

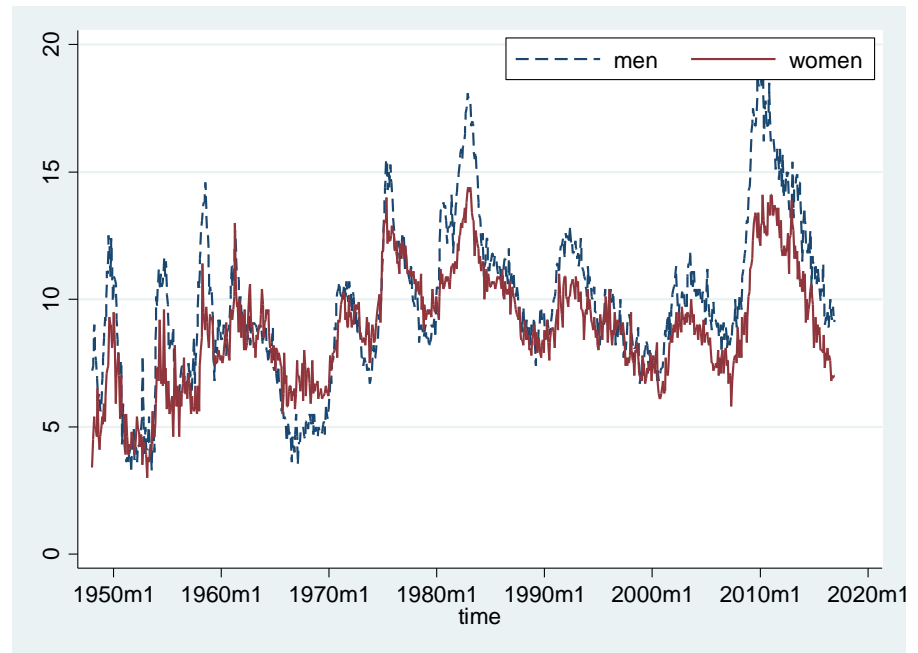
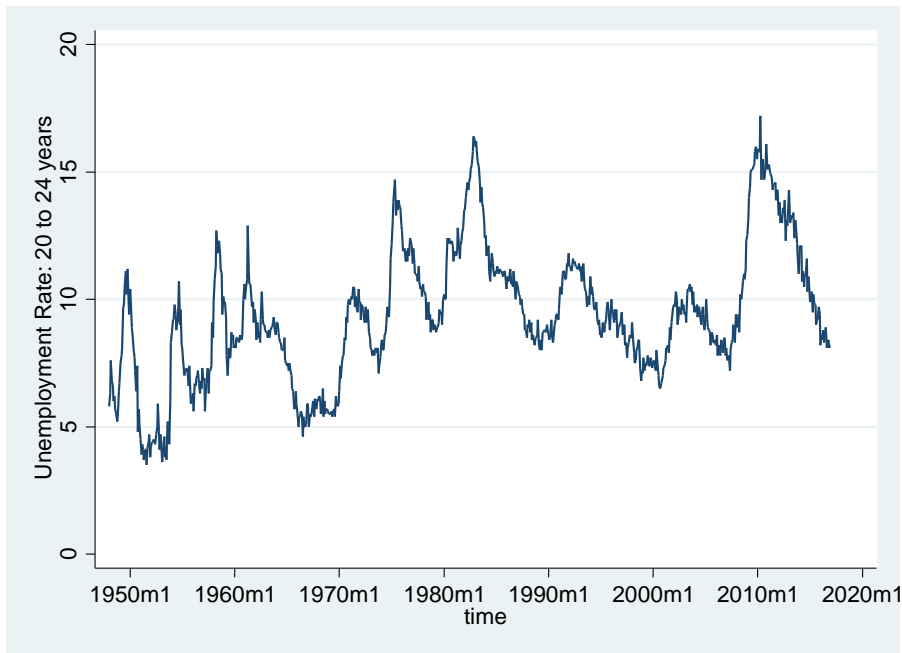
Unemployment Rate Example

- Find unemployment rates for men and women in your age bracket
- Go to FRED
- Categories/Population/Current Population Survey/Unemployment Rate
- Release Tables/Selected unemployment indicators
- Unemployment Rate
- For more detail, go to bottom: Table A-10
 - 20 to 24 years
 - LNS14000036
 - Men, 20 to 24 years
 - LNS14000037
 - Women, 20 to 24 years
 - LNS14000038

Unemployment Rate Example

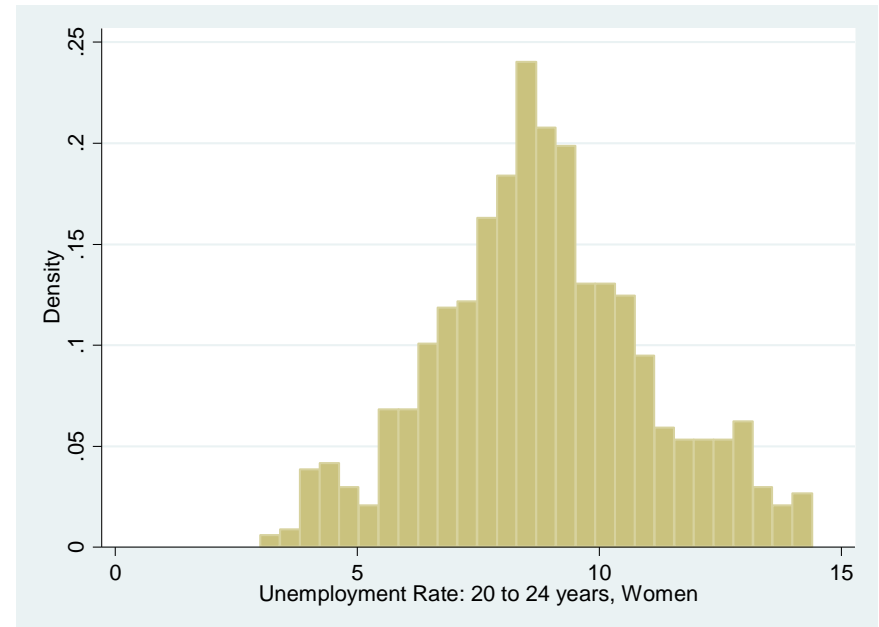
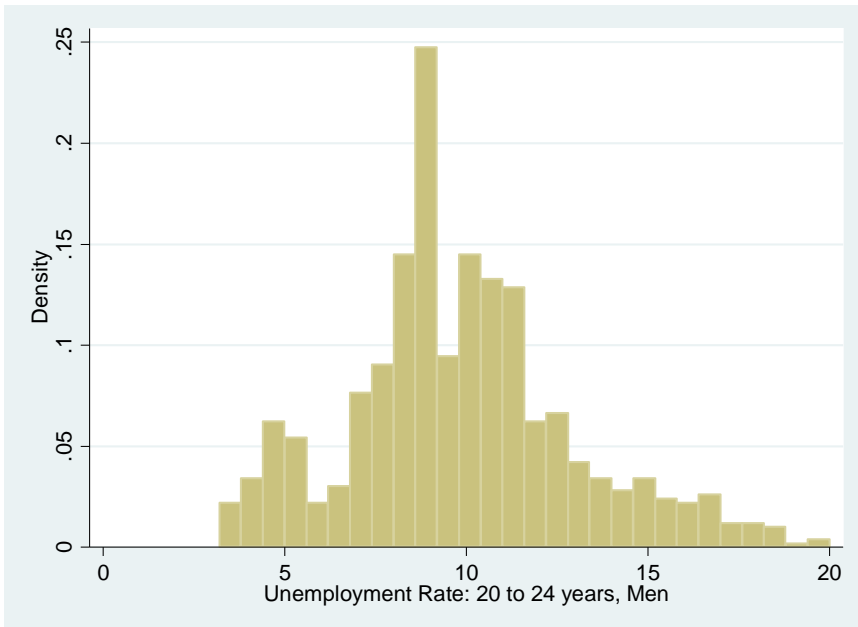
- Stata
 - `clear all`
 - `freduse LNS14000036 LNS14000037
LNS14000038`
 - `rename LNS14000036 ur`
 - `rename LNS14000037 urmen`
 - `rename LNS14000038 urwomen`
 - `tsmktim time, start(1948m1)`
 - `tsset time`
 - `tsline ur`
 - `tsline urmen urwomen`

U.S. Monthly Unemployment Rates, Ages 20-24 Total, and by Sex



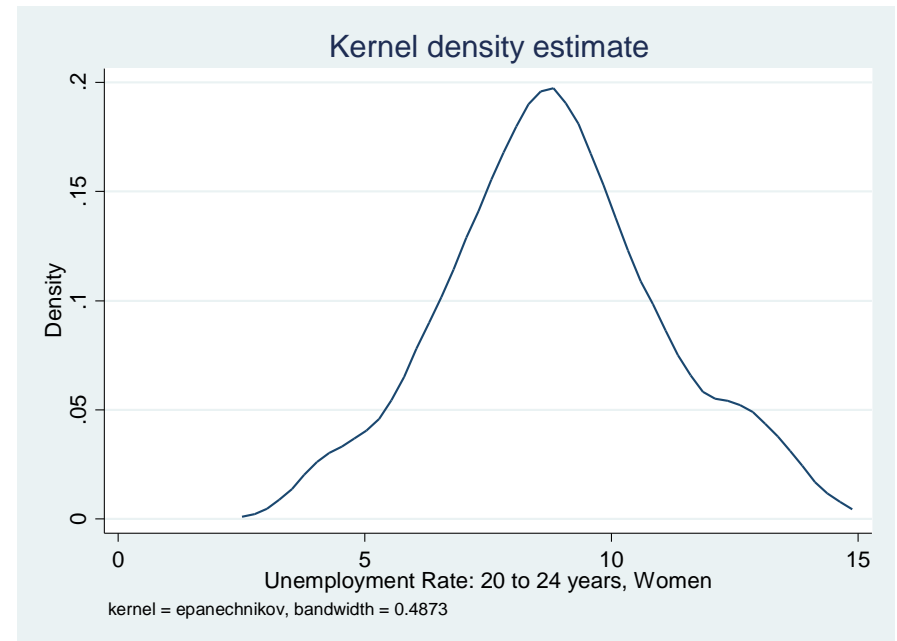
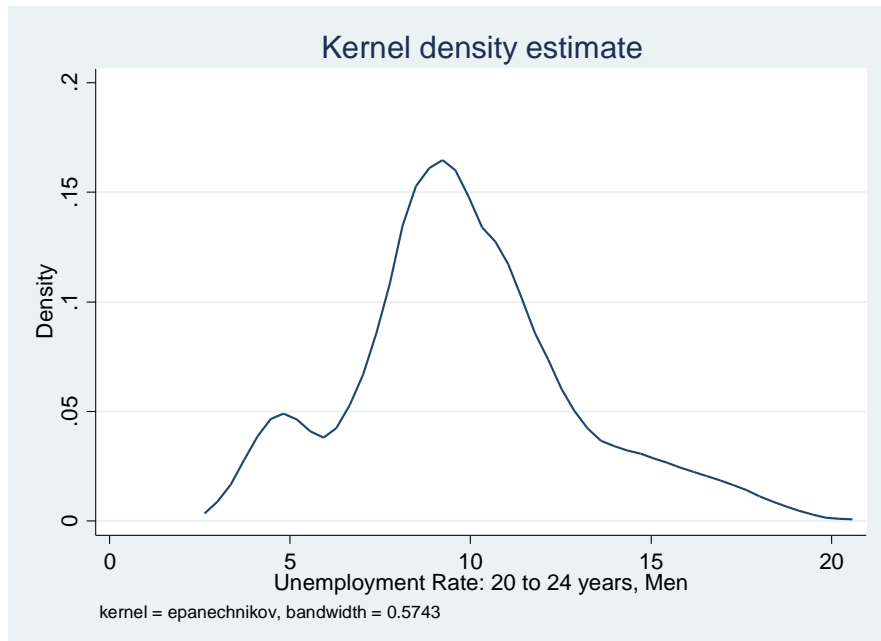
Histogram

- Sometimes a histogram is useful to view
 - histogram urmen



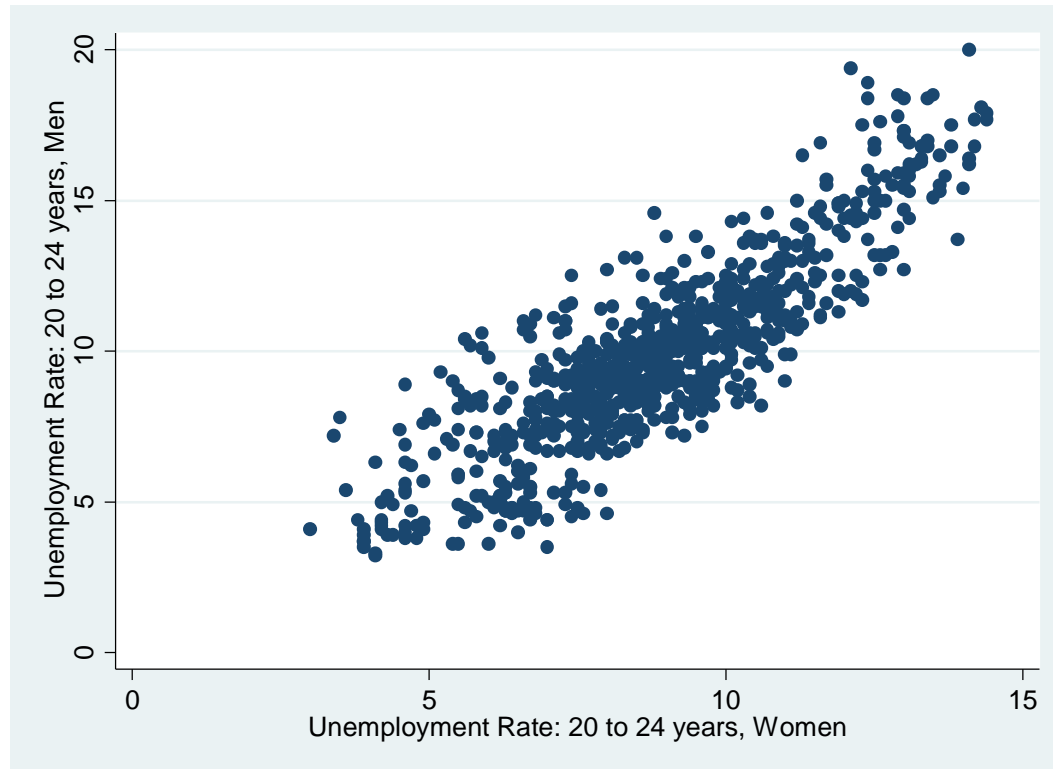
Density Estimate

- A density is a smoothed version of a histogram
 - `kdensity urmen`



Scatter Plot

- Sometimes a scatter plot can be insightful
 - scatter urmen urwomen



Lags and Leads

- We will often talk about lags and leads
- The first lag of y_t is written y_{t-1}
 - It is the observation from the previous period
 - For example, the lag of November is October
- The second lag is y_{t-2} , the k 'th lag is y_{t-k}
- The first lead of y_t is y_{t+1}
- The k 'th lead is y_{t+k}

Time Series Samples

- A historical sample is a set of observations in contiguous time, written as

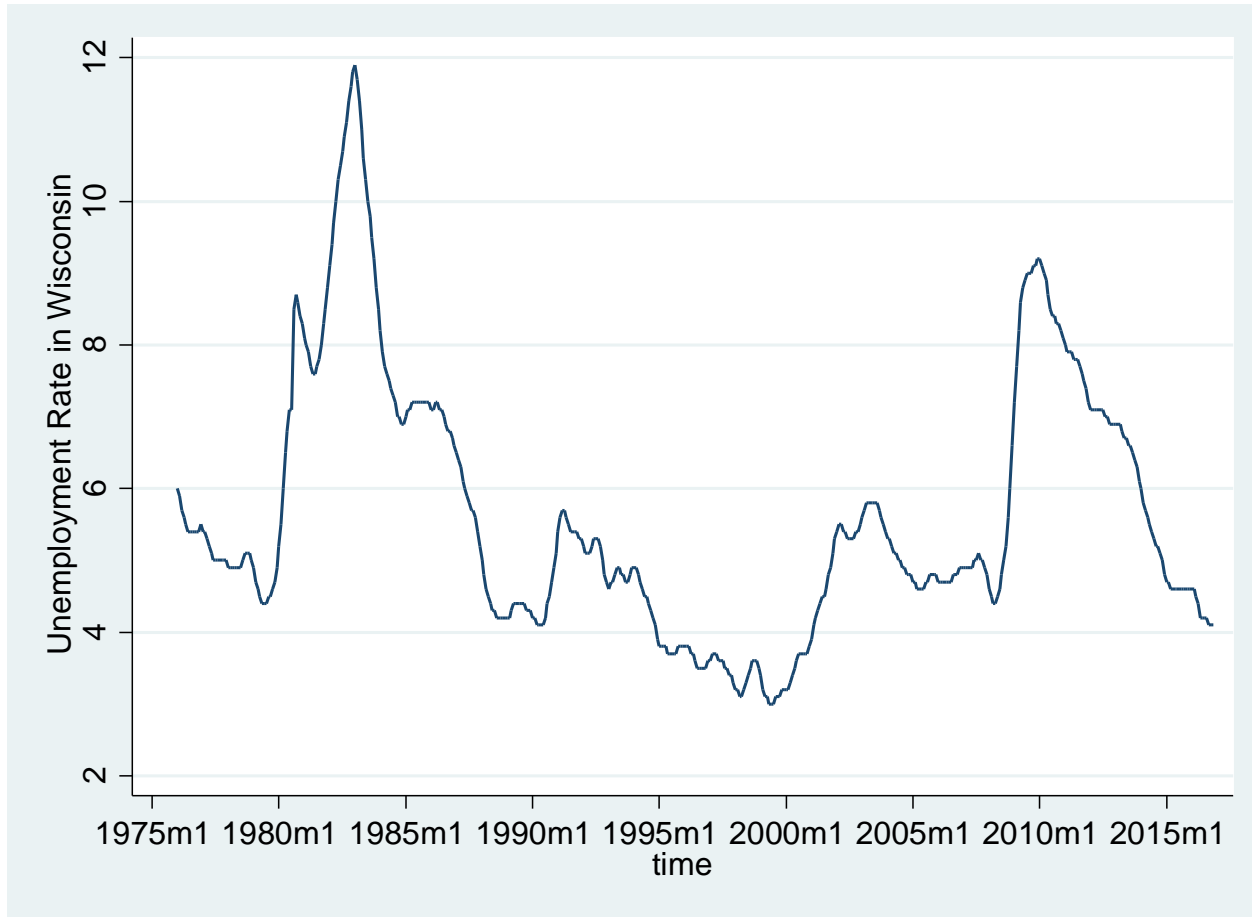
$$\{y_1, y_2, \dots, y_T\}$$

- T is the number of observations in-sample.
- The number of observations does not equal the number of years, unless the frequency is annual

Examples

- Number of Unemployed
- Unemployment Rate
- GDP
- Real GDP
- Price Level
- Inflation Rate
- New Housing Starts

Wisconsin Unemployment Rate



Forecast Period

- In-sample observations: $\{y_1, y_2, \dots, y_T\}$
- Out-of-sample period: $\{y_{T+1}, y_{T+2}, \dots, y_{T+h}\}$
- h is called the forecast horizon

Forecast Notation

- We denoted \hat{y} as the point forecast for y .
- This suggests \hat{y}_{T+h} as the point forecast for y_{T+h} .
- But is not enough. While it is the forecast for the time series at time period $T+h$, it is not clear when the forecast is made.
 - At time period T
 - At time period $T+1$
 - At time period $T+h-1$

Notation

- We will use the notation

$$\hat{y}_{t+h|t}$$

to refer to the forecast of y_{t+h} made at time t .

- Thus $\hat{y}_{T+h|T}$, $\hat{y}_{T+h|T+1}$, $\hat{y}_{T+h|T+2}$, etc. are the sequence of forecasts of y_{T+h} made in time periods T , $T+1$, $T+2$, etc.

Notation

- Similarly, the forecast distribution and density for y_{t+h} made at time t will be written as

$$F_{t+h|t}(y)$$

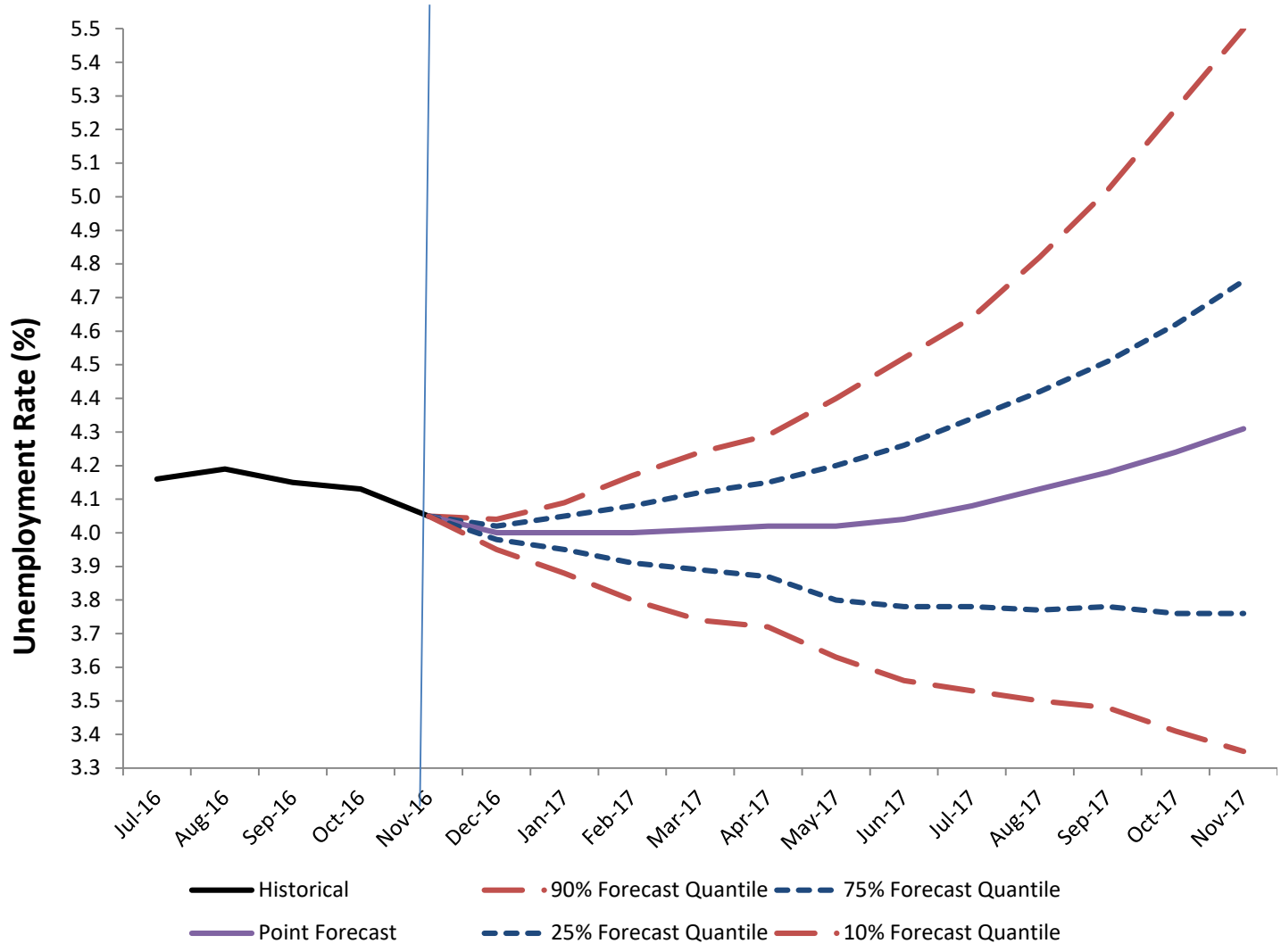
and

$$f_{t+h|t}(y)$$

Extrapolative Forecasts and Fan Charts

- At time T , make a sequence of forecasts for time periods $T+1, T+2, T+3, \dots, T+h$
- Point forecasts: $\hat{y}_{T+1|T}, \hat{y}_{T+2|T}, \dots, \hat{y}_{T+h|T}$
- Add interval forecasts
- Plot over forecast horizon
- This is called a *fan chart*.
- The intervals tend to fan out with the forecast horizon.

Wisconsin Unemployment Rate



Extrapolative Forecasts

	Point Forecast	50% Interval Forecast	80% Interval Forecast
2016:12	4.0%	(4.05%, 4.0%)	(4.0%, 4.0%)
2017: 1	4.0%	(4.0%, 4.1%)	(3.9%, 4.1%)
2017: 2	4.0%	(3.9%, 4.1%)	(3.8%, 4.2%)
2017: 3	4.0%	(3.9%, 4.1%)	(3.7%, 4.2%)
2017: 4	4.0%	(3.9%, 4.2%)	(3.7%, 4.3%)
2017: 5	4.0%	(3.8%, 4.2%)	(3.6%, 4.4%)
2017: 6	4.0%	(3.8%, 4.3%)	(3.6%, 4.5%)
2017: 7	4.1%	(3.8%, 4.3%)	(3.5%, 4.6%)
2017: 8	4.1%	(3.8%, 4.4%)	(3.4%, 4.8%)
2017: 9	4.2%	(3.8%, 4.5%)	(3.5%, 5.0%)
2017:10	4.2%	(3.8%, 4.6%)	(3.4%, 5.3%)
2017:11	4.3%	(3.8%, 4.8%)	(3.4%, 5.5%)

Information Set

- To forecast y_{t+h} at time t we use relevant information.
- Most information is values of other economic variables.
- Those observed up to time t .
- This includes previous values of the variable y_t .
- It can also include other relevant variables.
- All this information is the *Information Set*, written as Ω_t .
- For example $\Omega_t = \{y_1, y_2, y_3, \dots, y_t\}$ is the set of previous values.
- $\Omega_t = \{y_1, x_1, y_2, x_2, y_3, x_3, \dots, y_t, x_t\}$ includes another variable x_t

Conditional Mean Forecasts

- Under squared error loss, the optimal *unconditional* point forecast is the mean Ey

$$E(y) = \int yf(y)dy$$

- Under squared error loss, the optimal *conditional* point forecast is the *conditional mean*

$$E(y | \Omega_t) = \int yf(y | \Omega_t)dy$$

where $f(y | \Omega_t)$ is the conditional density of y given Ω_t

Conditional Mean

- The unconditional mean $E(y)$ is the average value of y in the entire population
- The conditional mean $E(y|x)$ given a set of variables x is the average value of y for the subpopulation with the variables x .
- The conditional mean $E(y|\Omega_t)$ given an information set is the average value of y in the subhistory with the previous history Ω_t .

Conditional Mean

- What is $E(y | \Omega_t)$?
- When Ω_t is discrete, then $E(y | \Omega_t)$ is simply the mean in the subpopulation.
- For example, the average wage among white male college graduates.
- When Ω_t large and/or is continuously distributed, this definition does not work.

Conditional Mean

- Let $f(y, \Omega_t)$ denote the joint density of (y, Ω_t) . Then the conditional density of y given Ω_t is

$$f(y | \Omega_t) = \frac{f(y, \Omega_t)}{f(\Omega_t)}$$

- It is the distribution of y holding Ω_t fixed.
- It is a slice of the joint density
- Then the conditional mean is

$$E(y | \Omega_t) = \int y f(y | \Omega_t) dy$$

Forecast Intervals

- Forecast intervals are constructed from the conditional distribution $F(y | \Omega_t)$.
- The endpoints are the conditional quantiles.
- Definition: The α 'th conditional quantile of y given Ω_t is the number $q_\alpha(\Omega_t)$ which satisfies

$$\alpha = F(q_\alpha(\Omega_t) | \Omega_t)$$

- It looks more complicated, but it is identical with the case with no Ω_t .

When does Conditioning Help?

- It helps if y and Ω_t are *dependent* or *correlated*.
- If y and Ω_t are independent, then $f(y | \Omega_t) = f(y)$ and $E(y | \Omega_t) = E(y)$.
 - There is no gain from conditioning.
- To optimally forecast unknown y we want to use observable variables Ω_t which are highly correlated with y .

Summary – Conditional Forecasts

- Information improves forecasts
- Conditioning on relevant variables reduces the risk (expected loss) of forecasts
- Point and interval forecasts are functions of the conditioning variables

Actual Forecasting

- Even if the variables in the information set Ω_t are known, the conditional mean function $E(y_{t+h} | \Omega_t)$ is unknown
 - The functional form is unknown
 - The parameters of the function are unknown
- Thus to make an actual forecast, we need to:
 - Create an approximate model for $E(y_{t+h} | \Omega_t)$
 - Estimate the model parameters from data.

Time-Series Components

- Recall that the optimal point forecast of a series y_{t+h} is its conditional mean.
- It is useful to decompose this mean into components

$$E(y_{t+h} | \Omega_t) = T_t + S_t + C_t$$

- T_t = Trend
- S_t = Seasonal
- C_t = Cycle

Components

- Trend
 - Very long term (decades)
 - Smooth
- Seasonal
 - Patterns which repeat annually
 - May be constant or variable
- Cycle
 - Business cycle
 - Correlation over 2-7 years
- It is useful to consider the components separately

Mean Forecasting

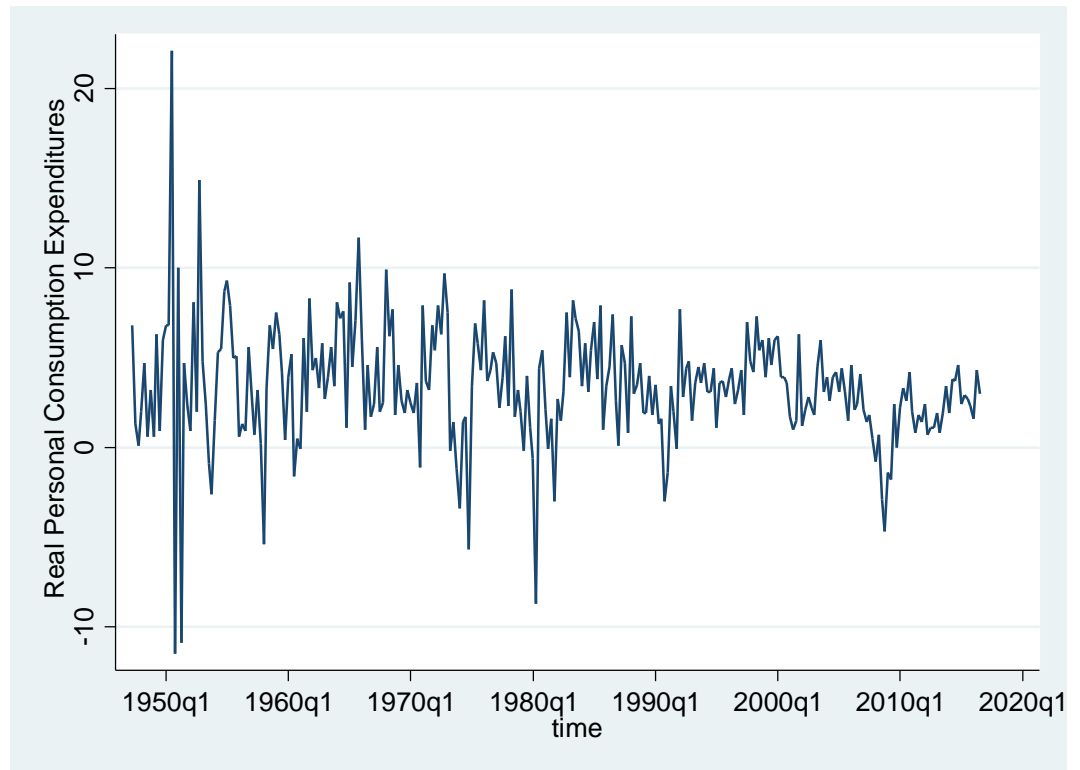
- The simplest forecasting model has no trend, seasonal or cycle, only a constant mean

$$E(y_{t+h} | \Omega_t) = \beta_0$$

- This might seem overly simple, but can be appropriate for a random **stationary** time-series
 - A series not growing or changing over time
 - Many series reported as percentage changes
- In this model, the optimal point forecast for y_{t+h} is $E(y_{t+h} | \Omega_t) = \beta_0$.
- An actual forecast is an estimate of β_0 .

Example: PCE

- U.S. Real Personal Consumption Expenditure
- From National Income Accounts
- Quarterly, percentage change from previous period, annualized
- 1947q2 to 2016q3
 - use `realgdpgrowth`
 - `tsline pce`



Estimation

- If $E(y_{t+h} | \Omega_t) = \beta_0$ then the optimal forecast is the mean $\beta_0 = E(y_{t+h})$
- The estimate of β_0 is the sample mean

$$b_0 = \frac{1}{T} \sum_{t=1}^T y_{t+h}$$

- This is the estimate of the optimal point forecast when $E(y_{t+h} | \Omega_t) = \beta_0$
- b_0 is also the least-squares estimate in an intercept-only model

Estimation

- In STATA, use the **regress** command
- See *STATA Handout* on website
- Sample mean is estimated “constant”
 - regress pce

Source	SS	df	MS	Number of obs	=	278
Model	0	0	.	F(0, 277)	=	0.00
Residual	3129.54144	277	11.2979835	Prob > F	=	.
Total	3129.54144	277	11.2979835	R-squared	=	0.0000
				Adj R-squared	=	0.0000
				Root MSE	=	3.3612

pce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
_cons	3.368345	.2015942	16.71	0.000	2.971494	3.765197

Fitted Values

- Fitted values are the sample mean

$$\hat{y}_t = b_0$$

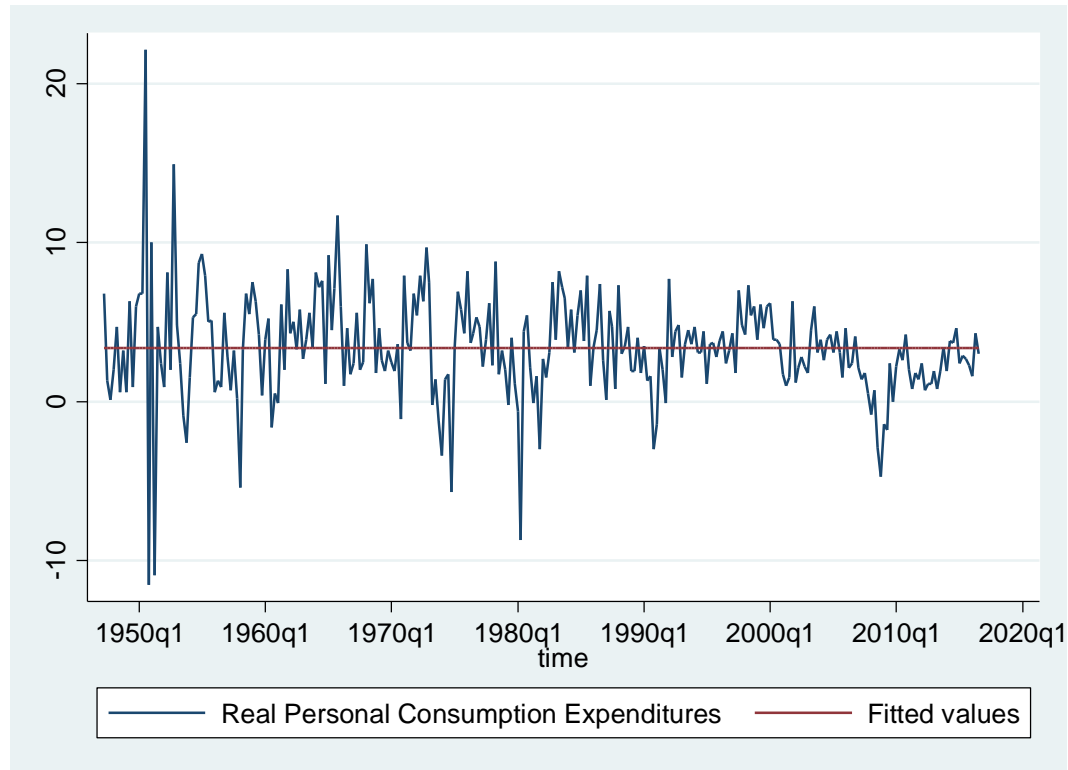
- In STATA use the **predict** command

```
. predict yp  
  (option xb assumed; fitted values)
```

- This creates a variable “yp” of fitted values

Plot actual against fitted

```
. tsline pce yp
```



Out-of-Sample

- Point forecasts are the sample mean

$$\hat{y}_{T+h} = b_0$$

- In STATA, use **tsappend** to expand sample, and **predict** to generate point forecasts.

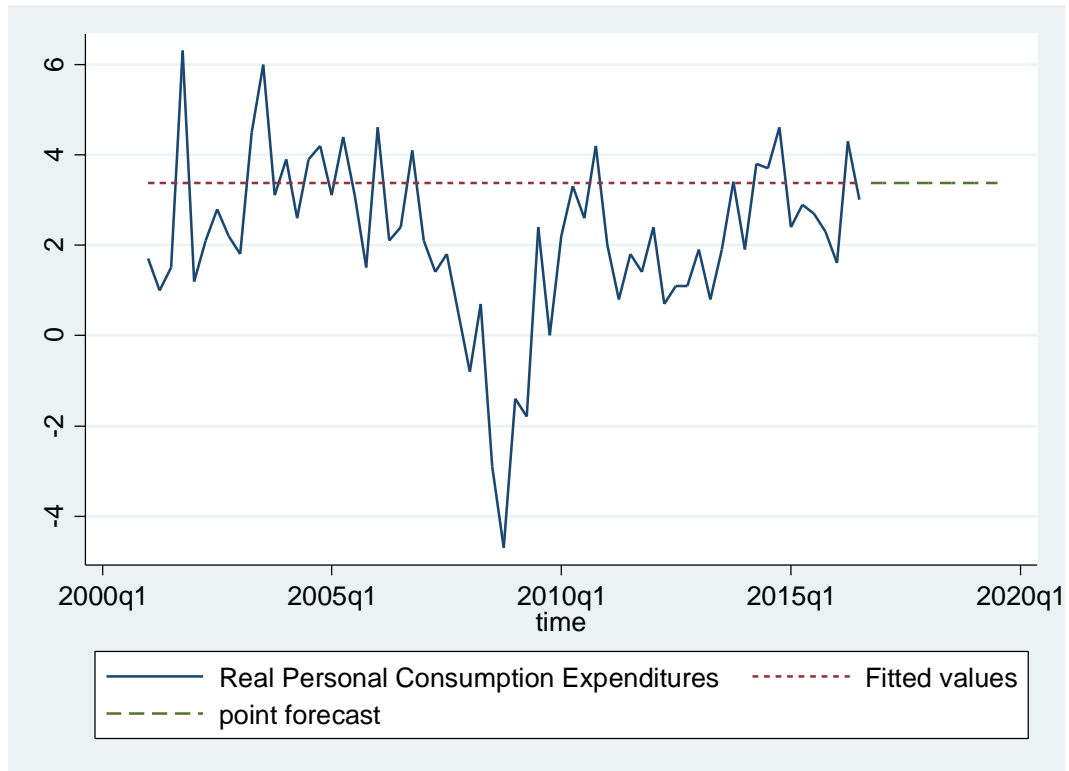
```
. tsappend, add(12)
```

```
. predict p if time>tq(2016q3)  
(option xb assumed; fitted values)  
(278 missing values generated)
```

```
. label variable p "point forecast"
```

```
. tsline pce yp p if time>tq(2000q4), lpattern (solid shortdash dash)
```

Out-of-Sample



Forecast Errors

- The forecast error e_t is the difference between the realized value and the conditional mean.

$$e_t = y_{t+h} - \mathbf{E}(y_{t+h} | \Omega_t)$$

or equivalently

$$y_{t+h} = \mathbf{E}(y_{t+h} | \Omega_t) + e_t$$

- We call e_t the forecast error.

Residuals

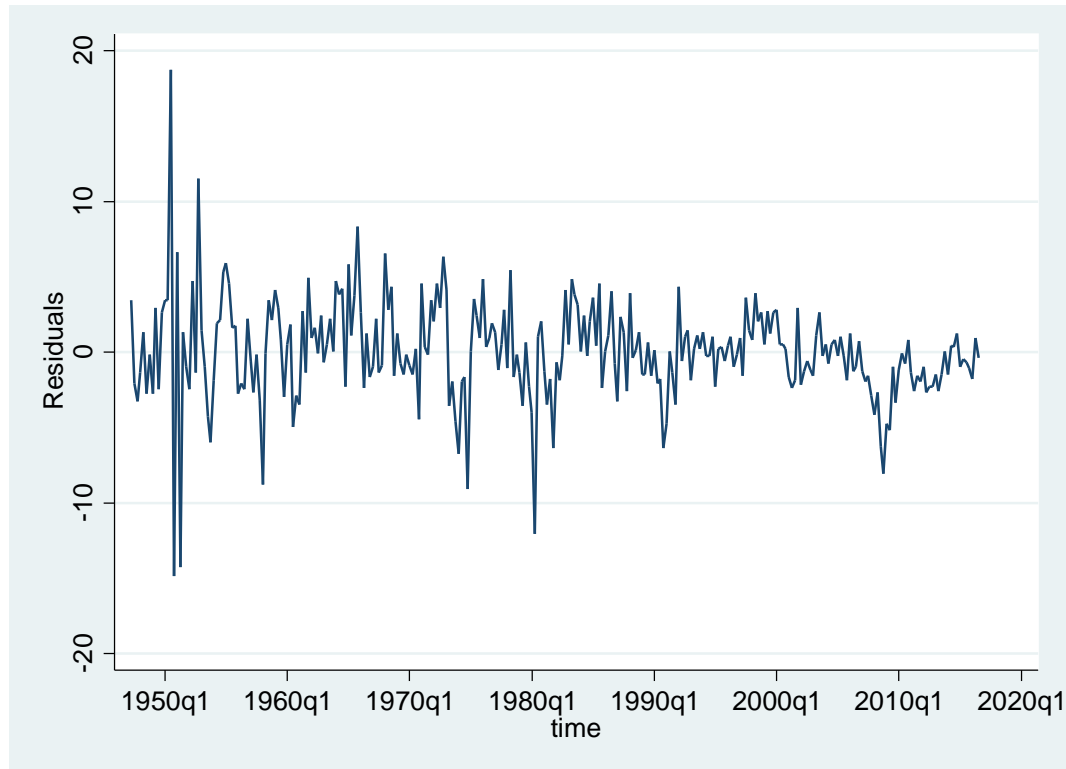
- The residuals are the in-sample fitted errors.
- The difference between the realized value and the in-sample forecast.

$$\begin{aligned}\hat{e}_t &= y_{t+h} - \hat{y}_{t+h} \\ &= y_{t+h} - b_0\end{aligned}$$

- In general, it is useful to plot the residuals against time, to see if any time series pattern remains.

Calculate and Plot Residuals

- `. predict e, residuals`
(12 missing values generated)
- `. tsline e`



Assignments

- Read Chapters 3-4 from Diebold
- Next Class: Chapter 5.1, 5.2 from Diebold
- Problem Set # 2
 - Due Tuesday (1/31)
- Read Chapter 2 from *The Signal and the Noise*
 - Reading Reflection
 - Due Thursday (2/2)