Eco 4306 Economic and Business Forecasting

Lecture 16

Chapter 10: Forecasting the Long Term: Deterministic and Stochastic Trends

Motivation

- ARMA models require the data are to be second order weakly stationary
- they thus can not be used for time series that grow over time, unless we transform them (by taking first differences, or using log and then taking first differences)
- our next goal is to learn how to analyze nonstationary data - account for the persistent upward or downward tendency in many economic and business time series

Overview

main objectives of Chapter 10

1. understand deterministic and stochastic trends, construct models that produce these trends and analyze their properties and forecasts
2. design statistical procedures to detect deterministic and stochastic trends in the data

10.1 Deterministic Trends

- simple linear model with a deterministic trend

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

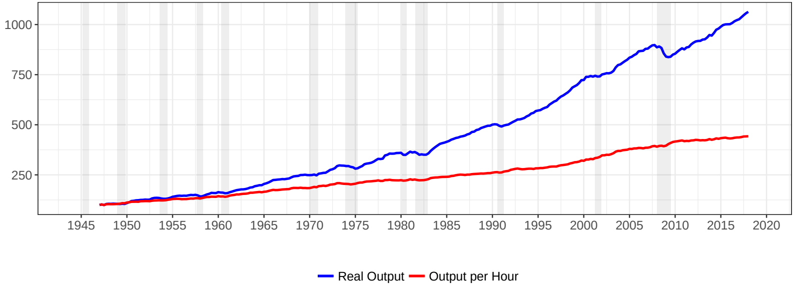
  where β_0 is the intercept and β_1 slope and ε_t a white noise error
- but trends can have different shapes, linear trend is just one particular case
- more generally, a model with deterministic trend can be written as

$$Y_t = g(t) + \varepsilon_t$$

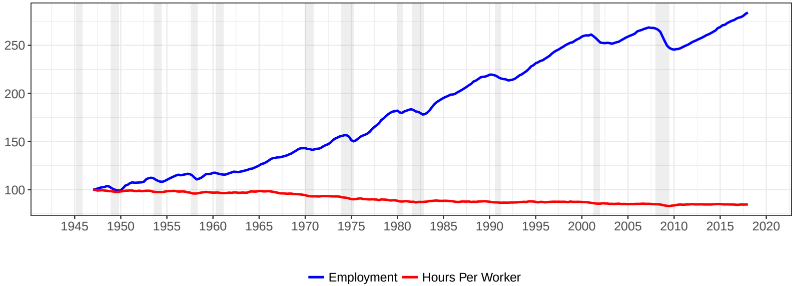
  where $g(t)$ is some function that specifies the deterministic trend

10.1.1 Trend Shapes

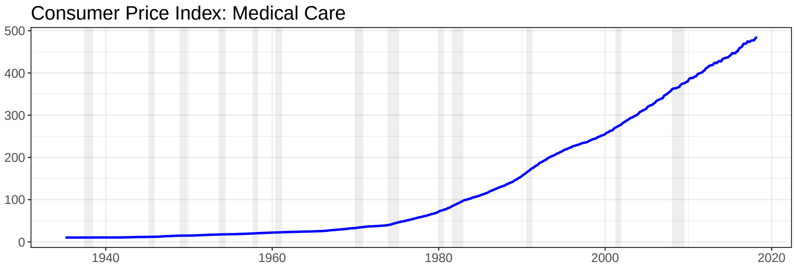
U.S. Nonfarm Business Sector, Output and Productivity, Indices, 1947Q1=100

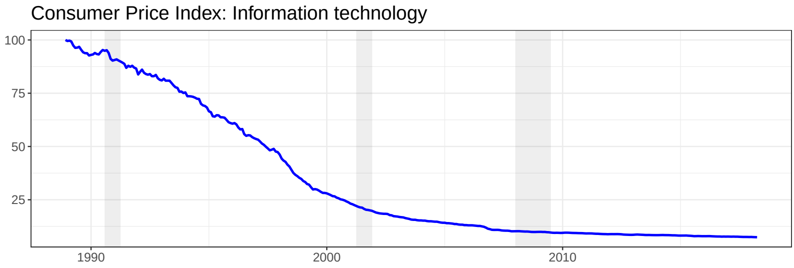


U.S. Nonfarm Business Sector, Employment and Hours, Indices, 1947Q1=100



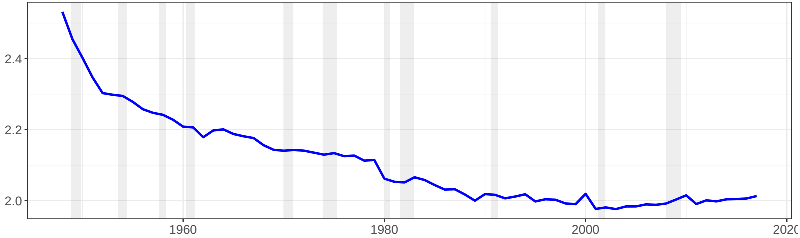
10.1.1 Trend Shapes



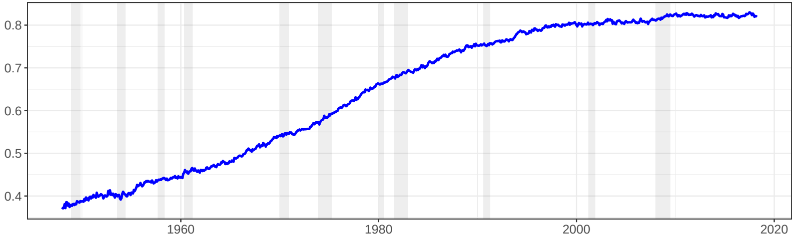


10.1.1 Trend Shapes

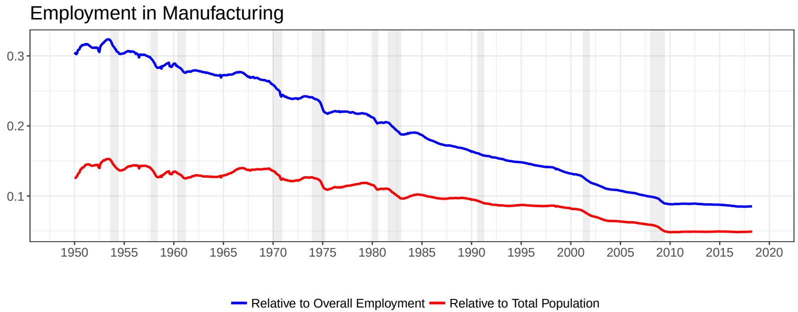
Average Number of People Living in U.S.a Household

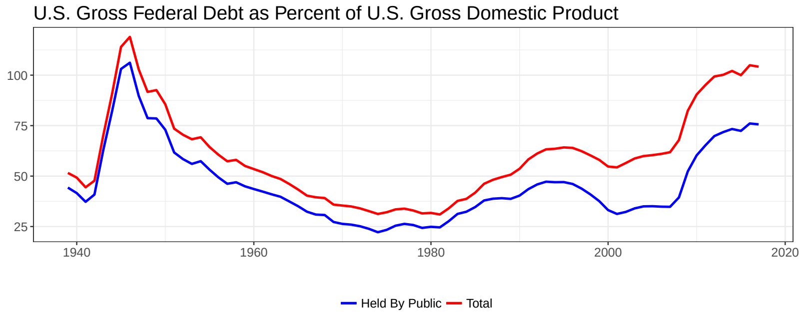


Relative Labor Force Participation Rate: Women/Men

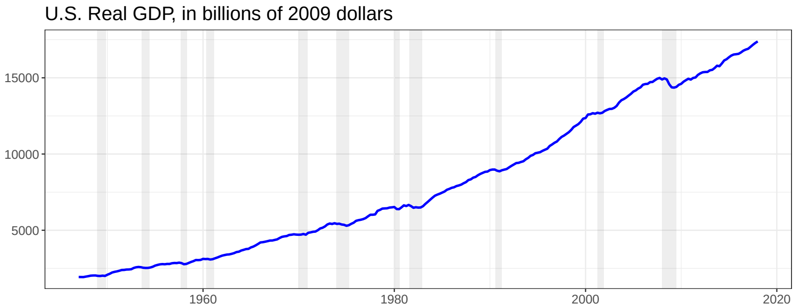


10.1.1 Trend Shapes





10.1.1 Trend Shapes



10.1.1 Trend Shapes

most common trend specifications

- linear trend, $g(t) = \beta_1 t$
- quadratic trend, $g(t) = \beta_0 + \beta_1 t + \beta_2 t^2$
- polynomial trend, $g(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_n t^n$
- exponential trend, $g(t) = \beta_0 e^{\beta_1 t}$
- logistic trend, $g(t) = \frac{\beta_2}{1+\beta_0 e^{\beta_1 t}}$

10.1.1 Trend Shapes

simulated series $Y_t = g(t) + \varepsilon_t$ for different trend specifications

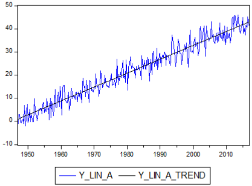
linear trend

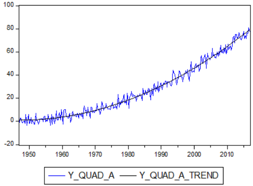
quadratic trend

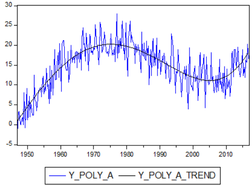
polynomial trend

$\beta_1 > 0$

$\beta_2 > 0$

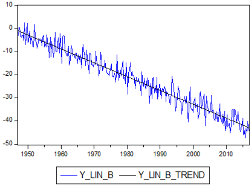


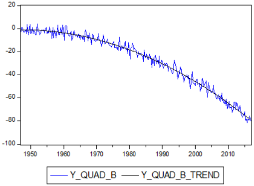


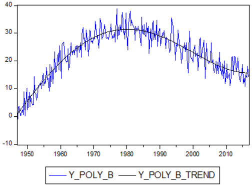


$\beta_1 < 0$

$\beta_2 < 0$





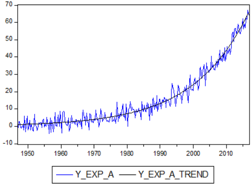


10.1.1 Trend Shapes

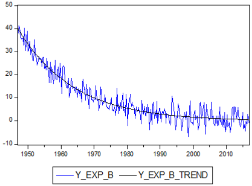
simulated series $Y_t = g(t) + \varepsilon_t$ for different trend specifications

exponential trend

$\beta_1 > 0$

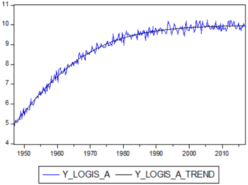


$\beta_1 < 0$

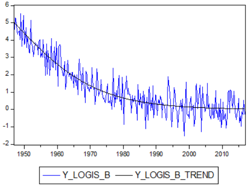


logistic trend

$\beta_0 = 1,\ \beta_1 < 0,\ \beta_2 = 10$

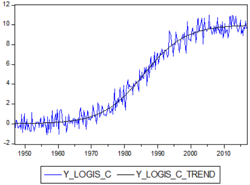


$\beta_0 = 1,\ \beta_1 > 0,\ \beta_2 = 10$

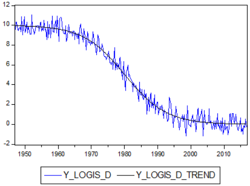


logistic trend

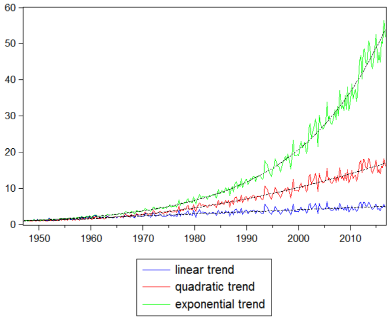
$\beta_0 > 1,\ \beta_1 < 0,\ \beta_2 = 10$

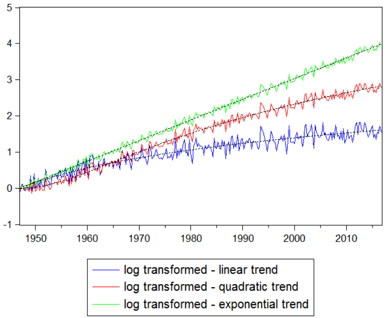


$\beta_0 < 1,\ \beta_1 > 0,\ \beta_2 = 10$



10.1.1 Trend Shapes





10.1.2 Trend Stationarity

- consider a process with a deterministic trend, $Y_t = g(t) + \varepsilon_t$, where ε_t is white noise
- the unconditional mean is

$$\mu_t = E(Y_t) = E(g(t) + \varepsilon_t) = E(g(t)) + E(\varepsilon_t) = g(t)$$

- the unconditional variance is

$$\gamma_0 = var(Y_t) = E[(Y_t - \mu_t)^2] = E[\varepsilon_t^2] = \sigma_\varepsilon^2$$

- autocovariance of order k is

$$\gamma_k = E[(Y_t - \mu_t)(Y_{t-k} - \mu_{t-k})] = E[\varepsilon_t \varepsilon_{t-k}] = 0$$

- autocorrelation of order k is

$$\rho_k = \frac{\gamma_k}{\gamma_0} = 0$$

10.1.2 Trend Stationarity

- thus the unconditional mean μ_t is time varying, but the unconditional variance is not, and the auto-covariance and autocorrelation functions do not depend on time
- because the mean of a process Y_t is not constant over time, it is not first order weakly stationary process
- but because the variance and autocovariances satisfy the requirements for second order weak stationarity, the detrended process, $\tilde{y}_t = y_t - \mu_t$, that is, $\tilde{y}_t = y_t - g(t)$, is second order weakly stationary, and we say that y_t is **trend-stationary**

10.1.3 Optimal Forecast

- recall: under quadratic loss function the optimal forecast is the conditional mean $f_{t,h} = \mu_{t+h|t} = E(Y_{t+h}|I_t)$ for $h = 1, 2, \ldots, s$
- we next analyze this optimal forecast under quadratic loss function for $h = 1, 2, \ldots$

10.1.3 Optimal Forecast

if $Y_t = g(t) + \varepsilon_t$ where ε_t is white noise, for the forecasting horizon $h = 1$ we have

1. optimal point forecast

$$f_{t,1} = \mu_{t+1|t} = E(Y_{t+1}|I_t) = E(g(t+1) + \varepsilon_{t+1}|I_t) = g(t+1)$$

2. forecast error

$$e_{t,1} = Y_{t+1} - f_{t,1} = g(t+1) + \varepsilon_{t+1} - g(t+1) = \varepsilon_{t+1}$$

3. uncertainty associated with the forecast is summarized by the variance of the forecast error

$$\sigma^2_{t+1|t} = var(e_{t,1}|I_t) = var(\varepsilon_{t+1}) = \sigma^2_\varepsilon$$

4. the density forecast is the conditional probability density function $f(Y_{t+1}|I_t)$, assuming ε_{t+1} is normally distributed white noise, we have

$$Y_{t+1}|I_t \sim N(g(t+1), \sigma^2_\varepsilon)$$

10.1.3 Optimal Forecast

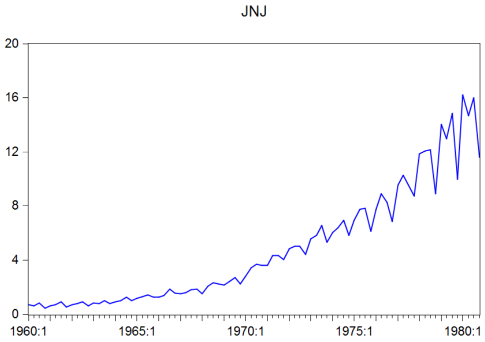
- in general if $Y_t = g(t) + \varepsilon_t$ where ε_t is white noise, for the forecast at horizon $h = s$ we have

$$
\begin{aligned}
f_{t,s} &= g(t+s) \\
e_{t,s} &= \varepsilon_{t+s} \\
\sigma^2_{t+s|t} &= \sigma^2_\varepsilon \\
Y_{t+s}|I_t &\sim N(g(t+s), \sigma^2_\varepsilon)
\end{aligned}
$$

- note that the uncertainty of the forecast is thus *same* regardless of the forecasting horizon, because we assumed that ε_t is white noise
- in general, model with deterministic trend can accommodate linear dependence in its stochastic component: instead of $Y_t = g(t) + \varepsilon_t$ where ε_t is white noise, we then have $Y_t = g(t) + u_t$ where u_t follows some ARMA(p, q) model

Example: Earnings per Share of Johnson and Johnson

- time series plot shows that earnings per share of Johnson and Johnson grew exponentially in the period from 1960Q1 to 1980Q4
- in addition, there is a seasonal pattern that will need to be incorporated into the estimated model



Example: Earnings per Share of Johnson and Johnson

- to build a model for forecasting, we start by generating time series for trend: choose **Object** → **Generate Series** and enter `t = @trend`
- to estimate a model with exponential trend

$$JNJ_t = \beta_0 + \beta_1 e^{\beta_2 t} + \varepsilon_t$$

choose **Object** → **New Object** → **Equation**, in the Equation specification box enter `JNJ = c(1) + c(2)*exp(c(3)*t)` and in Sample `1960Q1 1978Q4`

Dependent Variable: JNJ
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76
Convergence achieved after 23 iterations
Coefficient covariance computed using outer product of gradients
JNJ = C(1) + C(2)*EXP(C(3)*T)

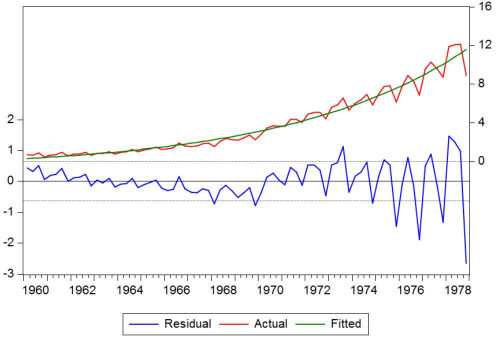
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.805113	0.374703	-2.148670	0.0350
C(2)	1.079629	0.232709	4.639389	0.0000
C(3)	0.032518	0.002683	12.12037	0.0000

R-squared	0.962673	Mean dependent var	3.853158
Adjusted R-squared	0.961650	S.D. dependent var	3.253986
S.E. of regression	0.637234	Akaike info criterion	1.975315
Sum squared resid	29.64294	Schwarz criterion	2.067318
Log likelihood	-72.06197	Hannan-Quinn criter.	2.012084
F-statistic	941.3322	Durbin-Watson stat	1.711509
Prob(F-statistic)	0.000000		

Example: Earnings per Share of Johnson and Johnson

the plot with actual vs fitted data and the regression residuals, which can be obtained by selecting **View** $\rightarrow$ **Actual, Fitted, Residual** $\rightarrow$ **Actual, Fitted, Residual Graph** reveals two problems with the estimated model:

- it can match the trend, but not the seasonal pattern
- the variance of residuals does not appear to be constant, it is increasing over time



Example: Earnings per Share of Johnson and Johnson

- to deal with issue of variance of residuals increasing over time we reestimate the model using log transformed data

$$\log JNJ_t = \beta_0 + \beta_1 t + \varepsilon_t$$

- note that this is equivalent to a multiplicative model $JNJ_t = \tilde{\beta}_0 e^{\beta_1 t} \tilde{\varepsilon}_t$ where $\tilde{\beta}_0 = e^{\beta_0}$ and $\tilde{\varepsilon}_t = e^{\varepsilon_t}$
- to estimate this model choose **Object** $\rightarrow$ **New Object** $\rightarrow$ **Equation**, in the Equation specification box write `log(JNJ) c t` and in Sample `1960Q1 1978Q4`

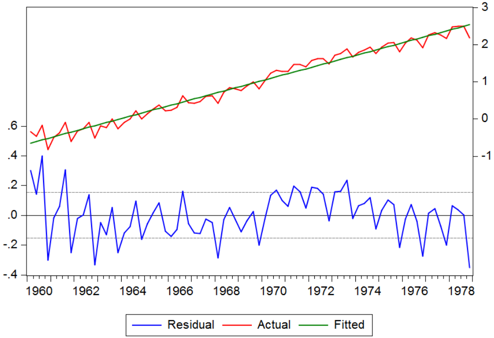
Dependent Variable: LOG(JNJ)
Method: Least Squares
Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.646089	0.034721	-18.60775	0.0000
T	0.042448	0.000799	53.11276	0.0000

R-squared	0.974438	Mean dependent var	0.945698
Adjusted R-squared	0.974093	S.D. dependent var	0.949594
S.E. of regression	0.152843	Akaike info criterion	-0.892842
Sum squared resid	1.728722	Schwarz criterion	-0.831507
Log likelihood	35.92800	Hannan-Quinn criter.	-0.868330
F-statistic	2820.966	Durbin-Watson stat	1.674157
Prob(F-statistic)	0.000000		

Example: Earnings per Share of Johnson and Johnson

residuals shows that the variance is now roughly same over time, but the model still can not match the seasonal pattern



Example: Earnings per Share of Johnson and Johnson

- this finding is supported by the correlogram, obtained using **View** $\rightarrow$ **Residual Diagnostics** $\rightarrow$ **Correlogram - Q-statistics**, which shows large significant component of PACF at lag 4, and significant components of ACF at lags 4, 8, 12
- the residuals are thus not white noise

Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.101	0.101	0.8026	0.370
		2	0.058	0.048	1.0678	0.586
		3	0.004	-0.006	1.0693	0.784
		4	0.655	0.660	36.340	0.000
		5	0.029	-0.179	36.409	0.000
		6	-0.006	-0.066	36.412	0.000
		7	0.005	0.120	36.414	0.000
		8	0.476	0.045	56.150	0.000
		9	-0.085	-0.204	56.786	0.000
		10	-0.111	-0.073	57.896	0.000
		11	-0.127	-0.157	59.369	0.000
		12	0.281	-0.066	66.665	0.000
		13	-0.158	-0.033	69.003	0.000
		14	-0.184	-0.097	72.224	0.000
		15	-0.159	0.008	74.674	0.000
		16	0.145	-0.033	76.747	0.000

Example: Earnings per Share of Johnson and Johnson

- we thus include the seasonal term in the specification of the innovation u_t

$$\log JNJ_t = \beta_0 + \beta_1 t + u_t$$
$$u_t = \phi_4 u_{t-4} + \varepsilon_t$$

- to do this choose **Object** $\rightarrow$ **New Object** $\rightarrow$ **Equation** , in the Equation specification box enter `log(JNJ) c t sar(4)` and in Sample `1960Q1 1978Q4`

Dependent Variable: LOG(JNJ)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76
Convergence achieved after 20 iterations
Coefficient covariance computed using outer product of gradients

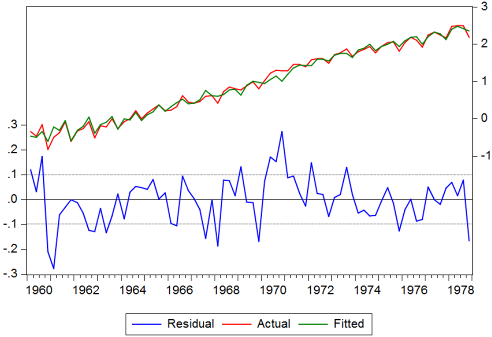
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.558737	0.056365	-9.912881	0.0000
T	0.040461	0.001536	26.34756	0.0000
AR(4)	0.834896	0.054380	15.35290	0.0000
SIGMASQ	0.009057	0.001312	6.906031	0.0000

R-squared	0.989822	Mean dependent var	0.945698
Adjusted R-squared	0.989397	S.D. dependent var	0.949594
S.E. of regression	0.097779	Akaike info criterion	-1.698175
Sum squared resid	0.688366	Schwarz criterion	-1.575505
Log likelihood	68.53066	Hannan-Quinn criter.	-1.649150
F-statistic	2333.918	Durbin-Watson stat	1.317320
Prob(F-statistic)	0.000000		

Inverted AR Roots	.96	.00-.96i	.00+.96i	-.96

Example: Earnings per Share of Johnson and Johnson

residuals no longer have any recognizable seasonal pattern



Example: Earnings per Share of Johnson and Johnson

but the first lag in the ACF and PACF is significant, resulting in p-values for Ljung-Box test that are lower than 0.05

Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.310	0.310	7.6158	
		2	0.142	0.050	9.2285	0.002
		3	-0.017	-0.083	9.2531	0.010
		4	-0.130	-0.123	10.646	0.014
		5	0.088	0.194	11.287	0.024
		6	0.091	0.047	11.992	0.035
		7	0.256	0.199	17.628	0.007
		8	0.089	-0.086	18.316	0.011
		9	-0.012	-0.029	18.328	0.019
		10	-0.034	-0.010	18.431	0.030
		11	-0.164	-0.104	20.886	0.022
		12	-0.012	0.020	20.899	0.034
		13	-0.126	-0.168	22.403	0.033
		14	-0.063	-0.044	22.786	0.044
		15	-0.097	-0.094	23.701	0.050
		16	-0.099	0.017	24.663	0.055

*Probabilities may not be valid for this equation specification.

Example: Earnings per Share of Johnson and Johnson

- to fix this issue we include the first regular AR lag in the model, so that u_t is now given by a multiplicative AR(1)+SAR(1) specification

$$\log JNJ_t = \beta_0 + \beta_1 t + u_t$$
$$u_t = \phi_1 u_{t-1} + \phi_4 u_{t-4} + \phi_1 \phi_4 u_{t-5} + \varepsilon_t$$

- to estimate this model choose **Object** → **New Object** → **Equation**, in the Equation specification box enter `log(JNJ) c t ar(1) sar(4)` and in Sample `1960Q1 1978Q4`

Dependent Variable: LOG(JNJ)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76
Convergence achieved after 60 iterations
Coefficient covariance computed using outer product of gradients

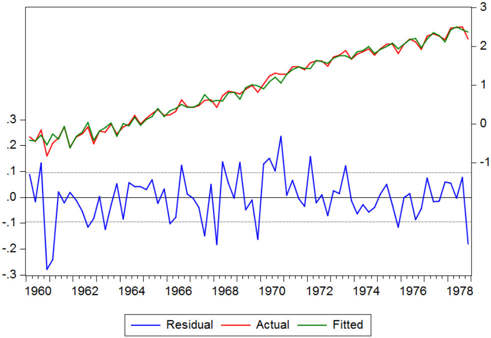
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.525757	0.115973	-4.533449	0.0000
T	0.039723	0.002461	16.14400	0.0000
AR(1)	0.297928	0.124682	2.389497	0.0195
SAR(4)	0.863483	0.051306	16.83009	0.0000
SIGMASQ	0.008312	0.001298	6.401631	0.0000

R-squared	0.990659	Mean dependent var	0.945698
Adjusted R-squared	0.990133	S.D. dependent var	0.949594
S.E. of regression	0.094326	Akaike info criterion	-1.747155
Sum squared resid	0.631709	Schwarz criterion	-1.593817
Log likelihood	71.39189	Hannan-Quinn criter.	-1.685874
F-statistic	1882.531	Durbin-Watson stat	1.871645
Prob(F-statistic)	0.000000		

Inverted AR Roots	.96	.30	.00-.96i	.00+.96i
	-.96			

Example: Earnings per Share of Johnson and Johnson

residuals do not show any systematic pattern



Example: Earnings per Share of Johnson and Johnson

correlogram also suggests that the residuals are white noise

Date: 03/26/17 Time: 16:40
Sample: 1960Q1 1978Q4
Included observations: 76
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.032	0.032	0.0795	
		2	0.075	0.074	0.5327	
		3	-0.010	-0.014	0.5404	0.462
		4	-0.191	-0.198	3.5556	0.169
		5	0.119	0.138	4.7300	0.193
		6	0.008	0.032	4.7356	0.316
		7	0.262	0.249	10.641	0.059
		8	0.018	-0.050	10.670	0.099
		9	-0.015	-0.004	10.689	0.153
		10	0.041	0.034	10.842	0.211
		11	-0.177	-0.094	13.712	0.133
		12	0.079	0.027	14.283	0.160
		13	-0.111	-0.127	15.442	0.163
		14	0.007	-0.031	15.447	0.218
		15	-0.073	-0.131	15.964	0.251
		16	-0.086	-0.019	16.691	0.273

*Probabilities may not be valid for this equation specification.

Example: Earnings per Share of Johnson and Johnson

- to create h-quarter ahead forecasts for $h = 1, 2, \ldots, 12$, so 1979Q1-1981Q4: choose **Forecast** and set "Series to forecast" to "JNJ""","Method" to "Dynamic forecast" and "Forecast sample" to "1979Q1 1981Q4"
- to create a sequence of 1-quarter ahead forecasts, from 1979Q1-1981Q4: choose **Forecast** and set "Series to forecast" to "JNJ""","Method" to "Static forecast" and "Forecast sample" to "1979Q1 1981Q4"

Example: Earnings per Share of Johnson and Johnson

- sequence of 1-step ahead forecasts is more precise than the multistep forecast - RMSE is 0.8480 for the former and 0.9913 for the latter
- confidence interval is narrower in the case of the 1-step ahead forecasts
- multistep forecast is not able to account for a change in the seasonal pattern, 1-step ahead forecasts are eventually able to do that though with a one year delay