9 Regression with Time Series

9.1 Some Basic Concepts

Static Models

(1) $$y_t = \beta_0 + \beta_1 x_t + u_t$$

$t = 1, 2, \ldots, T$, where T is the number of observation in the time series. The relation between y and x is contemporaneous.

Example: Static Phillips Curve:

$$\text{inflation}_t = \beta_0 + \beta_1 \text{unemployment}_t + u_t.$$

Finite Distributed Lag Model (FDL)

In FDL models earlier values of one or more explanatory variables affect the current value of y.

(2) $$y_t = \alpha_0 + \delta_0 x_t + \delta_1 x_{t-1} + \delta_2 x_{t-2} + u_t$$

is a FDL *of order two*.

Multipliers

Multipliers indicate the impact of a unit change in x on y.

Impact Multiplier: Indicates the immediate one unit change in x on y. In (2) δ_0 is the impact multiplier.

To see this, suppose x_t is constant, say c, before time point t, increases by one unit to $c+1$ at time point t and returns back to c at $t+1$. That is

$$\cdots, x_{t-2} = c, x_{t-1} = c, x_t = c+1, x_{t+1} = c, x_{t+2} = c, \ldots$$

Suppose for the sake of simplicity that the error term is zero, then

$$\begin{aligned}
y_{t-1} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c \\
y_t &= \alpha_0 + \delta_0 (c+1) + \delta_1 c + \delta_2 c \\
y_{t+1} &= \alpha_0 + \delta_0 c + \delta_1 (c+1) + \delta_2 c \\
y_{t+2} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 (c+1) \\
y_{t+3} &= \alpha_0 + \delta_0 c + \delta_1 c + \delta_2 c
\end{aligned}$$

from which we find

$$y_t - y_{t-1} = \delta_0,$$

which is the immediate change in y_t.

In the next period, $t+1$, the change is

$$y_{t+1} - y_{t-1} = \delta_1,$$

after that

$$y_{t+2} - y_{t-1} = \delta_2,$$

after which the series returns to its initial level $y_{t+3} = y_{t-1}$. The series $\{\delta_0, \delta_1, \delta_2\}$ is called the lag distribution, which summarizes the dynamic effect that a temporary increase in x has on y.

Lag Distribution: A graph of δ_j as a function of j. Summarizes the distribution of the effects of a one unit change in x on y as a function of j, $j = 0, 1, \ldots$.

Particularly, if we standardize the initial value of y at $y_{t-1} = 0$, the lag distribution traces out the subsequent values of y due to a one-unit, temporary change in x.

Interim multiplier of order J:

(3) $$\delta(J) = \sum_{j=0}^{J} \delta_j.$$

Indicates the cumulative effect up to J of a unit change in x on y. In (2) e.g., $\delta(1) = \delta_0 + \delta_1$.

Total Multiplier: (Long-Run Multiplier)

Indicates the total (long-run) change in y as a response of a unit change in x.

(4) $$\delta_\infty = \sum_{j=0}^{\infty} \delta_j.$$

Example 9.1: Suppose that in annual data

$$\texttt{int}_t = 1.6 + 0.48\,\texttt{inf}_t - 0.15\,\texttt{inf}_{t-1} + 0.32\,\texttt{inf}_{t-2} + u_t,$$

where int is an interest rate and inf is inflation rate. Impact and long-run multipliers?

Assumptions

Regarding the Classical Assumption, we need to account for the dependencies in time dimension.

Assumption (4) $\mathbb{E}[u_i|\mathbf{x}_i] = 0$ is replaced by

(5) $\mathbb{E}[u_t|\mathbf{x}_s] = 0$ for all $t, s = 1, \ldots, T$.

In such a case we say that $\mathbf{x}$ is strictly exogenous. Explanatory variables that are strictly exogeneous cannot react to what happened in the past. The assumption of strict exogeneity implies unbiasedness of OLS-estimates.

A weaker assumption is contemporaneous exogeneity:

(6) $\mathbb{E}[u_t|\mathbf{x}_t] = 0$

It implies only consistency of OLS-estimates.

Assumption

(7) $$\mathbb{C}\text{ov}[u_s, u_t] = 0, \text{ for all } s \neq t$$

is the assumption of no serial correlation or no autocorrelation.

Lack of serial correlation is required for the standard errors and the usual t- and F-statistics to be valid.

Unfortunately, this assumption is often violated in economic time series.

9.2 Trends and Seasonality

Economic time series have a common tendency of growing over time. Some series contain a time trend.

Usually two or more series are trending over time for reasons related to some unobserved common factors. As a consequence correlation between the series may be for the most part of the trend.

Linear time trend:

(8) $$y_t = \alpha_0 + \alpha_1 t + e_t.$$

(9) $$\mathbb{E}[y_t] = \alpha_0 + \alpha_1 t.$$

$\alpha_1 > 0$, upward trend,
$\alpha_1 < 0$, downward trend.

Exponential trend:

If the growth rate $\Delta y/y$ of an economy is β_1. That is

$$\frac{dy(t)/dt}{y(t)} = \beta_1, \tag{10}$$

then

$$y(t) = y(0)e^{\beta_1 t}. \tag{11}$$

Thus a constant growth rate leads to exponential trend model (c.f. continuously compounded interest rate).

Typical such series are GDP, Manufacturing production, and CPI.

Exponential trend is modeled in practice as

$$\log(y_t) = \beta_0 + \beta_1 t + e_t, \tag{12}$$

$t = 1, 2, \ldots,$ where β_1 is the growth rate.

Example 9.2: U.S. GDP growth 1950–1987.

```
Dependent Variable: LOG(USGNP)
Method: Least Squares
Sample: 1950 1987
Included observations: 38
========================================================
Variable   Coefficient  Std. Error  t-Statistic    Prob.
--------------------------------------------------------
C             7.1549     0.012264      583.43     0.0000
@TREND        0.0304     0.000570       53.31     0.0000
========================================================
R-squared           0.987  Mean dependent var       7.717
Adjusted R-squared 0.987  S.D. dependent var       0.340
S.E. of regression 0.039  Akaike info criterion   -3.623
Sum squared resid  0.053  Schwarz criterion       -3.536
Log likelihood    70.830  F-statistic           2841.678
Durbin-Watson stat 0.446  Prob(F-statistic)        0.000
========================================================
```

According to the estimation results the average growth has been about 3 percent per year.

Trending Variables in Regression

Common growth over time of series in the regression model may cause spurious regression relationships. Adding a time trend to the regression eliminates usually this problem.

Example 9.3: Housing Investment and Prices:

Regressing anual observations on housing investment per capita invpc on a housing price index price in constant elasticity form yields

$$\text{(13a)} \qquad \log(\widehat{\text{invpc}}) = -0.550 + 2.241 \log(\text{price})$$

The standard error on the slope coefficient of log(price) is 0.382, so it is statistically significant. However, both invpc and price have upward trends. Adding a time trend yields

(13b)

$$\log(\widehat{\text{invpc}}) = -0.913 - 0.381 \log(\text{price}) + 0.0098t$$

with a standard error of 0.679 on the price elasticity and 0.0035 on time. The time trend is statistically significant and it implies an approximate 1% increase in invpc per year. The estimated price elasticity is now negative and not statistically different from zero.

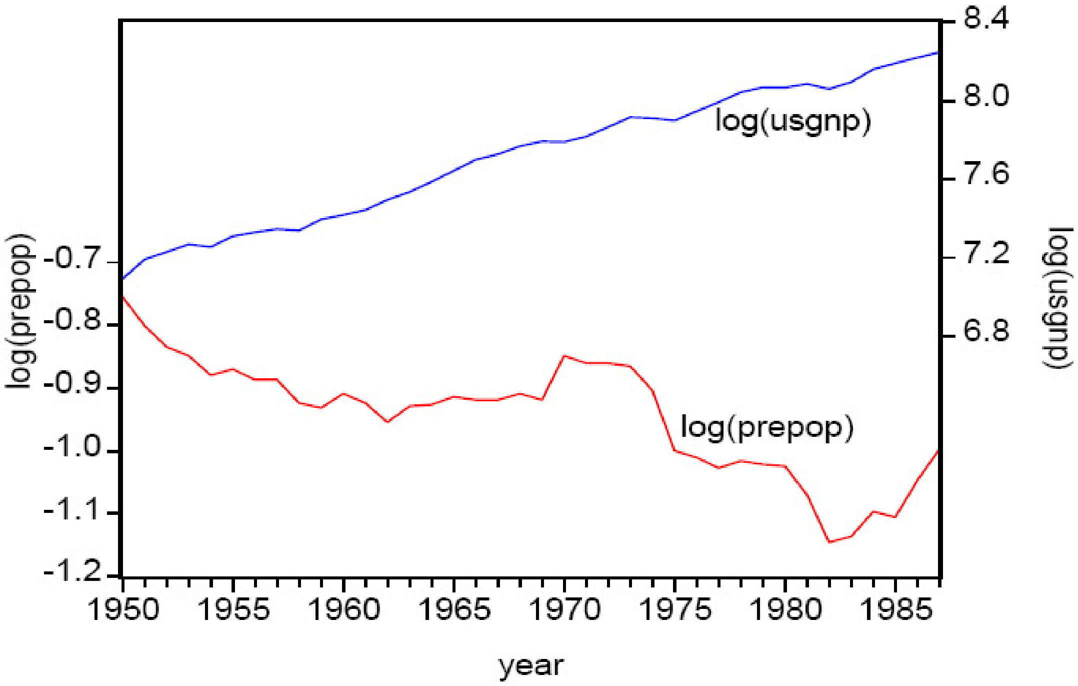
Example 9.4: Puerto Rican Employment and the U.S. Minimum Wage.

prepop = Puerto Rican employment/popul ratio

mincov = (average minimum wage/average wage)*avgcov, where avgcov is the proportion of workers covered by the minimum wage law.

mincov measures the importance of minimum wage relative to average wage.

Sample period 1950–1987



The specified model is
(14a)
$\log(\text{prepop}_t) = \beta_0 + \beta_i \log(\text{mincov}_t) + \beta_2 \log(\text{usgnp}_t) + u_t.$

```
Dependent Variable: LOG(PREPOP)
Method: Least Squares
Sample: 1950 1987
Included observations: 38
============================================================
Variable     Coefficient  Std. Error  t-Statistic    Prob.
------------------------------------------------------------
C              -1.054        0.765      -1.378         0.177
LOG(MINCOV)    -0.154        0.065      -2.380         0.023
LOG(USGNP)     -0.012        0.089      -0.138         0.891
============================================================
R-squared                 0.660  Mean dependent var    -0.944
Adjusted R-squared        0.641  S.D. dependent var     0.093
S.E. of regression        0.056  Akaike info criterion -2.862
Sum squared resid         0.109  Schwarz criterion     -2.733
Log likelihood           57.376  F-statistic           34.043
Durbin-Watson stat        0.340  Prob(F-statistic)      0.000
============================================================
```

Elasticity estimate of mincov is -0.154 and is statistically significant. This suggests that a higher minimum wage lowers the employment rate (as expected). The US GNP is not statistically significant.

Example 9.4: Puerto Rican Employment:

Adding a trend to (14a):
(14b)

$$\log(\text{prepop}_t) = \beta_0 + \beta_i \log(\text{mincov}_t) + \beta_2 \log(\text{usgnp}_t) + \beta_3 t + u_t$$

produces estimation results:

```
Dependent Variable: LOG(PREPOP)
Method: Least Squares
Sample: 1950 1987
Included observations: 38
===========================================================
Variable          Coefficient  Std. Error  t-Statistic  Prob.
-----------------------------------------------------------
C                   -8.729       1.300       -6.712     0.0000
LOG(MINCOV)         -0.169       0.044       -3.813     0.0006
LOG(USGNP)           1.057       0.177        5.986     0.0000
@TREND              -0.032       0.005       -6.442     0.0000
===========================================================
R-squared             0.847   Mean dependent var     -0.944
Adjusted R-squared    0.834   S.D. dependent var      0.093
S.E. of regression    0.038   Akaike info criterion  -3.607
Sum squared resid     0.049   Schwarz criterion      -3.435
Log likelihood       72.532   F-statistic            62.784
Durbin-Watson stat    0.908   Prob(F-statistic)       0.000
===========================================================
```

Seasonality

Monthly or quarterly series include often *seasonality* which shows up as regular cycles in the series.

A common way to account for the seasonality is to include a set of *seasonal dummy variables* into the model.

For example, monthly data:
(15)

$$\begin{aligned} y_t &= \beta_0 + \delta_1 \mathrm{feb}_t + \delta_2 \mathrm{mar}_t + \cdots + \delta_{11} \mathrm{dec}_t \\ &\quad + \beta_1 x_{t1} + \cdots + \beta_k x_{tk} + u_t \end{aligned}$$

$\mathrm{feb}_t, \ldots, \mathrm{dec}_t$ are dummy variables. January is the base month.

9.3 Time Series Models

Stationarity:

A stochastic process $\{y_t : t = 1, 2, \ldots\}$ is (co-variance) stationary, if

(i) $\mathbb{E}[y_t] = \mu$ for all t
(ii) $\mathbb{V}\text{ar}[y_t] = \sigma^2 < \infty$ for all t
(iii) $\mathbb{C}\text{ov}[y_t, y_{t+h}] = \gamma_h$ for all t, i.e., the co-variance depends only on the lag length h, not time.

Example:
A process with a time trend is not stationary, because its mean changes through time.

Establishing stationarity can be very difficult. However, we often must assume it since nothing can be learnt from time series regressions when the relationship between y_t and x_t is allowed to change arbitrarily out of sample.

Weakly Dependent Series

y_t is weakly dependent if y_t and y_{t+h} are "almost independent" as $h \to \infty$.

Covariance stationary sequences are said to be *asymptotically uncorrelated* if $\mathbb{C}\text{ov}[x_t, x_{t+h}] \to 0$ as $h \to \infty$. (Intuitive characterization of weak dependence.)

The weak dependence replaces the notion of random sampling implying law of large numbers (LLN) and the central limit theorem (CLT) holds.

Basic Time Series Processes

White Noise: Series y_t is (weak) white noise (WN) if

(i) $\mathbb{E}[y_t] = \mu$ for all t
(ii) $\mathbb{V}\text{ar}[y_t] = \sigma^2 < \infty$ for all t
(iii) $\mathbb{C}\text{ov}[y_s, y_t] = 0$ for all $s \neq t$.

Remark 9.1: (i) Usually $\mu = 0$. (ii) WN-process is stationary.

Random Walk (RW): y_t is a random walk process if

$$y_t = y_{t-1} + e_t, \tag{13}$$

where $e_t \sim \text{WN}(0, \sigma_e^2)$.

If $y_t \sim \text{RW}$ and assuming $y_0 = 0$, it can be easily shown $\mathbb{E}[y_t] = 0$,

$$\mathbb{V}\text{ar}[y_t] = t\sigma_e^2, \tag{14}$$

and

$$\mathbb{C}\text{orr}[y_t, y_{t+h}] = \sqrt{\frac{t}{t+h}}. \tag{15}$$

Remark 9.2: RW is a nonstationary process.

Random walk with drift:

$$y_t = \mu + y_{t-1} + e_t, e_t \sim \text{WN}(0, \sigma_e^2). \tag{16}$$

AR(1)-process:

(17) $$y_t = \phi_0 + \phi_1 y_{t-1} + e_t,$$

where $e_t \sim \text{WN}(0, \sigma_e^2)$ and $|\phi_1| < 1$.

The condition $|\phi_1| < 1$ is the condition for y_t to be stationary.

Integrated process:

We say that y_t is integrated of order one, denoted as $I(1)$, if $\Delta y_t = y_t - y_{t-1}$ is stationary (and weakly dependent).

Remark 9.3: A series is trend-stationary if it is of the form (8) $y_t = \alpha_0 + \alpha_1 t + e_t$, where e_t is stationary. A trend-stationary process is $I(1)$.

MA(1)-process:

(18)

$$y_t = \theta_0 + \theta_1 e_{t-1} + e_t, \quad e_t \sim \text{WN}(0, \sigma_e^2).$$

All MA processes are covariance stationary.

9.4 Serial Correlation and Heteroscedasticity in Time Series Regression

Assumption 3 $\mathbb{C}\text{ov}[u_t, u_{t+h}] = 0$ is violated if the error terms are correlated.

This problem is called the autocorrelation problem.

Consequences of error term autocorrelation in OLS:

(i) OLS is no more BLUE
(ii) Standard errors are (downwards) biased (t-statistics etc. become invalid), and the situation does not improve for $n \to \infty$.

However
(iii) OLS estimators are still unbiased for strictly exogeneous regressors.
(iv) OLS estimators are still consistent for stationary, weekly dependent data with contemporaneously exogeneous regressors.

Testing for Serial Correlation

$$y_t = \beta_0 + \beta_1 x_{t1} + \cdots \beta_k x_{tk} + u_t. \tag{22}$$

AR(1) errors

$$u_t = \rho u_{t-1} + e_t, \quad e_t \sim \mathrm{WN}(0, \sigma_e^2). \tag{23}$$

Typical procedure to test the first order autocorrelation is to obtain OLS residuals $\hat{u}_t$ by estimating (22), fit AR(1) in the $\hat{u}_t$ series, and use the resulting t-statistic to infer to test $H_0 : \rho = 0$.

An alternative is to use the traditional Durbin-Watson (DW) test.

$$DW = \frac{\sum_{t=2}^{T}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T}\hat{u}_t^2}. \tag{24}$$

It can be shown that

$$DW \approx 2(1 - \hat{\rho}). \tag{25}$$

Ljung-Box test

A general test for serial correlation is the Ljung-Box Q-statistic,

$$Q = T(T+2)\sum_{j=1}^{k}\frac{\hat{\rho}_j^2}{(T-j)}. \tag{26}$$

If the null hypothesis

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0 \tag{27}$$

is true Q has the asymptotic χ_k^2 distribution.

Serial Correlation-Robust Inference with OLS

The idea is to find robust standard errors for the OLS estimates.

Again using matrix notations simplifies considerably exposure. Let

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \tag{28}$$

where

$$\mathbb{C}\text{ov}[\mathbf{u}] = \Sigma_u, \tag{29}$$

Again, write

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}. \tag{30}$$

Then

$$\mathbb{C}\text{ov}[\hat{\boldsymbol{\beta}}] = (\mathbf{X}'\mathbf{X})^{-1}\Omega(\mathbf{X}'\mathbf{X})^{-1}. \tag{31}$$

The problem is how to estimate $\Omega = \mathbf{X}'\Sigma_u\mathbf{X}$.

Newey and West (1987)suggest and estimator

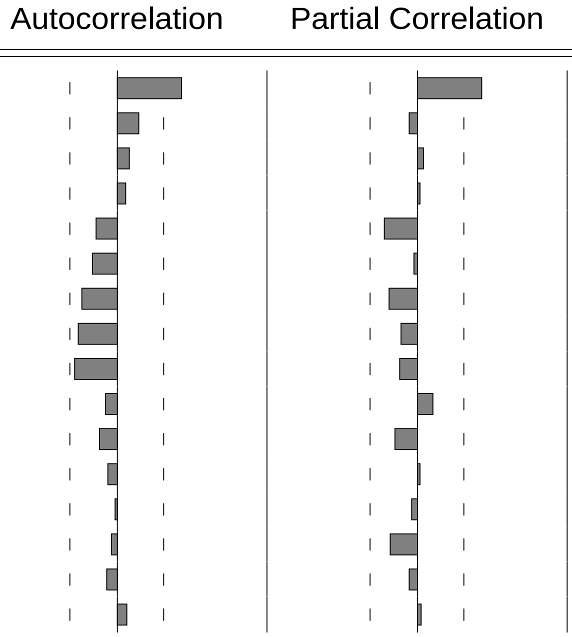
$$\hat{\Omega} = \frac{T}{T-k}\left[\sum_{t=1}^{T}\hat{u}_t^2\mathbf{x}_t\mathbf{x}_t' + \sum_{v=1}^{q}\left(\left(1-\frac{v}{q+1}\right)\sum_{t=v+1}^{T}(\mathbf{x}_t\hat{u}_t\hat{u}_{t-v}x_{t-v}' + \mathbf{x}_{t-v}\hat{u}_{t-v}\hat{u}_t\mathbf{x}_t')\right)\right] \tag{32}$$

which is supposed to be robust both against heteroscedasticity and autocorrelation. The q-variable is determined as a function of the number of observations.

Example 9.5: Puerto Rican Wage:

Correlogram of Residuals

Sample: 1950 1987
Included observations: 38



	AC	PAC	Q-Stat	Prob
1	0.443	0.443	8.0658	0.005
2	0.153	-0.054	9.0537	0.011
3	0.085	0.047	9.3684	0.025
4	0.065	0.020	9.5595	0.049
5	-0.143	-0.228	10.500	0.062
6	-0.167	-0.020	11.821	0.066
7	-0.243	-0.189	14.707	0.040
8	-0.270	-0.115	18.397	0.018
9	-0.288	-0.120	22.740	0.007
10	-0.081	0.112	23.098	0.010
11	-0.117	-0.153	23.865	0.013
12	-0.063	0.019	24.099	0.020
13	-0.008	-0.035	24.104	0.030
14	-0.035	-0.188	24.179	0.044
15	-0.070	-0.050	24.503	0.057
16	0.068	0.027	24.819	0.073

The residuals are obviously autocorrelated. Using the the above autocorrelation robust standard errors yields the following results.

```
Dependent Variable: LOG(PREPOP)
Method: Least Squares
Sample: 1950 1987
Included observations: 38
Newey-West HAC Standard Errors & Covariance (lag truncation=3)
==========================================================
Variable        Coefficient Std. Error  t-Statistic   Prob.
----------------------------------------------------------
C                -8.728657   1.589053   -5.492994    0.0000
LOG(MINCOV)      -0.168695   0.038469   -4.385230    0.0001
LOG(USGNP)        1.057351   0.219516    4.816736    0.0000
@TREND           -0.032354   0.006605   -4.898514    0.0000
==========================================================
R-squared               0.847089  Mean dep. var     -0.944074
Adjusted R-squared      0.833597  S.D. dep. var      0.092978
S.E. of regression      0.037928  Akaike info crit. -3.606957
Sum squared resid       0.048910  Schwarz criterion -3.434580
Log likelihood          72.53218  F-statistic        62.78374
Durbin-Watson stat      0.907538  Prob(F-statistic)  0.000000
==========================================================
```

We observe that particularly the standard error of log(usgnb) increases compared to Example 9.4.

Estimating the residual autocorrelation as an AR-process

An alternative to the robustifying of the standard errors with the Newy-White procedure (32), is to explicitly model the error term as an autoregressive process as is done in equation (23) and estimate it.

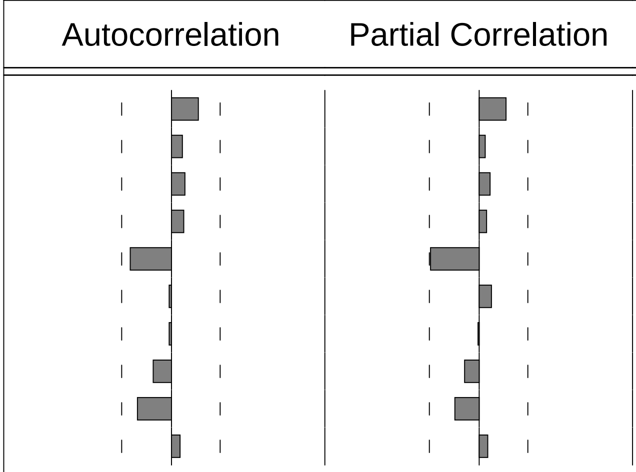
Example 9.6: Estimation the Puerto Rican example with AR(1) errors yields:

```
================================================================
Dependent Variable: LOG(PREPOP)
Method: Least Squares
Sample (adjusted): 1951 1987
Included observations: 37 after adjustments
Convergence achieved after 14 iterations
================================================================
Variable          Coefficient      Std. Error    t-Statistic Prob.
----------------------------------------------------------------
C                  -5.256865         1.492220   -3.522850   0.0013
LOG(MINCOV)        -0.090233         0.048103   -1.875824   0.0698
LOG(USGNP)          0.588807         0.207160    2.842278   0.0077
@TREND             -0.019385         0.006549   -2.959877   0.0058
AR(1)               0.701498         0.123959    5.659112   0.0000
================================================================
R-squared           0.907075     Mean dependent var      -0.949183
Adjusted R-squared 0.895459     S.D. dependent var       0.088687
S.E. of regression 0.028675     Akaike info criterion   -4.140503
Sum squared resid   0.026312     Schwarz criterion       -3.922811
Log likelihood      81.59930     F-statistic              78.09082
Durbin-Watson stat 1.468632     Prob(F-statistic)        0.000000
================================================================
```

Residual autocorrelations:

Correlogram of Residuals

Sample: 1951 1987
Included observations: 37
Q-statistic probabilities adjusted for 1 ARMA term(s)



	AC	PAC	Q-Stat	Prob
1	0.187	0.187	1.4071	
2	0.078	0.044	1.6551	0.198
3	0.097	0.078	2.0557	0.358
4	0.085	0.053	2.3723	0.499
5	-0.273	-0.320	5.7274	0.220
6	-0.016	0.088	5.7395	0.332
7	-0.013	-0.005	5.7475	0.452
8	-0.122	-0.096	6.4908	0.484
9	-0.226	-0.156	9.1276	0.332
10	0.061	0.064	9.3267	0.408

The results indicate no additional autocorrelation in the residuals.

Residual Autocorrelation and Common Factor

The lag operator L is defined as $Ly_t = y_{t-1}$ (generally $L^p y_t = y_{t-p}$).

Thus, equation (23) can be written as

$$(1 - \rho L)u_t = e_t \tag{33}$$

or

$$u_t = \frac{e_t}{1 - \rho L} \tag{34}$$

Using this in (22), we can write (for the simplicity, assume $k = 1$ and denote $x_t = x_{t1}$)

$$y_t = \beta_0 + \beta_1 x_t + \frac{e_t}{1 - \rho L} \tag{35}$$

or

(36)

$$(1 - \rho L)y_t = (1 - \rho L)\beta_0 + \beta_1(1 - \rho L)x_t + e_t$$

This implies that the dynamics of y_t and x_t share $1 - \rho L$ in common, called common factor.

Thus, the autocorrelation in u_t is equivalent that there is a common factor in the regression in (23).

This can be tested by estimating the unrestricted regression

(37) $$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 x_t + \alpha_3 x_{t-1} + e_t$$

and testing whether it satisfies restrictions implied by (36), which can be written as

(38) $$y_t = (1 - \rho)\beta_0 + \rho y_{t-1} + \beta_1 x_t - \beta_1 \rho x_{t-1} + e_t$$

That is, whether

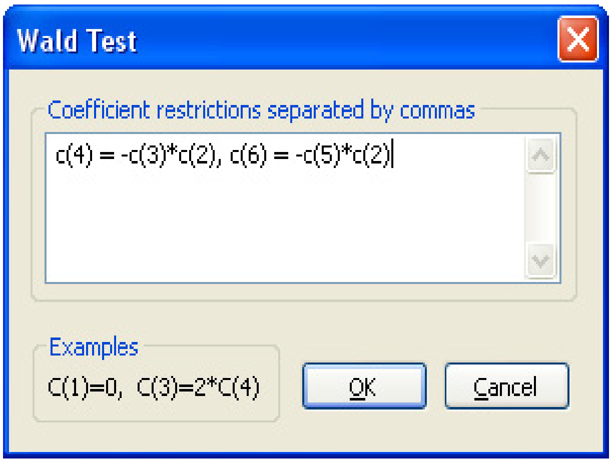
(39) $$\alpha_3 = -\alpha_1\alpha_2$$

If this hypothesis is not accepted, the question is of wrong dynamic specification of the model, not autocorrelation in the residuals.

Example 9.7: Puerto Rican example. The unrestricted model estimates are:

```
Dependent Variable: LOG(PREPOP)
Method: Least Squares
Sample (adjusted): 1951 1987
Included observations: 37 after adjustments
===========================================================
Variable          Coefficient  Std. Error t-Statistic  Prob.
-----------------------------------------------------------
C                   -5.612596    1.547354   -3.627223  0.0011
LOG(PREPOP(-1))      0.535366    0.124932    4.285246  0.0002
LOG(MINCOV)         -0.142230    0.047224   -3.011794  0.0052
LOG(MINCOV(-1))      0.033409    0.046831    0.713386  0.4811
LOG(USGNP)           0.561893    0.188654    2.978437  0.0057
LOG(USGNP(-1))       0.137353    0.226785    0.605651  0.5493
@TREND              -0.019768    0.005975   -3.308752  0.0024
===========================================================
R-squared            0.928815 Mean dependent var    -0.949183
Adj R-squared        0.914578 S.D. dependent var     0.088687
S.E. of reg          0.025921 Akaike info criterion -4.298904
Sum squared resid    0.020156 Schwarz criterion     -3.994136
Log likelihood       86.52972 Hannan-Quinn criter.  -4.191459
F-statistic          65.23934 Durbin-Watson stat     1.634219
Prob(F-statistic)    0.000000
===========================================================
```

Testing for the implied restriction by AR(1)-residuals by imposing restrictions in EViews (`View > Coefficient Tests > Wald-Coefficient Restrictions...`)



gives:

Wald Test:
Equation: EX97

Test Statistic	Value	df	Probability
F-statistic	4.436898	(2, 30)	0.0205
Chi-square	8.873795	2	0.0118

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)*C(3) + C(4)	-0.042736	0.033279
C(2)*C(5) + C(6)	0.438171	0.150355

The null hypothesis is rejected, which implies that rather than the error term is autocorrelated the model should be specified as

$$y_t = \beta_0 + \alpha_1 y_{t-1} + \beta_1 x_{t1} + \beta_{11} x_{t-1,1} + \beta_2 x_{t,2} + \beta_{21} x_{t-1,2} + e_t,$$

where $y = \log(\texttt{prepop})$, $x_1 = \log(\texttt{mincov})$, $x_2 = \log(\texttt{usgnp})$, and $x_3 = \texttt{trend}$.

Finally, because the coefficient estimates of $\log(\texttt{mincov}_{t-1})$ and $\log(\texttt{usgnp}_{t-1})$ are not statistically significant, they can be dropped from the model, such that the final model becomes:

```
Dependent Variable: LOG(PREPOP)
Method: Least Squares
Included observations: 37 after adjustments
==========================================================
Variable       Coefficient   Std. Error  t-Statistic  Prob.
----------------------------------------------------------
C                -5.104888    1.188371    -4.295701 0.0002
LOG(PREPOP(-1))   0.558090    0.103817     5.375713 0.0000
LOG(MINCOV)      -0.106787    0.031282    -3.413647 0.0018
LOG(USGNP)        0.630332    0.154822     4.071322 0.0003
@TREND           -0.017463    0.004849    -3.601696 0.0011
==========================================================
R-squared           0.926226  Mean dependent var   -0.949183
Adj R-squared       0.917004  S.D. dependent var    0.088687
S.E. of reg         0.025550 Akaike info criter.  -4.371286
Sum sqerd resid     0.020889 Schwarz criterion    -4.153594
Log likelihood      85.86878 Hannan-Quinn criter. -4.294539
F-statistic         100.4388 Durbin-Watson stat    1.468436
Prob(F-statistic)   0.000000
==========================================================
```

Autocorrelation has also disappeared from the residuals:

```
Sample: 1951 1987
Included observations: 37
==============================================================
Autocorr       Partial Corr            AC     PAC    Q-Stat  Prob
--------------------------------------------------------------
 . |*.    |        . |*.    |  1     0.203   0.203    1.6465 0.199
 . |*.    |        . | .    |  2     0.079   0.039    1.9008 0.387
 . |*.    |        . | .    |  3     0.092   0.072    2.2586 0.521
 . | .    |        . | .    |  4     0.018  -0.017    2.2724 0.686
***| .    |      ***| .    |  5    -0.346  -0.372    7.6823 0.175
 .*| .    |        . | .    |  6    -0.111   0.020    8.2560 0.220
 .*| .    |        . | .    |  7    -0.067  -0.010    8.4747 0.293
 .*| .    |        . | .    |  8    -0.124  -0.052    9.2354 0.323
 .*| .    |        .*| .    |  9    -0.143  -0.091   10.2920 0.327
 . |*.    |        . |*.    | 10     0.156   0.110   11.5860 0.314
==============================================================
```

Autoregressive Conditional Heteroscedasticity ARCH

Temporal volatility clustering is typical for speculative series like stock returns, interest rates, currencies, etc.

Example 9.7: The following time series plot shows Microsoft's weekly (log) returns ($r_t = 100 \log(P_t/P_{t-1})$) for the sample period from January 1990 through October 2006.

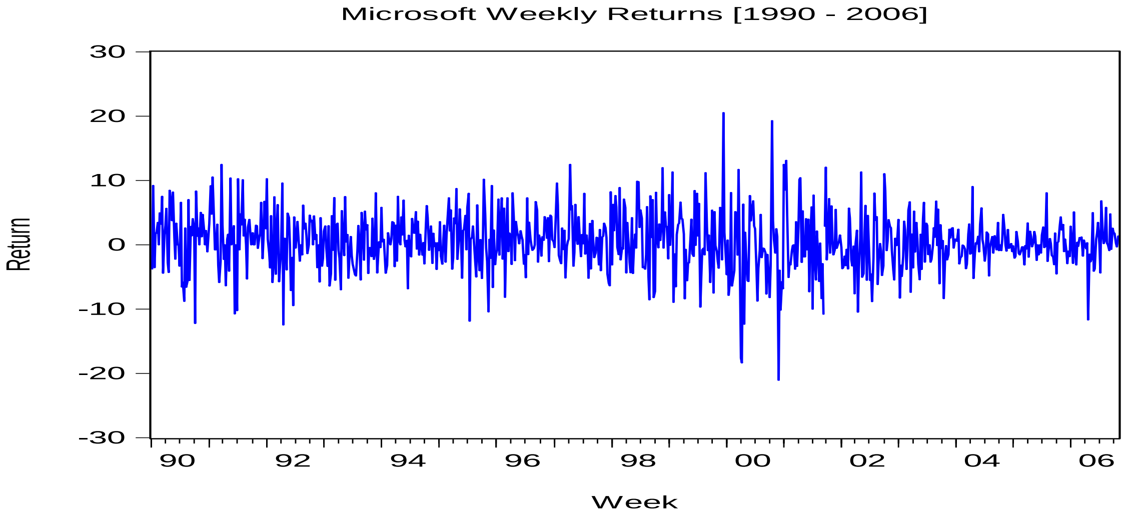


Figure 9.1: Microsoft weekly returns for the sample period January 1990 through November 2006.

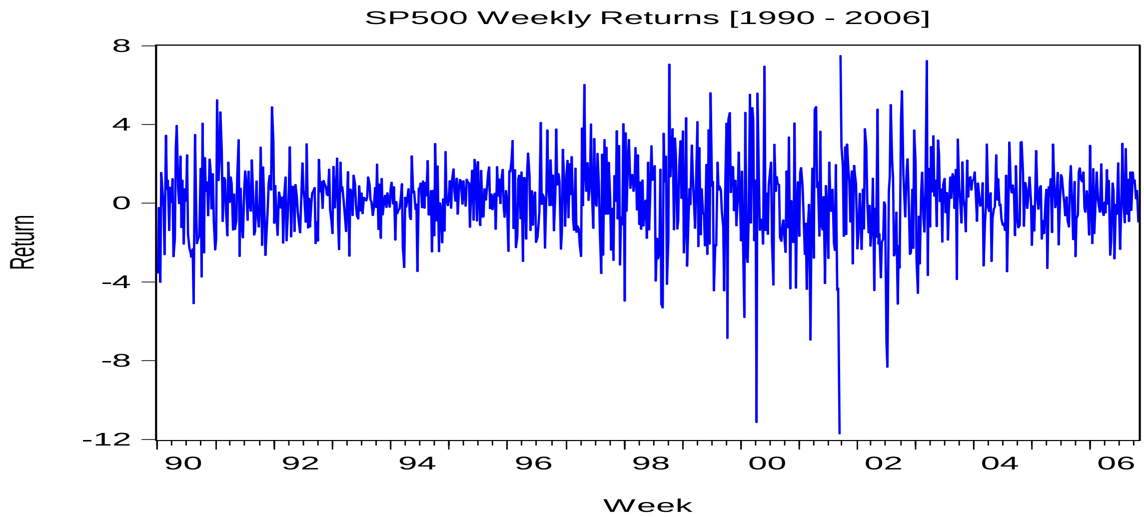


Figure 9.1: SP500 weekly returns for the sample period January 1990 through November 2006.

There is probably some volatility clustering present in both return series.

Engle (1982) suggested modeling the clustering volatility with the ARCH models.

A time series u_t follows an ARCH(1)-process (AutoRegressive Conditional Heteroscedasticity) if

$$\mathbb{E}[u_t] = 0, \tag{40}$$

$$\mathbb{C}\text{ov}[u_t, u_s] = 0, \text{ for all } t \neq s, \tag{41}$$

and

$$h_t = \mathbb{V}\text{ar}[u_t | u_{t-1}] = \alpha_0 + \alpha_1 u_{t-1}^2, \tag{42}$$

where $\alpha_0 > 0$ and $0 \leq \alpha_1 < 1$.

We say that u_t follows and ARCH(1)-process and denote $u_t \sim \text{ARCH}(1)$.

An important property of an ARCH process is that if $u_t \sim \text{ARCH}$, then

$$z_t = \frac{u_t}{\sqrt{h_t}} \sim \text{WN}(0,1). \tag{43}$$

That is the standardized variable z_t is white noise ($\mathbb{E}[z_t] = 0$, $\mathbb{V}\text{ar}[z_t] = 1$, and $\mathbb{C}\text{ov}[z_t, z_{t+k}] = 0$ for all $k \neq 0$).

Furthermore

$$\begin{aligned} \mathbb{V}\text{ar}[z_t|u_{t-1}] &= \mathbb{V}\text{ar}\left[\frac{u_t}{\sqrt{h_{t-1}}}|u_{t-1}\right] \\ &= \frac{1}{h_t}\mathbb{V}\text{ar}[u_t|u_{t-1}] = 1. \end{aligned} \tag{44}$$

That is, z_t has constant conditional variance.

This implies that there should not remain any volatility clustering in z_t.

This can be checked by investigating the autocorrelations of the squared z_t-series, z_t^2, for example with the Ljung-Box Q-statistic, defined in (26).

It may be noted further that if $z_t \sim N(0,1)$, then

$$u_t | u_{t-1} \sim N(0, h_t). \tag{45}$$

Remark 9.4: If (43) holds, using (43) we can always write

$$u_t = \sqrt{h_t}\, z_t, \tag{46}$$

where z_t are independent $N(0,1)$ random variables.

The parameters of the ARCH-process are estimated by the method of maximum likelihood (ML).

Example 9.8: Consider the SP500 returns. With EViews we can estimate by selecting `Quick` > `Estimate Equation` ..., specifying `rm c` in the model box, selecting Estimation Method:
`ARCH-Autoregressive Conditional Heteroscedasticity`, selecting ARCH option equal to 1 and GARCH option equal to 0.

Note that specifying in the model box `rm c` implies that we estimate $u_t = r_{m,t} - \mathbb{E}[r_{mt}]$.

These specifications yield the following results:

```
==================================================================
Dependent Variable: RM
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/08/1990 10/30/2006
Included observations: 877 after adjustments
Convergence achieved after 9 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2
==================================================================
           Coefficient     Std. Error      z-Statistic     Prob.
------------------------------------------------------------------
C             0.189986       0.062065         3.061069    0.0022
==================================================================
    Variance Equation
==================================================================
C             3.237585       0.135488         23.89579    0.0000
RESID(-1)^2   0.247032       0.039617         6.235534    0.0000
==================================================================
R-squared          -0.000294      Mean dependent var     0.154412
Adjusted R-squared -0.002583      S.D. dependent var     2.077382
S.E. of regression  2.080063      Akaike info criterion  4.240795
Sum squared resid   3781.503      Schwarz criterion      4.257134
Log likelihood     -1856.588      Durbin-Watson stat     2.149700
==================================================================
```

Thus the estimated ARCH(1) model is

$$\hat{h}_t = \underset{(0.135)}{3.238} + \underset{(0.0396)}{0.247} u_{t-1}^2, \tag{47}$$

where $u_t = r_{m,t} - \mu$ with $\mu = \mathbb{E}[r_{m,t}]$, estimated from the sample period as $\hat{\mu} = 0.18986$ (i.e., the average weekly return has been approximately 0.19%, or $\approx$ 9.9% per year).

Check next the autocorrelation of the squared standardized residuals $\widehat{z}_t = \widehat{u}_t / \sqrt{\widehat{h}_t}$.

Correlogram of Standardized Residuals Squared

Sample: 1/02/1990 10/24/2006
Included observations: 877

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.006	-0.006	0.0274	0.869
		2	0.028	0.028	0.7023	0.704
		3	0.098	0.098	9.1175	0.028
		4	0.083	0.084	15.178	0.004
		5	0.044	0.041	16.884	0.005
		6	0.104	0.093	26.502	0.000
		7	0.096	0.084	34.603	0.000
		8	0.037	0.023	35.787	0.000
		9	0.036	0.010	36.945	0.000
		10	0.065	0.033	40.685	0.000

It seems that there is still left some volatility clustering especially due to the longer lags. The p-value from the third order forwards is statistically significant. Thus the simple ARCH(1) does not seem to fully capture the volatility clustering in the series. We can try to improve the model.

GARCH

An important extension to the ARCH model is due to Bollerslev (1986), which is called GARCH, Generalized ARCH. GARCH(1,1) is of the form

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta h_{t-1}, \tag{48}$$

where it is assumed that $\alpha_0 > 0$ and $0 < \alpha + \beta < 1$.

Remark 9.5: If $\beta > 0$ then α must also be > 0.

The GARCH-term h_{t-1} essentially accumulates the historical volatility and β indicates the persistence of the volatility.

Example 9.9: GARCH(1,1) model for the SP500 returns

```
Dependent Variable: RM
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/02/1990 10/24/2006
Included observations: 877 after adjustments
Convergence achieved after 17 iterations
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)
==============================================================
             Coefficient     Std. Error  z-Statistic     Prob.
--------------------------------------------------------------
C               0.197385       0.058686     3.363410    0.0008
==============================================================
                      Variance Equation
==============================================================
C                   0.023269   0.013163     1.767719    0.0771
RESID(-1)^2         0.057345   0.012168     4.712861    0.0000
GARCH(-1)           0.937023   0.012920     72.52709    0.0000
==============================================================
R-squared          -0.000428  Mean dependent var      0.154412
Adjusted R-squared -0.003866  S.D. dependent var      2.077382
S.E. of regression  2.081394  Akaike info criterion   4.119464
Sum squared resid   3782.013  Schwarz criterion       4.141250
Log likelihood     -1802.385  Durbin-Watson stat      2.149410
==============================================================
```

The autocorrelations of the squared standardized residuals indicate still some possible first order autocorrelation remained in the series. Nevertheless the estimate of the first order autocorrelation, though statistically significant (p-value 0.022) at the 5% level, is small by magnitude (0.078). Thus, we can conclude that the GARCH(1,1) fits pretty well to the data.

Correlogram of Standardized Residuals Squared

Sample: 1/02/1990 10/24/2006
Included observations: 877

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.078	0.078	5.3952	0.020
		2	-0.007	-0.014	5.4446	0.066
		3	0.041	0.043	6.9264	0.074
		4	-0.007	-0.014	6.9687	0.138
		5	-0.017	-0.015	7.2345	0.204
		6	-0.011	-0.011	7.3509	0.290
		7	-0.026	-0.024	7.9459	0.337
		8	-0.040	-0.035	9.3347	0.315
		9	-0.017	-0.011	9.5916	0.385
		10	-0.009	-0.006	9.6665	0.470

There are several other extensions of the basic ARCH-model (see EViews manual or help, or for a comprehensive presentation Taylor, Stephen J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton University Press).

ARCH in regression

Consider the simple regression

$$y_t = \beta_0 + \beta_1 x_t + u_t. \tag{49}$$

If the error term u_t follows a GARCH-process then accounting for it and estimating the regression parameters with the method of maximum likelihood rather than the OLS yield more accurate estimates.

Particularly, if the ARCH-effect is strong, OLS may lead highly unstable estimates to β_0 and β_1 (usually, however, OLS works pretty well).

Example 9.11: Consider next the market model of Microsoft weekly returns on SP500.

Estimating the market model

$$r_t = \beta_0 + \beta_1 r_m + u_t, \tag{50}$$

with OLS yields

Estimation results:

```
==================================================================
Dependent Variable: R
Method: Least Squares
Sample (adjusted): 1/08/1990 10/30/2006
Included observations: 877 after adjustments
==================================================================
Variable       Coefficient    Std. Error   t-Statistic        Prob.
------------------------------------------------------------------
C                 0.272492      0.128328      2.123402       0.0340
RM                1.169974      0.061639     18.98110        0.0000
==================================================================
R-squared            0.291660   Mean dependent var         0.453150
Adjusted R-squared   0.290850   S.D. dependent var         4.500432
S.E. of regression   3.789860   Akaike info criterion      5.504813
Sum squared resid    12567.66   Schwarz criterion          5.515706
Log likelihood      -2411.861   F-statistic                360.2821
Durbin-Watson stat   2.013695   Prob(F-statistic)          0.000000
==================================================================
```

ML with GARCH(1,1) error specification yields the following results:

```
=================================================================
Dependent Variable: R
Method: ML - ARCH (Marquardt) - Normal distribution
Sample (adjusted): 1/02/1990 10/24/2006
Included observations: 877 after adjustments
Convergence achieved after 20 iterations
Variance backcast: ON
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*GARCH(-1)
=================================================================
                   Coefficient  Std. Error  z-Statistic    Prob.
-----------------------------------------------------------------
C                     0.262170    0.117688     2.227670    0.0259
RM                    1.138309    0.068997     16.49795    0.0000
=================================================================
                          Variance Equation
=================================================================
C                     0.153567    0.058060     2.644976    0.0082
RESID(-1)^2           0.038894    0.009150     4.250488    0.0000
GARCH(-1)             0.949674    0.012409     76.52862    0.0000
=================================================================
R-squared             0.291435  Mean dependent var       0.453150
Adjusted R-squared    0.288184  S.D. dependent var       4.500432
S.E. of regression    3.796977  Akaike info criterion    5.415744
Sum squared resid     12571.65  Schwarz criterion        5.442976
Log likelihood       -2369.804  F-statistic              89.66397
Durbin-Watson stat    2.010866  Prob(F-statistic)        0.000000
=================================================================
```

The results show that in terms of standard errors there are no material differences.

End of the notes: