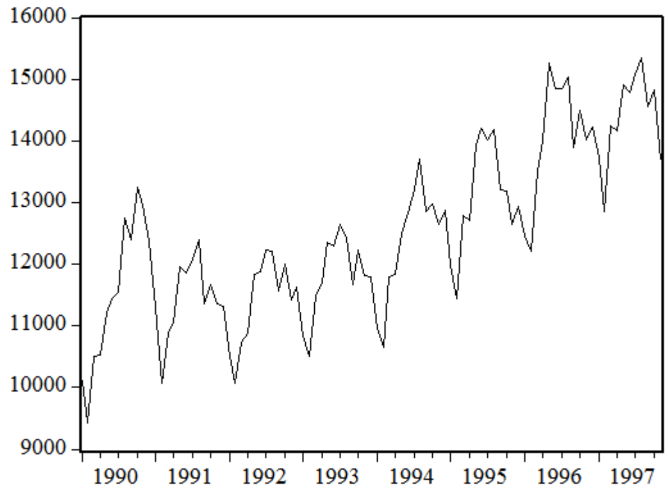
Computer Handout 7: Modeling and Forecasting Seasonality
Diego Escobari
Econ 3342

This Computer Handout 7 will show the use of dummy variables to model and forecast seasonality.

Let the variable of interest be monthly gasoline sales (gs) measured in dollars. This is a typical variable that has seasonal fluctuations; consumption is higher during the summer months.

The following time series graph of gasoline sales from January 1990 through November 1997 illustrates the importance of seasonal component in this variable:



A part of the data set showing the monthly dummy variables is:

obs	GS	D1	D2	D3	D4
1990M01	10120.00	1.000000	0.000000	0.000000	0.000000
1990M02	9434.000	0.000000	1.000000	0.000000	0.000000
1990M03	10497.00	0.000000	0.000000	1.000000	0.000000
1990M04	10537.00	0.000000	0.000000	0.000000	1.000000
1990M05	11210.00	0.000000	0.000000	0.000000	0.000000
1990M06	11442.00	0.000000	0.000000	0.000000	0.000000

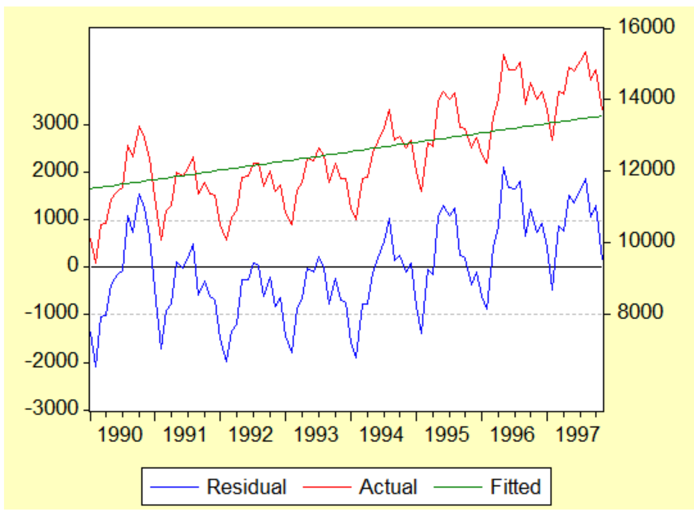
The naive econometric model that just accounts for a linear trend would be: ls gs @trend

Dependent Variable: GS
Method: Least Squares
Date: 10/10/10 Time: 23:51
Sample: 1990M01 1997M11
Included observations: 95

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	21.78026	0.174939	124.5018	0.0000

R-squared	0.499578	Mean dependent var	12499.87
Adjusted R-squared	0.499578	S.D. dependent var	1387.530
S.E. of regression	981.5455	Akaike info criterion	16.62660
Sum squared resid	90562568	Schwarz criterion	16.65349
Log likelihood	-788.7637	Durbin-Watson stat	0.425772

with the following actual, fitted and residuals graph:



Notice how poorly this model fits the data without explaining any of the variation within each year.

The econometric model to account for the seasonality (and a linear time trend) using dummy variables can be estimated with the following EViews command:

```
ls gs @trend d1 d2 d3 d4 d5 d6 d7 d8 d9 d10 d11 d12
```

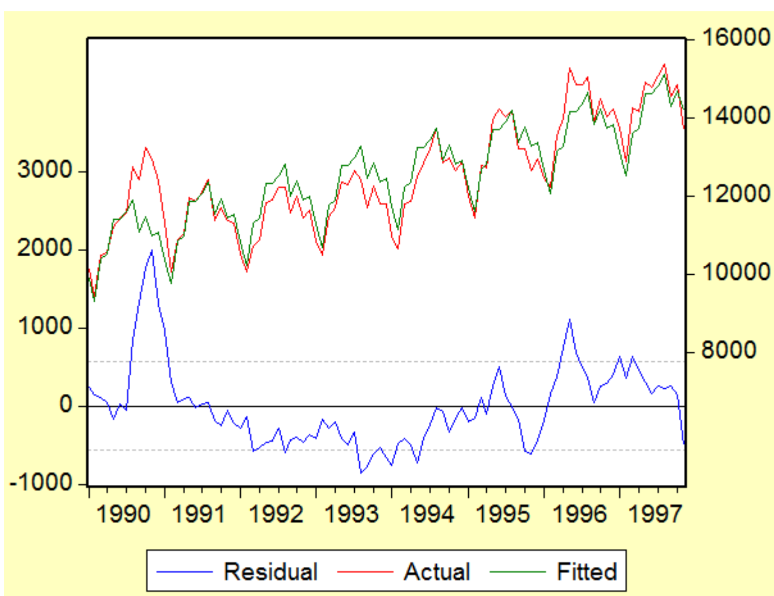
The estimation output is:

Dependent Variable: GS
Method: Least Squares
Date: 10/10/10 Time: 23:56
Sample: 1990M01 1997M11
Included observations: 95

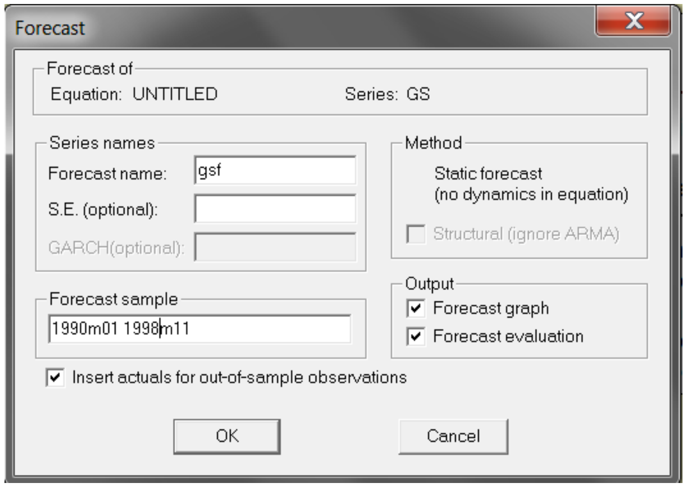
Variable	Coefficient	Std. Error	t-Statistic	Prob.
@TREND	38.28112	2.138905	17.89753	0.0000
D1	-10332.86	1235.613	-8.362539	0.0000
D2	-10955.02	1237.724	-8.850941	0.0000
D3	-9905.677	1239.835	-7.989516	0.0000
D4	-9827.333	1241.945	-7.912856	0.0000
D5	-8979.614	1244.056	-7.218015	0.0000
D6	-8996.145	1246.167	-7.219054	0.0000
D7	-8845.177	1248.278	-7.085904	0.0000
D8	-8580.833	1250.389	-6.862531	0.0000
D9	-9438.614	1252.500	-7.535819	0.0000
D10	-9083.895	1254.611	-7.240405	0.0000
D11	-9637.426	1256.723	-7.668697	0.0000
D12	-9565.360	1248.478	-7.661619	0.0000

R-squared	0.853734	Mean dependent var	12499.87
Adjusted R-squared	0.832329	S.D. dependent var	1387.530
S.E. of regression	568.1602	Akaike info criterion	15.64921
Sum squared resid	26470090	Schwarz criterion	15.99869
Log likelihood	-730.3375	Durbin-Watson stat	0.237427

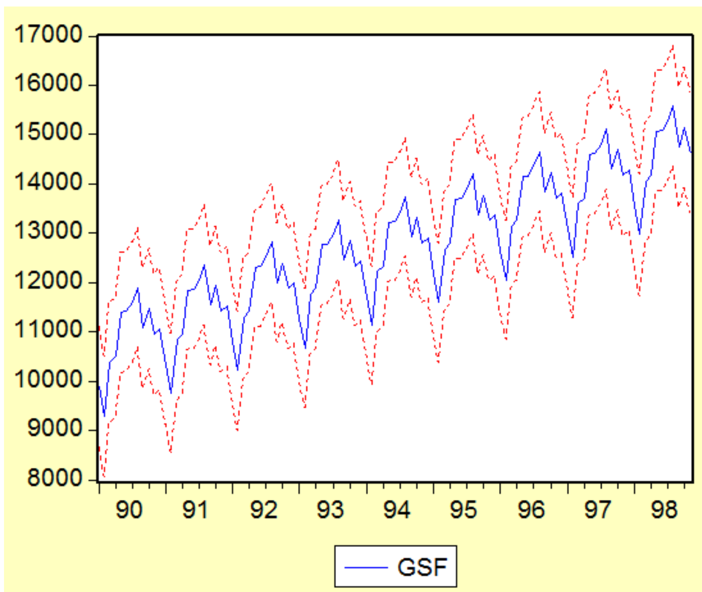
with the following graph for the actual, fitted and residual (actual data, in-sample forecast, forecast error):



To make an out-of-sample forecast, we simply follow the same steps as in Computer Handout 6 to forecast the trend. Just make sure you modify the "Forecast sample" and include some observations beyond time T. In this case from December 1997 through November 1998.



That yields:



Notice that the forecast includes one additional year (from Dec 97 to Nov 98) and that each year has the same forecasted values plus the upward slopping trend.

How to create dummy variables

If the dummy variables d1, d2, ..., d12 are not readily available in the dataset, they can easily be created using the following command:

genr dum1 = @seas(1)

This generates the dummy for the first month. You have to repeat this for all 12 months in the sample: genr dum2 = @seas(2)... until genr dum12 = @seas(12).