Estimation Uncertainty

- The sample mean

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

is an estimate of $\beta_0 = E(y_{t+h})$

- The estimation error is

$$\begin{aligned} b_0 - \beta_0 &= \frac{1}{T}\sum_{t=1}^{T} y_{t+h} - \beta_0 \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(y_{t+h} - \beta_0\right) \\ &= \frac{1}{T}\sum_{t=1}^{T} e_t \end{aligned}$$

Estimation Variance

- Under classical conditions,

$$\mathrm{var}(b_0) = \frac{\sigma^2}{T}$$

  where $\sigma^2 = \mathrm{var}(e_t)$

- The standard error for b_0 is an estimate of the standard deviation

$$sd(b_0) = \sqrt{\frac{\hat{\sigma}^2}{T}}$$

Forecast Variance

- When the sample mean b_0 is used as the forecast for y_{T+h} then the prediction error is

$$y_{T+h} - b_0 = e_{T+h} + \beta_0 - b_0$$

  which is the sum of the forecast error e_{T+h} and the estimation uncertainty β_0-b_0.
- The forecast variance is

$$\begin{aligned} \operatorname{var}(y_{T+h} - b_0) &= \operatorname{var}(e_{T+h}) + \operatorname{var}(\beta_0 - b_0) \\ &= \sigma^2 + \frac{\sigma^2}{T} \\ &= \left(1 + \frac{1}{T}\right)\sigma^2 \end{aligned}$$

Standard Deviation of Forecast

- The standard deviation of the forecast is the estimate

$$s_{T+h} = \sqrt{\left(1 + \frac{1}{T}\right)\hat{\sigma}^2}$$

- This is slightly larger than the regression standard deviation $\hat{\sigma}$
- Calculated in STATA after a regression using the stdf option to the predict command:
- predict s, stdf
- This creates variable “s”

Normal Forecast Intervals

- Let $\hat{y}_{T+h}$ be a forecast for y_{T+h}
- The prediction error is $y_{T+h} - \hat{y}_{T+h}$
- Let s_{T+h} be the st. deviation of the forecast
- If the prediction errors are normally distributed, the (1-α)% forecast interval endpoints are

$$L_{T+h} = \hat{y}_{T+h} + s_{T+h} z_{\alpha/2}$$

$$U_{T+h} = \hat{y}_{T+h} + s_{T+h} z_{1-\alpha/2}$$

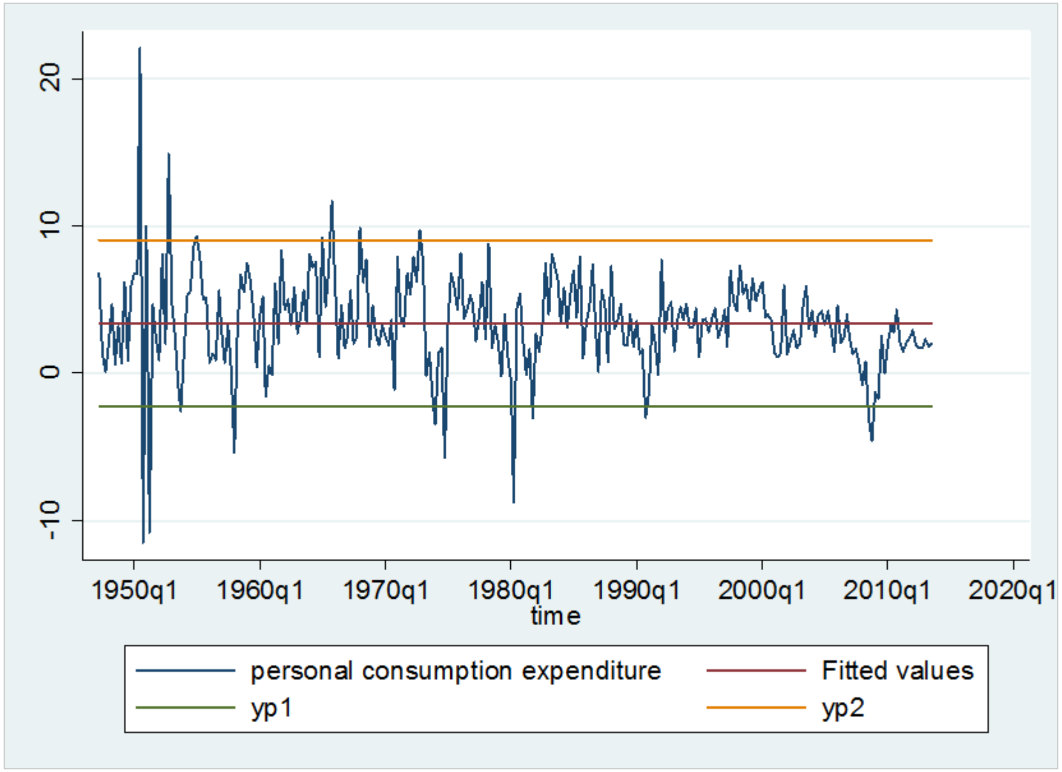
  where $z_{\alpha/2}$ and $z_{1-\alpha/2}$ are the α/2 and 1-α /2 quantiles of the normal distribution
- *e.g.* $\hat{y}_{T+h} \pm 1.64\ s_{T+h}$ for a 90% interval

```
. predict s, stdf

. generate yp1=yp-1.645*s
(12 missing values generated)

. generate yp2=yp+1.645*s
(12 missing values generated)

. tsline pce yp yp1 yp2
```



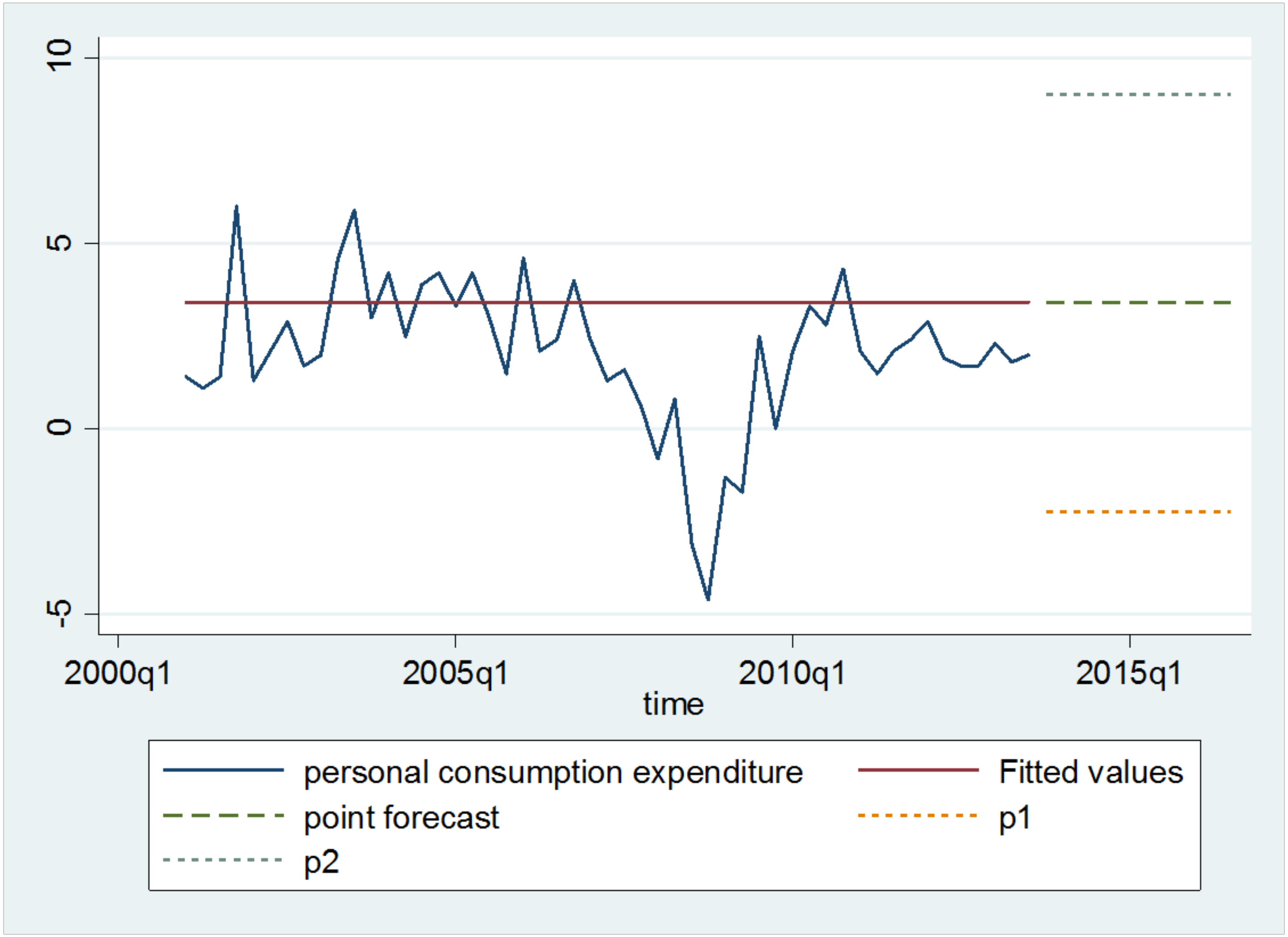
Out-of-Sample

```
. generate p1=p-1.645*s
(266 missing values generated)

. generate p2=p+1.645*s
(266 missing values generated)

. tsline pce yp p p1 p2 if time>tq(2000q4)
```

Out-of-Sample

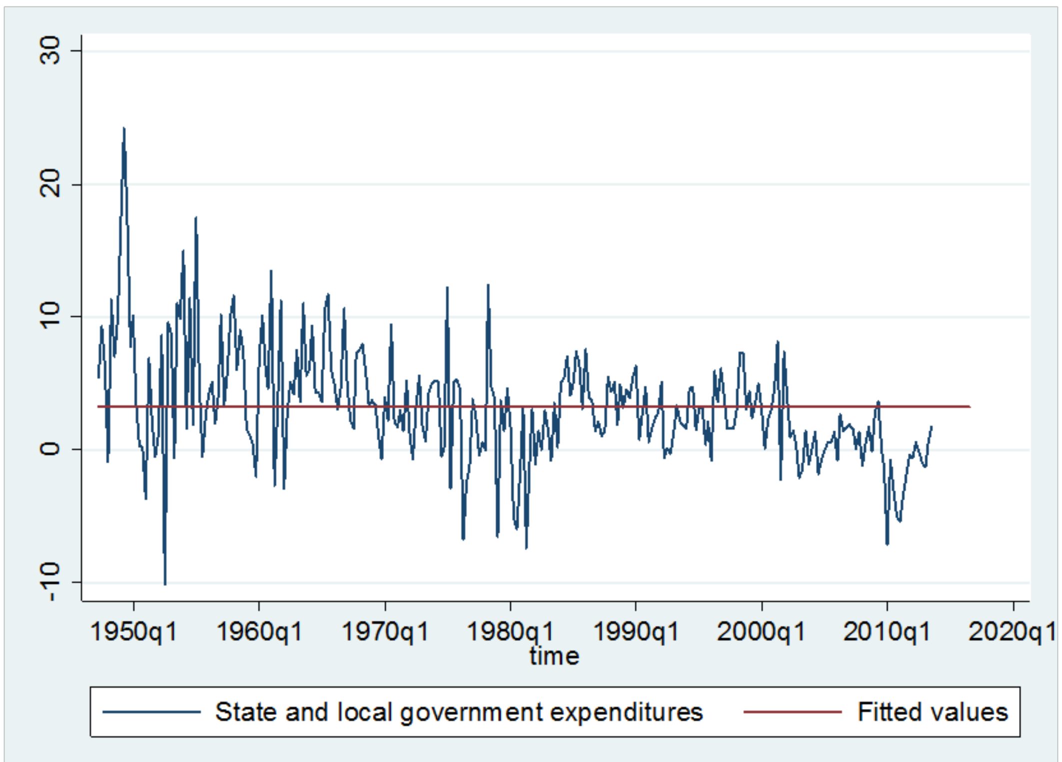


Mean Shifts

- Sometimes the mean of a series changes over time
- It can drift slowly, or change quickly
  - Possibly due to a policy change
- In this case, forecasting based on a constant mean model can be misleading

State and Local Government Spending Percentage Growth Rate (Quarterly)

- Average for 1947-2013: 3.24%
- But this has not been typical in recent years.



Alternatives

- Subsample estimation
  - Estimate the mean on subsamples
  - Forecasts are based on the most recent
- Dummy Variable formulation

$$\mathrm{E}\left(y_{t+h} \mid \Omega_t\right) = \beta_0 + \beta_1 d_t$$

$$d_t = 1\left(t \geq \tau\right)$$

- τ is the breakdate
  - The date when the mean shifts
  - The coefficient β_0 is the mean before $t=\tau$
  - The coefficient β_1 is the shift at $t=\tau$
  - The sum $\beta_0+\beta_1$ is the mean after $t=\tau$

Forecast

- Linear Regression y_{t+h} on d_t
- Example
  - State and Local Government Percentage Growth
  - Mean breaks in 1970q1 and 2002q1

```
. generate d1=(time>=tq(1970q1))

. generate d2=(time>=tq(2002q1))

. regress state d1 d2
```

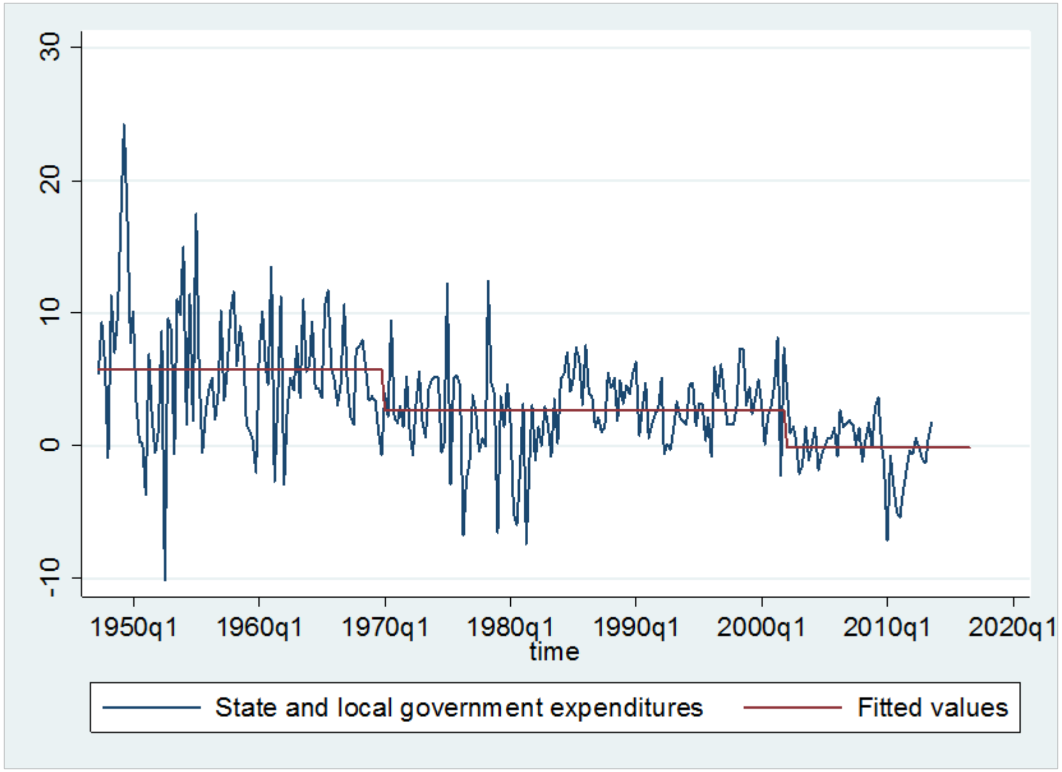
Source	SS	df	MS
Model	1153.92979	2	576.964895
Residual	3996.4187	263	15.1955084
Total	5150.34849	265	19.4352773

Number of obs =	266
F(2, 263) =	37.97
Prob > F =	0.0000
R-squared =	0.2240
Adj R-squared =	0.2181
Root MSE =	3.8981

state	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d1	-3.065908	.5345077	-5.74	0.000	-4.118367	-2.013449
d2	-2.837084	.6648486	-4.27	0.000	-4.146188	-1.527981
_cons	5.76044	.4086363	14.10	0.000	4.955825	6.565055

Fitted

- Out-of-sample forecast falls from 3.2% to -0.14%!



Should you use Mean Shifts?

- Only after great hesitation and consideration.
- Should use shifts and breaks reluctantly and with care.
- Do you have a model or explanation?
- What is the forecasting power of a mean shift?
  - If they have happened in the past, will there be more in the future?
- Yet, if there has been an obvious shift, a simple constant mean model will forecast terribly.

How to Select Breakdates

- Judgmental
  - Dates of known policy shifts
  - Important events
  - Economic crises
- Informal data-based
  - Visual inspection
- Formal data-based
  - Estimate regression for many possible breakdates
  - Select one which minimizes sum of squared error
  - This is the least-squares breakdate estimator

Trend Models

- A trend model is

$$T_t = g(Time_t)$$

  where $Time_t$ is the time index.
- In STATA, $Time_t$ is an integer sequence, normalized to be zero at first observation of 1960.
- Most common models
  - Linear Trend
  - Exponential Trend
  - Quadratic Trend
  - Trends with Changing SLope

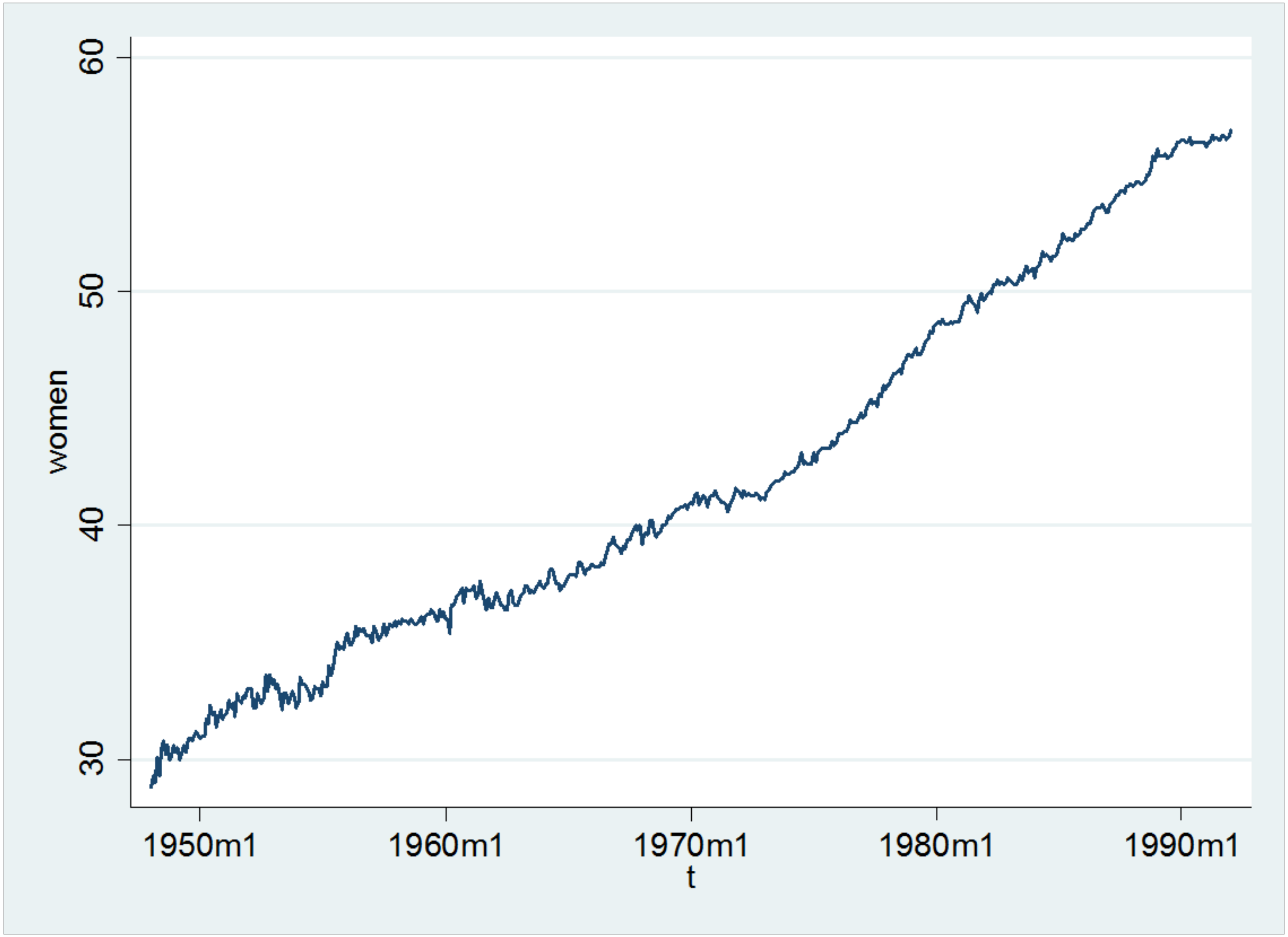
Warning:
Be skeptical of Trend Models

- While in some cases, trend forecasting can be useful.
- In many cases, it can be hazardous.
- We will examine some examples from another textbook (Diebold: Elements of Forecasting)
- They did not forecast well out of sample.
- A constructive alternative is to forecast growth rates, as we did for consumption expenditure.

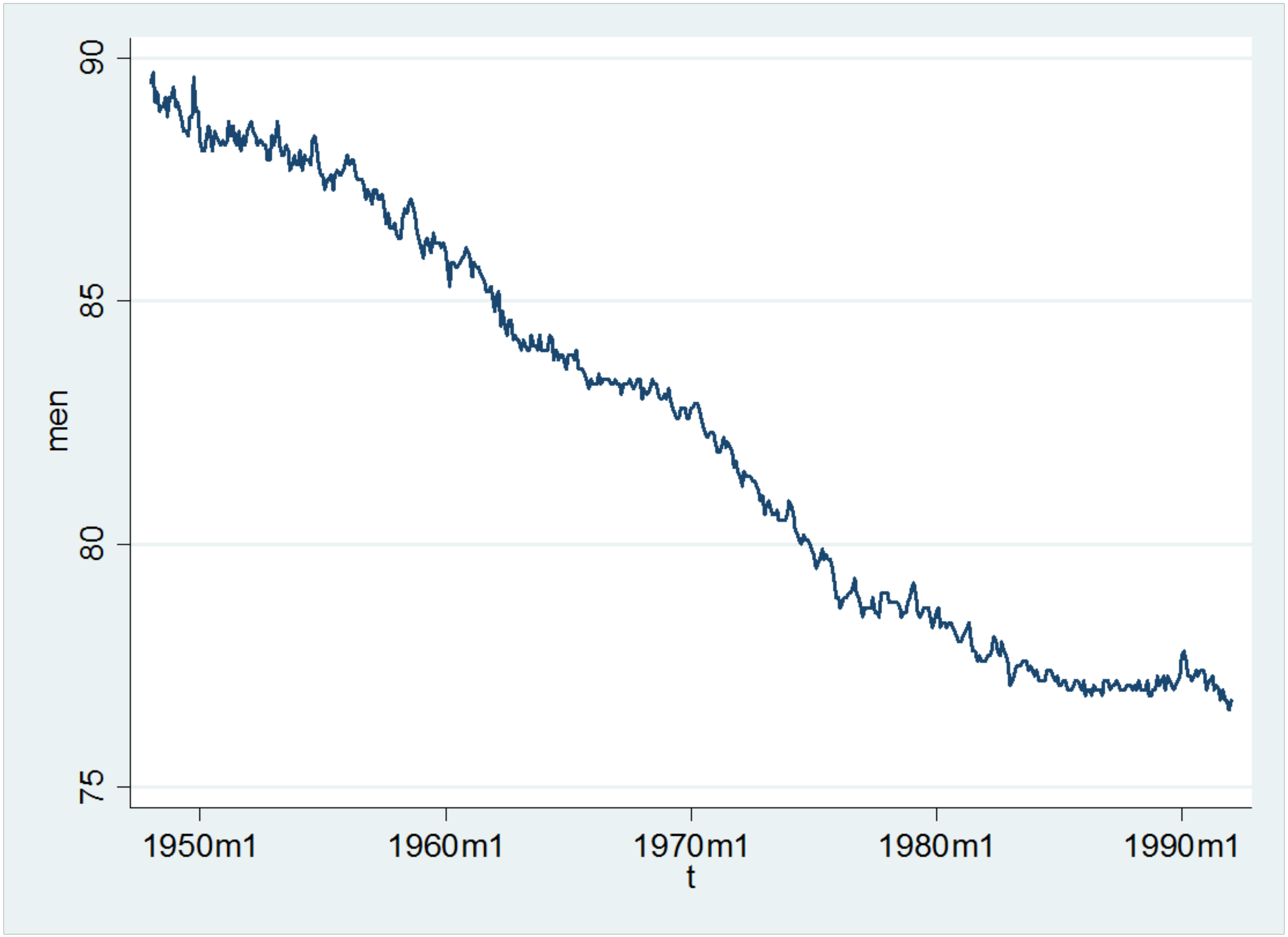
Example 1
Labor Force Participation Rate

- From BLS
- Monthly, 1948-2009, Seasonally adjusted
- Men and Women, ages 25+
- Percentage of population in labor force (employed plus unemployed divided by population)
- We will estimate on 1948-1992
- Forecast 1993-2009

Women's Labor Participation Rate 1948-1992



Men's Labor Participation Rate 1948-1992



Linear Trend Model

- The labor force participation rates have been smoothly and linearly increasing (for women) and smoothly and linearly decreasing (for men) over 1948-1992
- This suggests a linear trend

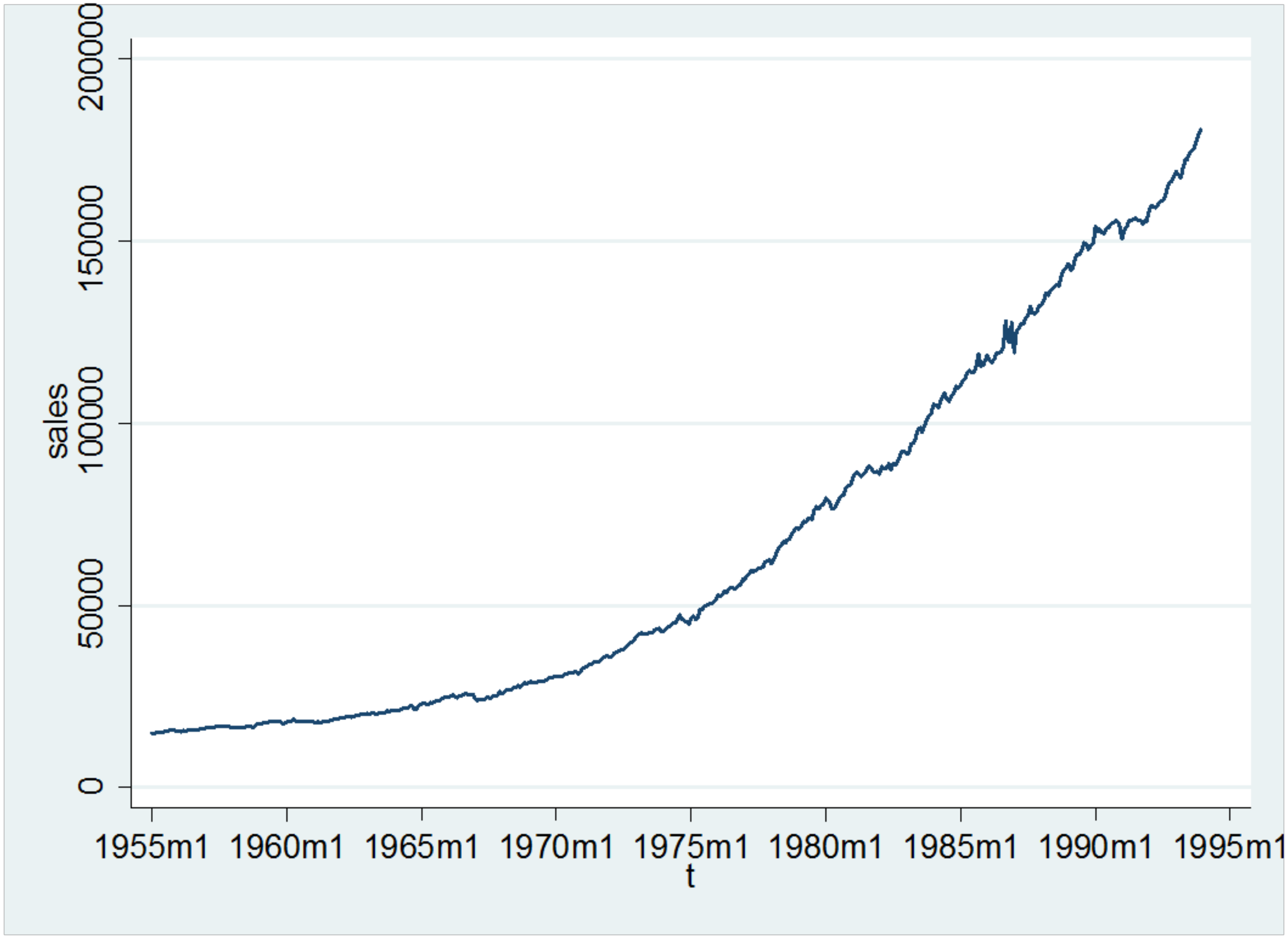
$$T_t = \beta_0 + \beta_1 Time_t$$

- In this model, β_1 is the expected period-to-period change in the trend T_t

Example 2
Retail Sales, Current Dollars

- From Census Bureau
- Monthly, 1955-2001, seasonally adjusted
  - This particular series discontinued after 2001
- We will 1955-1991
- Forecast 1992-current

Retail Sales 1955-1993



Quadratic Trends

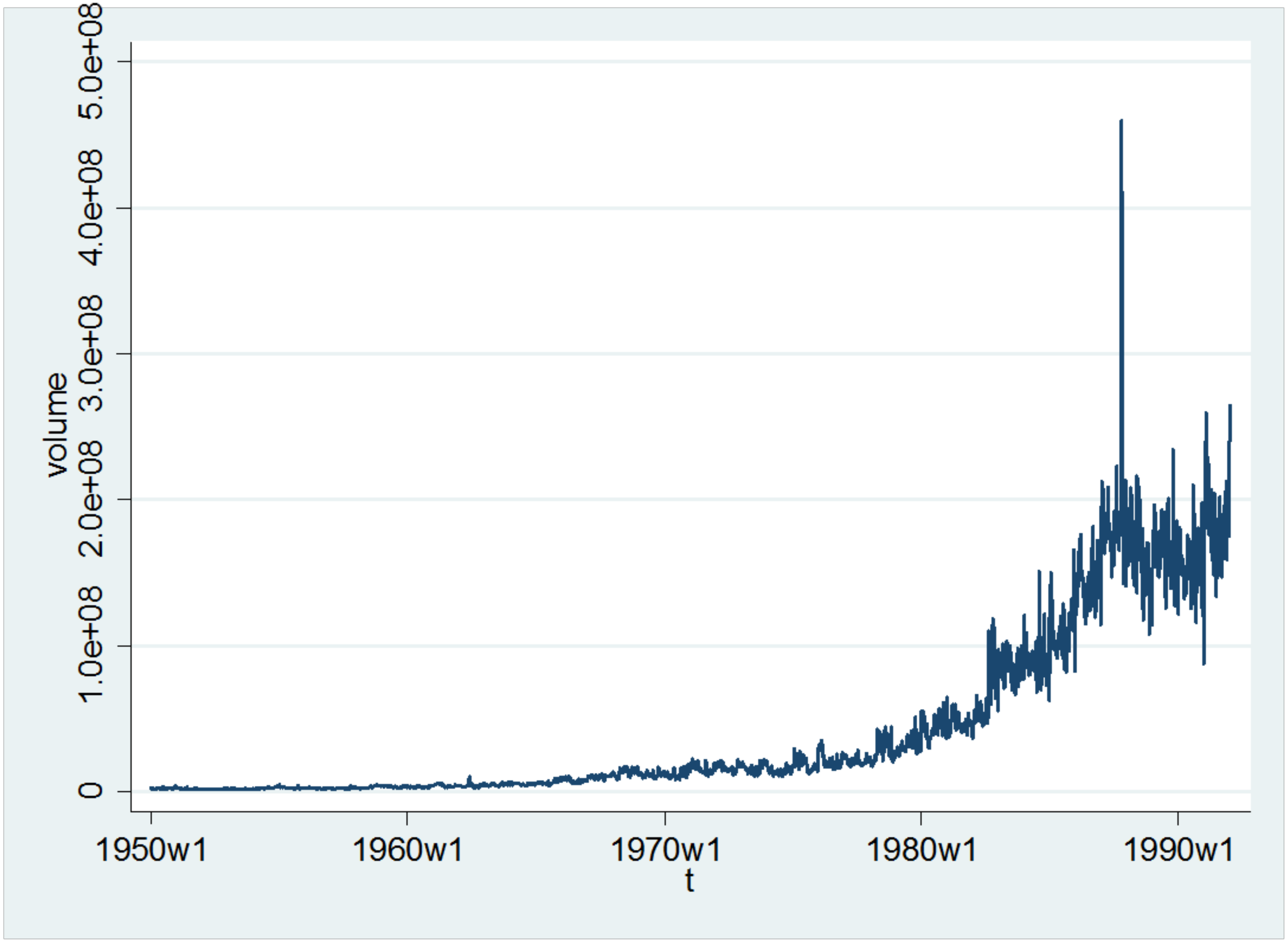
- The retail sales series has been increasing smoothly over 1955-1993, but not linearly.
- To model this we will use a quadratic trend

$$T_t = \beta_0 + \beta_1 Time_t + \beta_2 Time_t^2$$

Example 3
Transaction Volume, S&P Index

- From Yahoo Finance
- Weekly, 1950-current
- We estimate on 1955-1993
- Forecast 1994-2001

Transaction Volume



Exponential Trend

- To model this we will use an exponential trend

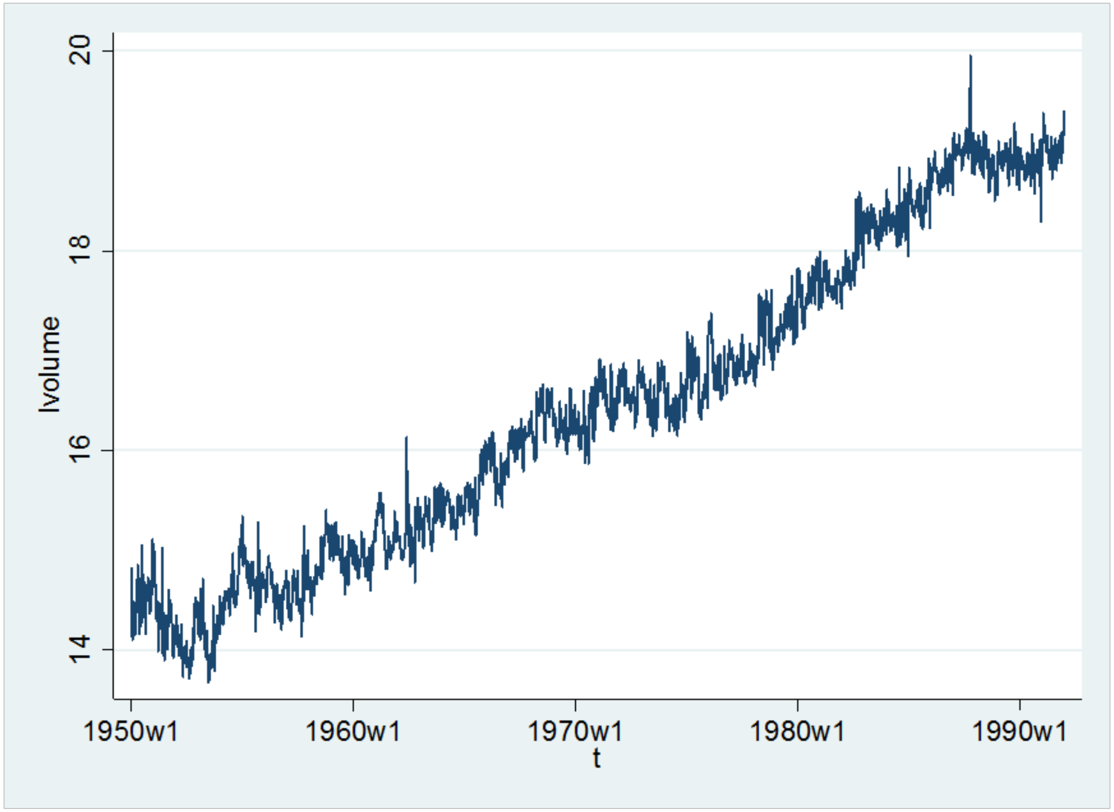
$$T_t = e^{\beta_0 + \beta_1 Time_t}$$

- The exponential trend is linear after taking (natural) logarithms

$$\ln(T_t) = \beta_0 + \beta_1 Time_t$$

- This is typically estimated by a linear model after taking logs of the variable to forecast

Ln(Volume)



- In logarithms, trend is roughly linear.

Exponential Trends

- Most economic series which are growing (aggregate output, such as GDP, investment, consumption) are exponentially increasing
  - Percentage changes are stable in the long run
- These series cannot be fit by a linear trend
- We can fit a linear trend to their (natural) logarithm

Linear Models

- The linear and quadratic trends are both linear regression models of the form

$$T_t = \beta_0 + \beta_1 x_{1t}$$

or

$$T_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t}$$

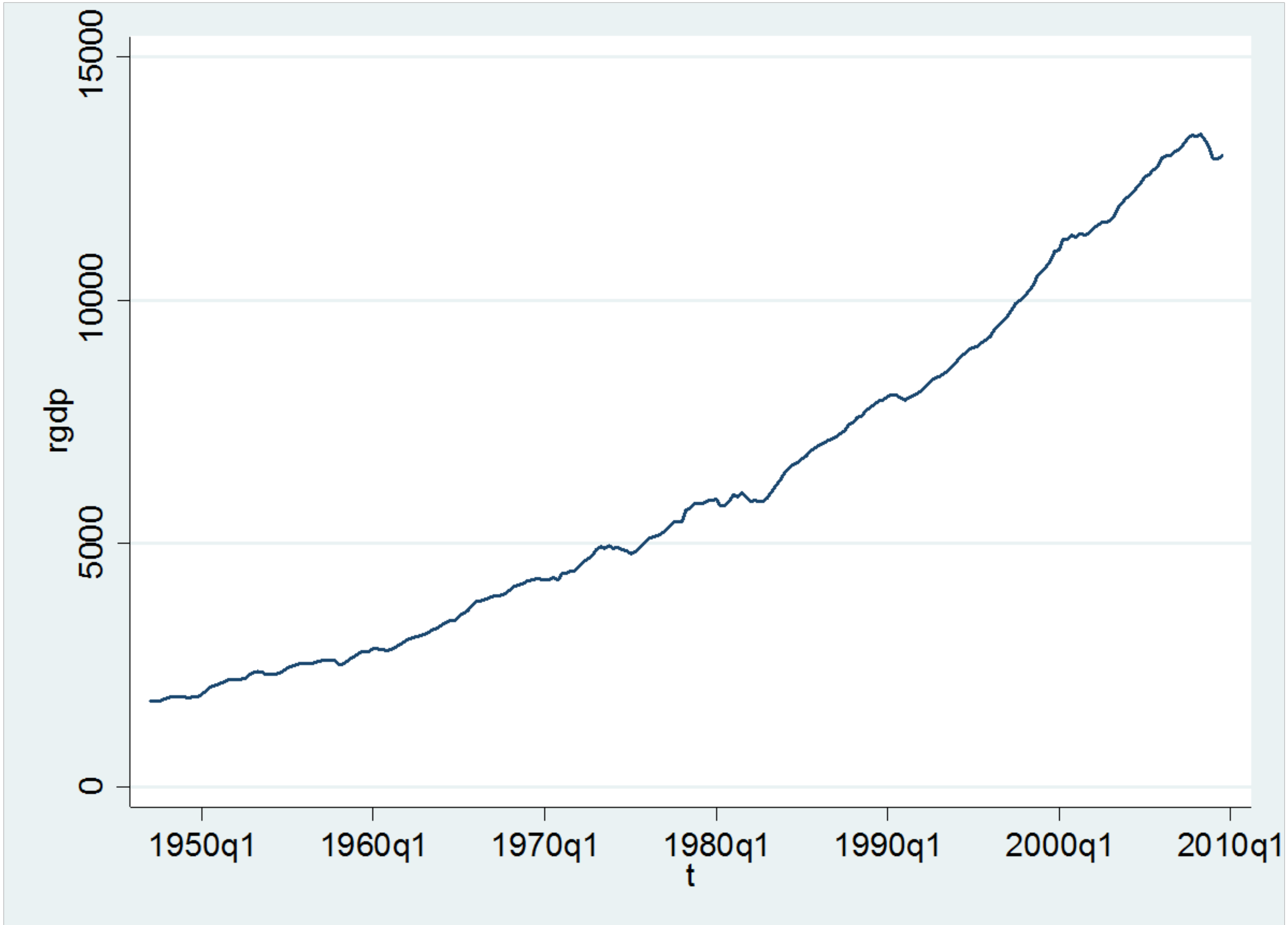
where

  - $x_{1t} = Time_t$
  - $x_{2t} = Time_t^2$

Example 4
Real GDP

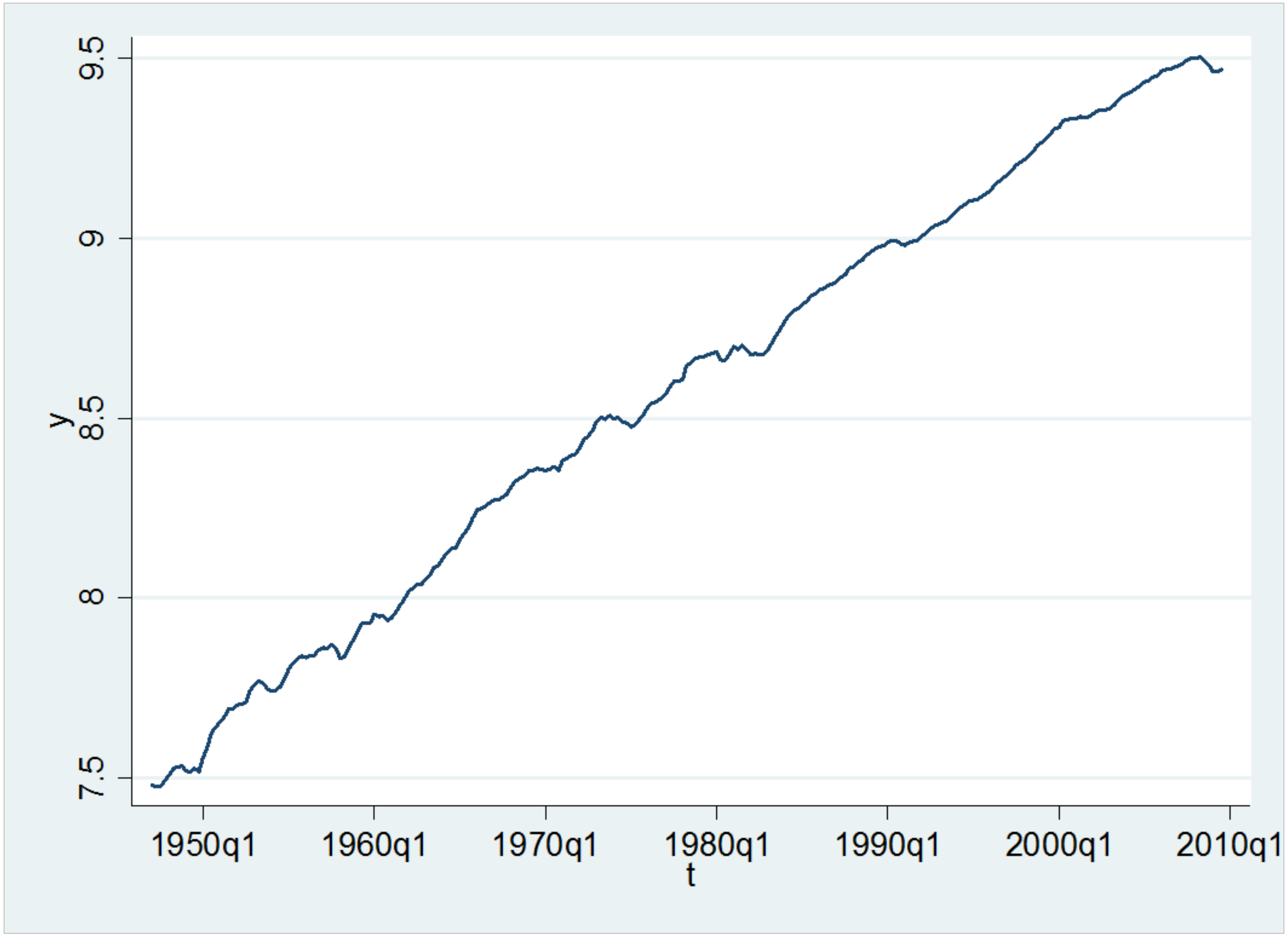
- From BEA
- Quarterly, 1947-2009
- We will estimate on 1947-1990, forecast 1991-2009
- Also use an exponential trend

Real GDP



15000
10000
5000
0
rgdp
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
t

Ln(Real GDP)



9.5
9
8.5
8
7.5
y
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
t

Linear Forecasting

- The goal is to forecast future observations given a linear function of observables
- In the case of trend estimation, these observables are functions of the time index
- In other cases, they will be other functions of the data
- In the model $T_t = \beta_0 + \beta_1 x_t$

  the forecast for y_{t+h} is $\hat{y}_{t+h} = b_0 + b_1 x_t$ where b_0 and b_1 are estimates

Estimation

- How should we select b_0 and b_1 ?
- The goal is to produce a forecast with low mean square error (MSE)
- The best linear forecast is the linear function

$\beta_0+\beta_1 x_t$ that minimizes the MSE

$$E\left(y_{t+h}-\hat{y}_{t+h}\right)^2 = E\left(y_{t+h}-\beta_0-\beta_1 x_t\right)^2$$

- We do not know the MSE, but we can estimate it by a sample average

Sum of Squared Errors

- Sample estimate of mean square error is the sum of squared errors

$$S_n(\beta_0, \beta_1) = \frac{1}{n}\sum_{t=1}^{n}\left(y_{t+h} - \beta_0 - \beta_1 x_t\right)^2$$

- The best linear forecast is the linear function $\beta_0 + \beta_1 x_t$ that minimizes the MSE, or expected sum of squared errors.
- Our sample estimate of the best linear forecast is the linear function which minimizes the (sample) sum of squared errors.
- This is called the least-squares estimator

Least Squares

- The least-squares estimates (b_0,b_1) are the values which minimize the sum of squared errors

$$S_n(\beta_0, \beta_1) = \frac{1}{n}\sum_{t=1}^{n}\left(y_{t+h} - \beta_0 - \beta_1 x_t\right)^2$$

- This produces estimates of the best linear predictor – the linear function $\beta_0+\beta_1 x_t$ that minimizes the MSE

Multiple Regressors

- There are multiple regressors

$$T_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t}$$

- For example, the quadratic trend

$$T_t = \beta_0 + \beta_1 Time_t + \beta_2 Time_t^2$$

- The best linear predictor is the linear function $\beta_0+\beta_1 x_{1t}+\beta_2 x_{2t}$ that minimizes the MSE

$$E\left(y_{t+h} - \hat{y}_{t+h}\right)^2 = E\left(y_{t+h} - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t}\right)^2$$

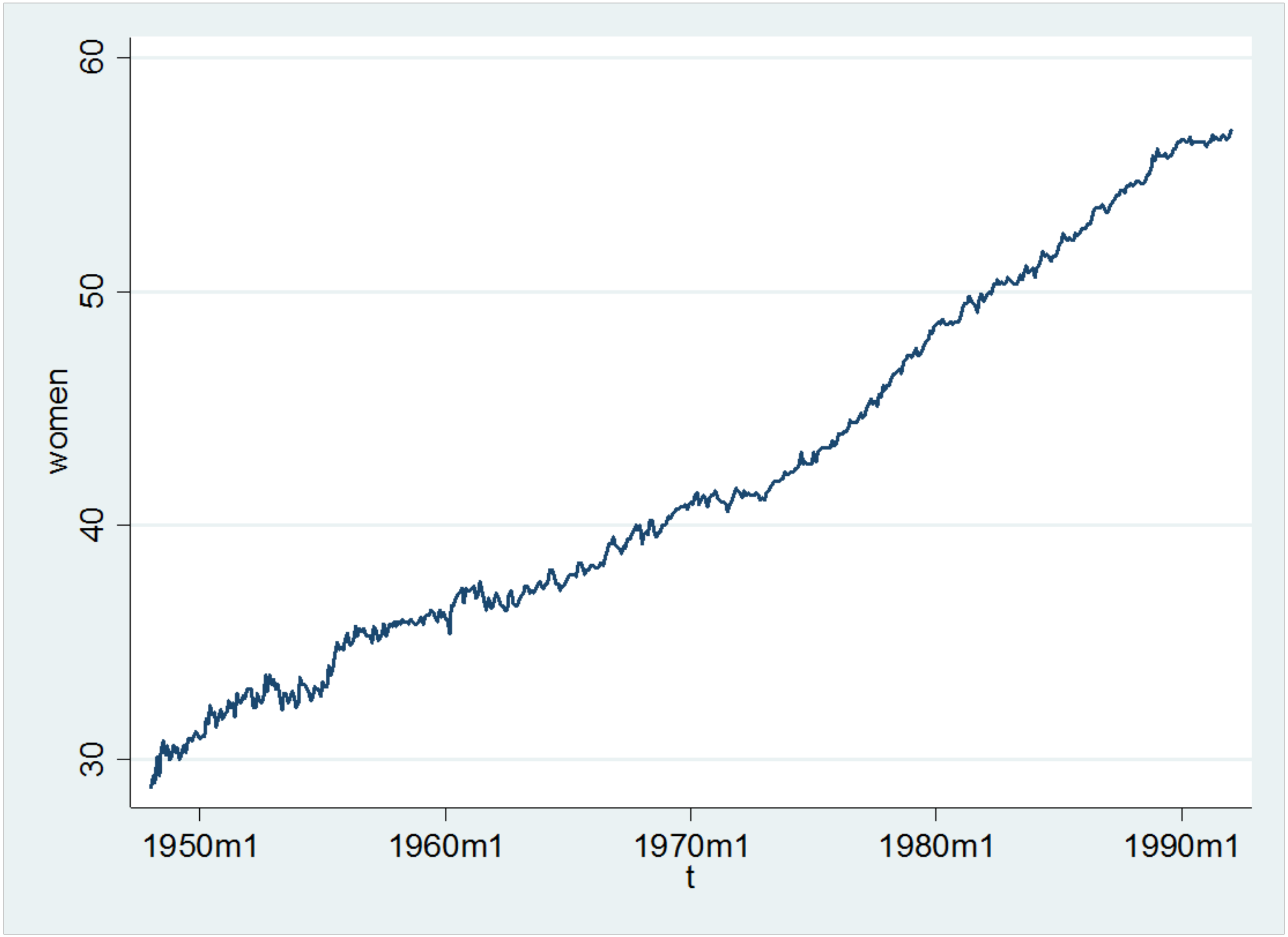
Multiple Regression

- The sample estimate of the best linear predictor are the values (b_0,b_1,b_2) which minimize the sum of squared errors

$$S_n(\beta_0, \beta_1, \beta_2) = \frac{1}{n}\sum_{t=1}^{n}\left(y_{t+h} - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t}\right)^2$$

- In STATA, use the **regress** command

Example 1
Women's Labor Force Participation Rate



Regression Estimation

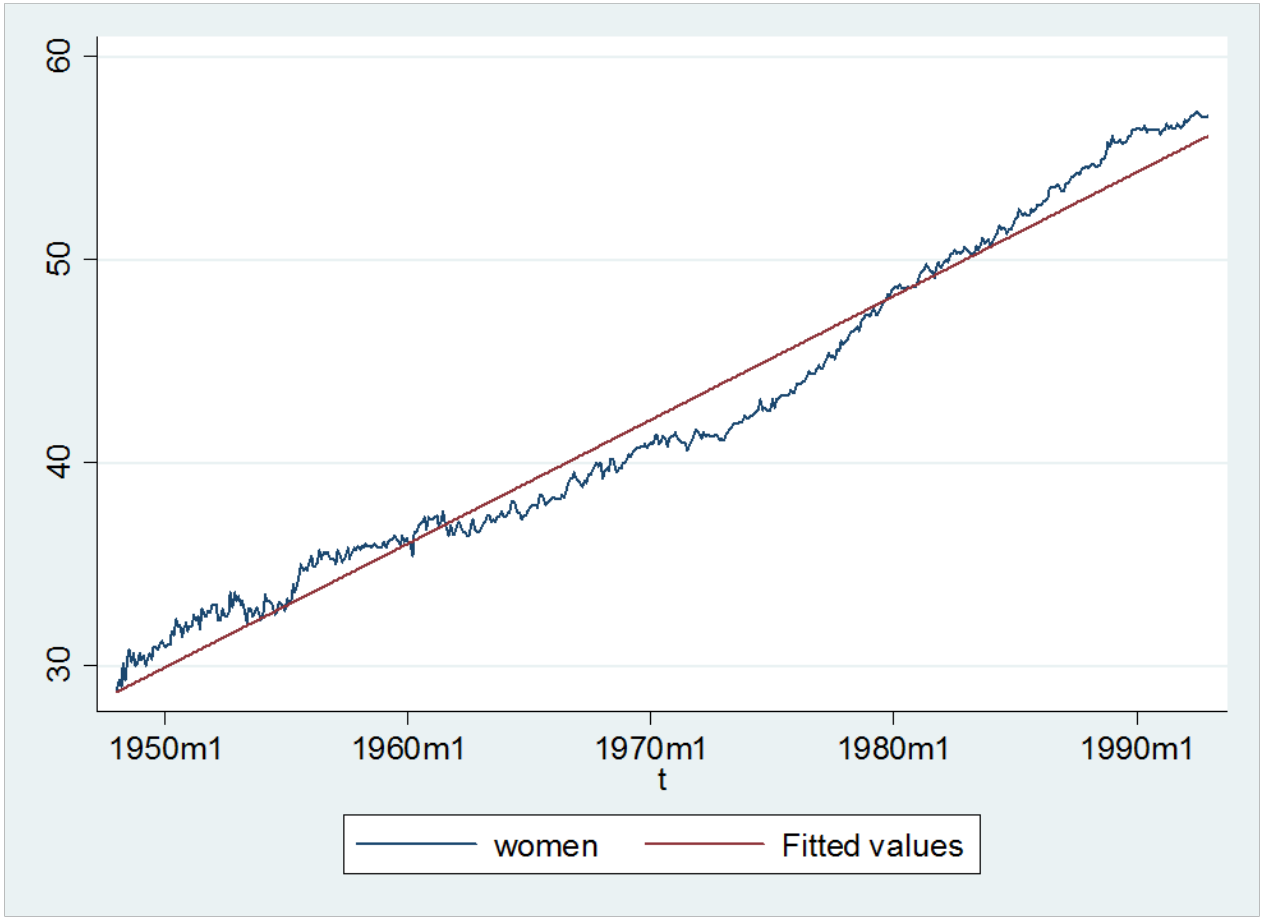
```
. use "C:\Users\Bruce Hansen\Documents\docs\classdocs\390\participation.dta"

. regress women t if t<=tm(1992m12)

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  1,   538) =19575.55
       Model |  33879.9082     1  33879.9082           Prob > F      =  0.0000
    Residual |  931.130559   538  1.73072595           R-squared     =  0.9733
-------------+------------------------------           Adj R-squared =  0.9732
       Total |  34811.0387   539  64.5844874           Root MSE      =  1.3156

------------------------------------------------------------------------------
       women |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .0508126   .0003632   139.91   0.000     .0500992    .0515261
       _cons |   36.02153   .0726804   495.62   0.000     35.87876     36.1643
------------------------------------------------------------------------------
```

In-Sample Fit



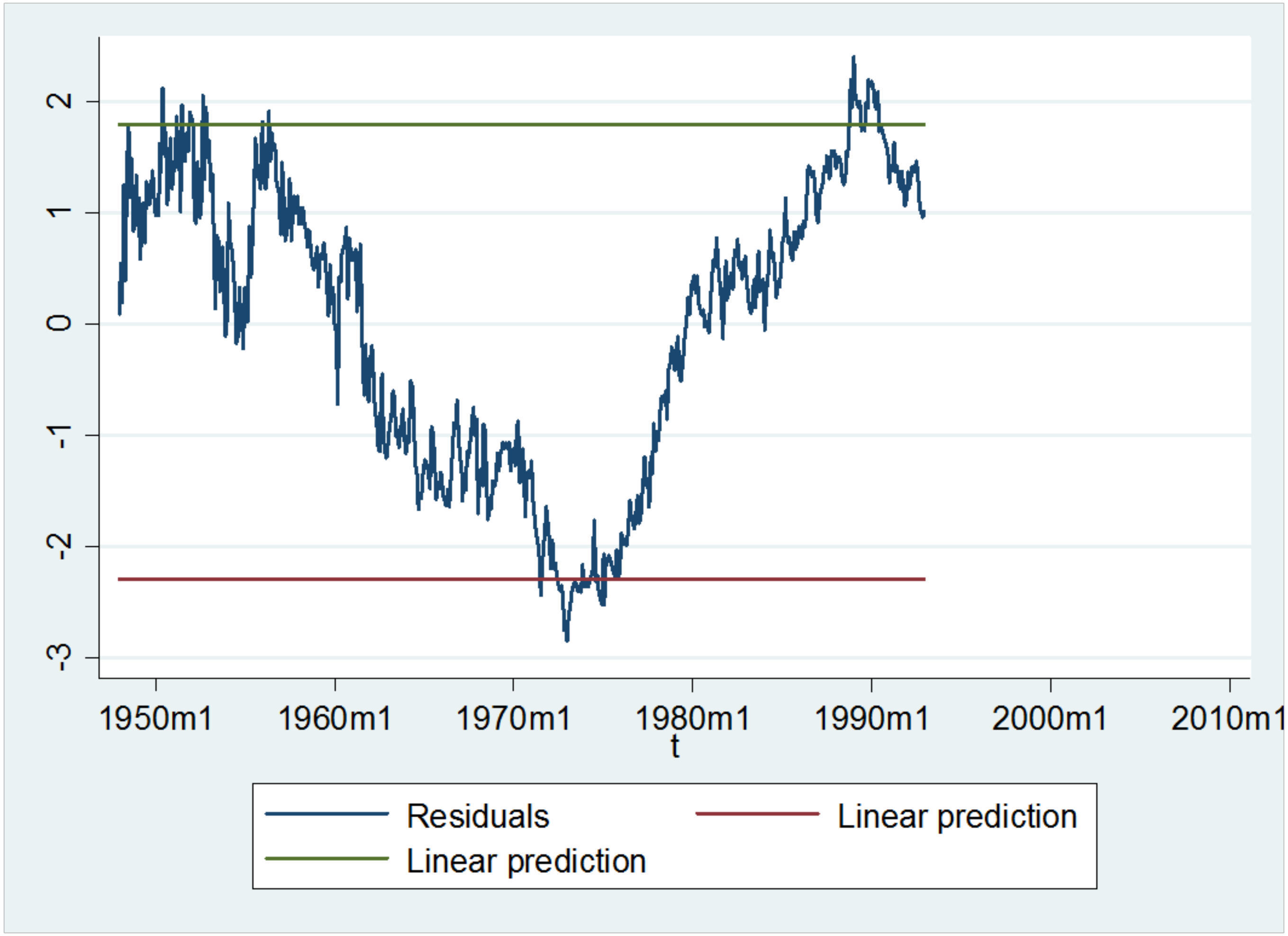
Residuals

- Residuals are difference between data and fitted regression line

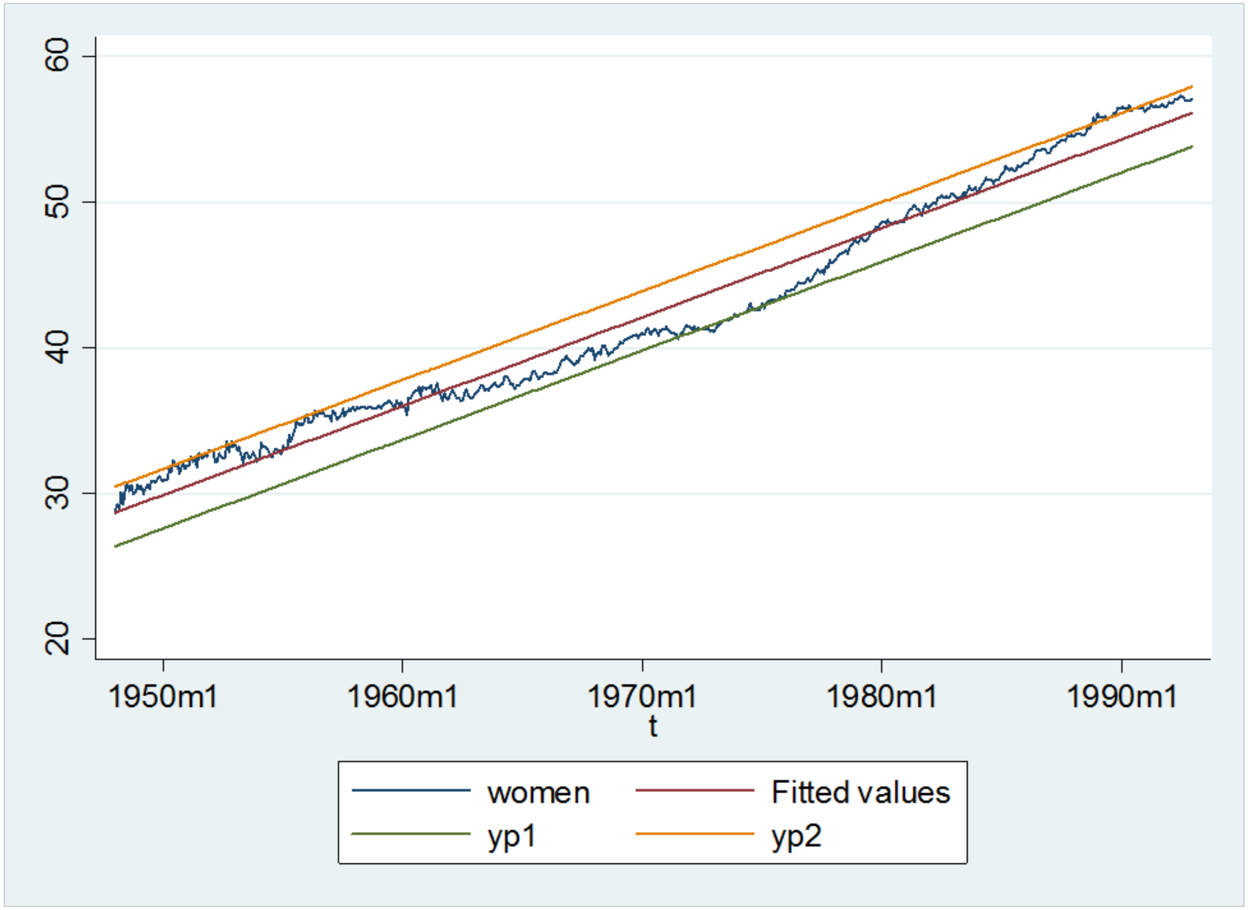
$$\hat{e}_t = y_{t+h} - T_t$$
$$= y_{t+h} - b_0 - b_1 Time_t$$

```
. predict e if t<=tm(1992m12), residuals
(204 missing values generated)
```

Residual Plot



In-Sample Fit



- Compute with **predict** command
- Fit looks good