Seasonality

- Recall that we said that it can be useful to describe the mean of a time series as the sum of components

$$\mu_t = T_t + S_t + C_t$$

  where S_t is the seasonal component.
- The seasonal component S_t is a repetitive cycle over the calendar year
- Seasonality S_t can be deterministic (predictable) or stochastic

Seasonality – Examples

- Gasoline consumption rises in summer due to increased auto travel
- International airline prices rise in summer due to increased tourism
- Natural gas consumption and prices rise in winter due to heating
- Electricity consumption increases in summer due to air conditioning
- Construction activity and jobs decrease in winter in the Midwest
- Consumer spending increases in November and December due to holiday shopping

Deterministic vs Stochastic Seasonality

- If the seasonal pattern repeats year after year, it is deterministic and predictable.
  - Christmas is always in December
- If the seasonal pattern roughly repeats itself, but evolves over the years, it is stochastic and only partially predictable
  - Holiday shopping as a percentage of income is not a fixed constant
- Seasonal patterns can change dramatically as the economy evolves
  - The spread of air conditioning shifted the seasonal pattern of residential electricity consumption from winter to summer

Seasonal Adjustment

- Most economic indicators reported by the government are **seasonally adjusted**.
- Roughly, the component S_t is estimated, and then what is reported is

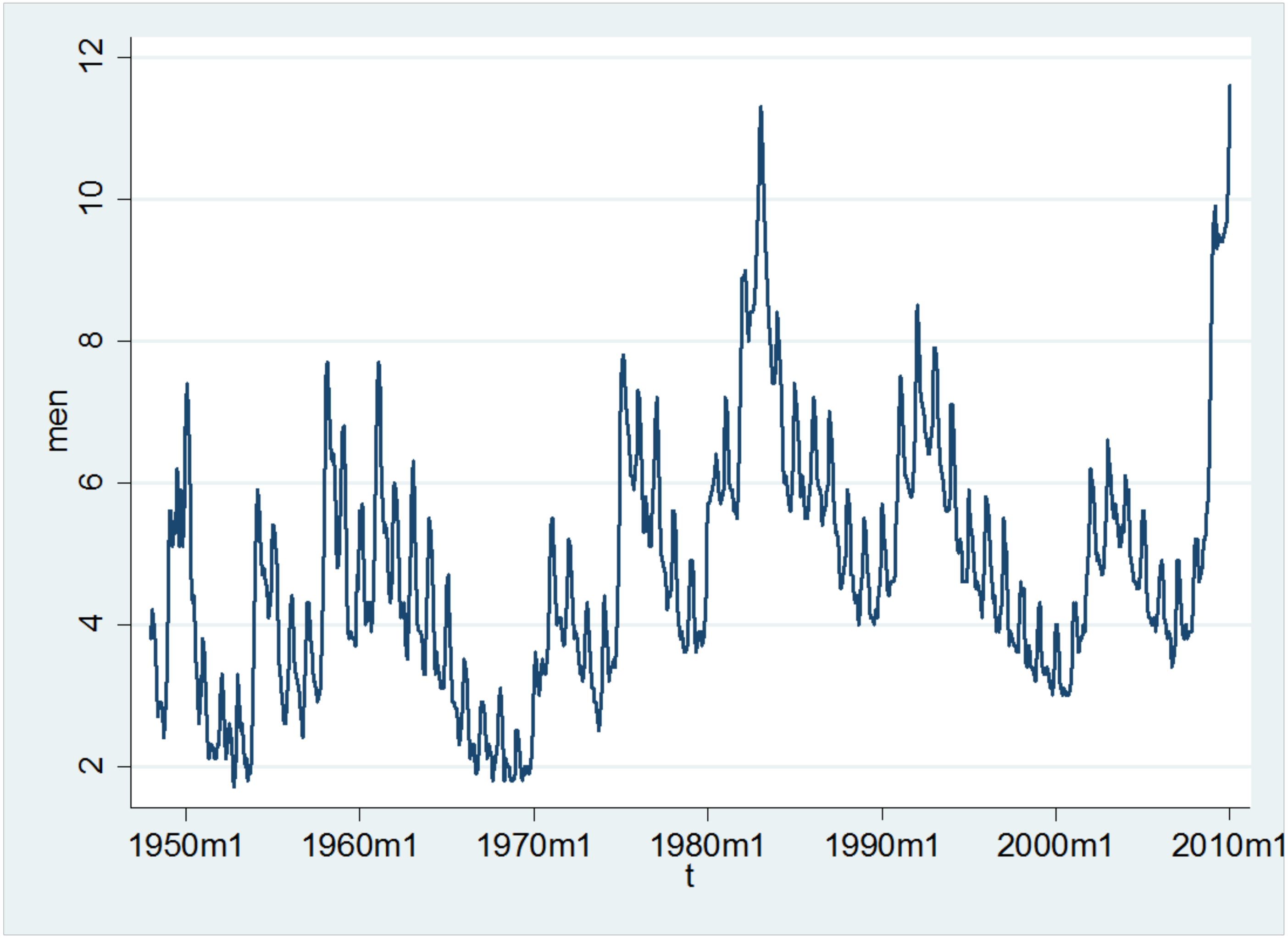
$$\begin{aligned} y_t^* &= y_t - S_t \\ &= T_t + C_t \end{aligned}$$

- The idea is that seasonality distracts from the main reporting purpose
  - Seasonally adjusted data allows users to focus on trend and business cycle movements
- Seasonal adjustment by central statistical agencies is sophisticated, allowing for evolving seasonal patterns.

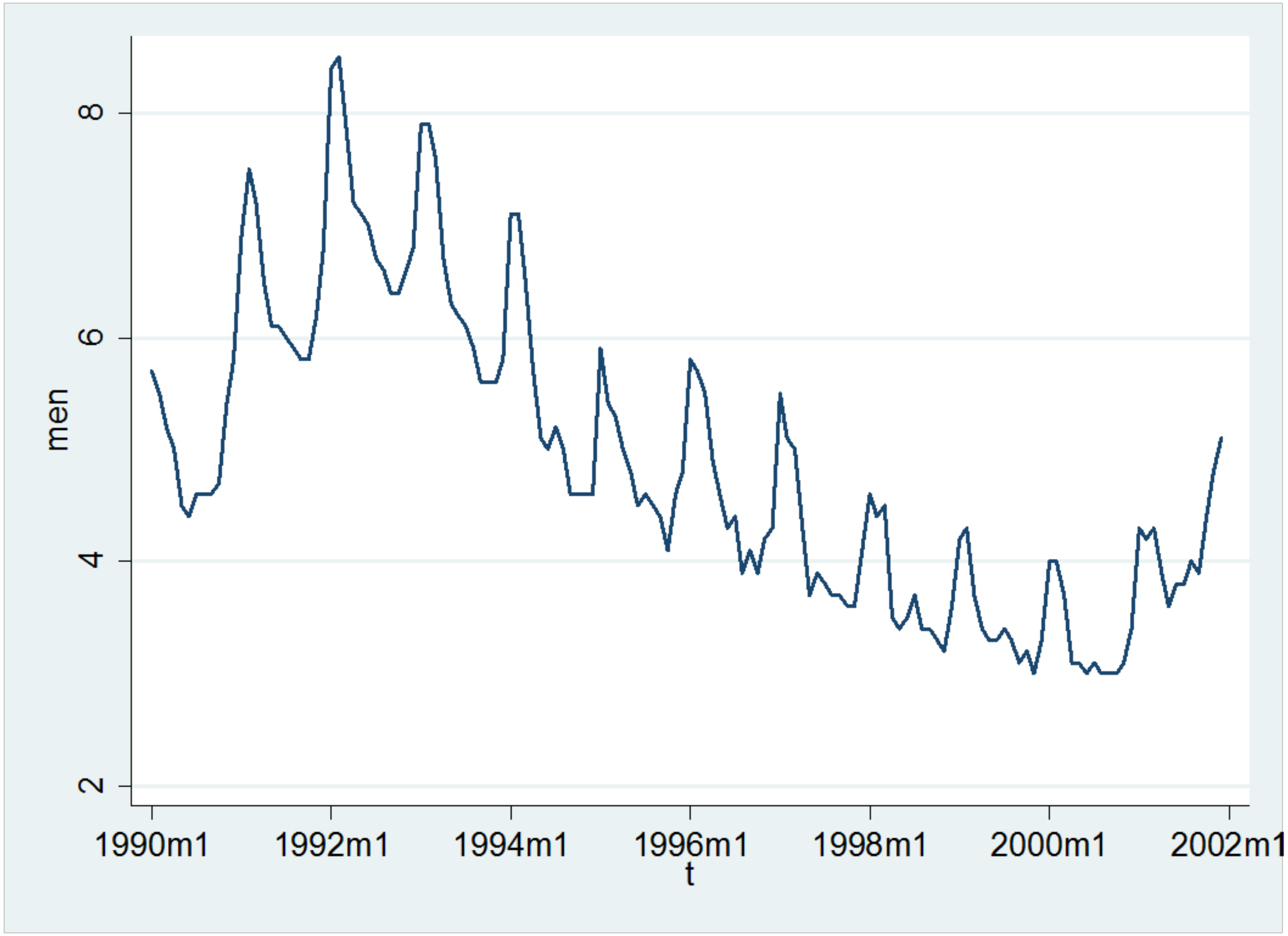
Examples of Seasonal Time Series

- First Example:
- U.S. Unemployment Rate
  - Men, 20+ years
  - 1948-present
  - Not seasonally adjusted

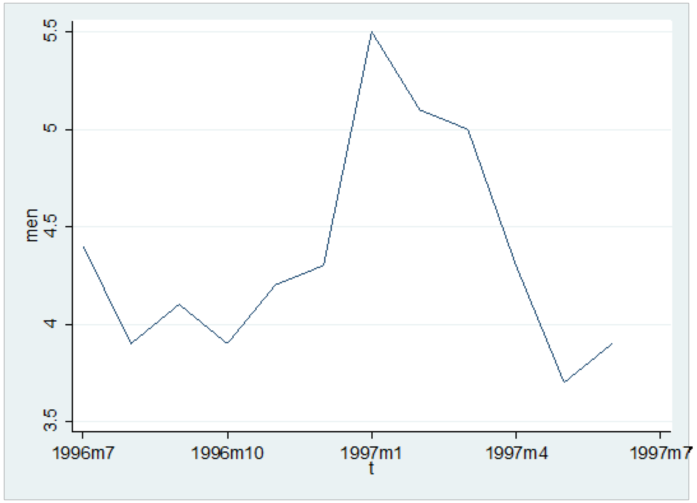
U.S. Unemployment Rate
Men, 20+ years, 1948-2009

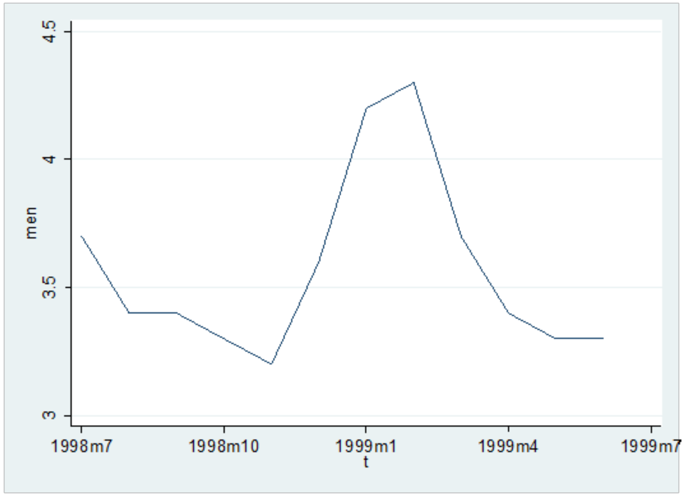


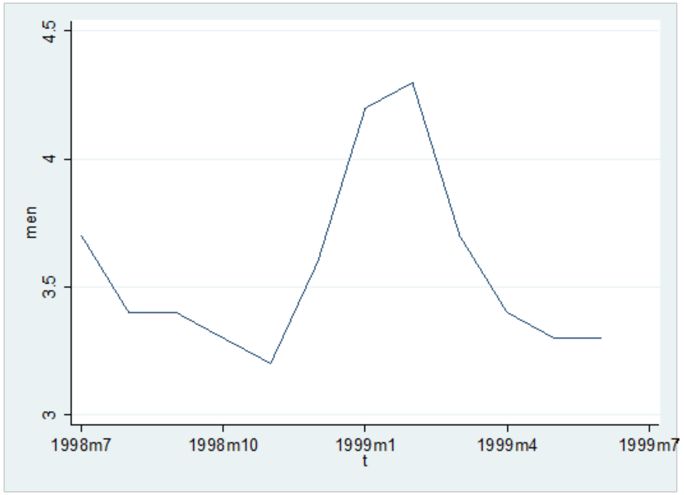
Unemployment Rate, 1990-2001

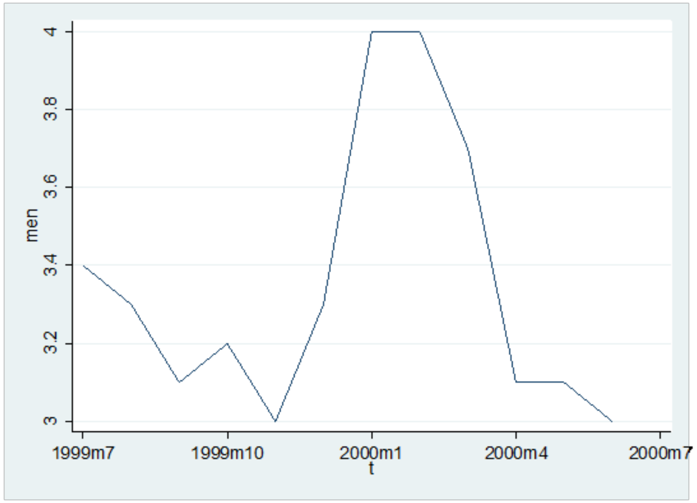


Unemployment Rate, by year

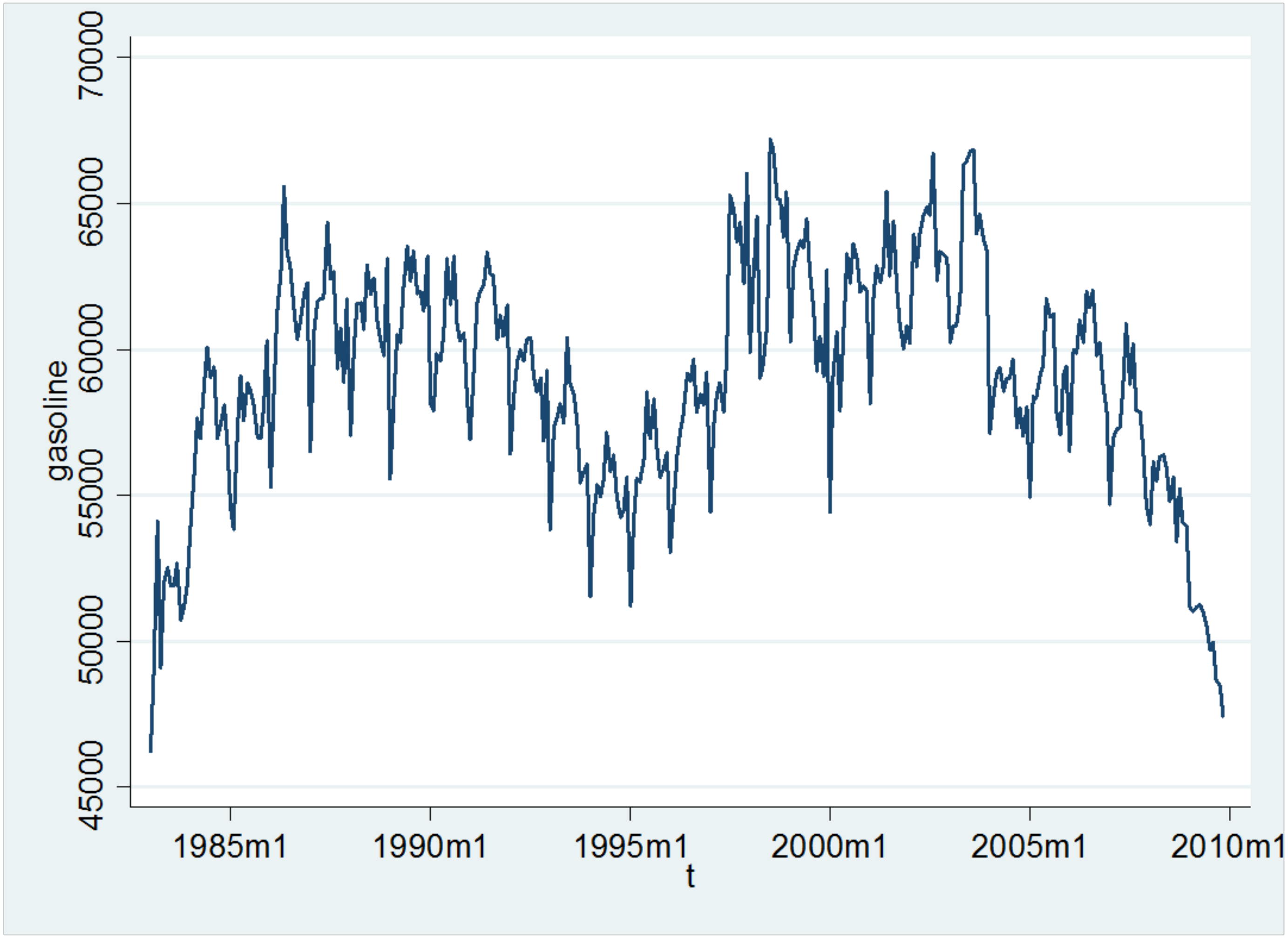




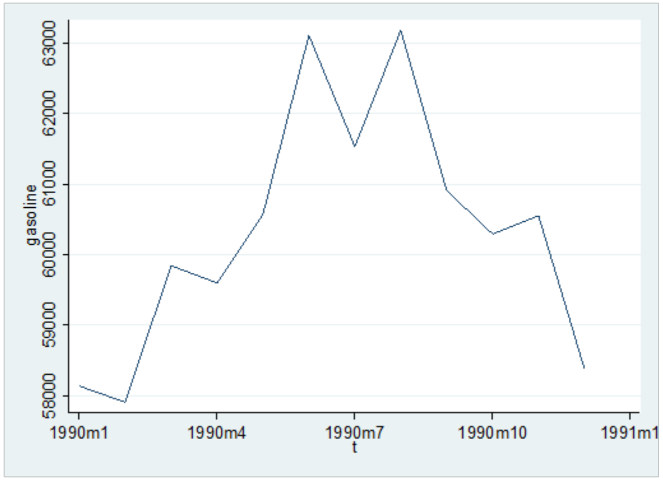


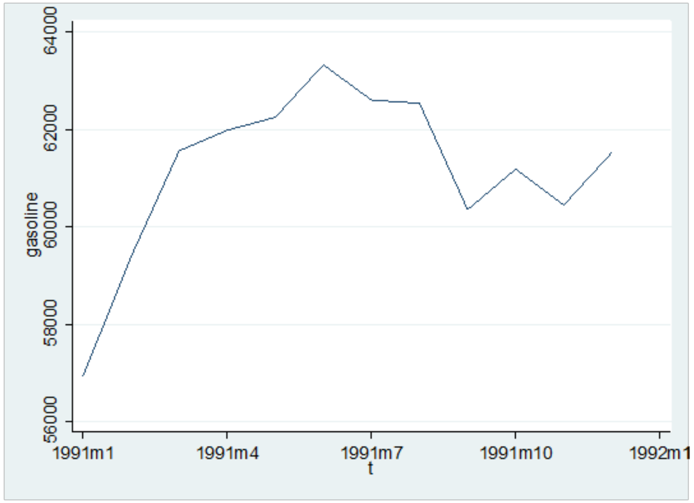


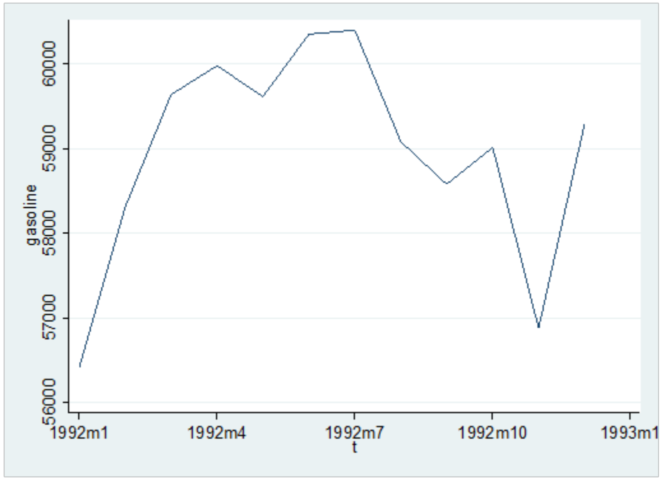
Example 2
U.S. Gasoline Sales Volume

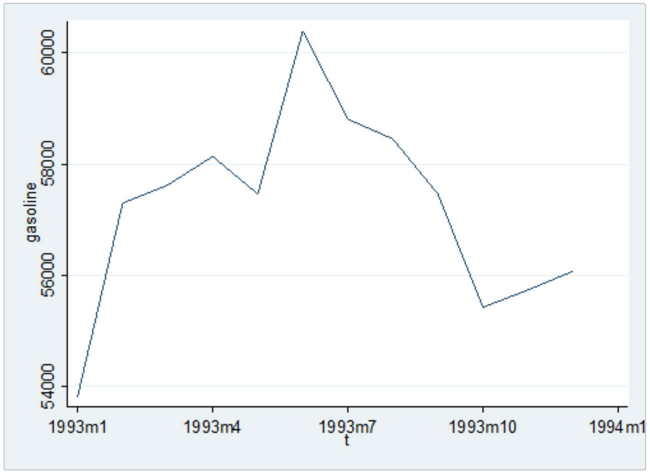


Gasoline Sales, by year

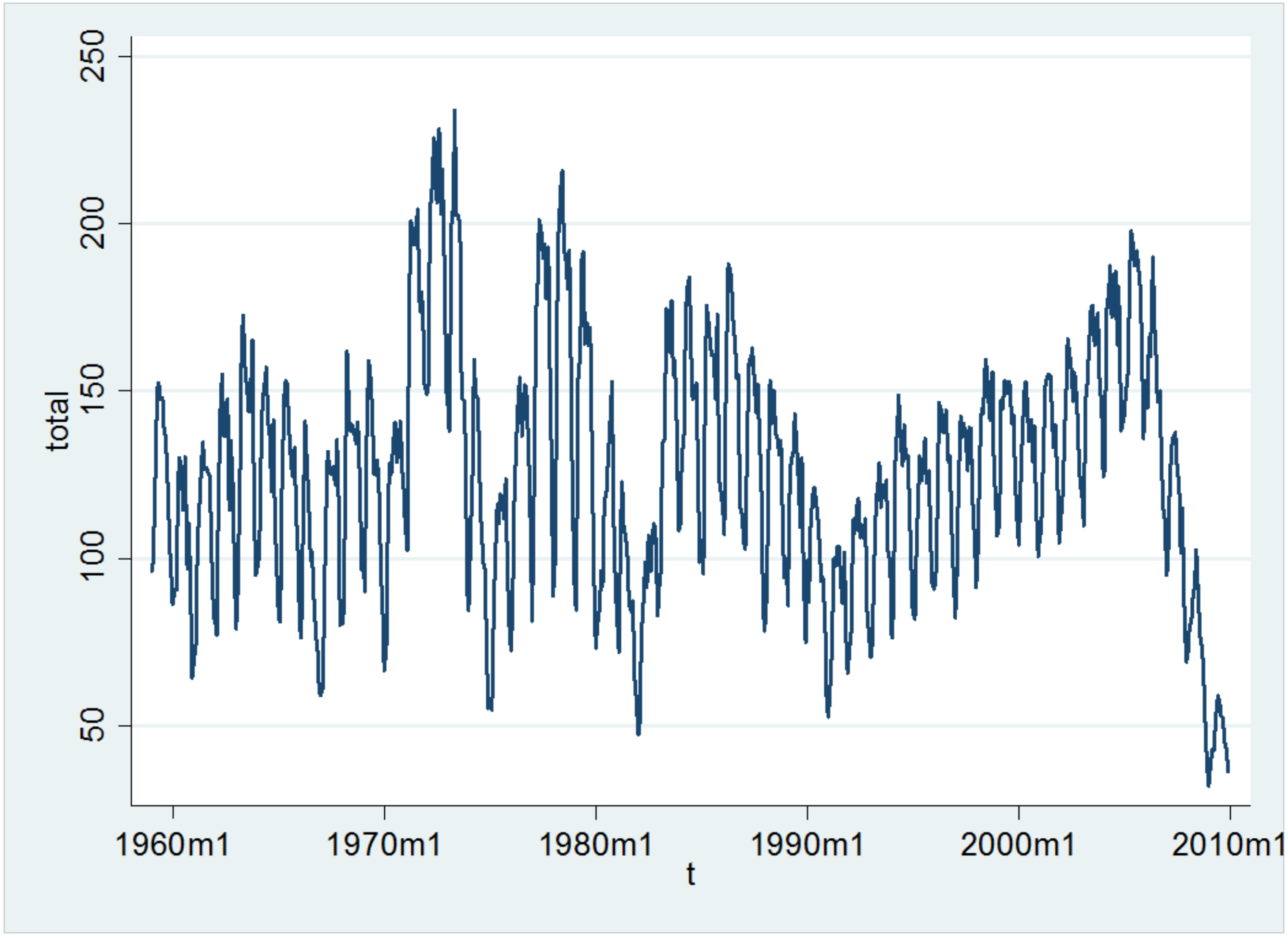




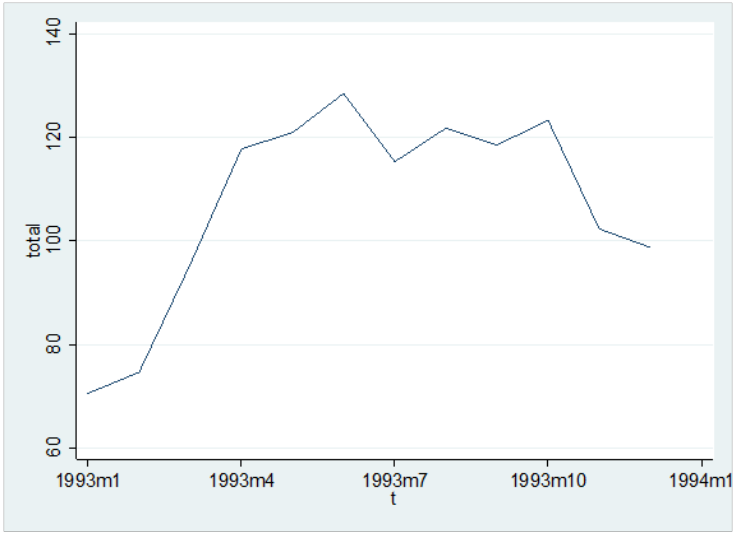


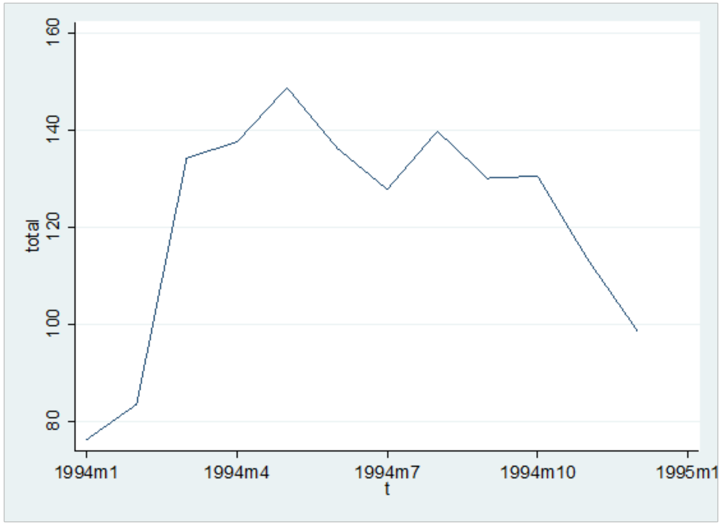


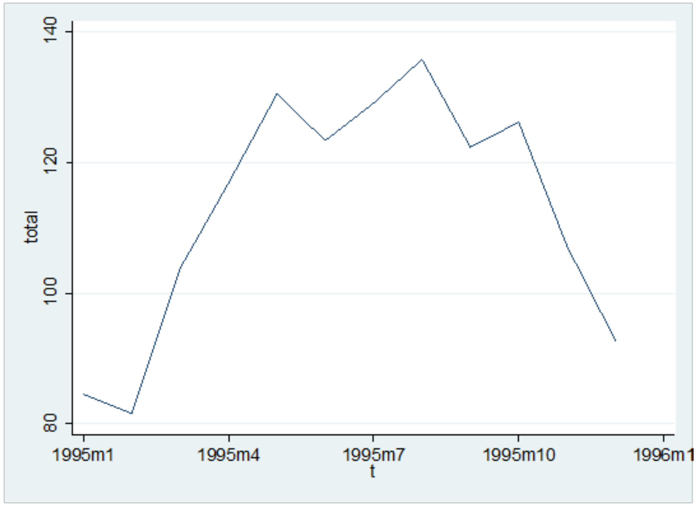
Example 3
U.S. Housing Starts
(New Privately Owned Housing Units)

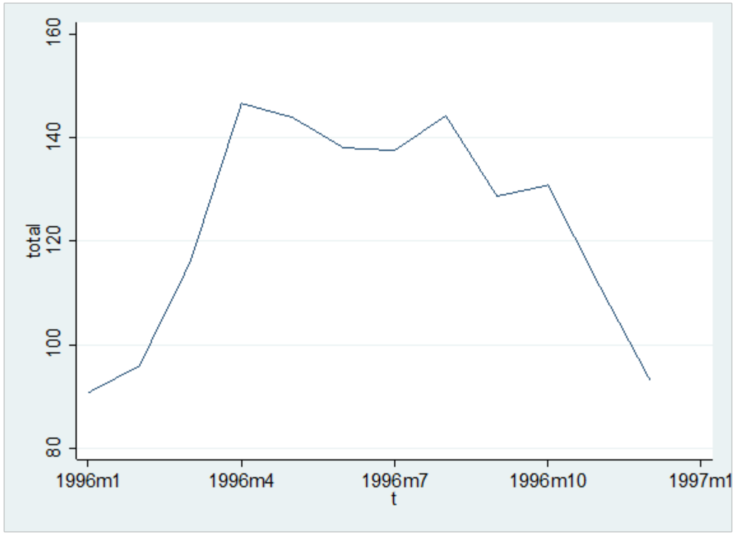


Housing Starts, by year









Deterministic Seasonality

- If seasonality is constant and deterministic then S_t is simply a different constant for each period
- For example, for monthly data

$$S_t = \begin{cases} \gamma_1 & if \quad t = January \\ \gamma_2 & if \quad t = February \\ \vdots & \qquad \vdots \\ \gamma_{12} & if \quad t = December \end{cases}$$

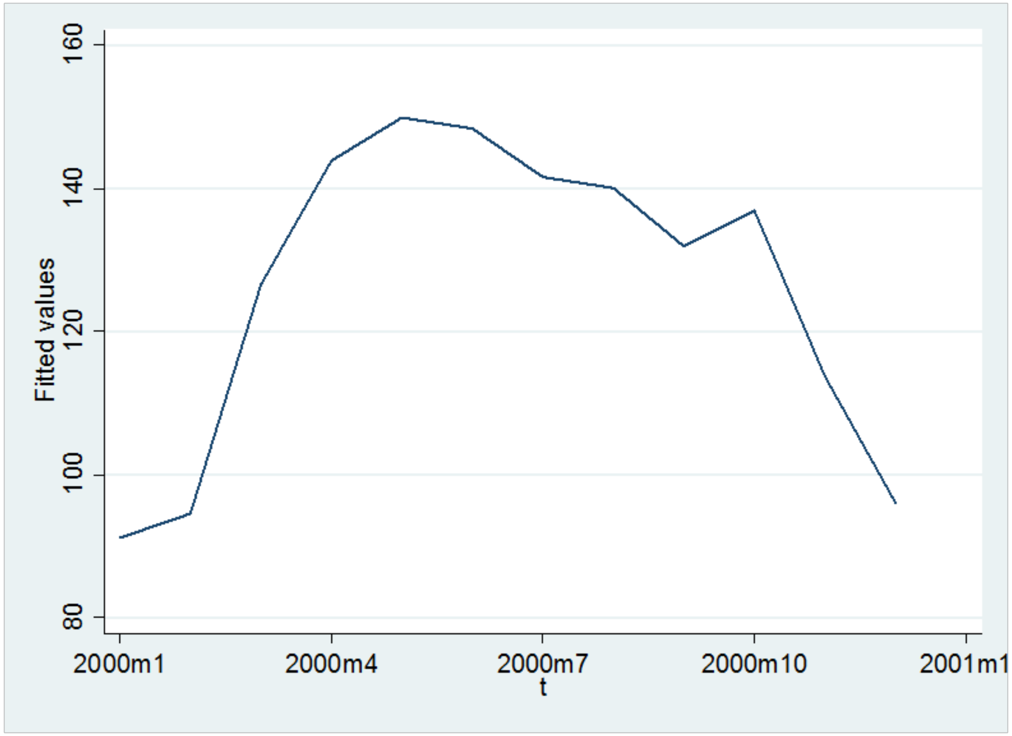
- Seasonality is a constant which varies by the calendar period (quarter, month, week, day, or time of day)

Fitted Values and Forecasts
Pure Deterministic Seasonality

- In the simple pure deterministic seasonality model, fitted values and forecasts are the simple seasonal pattern

Example – Housing Starts

January	91
February	95
March	127
April	144
May	150
June	148
July	142
August	140
September	132
October	137
November	114
December	96



Seasonal Dummy Model

- Deterministic seasonality S_t can be written as a function of seasonal dummy variables
- Let s be the seasonal frequency
  - s=4 for quarterly
  - s=12 for monthly
- Let D_{1t} , D_{2t} , D_{3t} ,..., D_{st} be seasonal dummies
  - D_{1t}= 1 if s is the first period, otherwise D_{1t}= 0
  - D_{2t}= 1 if s is the second period, otherwise D_{2t}= 0
- At any time period t, one of the seasonal dummies D_{1t} , D_{2t} , D_{3t} ,..., D_{st} will equal 1, all the others will equal 0.

Seasonal Dummy Model

- Deterministic seasonality

$$S_t = \begin{cases} \gamma_1 & if \quad t = January \\ \gamma_2 & if \quad t = February \\ \vdots & \qquad \vdots \\ \gamma_{12} & if \quad t = December \end{cases}$$

$$= \sum_{i=1}^{s} \gamma_i D_{it}$$

a linear function of the dummy variables

Estimation

- Least squares regression

$$y_{t+h} = \sum_{i=1}^{s} \gamma_i D_{it} + e_t$$

$$= \alpha + \sum_{i=1}^{s-1} \beta_i D_{it} + e_t$$

- You can either
  - Regress y on all the seasonal dummies, omitting the intercept, or
  - Regress y on an intercept and the seasonal dummies, omitting one dummy (one season, e.g. December)
- You cannot regress on both the intercept plus all seasonal dummies, for they would be collinear and redundant.

Interpreting Coefficients

- In the model

$$S_t = \alpha + \sum_{i=1}^{s-1} \beta_i D_{it}$$

  the intercept $\alpha=\gamma_s$ is the seasonality in the omitted season.
- The coefficients $\beta_i=\gamma_i-\gamma_s$ are the difference in the seasonal component from the s'th period.

STATA Programming

- If the time index is *t* and is formatted as a time index, you can determine the period using the commands

```
generate m=month(dofm(t))
generate q=quarter(dofq(t))
```

for monthly and quarterly data, respectively

(See dates and times in STATA Data manual)

Creating Dummies

- If *m* is the month (1 for January, 2 for February, etc.), then
  - **generate m1=(m==1)**
  - This creates a dummy variable “m1” for January
  - Then
  - **regress y m1 m2 m3 m4 m5 m6 m7 m8 m9 m10 m11**
    or
  - **regress y m1 m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12, noconstant**
- Easier
  - Type “b12.m” in the regressor list
  - **regress y b12.m**
  - This includes dummies for months 1 through 11, omits 12
  - Same as mechanically listing the eleven dummies, but easier.
  - It is important that “m” be the numerical month (1 for January, 2 for February, etc.)

Estimation

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

. regress total b12.m
```

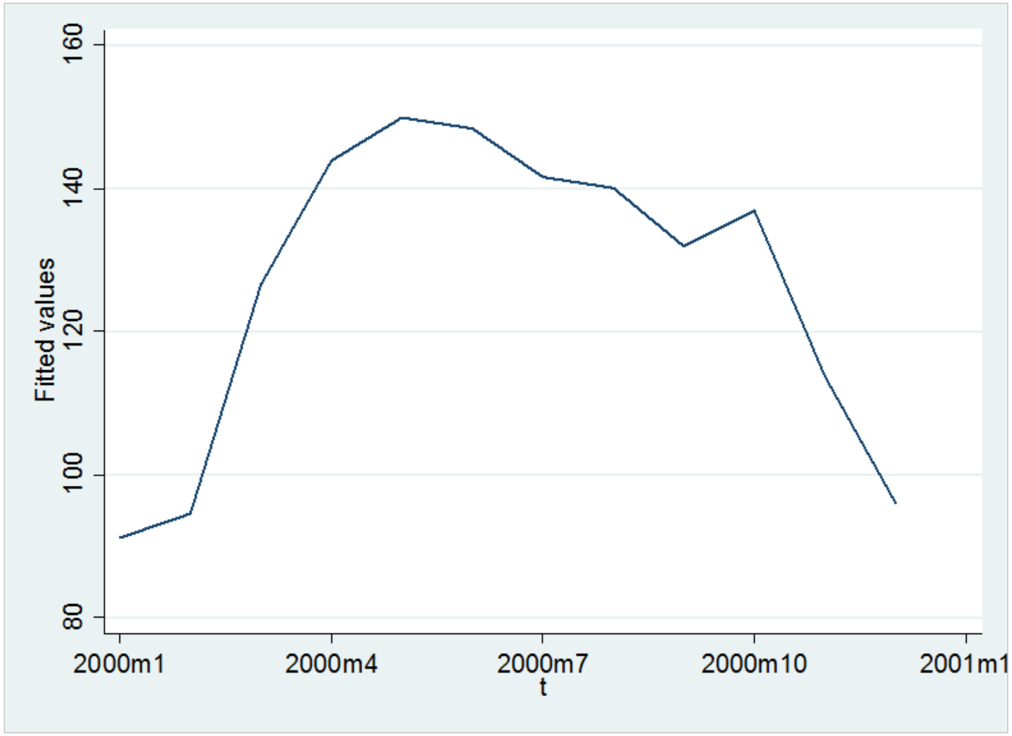
Source	SS	df	MS
Model	267331.386	11	24302.8533
Residual	557738.603	600	929.564339
Total	825069.989	611	1350.36005

Number of obs =	612
F(11, 600) =	26.14
Prob > F =	0.0000
R-squared =	0.3240
Adj R-squared =	0.3116
Root MSE =	30.489

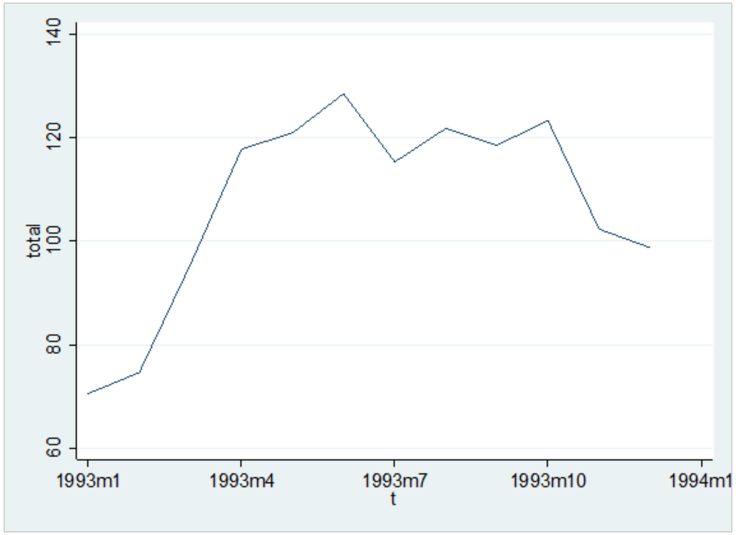
total	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-4.931373	6.037674	-0.82	0.414	-16.78891	6.92617
2	-1.547058	6.037674	-0.26	0.798	-13.4046	10.31048
3	30.51765	6.037674	5.05	0.000	18.66011	42.37519
4	47.82353	6.037674	7.92	0.000	35.96599	59.68107
5	53.87255	6.037674	8.92	0.000	42.01501	65.73009
6	52.31569	6.037674	8.66	0.000	40.45815	64.17323
7	45.55294	6.037674	7.54	0.000	33.6954	57.41048
8	43.95294	6.037674	7.28	0.000	32.0954	55.81048
9	35.82745	6.037674	5.93	0.000	23.96991	47.68499
10	40.84902	6.037674	6.77	0.000	28.99148	52.70656
11	17.64706	6.037674	2.92	0.004	5.789517	29.5046
_cons	96.07843	4.26928	22.50	0.000	87.69388	104.463

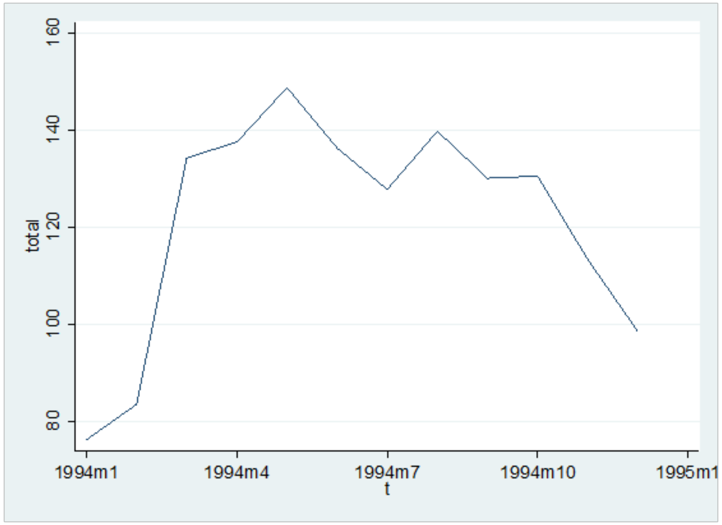
Estimated Seasonality – Housing Starts

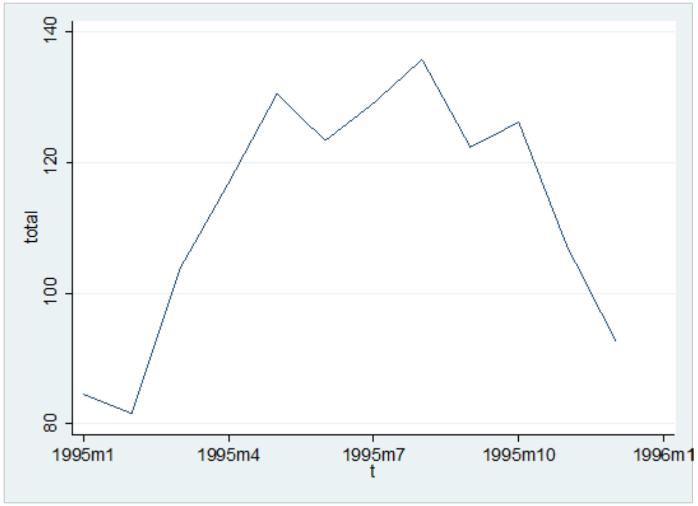
January	91
February	95
March	127
April	144
May	150
June	148
July	142
August	140
September	132
October	137
November	114
December	96



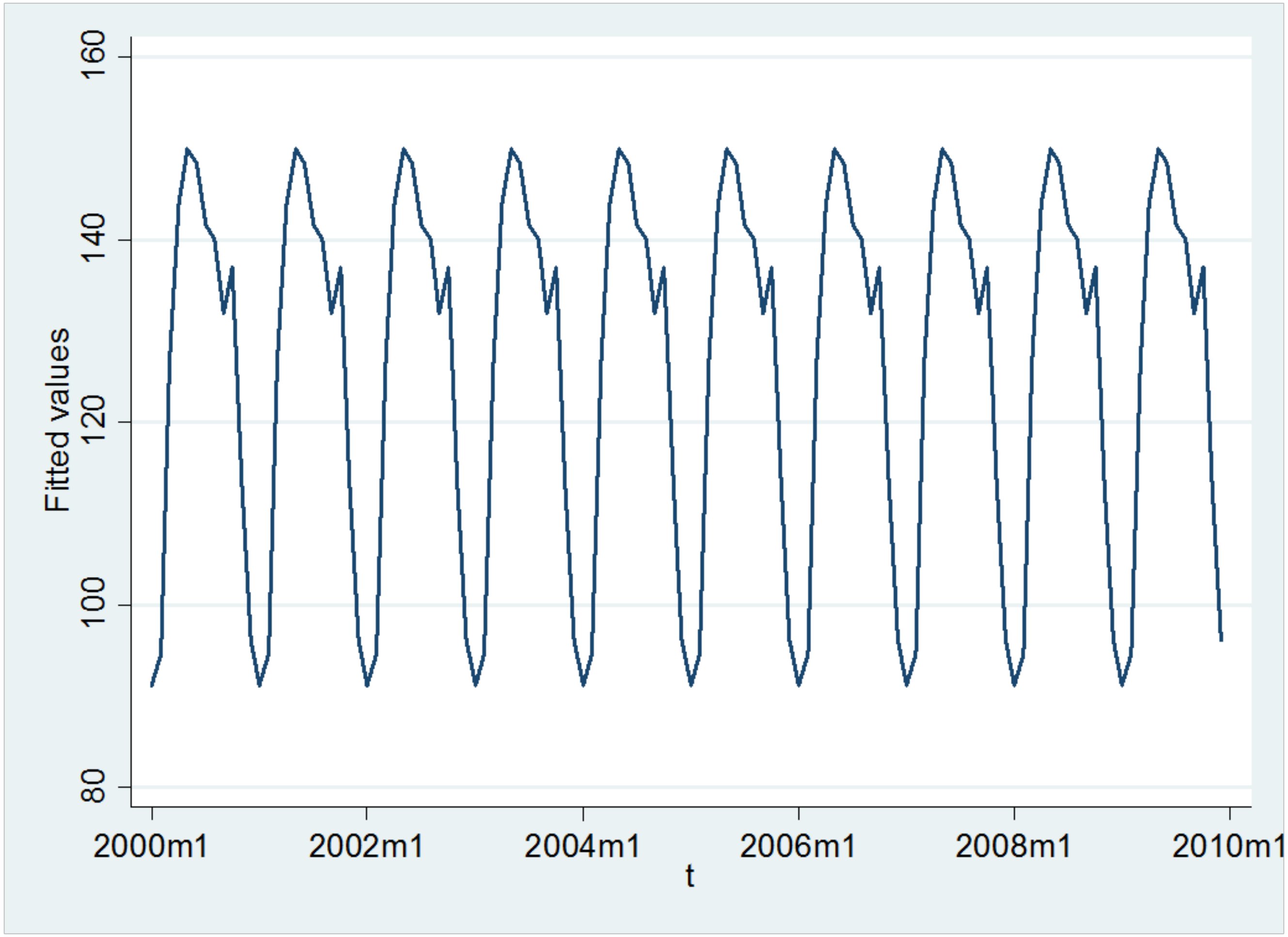
Housing Starts, by year, and estimated seasonality







Predicted Values



Example 1
Unemployment Rate

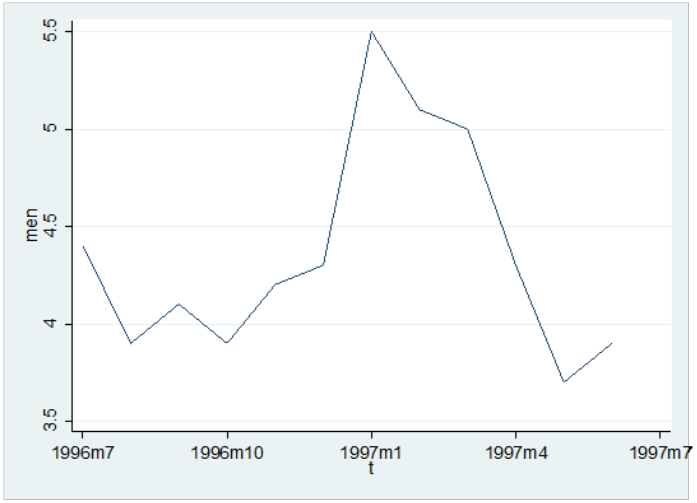
```
. use ur_nsa

. regress ur b12.m
```

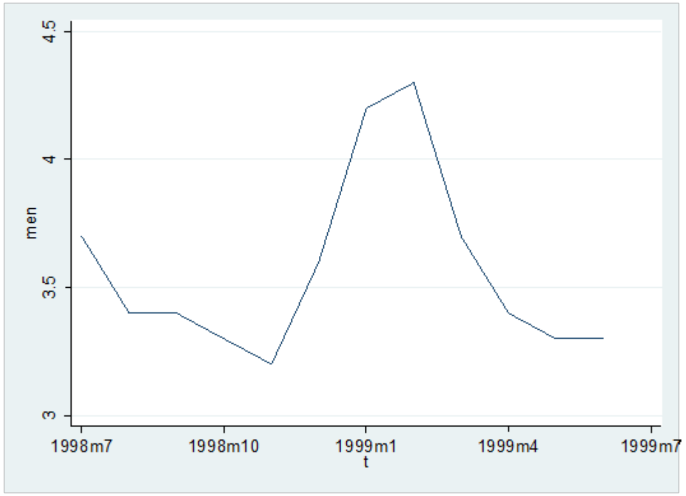
Source	SS	df	MS
Model	105.325822	11	9.57507469
Residual	1571.36666	733	2.14374715
Total	1676.69248	744	2.25361893

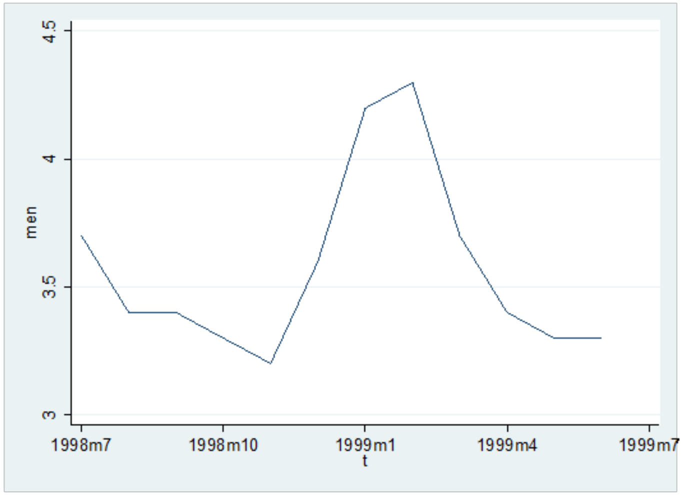
Number of obs	=	745
F(11, 733)	=	4.47
Prob > F	=	0.0000
R-squared	=	0.0628
Adj R-squared	=	0.0488
Root MSE	=	1.4642

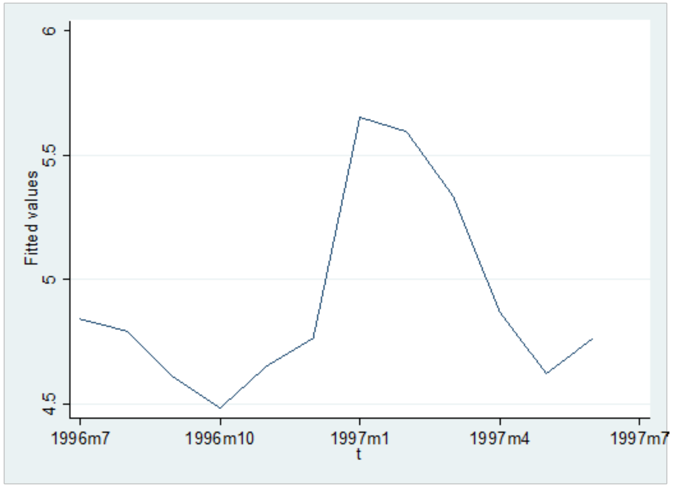
ur	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	.8878648	.2619242	3.39	0.001	.3736537	1.402076
2	.8306451	.2629698	3.16	0.002	.3143813	1.346909
3	.5709677	.2629698	2.17	0.030	.0547039	1.087232
4	.1048387	.2629698	0.40	0.690	-.4114251	.6211026
5	-.1435484	.2629698	-0.55	0.585	-.6598123	.3727154
6	-.0016129	.2629698	-0.01	0.995	-.5178768	.514651
7	.0758064	.2629698	0.29	0.773	-.4404574	.5920703
8	.0274194	.2629698	0.10	0.917	-.4888445	.5436832
9	-.1580645	.2629698	-0.60	0.548	-.6743284	.3581993
10	-.2822581	.2629698	-1.07	0.283	-.7985219	.2340058
11	-.1129032	.2629698	-0.43	0.668	-.6291671	.4033606
_cons	4.764516	.1859478	25.62	0.000	4.399462	5.12957

Unemployment Rate, by year, and estimated seasonality

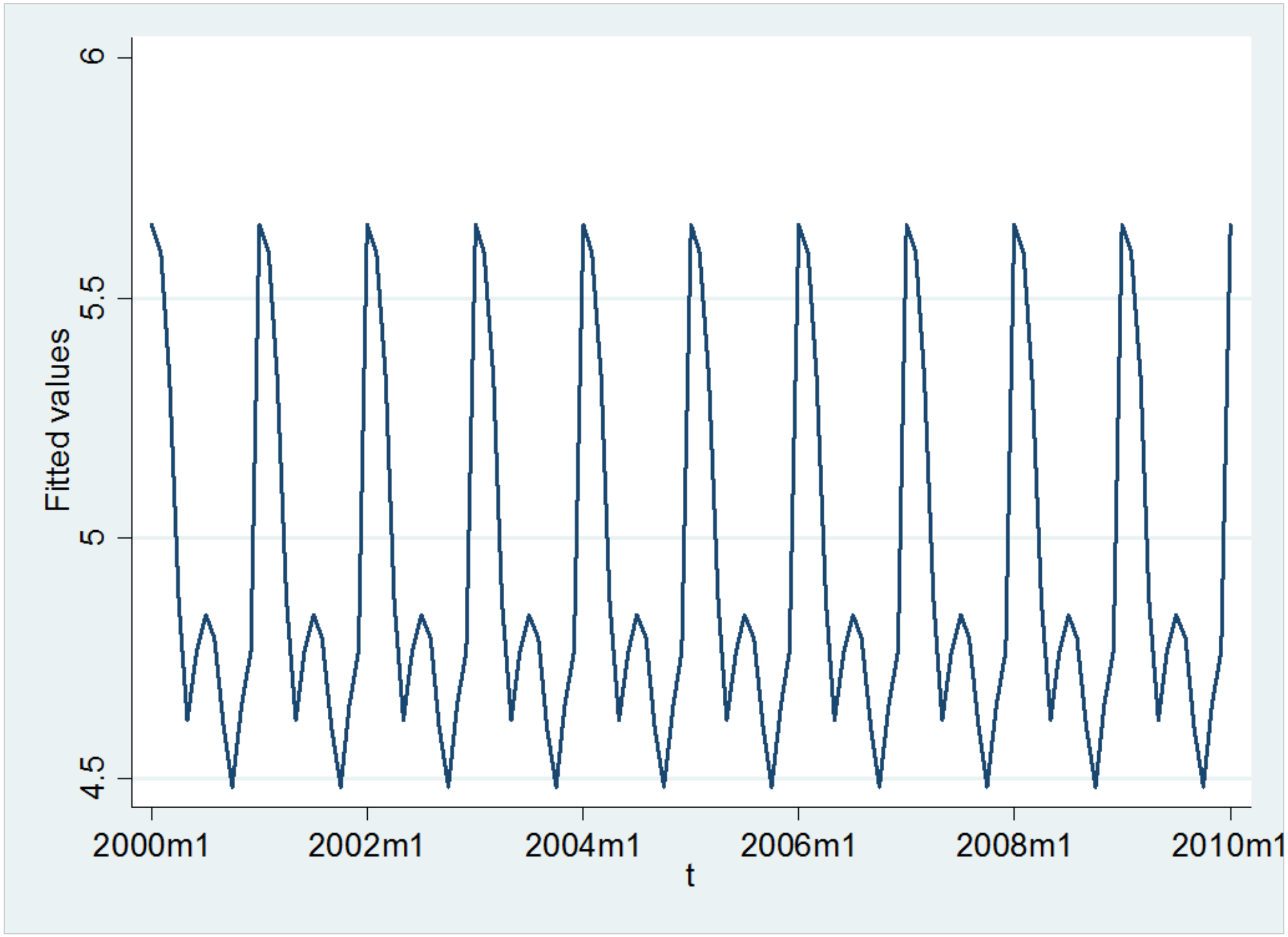








Predicted Values



Example 2
Gasoline Sales

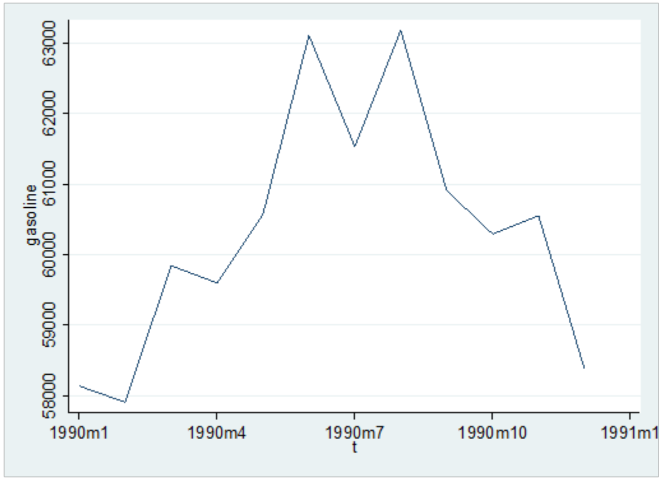
```
. use gasoline

. regress gasoline b12.m
```

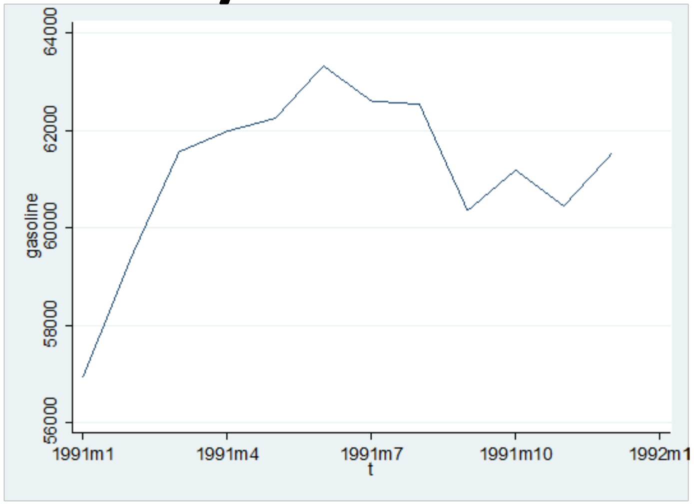
Source	SS	df	MS
Model	636131968	11	57830178.9
Residual	4.1919e+09	311	13478698.6
Total	4.8280e+09	322	14993811.3

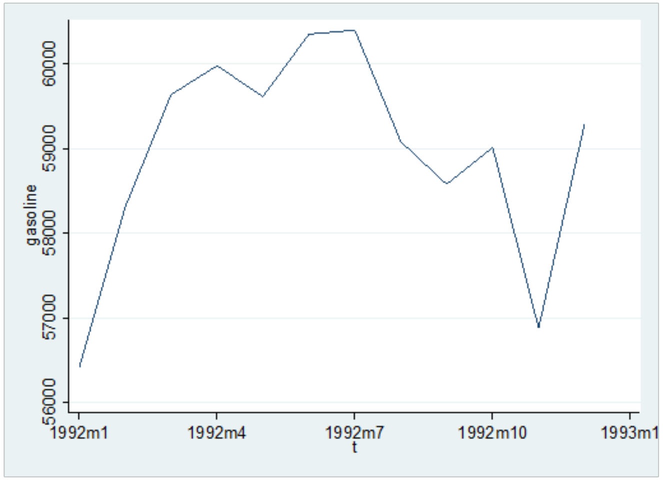
Number of obs	=	323
F(11, 311)	=	4.29
Prob > F	=	0.0000
R-squared	=	0.1318
Adj R-squared	=	0.1010
Root MSE	=	3671.3

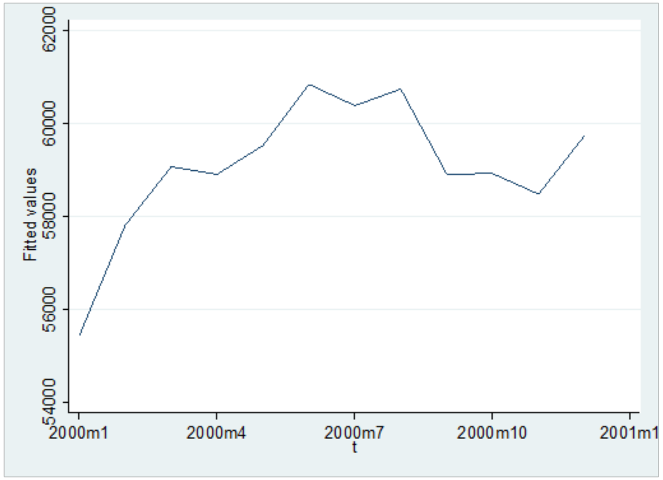
gasoline	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
m						
1	-4309.025	1008.773	-4.27	0.000	-6293.908	-2324.142
2	-1928.984	1008.773	-1.91	0.057	-3913.867	55.8983
3	-671.3583	1008.773	-0.67	0.506	-2656.241	1313.524
4	-829.025	1008.773	-0.82	0.412	-2813.908	1155.858
5	-210.0881	1008.773	-0.21	0.835	-2194.971	1774.795
6	1102.401	1008.773	1.09	0.275	-882.4817	3087.284
7	644.1563	1008.773	0.64	0.524	-1340.726	2629.039
8	1003.964	1008.773	1.00	0.320	-980.9188	2988.847
9	-822.2439	1008.773	-0.82	0.416	-2807.127	1162.639
10	-817.0473	1008.773	-0.81	0.419	-2801.93	1167.835
11	-1260.781	1008.773	-1.25	0.212	-3245.663	724.102
_cons	59735.28	720.008	82.96	0.000	58318.57	61151.98

Gasoline Sales, by year, and estimated seasonality

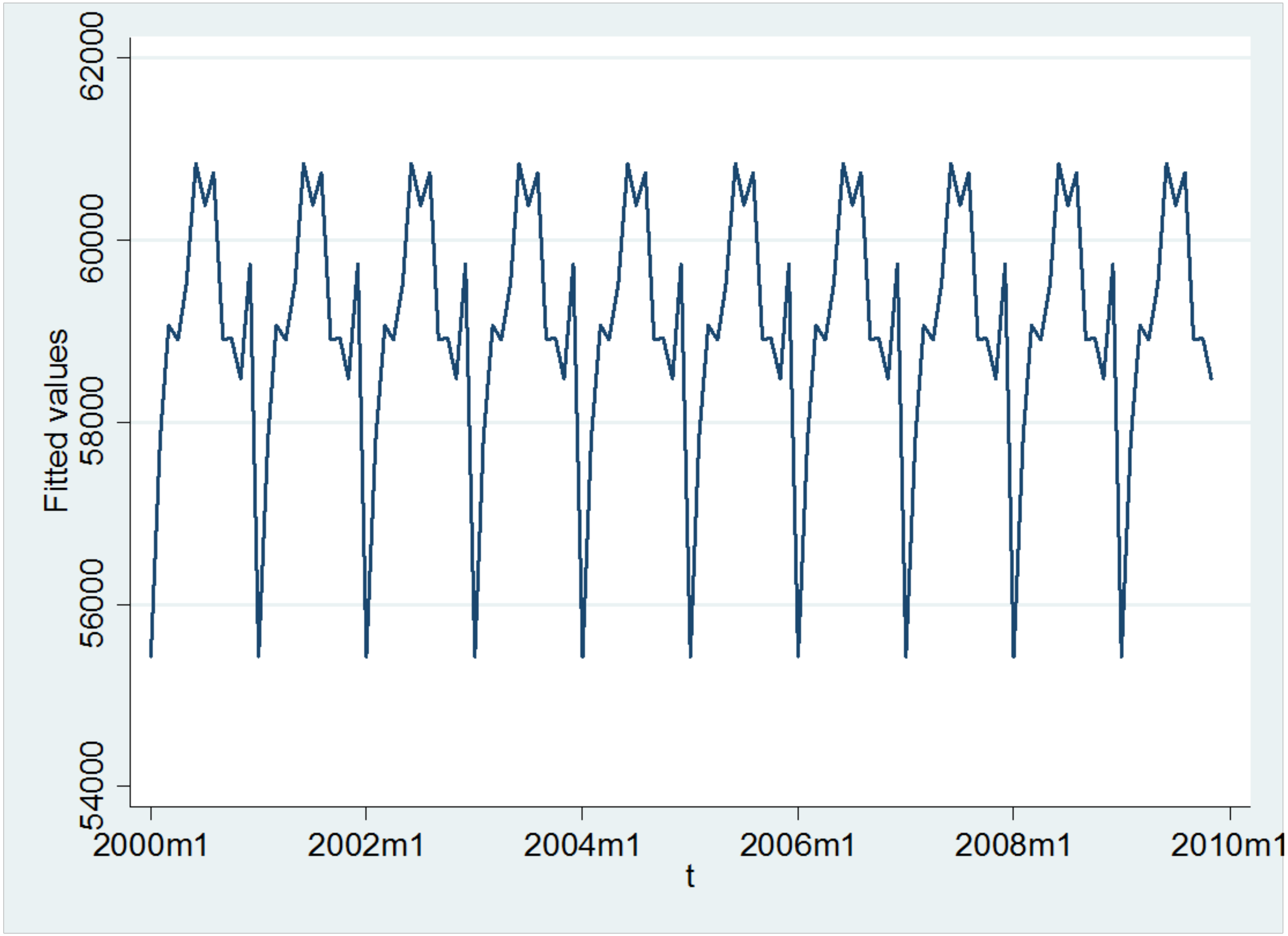








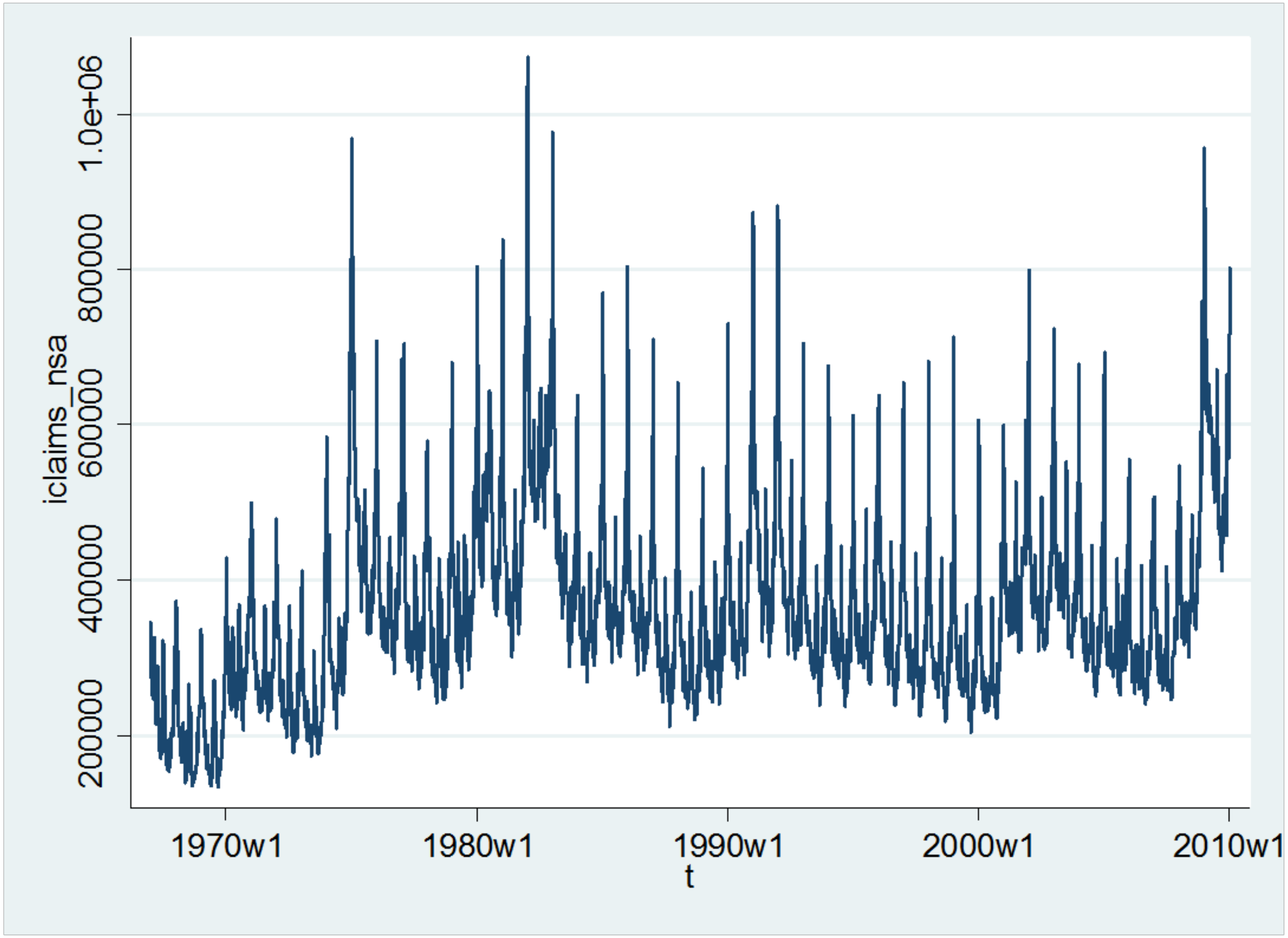
Predicted Values



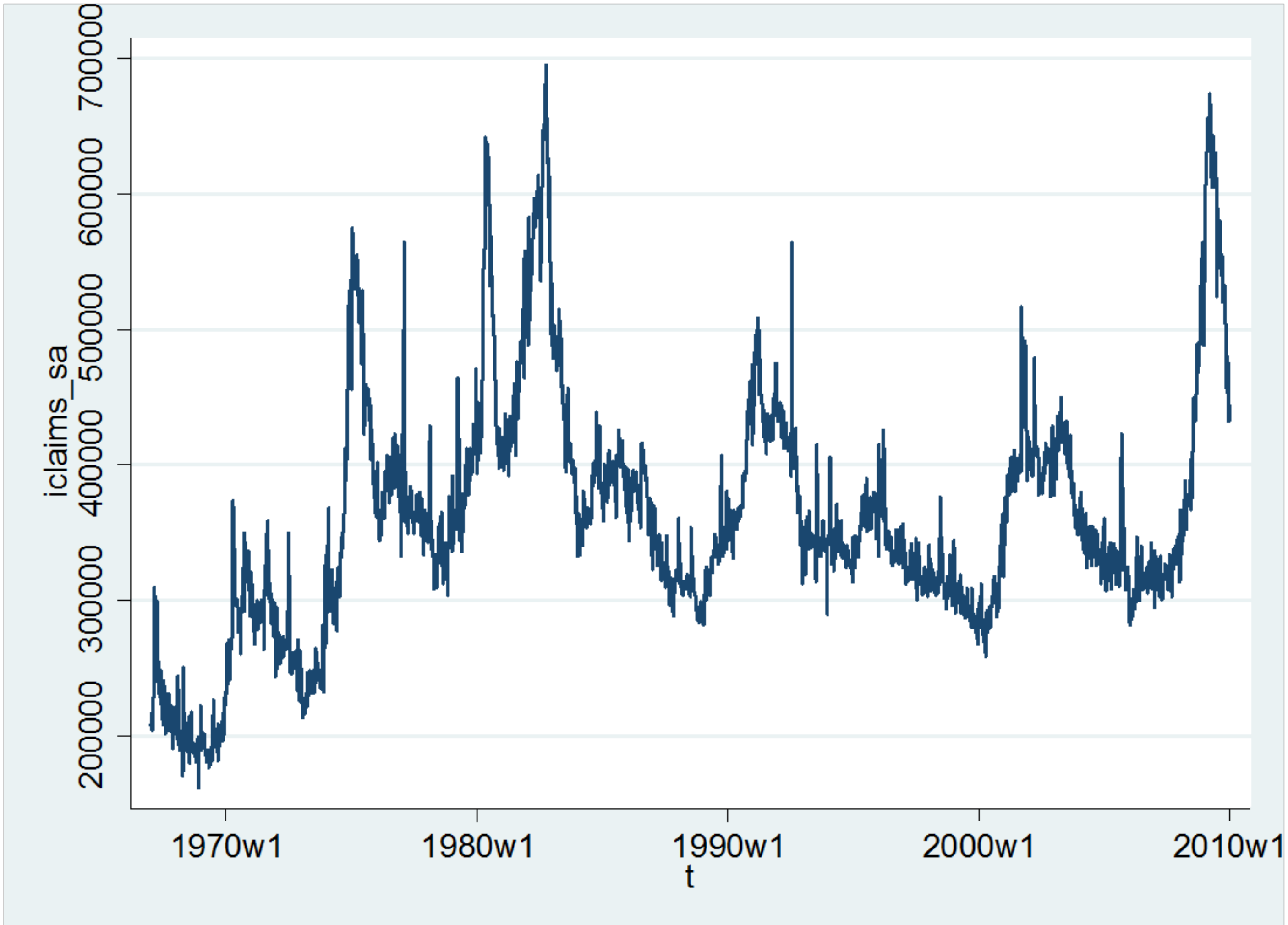
Application – Weekly Data

- Unemployment Insurance Claims
- Department of Labor
- Issued Weekly
- Important indicator for unemployment

Unemployment Claims
Not Seasonally Adjusted



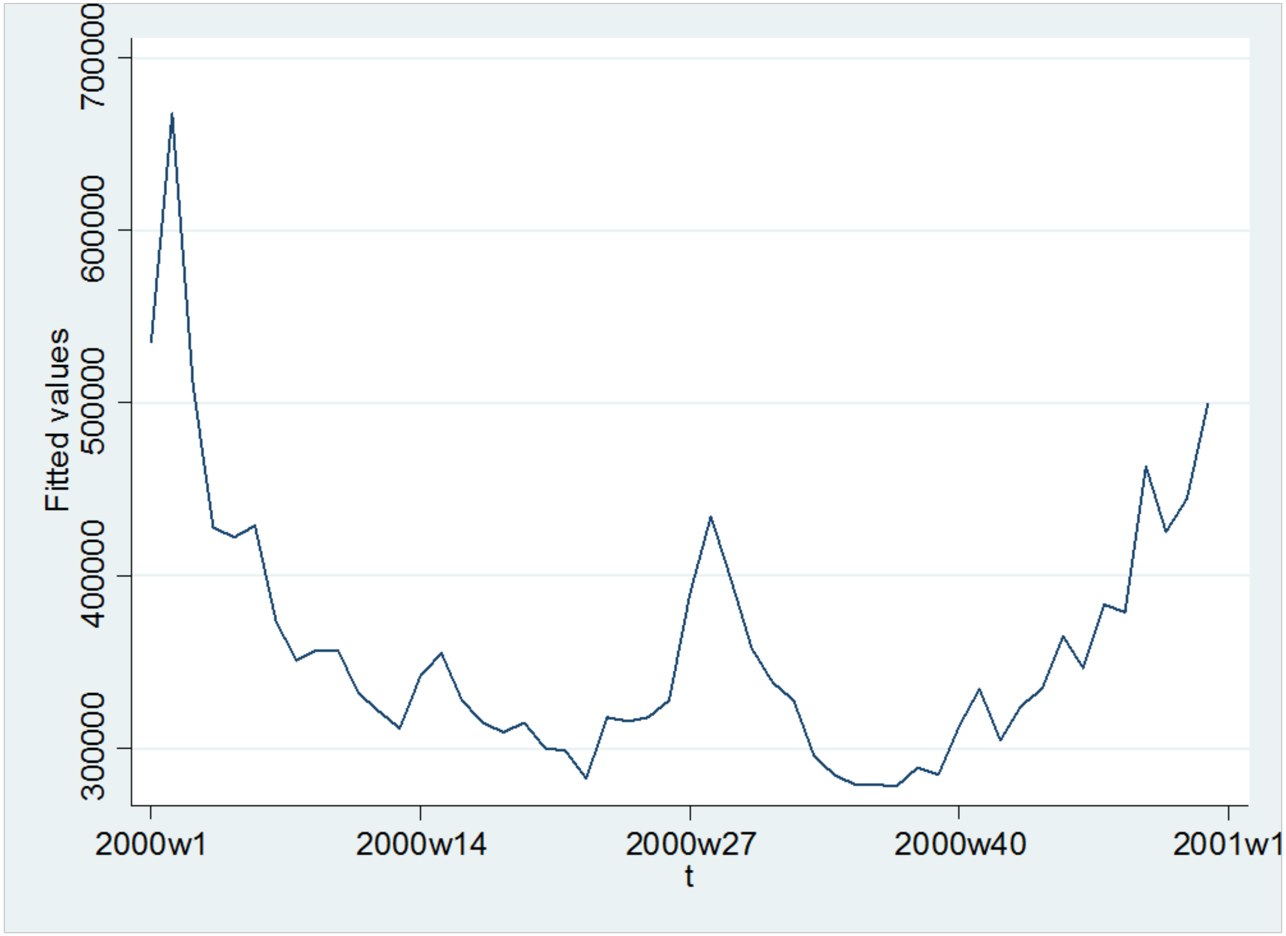
Unemployment Claims
Official Seasonally Adjusted Series



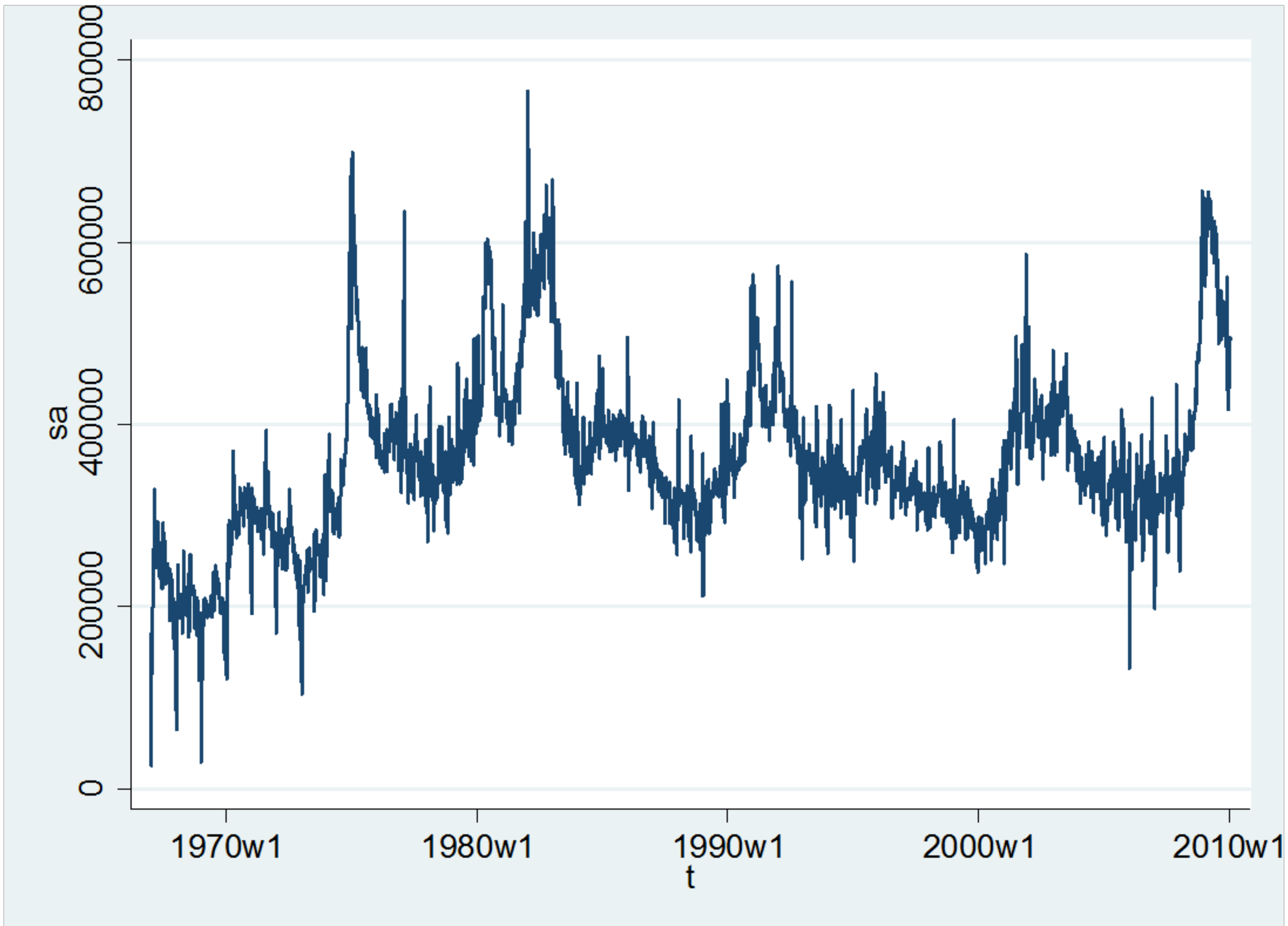
Estimation

```
. regress iclaims_nsa b52.w

      Source |       SS       df       MS              Number of obs =    2238
-------------+------------------------------           F( 51,  2186) =   29.09
       Model |  1.2804e+13    51  2.5105e+11           Prob > F      =  0.0000
    Residual |  1.8865e+13  2186  8.6297e+09           R-squared     =  0.4043
-------------+------------------------------           Adj R-squared =  0.3904
       Total |  3.1668e+13  2237  1.4157e+10           Root MSE      =   92896

------------------------------------------------------------------------------
 iclaims_nsa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           w |
          1  |   35954.91   19920.38     1.80   0.071    -3109.952    75019.78
          2  |   168185.3   19920.38     8.44   0.000     129120.4    207250.2
          3  |   11166.58   20034.54     0.56   0.577    -28122.15    50455.32
          4  |  -71213.05   20034.54    -3.55   0.000    -110501.8   -31924.31
          5  |  -77421.58   20034.54    -3.86   0.000    -116710.3   -38132.85
          6  |  -70397.88   20034.54    -3.51   0.000    -109686.6   -31109.15
          7  |  -125853.5   20034.54    -6.28   0.000    -165142.3    -86564.8
          8  |  -148580.9   20034.54    -7.42   0.000    -187869.6   -109292.2
          9  |  -142809.6   20034.54    -7.13   0.000    -182098.4   -103520.9
         10  |  -142668.5   20034.54    -7.12   0.000    -181957.2   -103379.8
         11  |  -167656.8   20034.54    -8.37   0.000    -206945.6   -128368.1
         12  |  -178125.4   20034.54    -8.89   0.000    -217414.1   -138836.7
         13  |  -187898.8   20034.54    -9.38   0.000    -227187.6   -148610.1
         14  |  -157631.2   20034.54    -7.87   0.000    -196919.9   -118342.5
         15  |  -144329.8   20034.54    -7.20   0.000    -183618.5   -105041.1
         16  |  -171520.5   20034.54    -8.56   0.000    -210809.2   -132231.8
```

Estimated Seasonal Process



Seasonally Adjusted (by Dummy Variable Method)



Other types of seasonality

- Daily data
  - Day of the week
  - Handle by including dummy variables for each day
- High-frequency data
  - Include hourly or time-of-day indicators
- Holiday effects
  - Flower sales big on Valentines Day, Mothers Day, Easter, yet these days can move around
  - Trading-day/business-day variation
    - Number of trading days/business days varies across months
    - Can divide by number of trading days, or include as a regressor