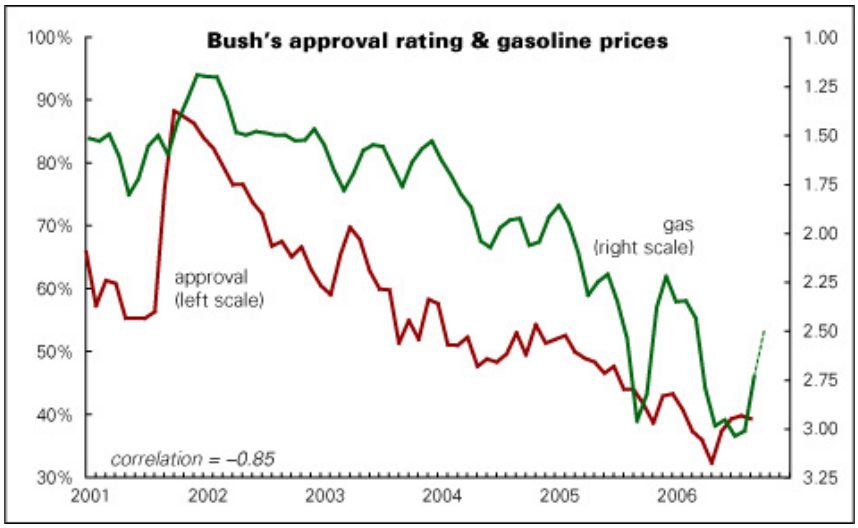
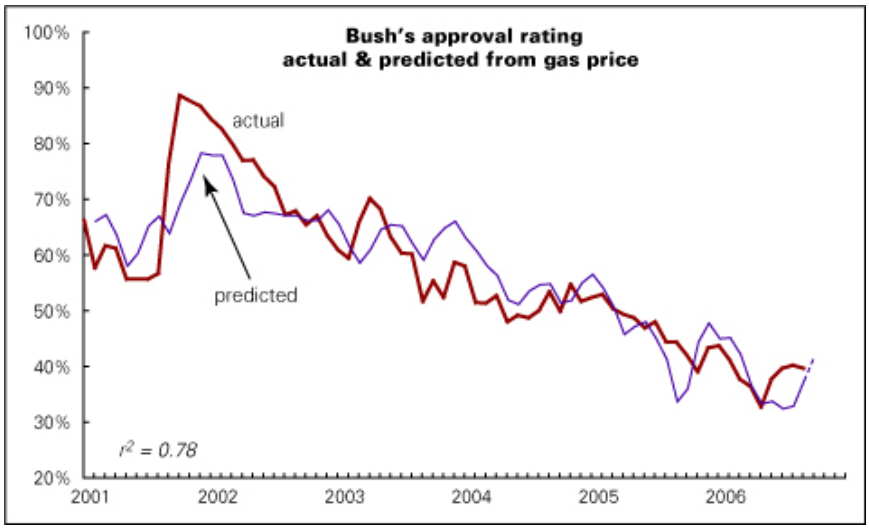
Stat 5100 Handout #33 – SAS: Time Series

Example 1: "Bush and the price of gas"

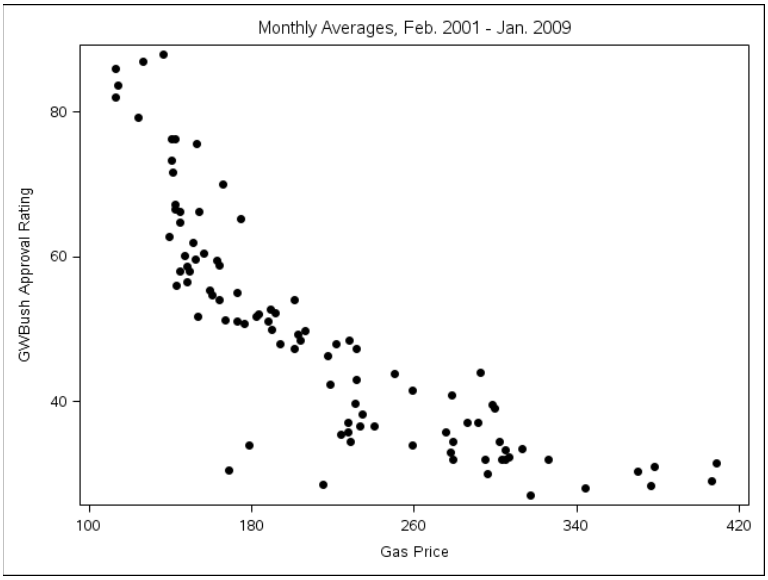
- http://www.leftbusinessobserver.com/BushNGas.html
- "…no occupant of the White House has ever seen his popularity so closely tied to the price of gas."
- "There's no precedent for this tight relationship."



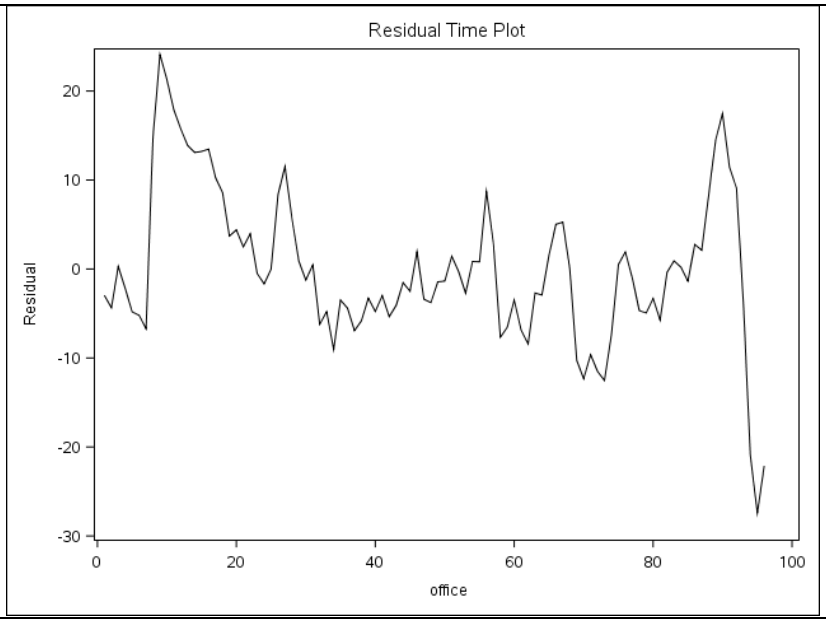


But – can we justify a conclusion that gas price significantly affects approval rating? (HW 7 will address this more completely)

Presidential Approval Ratings and Gas Price

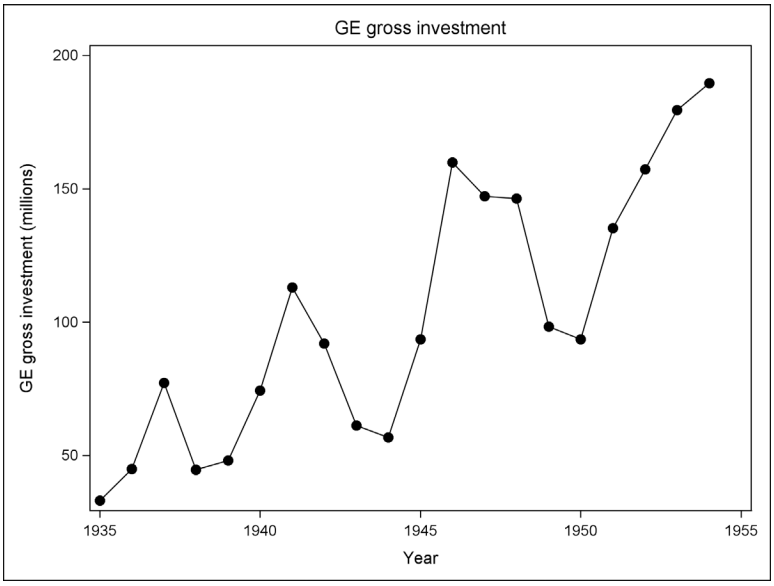


Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	**1**	88.80015	2.82573	31.43	<.0001
price	**1**	-0.18281	0.01242	-14.72	<.0001

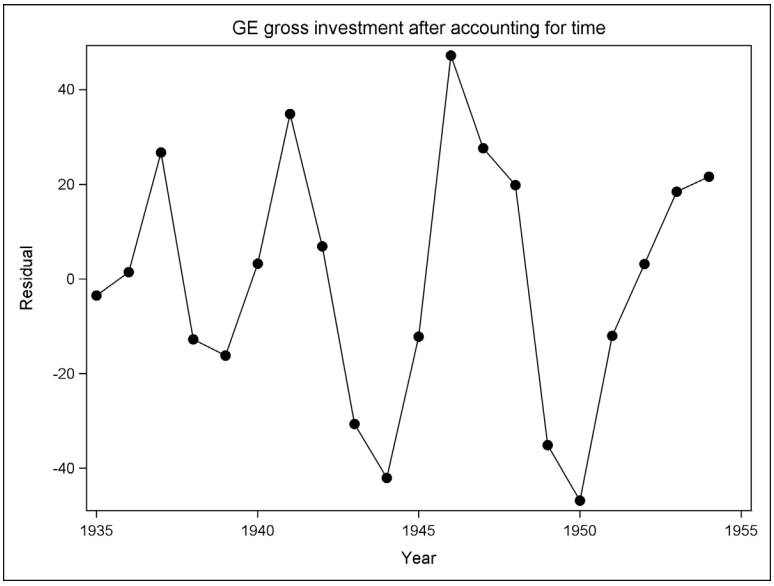


Example 2: General Electric's gross investment (in millions of dollars) for years 1935 – 1954. Originally presented in Grunfeld, Y. (1958), "The Determinants of Corporate Investment," Ph.D. dissertation, University of Chicago; discussed in Boot, J.C.G. (1960), "Investment Demand: An Empirical Contribution to the Aggregation Problem," International Economic Review, 1, 3-30. See also Damodar N. Gujarati, Basic Econometrics, Third Edition, 1995, McGraw-Hill, [1995, pp. 522-525].

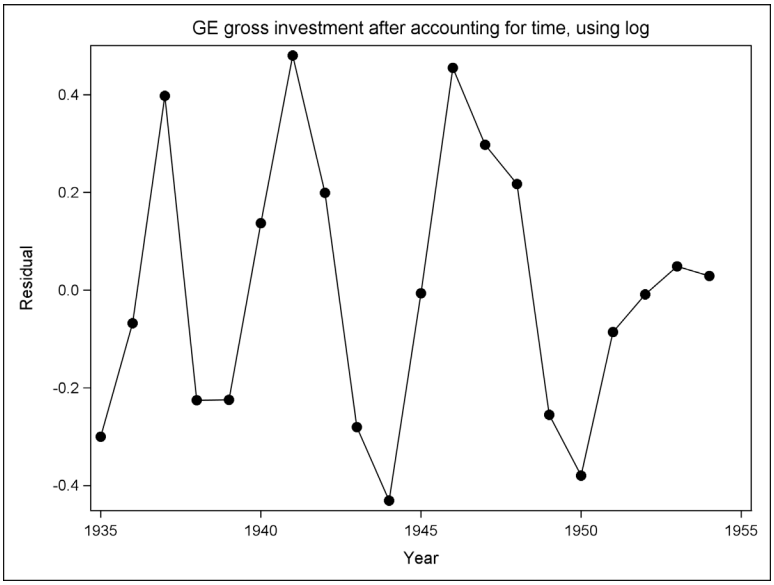
```
data GE; input year GEinv @@; cards;
  1935  33.1   1936   45.0   1937   77.2   1938   44.6   1939   48.1
  1940  74.4   1941  113.0   1942   91.9   1943   61.3   1944   56.8
  1945  93.6   1946  159.9   1947  147.2   1948  146.3   1949   98.3
  1950  93.5   1951  135.2   1952  157.3   1953  179.5   1954  189.6
   ;
proc sgplot data=GE noautolegend;
  scatter y=GEinv x=year /
      markerattrs=(symbol=CIRCLEFILLED size=8pt);
  series y=GEinv x=year / lineattrs=(pattern=solid);
  xaxis label='Year';
  yaxis label='GE gross investment (millions)';
  title1 'GE gross investment';
run;
```




```
/* 1. Make data stationary */
proc reg data=GE noprint;
  model GEinv=year;  output out=a1 r=resid;
  title1 'simple regression on time';
proc sgplot data=a1 noautolegend;
  scatter y=resid x=year /
      markerattrs=(symbol=CIRCLEFILLED size=8pt);
  series y=resid x=year / lineattrs=(pattern=solid);
  xaxis label='Year';    yaxis label='Residual';
  title1 'GE gross investment after accounting for time';
run;
```




```
data GE; set GE;  logGEinv=log(GEinv);
proc reg data=GE noprint;
  model logGEinv=year;  output out=a2 r=resid;
  title1 'simple regression on time, using log';
proc sgplot data=a2 noautolegend;
  scatter y=resid x=year /
      markerattrs=(symbol=CIRCLEFILLED size=8pt);
  series y=resid x=year / lineattrs=(pattern=solid);
  xaxis label='Year';    yaxis label='Residual';
  title1 'GE gross investment after accounting for time,
using log';
run;
```



```
/* 2. Test for independence and
   3. Investigate potential dependence structures */
data newuse; set a2;
   Z = resid;
proc arima data=newuse;
  identify var=Z nlag=12 ;
  title1 'Look at SPAC: AR(2)';
run;
```

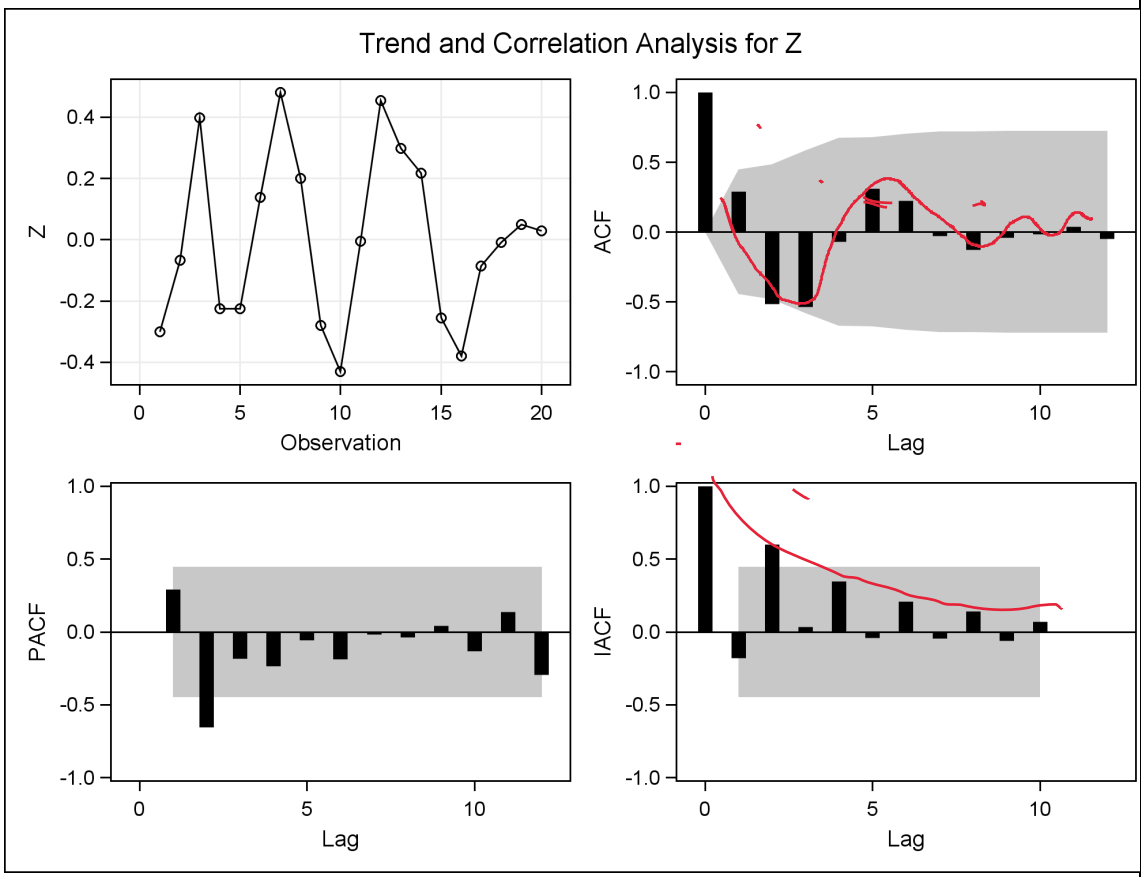
Ho: no auto-correlation
"white noise process"

Look at SPAC: AR(2)

The ARIMA Procedure

Autocorrelation Check for White Noise

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	20.42	6	0.0023	0.290	-0.517	-0.535	-0.070	0.310	0.225
12	21.33	12	0.0457	-0.030	-0.127	-0.040	-0.016	0.040	-0.049



```
/* 4. Fit dependence structure and assess model adequacy */
proc arima data=newuse;
  identify var=logGEinv crosscorr=(year) nlag=12;
  estimate p=2 input=(year) method=uls plot;
  title1 'AR(2) model fit to log of GE data';
run;
```

out

AR(2) model fit to log of GE data

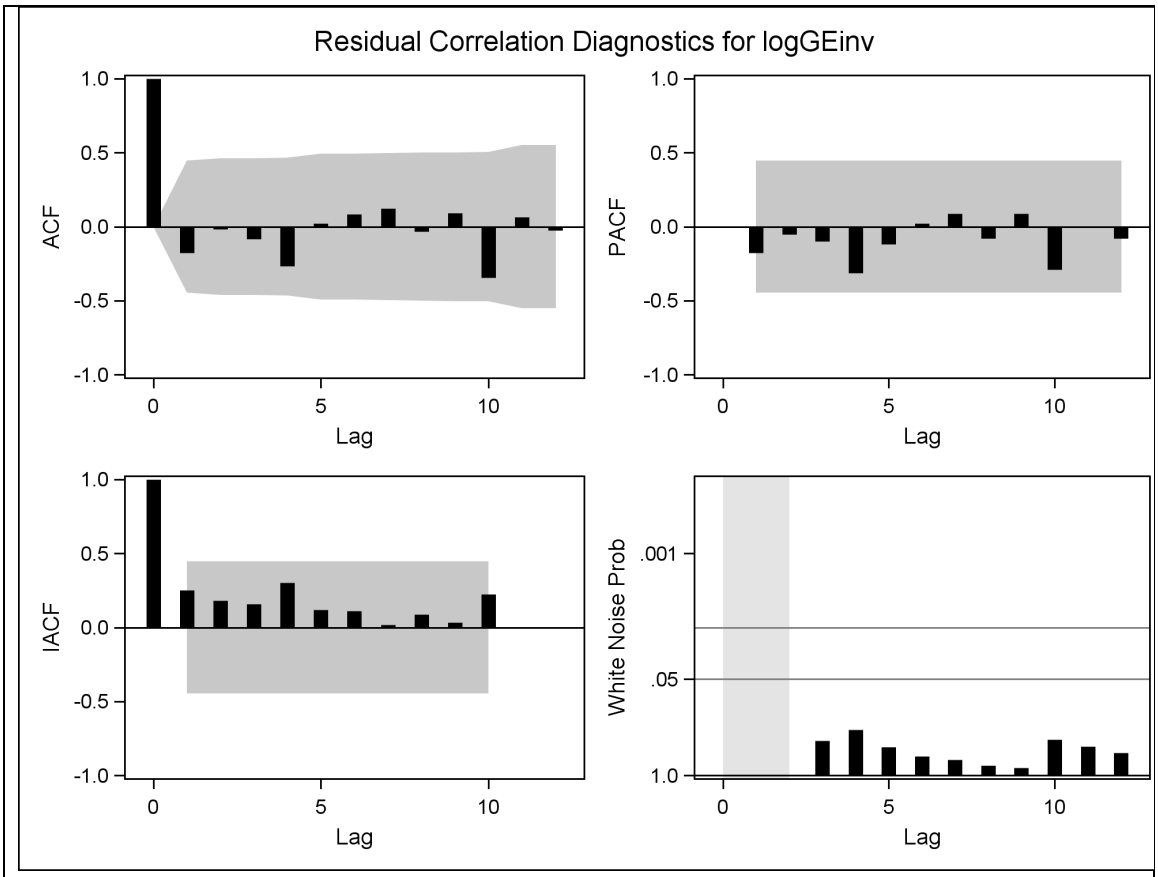
Unconditional Least Squares Estimation

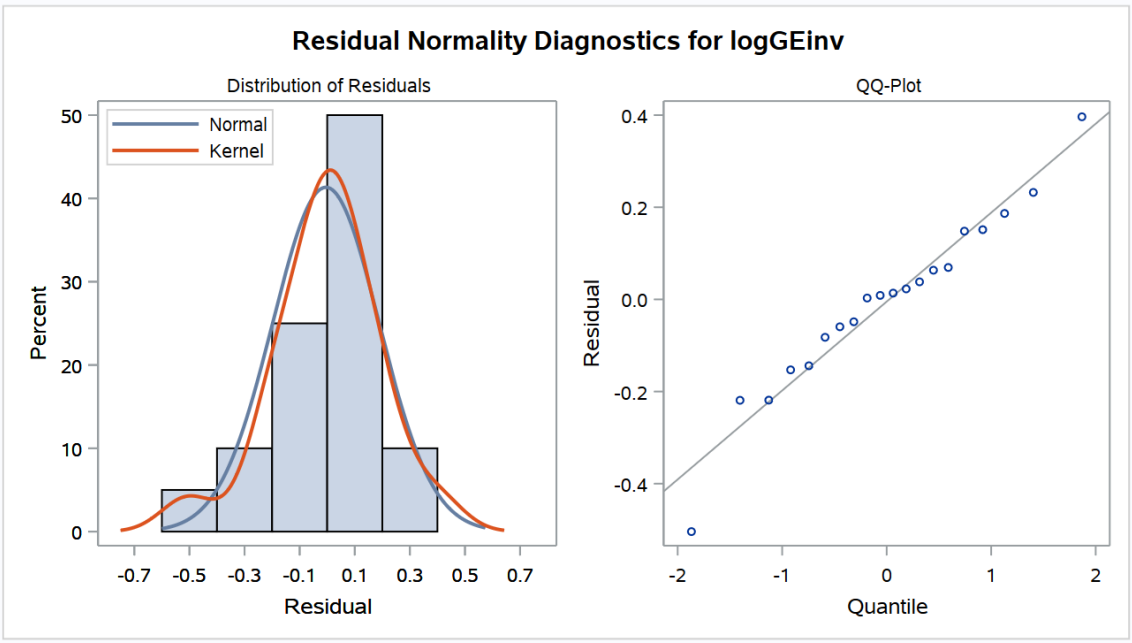
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	-135.17006	14.84188	-9.11	<.0001	0	logGEinv	0
AR1,1	0.51014	0.18639	2.74	0.0146	1	logGEinv	0
AR1,2	-0.71635	0.17516	-4.09	0.0009	2	logGEinv	0
NUM1	0.07183	0.0076327	9.41	<.0001	0	year	0

Constant Estimate	-163.042
Variance Estimate	0.044281
Std Error Estimate	0.210431

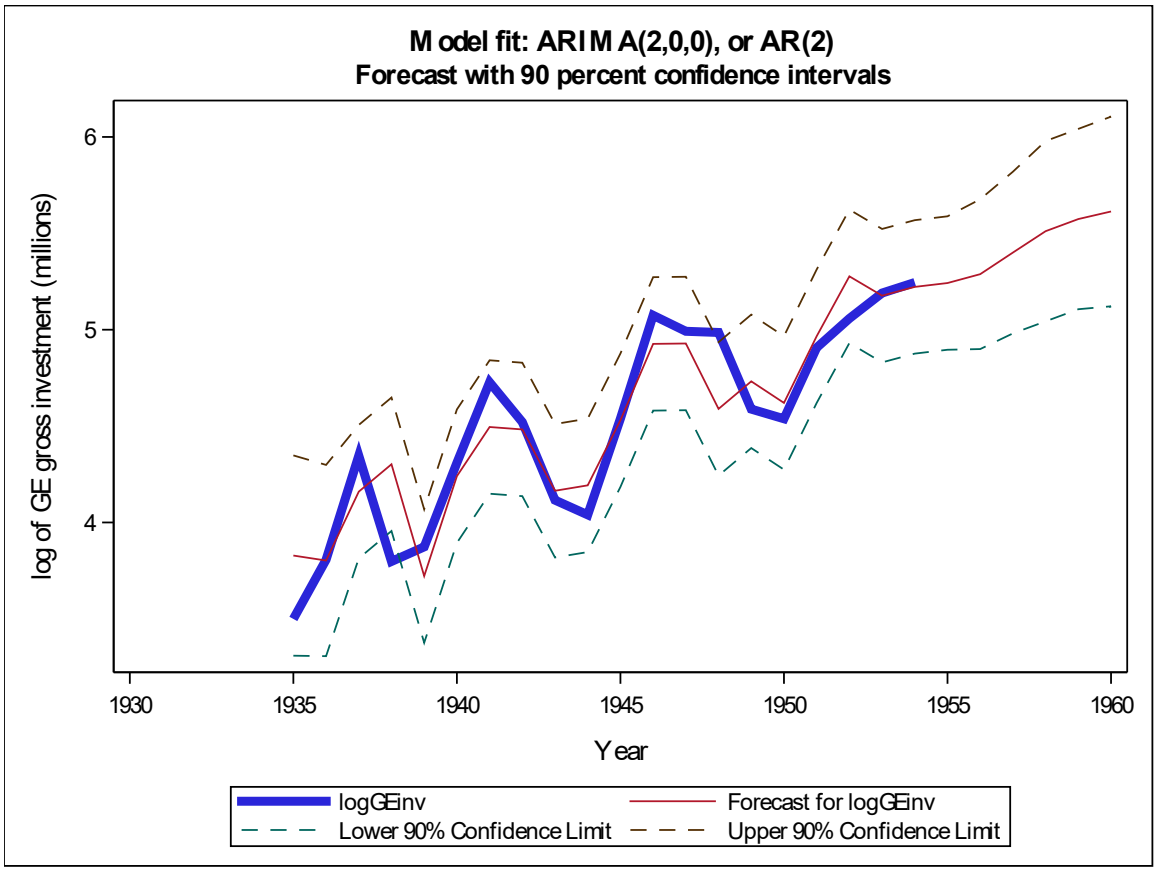
Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	3.11	4	0.5395	-0.176	-0.019	-0.086	-0.269	0.018	0.078
12	9.44	10	0.4910	0.122	-0.032	0.094	-0.343	0.065	-0.026
18	14.23	16	0.5815	-0.140	0.189	-0.037	0.005	0.074	-0.004


Residual Correlation Diagnostics for logGEinv
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
0
5
10
.001
.05
1.0


Residual Normality Diagnostics for logGEinv
Distribution of Residuals
Normal
Kernel
Percent
Residual
50
40
30
20
10
0
-0.7
-0.5
-0.3
-0.1
0.1
0.3
0.5
0.7
QQ-Plot
Residual
Quantile
0.4
0.2
0.0
-0.2
-0.4
-2
-1
0
1
2

```
/* 5. Forecast */
data fout1; set fout1;
  time = _n_ + 1934;
proc sgplot data=fout1;
  series x=time y=logGEinv /
              lineattrs=(pattern=solid thickness=5);
  series x=time y=forecast / lineattrs=(pattern=solid);
  series x=time y=l90 / lineattrs=(pattern=dash);
  series x=time y=u90 / lineattrs=(pattern=dash);
  xaxis label='Year' values=(1930 to 1960 by 5);
  yaxis label='log of GE gross investment (millions)';
  title1 'Model fit: ARIMA(2,0,0), or AR(2)';
  title2 'Forecast with 90 percent confidence intervals';
run;
```




```
proc arima data=newuse;
  identify var=logGEinv crosscorr=(year) nlag=12;
  estimate p=1 q=1 input=(year) method=uls plot;
  title1 'ARMA(1,1) model fit to log of GE data';
  title2 '(for comparison)';
run;
```

ARMA(1,1) model fit to log of GE data
(for comparison)

Unconditional Least Squares Estimation							
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Lag**	**Variable**	**Shift**
MU	-136.13361	27.32905	-4.98	0.0001	0	logGEinv	0
MA1,1	-0.99998	0.50174	-1.99	0.0636	1	logGEinv	0
AR1,1	-0.26677	0.25864	-1.03	0.3177	1	logGEinv	0
NUM1	0.07233	0.01405	5.15	<.0001	0	year	0

Constant Estimate	-172.45
Variance Estimate	0.057055
Std Error Estimate	0.238861

70.21

Autocorrelation Check of Residuals									
To Lag	**Chi-Square**	**DF**	**Pr > ChiSq**	**Autocorrelations**					
6	9.31	4	0.0537	-0.031	-0.243	-0.476	0.032	0.182	0.177
12	10.60	10	0.3893	-0.026	-0.100	0.016	-0.084	0.088	-0.062