for standard errors
  - Use **reg** and **predict sf, stdf** for forecast intervals

Joint Tests

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

- How do we assess if a subset of coefficients are jointly zero? Example: 3rd+4th lags

```
. reg gdp L(1/4).gdp, r

Linear regression                               Number of obs     =        275
                                                F(4, 270)         =       9.86
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1597
                                                Root MSE          =     3.6147
```

gdp	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
gdp						
L1.	.3363865	.0743514	4.52	0.000	.1900044	.4827686
L2.	.1422228	.0834467	1.70	0.089	-.0220661	.3065118
L3.	-.0682526	.0721168	-0.95	0.345	-.2102354	.0737302
L4.	-.0728112	.0741333	-0.98	0.327	-.218764	.0731416
_cons	2.131395	.4227888	5.04	0.000	1.299012	2.963777

Joint Hypothesis

- This is a joint test of

$$\beta_3 = 0$$

$$\beta_4 = 0$$

- This can be done with an “F test”
- In STATA, after **regress** (**reg**) or **newey**

  .test L3.gdp L4.gdp

- List variables whose coefficients are tested for zero.

Joint Tests

- "F test" named after R.A. Fisher
  - (1890-1962)
  - A founder of modern statistical theory
- Modern form known as a "Wald test", named after Abraham Wald (1902-1950)
  - Early contributor to econometrics

F test computation

```
. test L3.gdp L4.gdp

 ( 1)  L3.gdp = 0
 ( 2)  L4.gdp = 0

       F(  2,   270) =    1.09
            Prob > F =    0.3365
```

- You need to list each variable separately
- STATA describes the hypothesis
- The value of "F" is the F-statistic
- "Prob>F" is the p-value
  - Small p-values cause rejection of hypothesis of zero coefficients
  - Conventionally, reject hypothesis if p-value < 0.05

Example: 2-step-ahead GDP AR(4)

```
. newey gdp L(2/5).gdp, lag(2)

Regression with Newey-West standard errors      Number of obs     =        274
maximum lag: 2                                  F(  4,        269) =       3.33
                                                Prob > F          =     0.0110
```

gdp	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
gdp						
L2.	.2468265	.0740484	3.33	0.001	.1010384	.3926147
L3.	-.0100614	.0696302	-0.14	0.885	-.1471508	.127028
L4.	-.0710521	.075953	-0.94	0.350	-.2205902	.0784859
L5.	-.0933157	.064904	-1.44	0.152	-.2211001	.0344687
_cons	2.981623	.4950533	6.02	0.000	2.006952	3.956295

```
. test L3.gdp L4.gdp L5.gdp

 ( 1)  L3.gdp = 0
 ( 2)  L4.gdp = 0
 ( 3)  L5.gdp = 0

       F(  3,   269) =    1.13
            Prob > F =    0.3354
```

Testing after Estimation

- The commands **predict** and **test** are applied to the most recently estimated model
- The command test uses the standard error method specified by the estimation command
  - **reg y x** : classical F test
  - **reg r x, r**: heteroskedasticity-robust F test
  - **newey y x, lag(m)**: correlation-robust F test
    - (The robust tests are actually Wald statistics)

Assignments

- Read Wooldridge Chapter 12.1 and 12.5
  - An electronic copy is in files at Learn@UW
- Forecasting Project
  - Project Description (3/28)
- Read Chapter 8 from *The Signal and the Noise*
  - Reading Reflection
  - Thursday (3/30)