# Eco 4306 Economic and Business Forecasting

Lecture 26
Chapter 13: Forecasting Volatility I

# Motivation

a summary of what we did so far

- we have developed univariate linear models (AR, MA, ARIMA) and multivariate linear models (VAR, VEC) for stochastic process $\{Y_t\}$
- the $h$-step optimal forecast under quadratic loss function was the conditional expectation of $Y_{t+h}$, based on the information set $I_t$

$$f_{t,h} = E(Y_{t+h}|I_t)$$

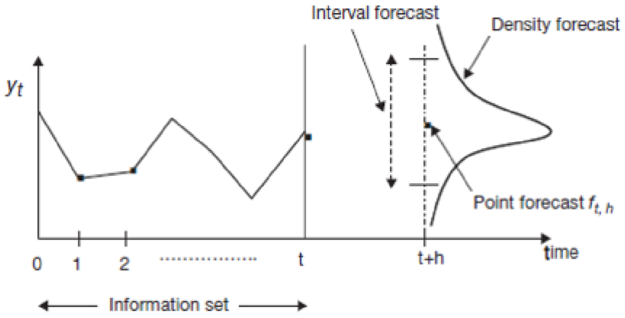
- so our focus was thus to model the behavior of the first conditional moment, the conditional mean $\mu_{t+h|t} = E(Y_{t+h}|I_t)$, of the random variable $Y_{t+h}$

## 13.1.1 The World is Concerned About Uncertainty

- in economic and business press, you will find many statements related to expected GDP growth, expected inflation, expected change in employment, ...
- these are essentially statements about the conditional mean
- but you will also find statements related to output volatility, price volatility, asset risk, market uncertainty, ...
  - "Volatility can wreak havoc on economies. Sudden, sharp ups and downs in business activity can make it difficult for consumers to plan their spending, workers to feel secure in their jobs and companies to determine their future investments. Because of their impact on expectations and business and consumer confidence, swings in the economy can become self-reinforcing. Volatility can also spill over into real and financial asset markets, where severe price movements can produce seemingly arbitrary redistributions of wealth."
  - "Illiquid assets become very risky in volatile times, as the preference for liquid assets creates an environment in which the negative effect of volatility is reflected more strongly on (the suddenly shunned) illiquid assets."
  - "The choice of exchange rate regime has important implications in terms of output volatility."
- these statements concern not just the level, but also the volatility of asset prices and output - the second moments of their time series

## 13.1.2 Volatility Within the Context of Our Forecasting Problem

- recall the graphical representation of the forecasting problem from Chapter 1
- point forecast under a quadratic loss function is conditional mean, $f_{t,h} = \mu_{t+h|t}$
- uncertainty of forecast is variance of forecast error, $\sigma^2_{t+h|t} = var(Y_{t+h} - f_{t,h}|I_t)$
- for an assumed density function and confidence level, we constructed the interval forecasts, e.g. under normality, a 95% confidence interval is $f_{t,h} \pm 1.96\sigma_{t+h|t}$



## 13.1.2 Volatility Within the Context of Our Forecasting Problem

- in MA, AR, and ARMA models optimal forecast depends on the information set but the variance of the forecast error does not depend on the information set because we have assumed that the variance of innovations $\varepsilon_t$ is constant
- for example
  - for AR(1) at horizon $s$
    $$\sigma^2_{t+s|t} = (1 + \phi_1^2 + \phi_1^4 + \ldots + \phi_1^{2(s-1)})\sigma^2_\varepsilon$$
  - for MA(1) at horizon $s \geq 2$
    $$\sigma^2_{t+s|t} = (1 + \theta_1^2)\sigma^2_\varepsilon$$
- so up to now our forecasting has exclusively focused on the modeling of the conditional mean of the stochastic process $\{Y_t\}$, because its conditional variance has been assumed constant over time

## 13.1.3 Setting the Objective

- we will now move to modeling and forecasting volatility that is changing over time
- we would like to make statements such as:

  given current information $I_t$, the probability of S&P500 index falling by 20% or more in the in the next 3 days is 5%, so $P(r_{t+3}^{SP500} \leq -20\% | I_t) = 5\%$

- important to distinguish uncertainty vs volatility vs risk
- uncertainty is a quite general concept - some future events are completely unknown to us today so that it is impossible to assign probabilities to something that is completely unknown
- volatility is more specific and aims to measure the dispersion of a random variable
- most commonly dispersion is measured using variance $\sigma^2$ or standard deviation $\sigma$
- but there are other measures of dispersion, e.g. absolute deviation $E|Y - \mu_Y|$, the range $\max Y - \min Y$, the interquartile range $P_{75} - P_{25}$, ...

## 13.1.3 Setting the Objective

- risk is attached to a potential loss, for example
  - for a portfolio long in stocks, if the stocks prices fall, the investor will suffer a loss
  - needs to quantify risk by assessing the probability of a down movement in the stock price
  - if the portfolio is short in stocks, a down movement will be beneficial, and the risk is in the upward movements
  - needs to assess the probability of a move upward to make the appropriate decision to cushion loss
  - based on this evaluation investor may consider buying insurance or increasing capital reserves

## 13.1.3 Setting the Objective

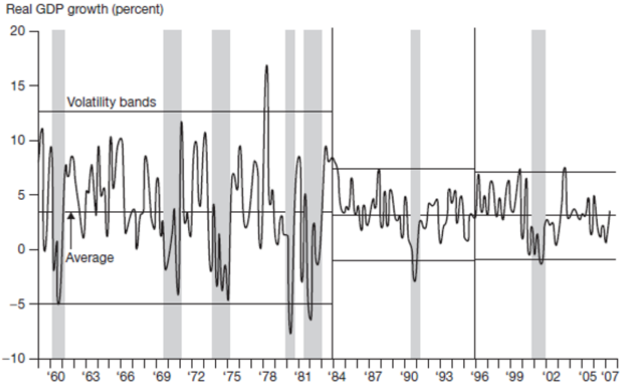
- our objective is to go beyond the conditional mean and model and forecast conditional variance
- modeling of the conditional variance $\sigma^2_{t+1|t} = var(Y_{t+1}|I_t)$ will be in the same spirit modeling the conditional mean $\mu_{t+1|t} = E(Y_{t+1}|I_t)$

# 13.2 Time-Varying Dispersion: Empirical Evidence

- the reason we need to relax the assumption of constant conditional variance is that in practice, many economic and business time series exhibit episodes of high, and low volatility
- for financial markets data **volatility clustering** where time series shows periods of high volatility and periods of low volatility is actually more common than constant volatility
- this means that volatility changes over time and the assumption $\sigma^2_{t+1|t} = \sigma^2_\varepsilon = const.$ does not capture the properties of these time series
- time varying volatility vs time dependence in volatility: to be able to model and forecast volatility, it is crucial to be able to determine how the volatility depends on the information set

## 13.2 Time-Varying Dispersion: Empirical Evidence

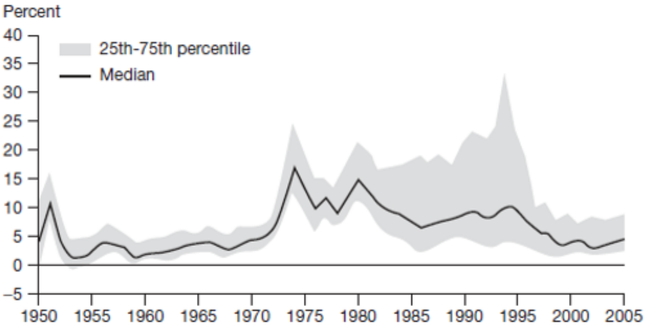
- annualized quarterly GDP growth from 1959 to 2007 exhibits a significant decline in volatility in the 1980s that has continued into the 2000s
- horizontal volatility bands enclose 95% of the observations in a given period of time
- between 1959 and 1983, the 95% interval was $(-5.3, 12.5)$, between 1984 and 1995, $(-1.1, 7.5)$; and between 1996 and 2007 $(-1.0, 7.2)$
- standard deviation of GDP growth was 4.47% from 1959 to 1983, 2.14% from 1984 to 1995, and 2.04% from 1996 to 2007




*Note: Shaded areas denote recessions.*

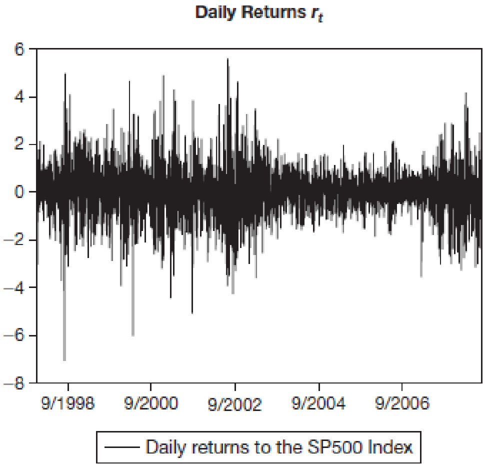
# 13.2 Time-Varying Dispersion: Empirical Evidence

- median global inflation rate (solid line),dropped from 15% in 1980 to around 5% in 2005
- dispersion of the world inflation rate measured by the interquartile range (the shaded area) expanded considerably, from roughly 5% in 1950s-1970s to more than 30% in mid-1990s, before narrowing down again to roughly 5% in 2000s



# 13.2 Time-Varying Dispersion: Empirical Evidence

- daily returns for S&P 500 show that volatility changes quite considerably over time
- **volatility clustering**: days with high volatility are usually followed by days with high volatility; days with low volatility are usually followed by days with low volatility



# 13.2 Time-Varying Dispersion: Empirical Evidence

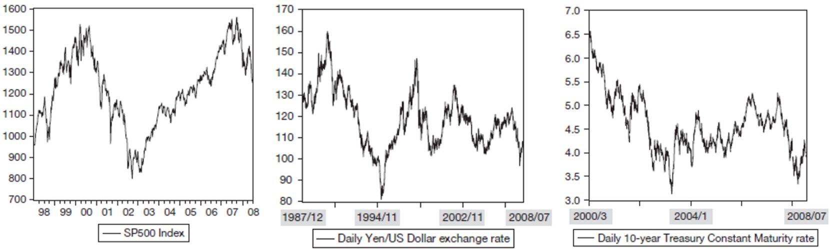
- preceding figures make a case in favor of time-varying volatilities
- the question now is to determine whether there is enough time dependence so that we can build a model to capture the dynamics of volatility
- time dependence would allow to make statements like "if today we are in a highly volatile market, the expected return of a given financial asset will be between -15% and 15% with 95% probability; on the contrary, if today we are in a low volatility market, the expected return will be between -1% and 1% with 95% probability
- in these two cases the confidence intervals attached to the forecast are very different, for those engaged in the business of risk management, these forecasts have different consequences
- for example, in high volatility times, the possibility of large losses in a portfolio may require extra insurance, which may not be all that necessary in low volatility times

# 13.3 Is There Time Dependence in Volatility?

- for a given stochastic process $\{Y_t\}$ we measured time dependence in the conditional mean $\mu_t$ by examining the autocorrelation functions of $\{Y_t\}$
- we now would like to measure time dependence in the conditional variance
- since we are dealing with second moments, it is logical to analyze the autocorrelation functions of $\{(Y_t - \mu_t)^2\}$
- other measures of dispersion that we can also analyze: the absolute deviation $|Y_t - \mu_t|$, or the difference between the highest and the lowest values $\max Y_t - \min Y_t$
- by examining the autocorrelation functions of these new transformed processes we will investigate whether there is time dependence in volatility

# 13.3 Is There Time Dependence in Volatility?

- weekly S&P500 index from January 1998 to July 2008
- daily yen/U.S. dollar exchange rate from December 1987 to July 2008
- daily 10-year Treasury constant maturity rates from March 2000 to July 2008



# 13.3 Is There Time Dependence in Volatility?

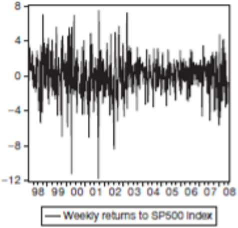
- all three series exhibit a high degree of persistence, we need to test them for the presence of non-stationarity using the unit root tests
- results of the Augmented Dickey-Fuller test show that we can not reject the null hypothesis of unit root for log transformed data (p-values are larger than the customary 5% level)

# 13.3 Is There Time Dependence in Volatility?

- consider next the first difference of each series in logs, so the returns $r_t = 100\Delta \log Y_t$
- figures below show $r_t$, $r_t^2$, $|r_t|$, and $\max r_t - \min r_t$ based on weekly S&P 500 time series, and their corresponding autocorrelation functions
- time series of weekly returns $r_t$ for the S&P 500 does not have any autocorrelation - this seem to point to $r_t$ being a white noise process
- but $r_t^2$, $|r_t|$, and $\max r_t - \min r_t$ all have positive autocorrelations which are statistically significant; $p$-values corresponding to the Q-statistics are zero, rejecting very strongly the null hypothesis of zero autocorrelation
- this means that we can build models for higher moments that will capture this time dependence

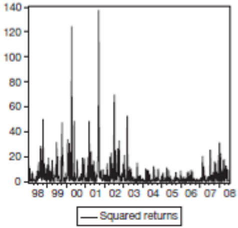
# 13.3 Is There Time Dependence in Volatility?

- returns $r_t$ and squared returns $r_t^2$ for weekly S&P500 index




Sample: 1/12/1998 7/07/2008
Included observations: 547

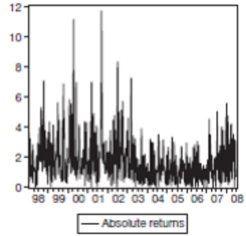
| Autocorrection | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.077 | -0.077 | 3.2986 | 0.069 |
| | | 2 | 0.040 | 0.035 | 4.1955 | 0.123 |
| | | 3 | -0.024 | -0.019 | 4.5169 | 0.211 |
| | | 4 | -0.039 | -0.044 | 5.3473 | 0.253 |
| | | 5 | 0.001 | -0.003 | 5.3481 | 0.375 |
| | | 6 | 0.048 | 0.051 | 6.6214 | 0.357 |
| | | 7 | -0.112 | -0.107 | 13.568 | 0.059 |
| | | 8 | -0.009 | -0.031 | 13.610 | 0.093 |
| | | 9 | -0.047 | -0.039 | 14.823 | 0.096 |
| | | 10 | 0.087 | 0.083 | 19.066 | 0.039 |
| | | 11 | 0.058 | 0.065 | 20.964 | 0.034 |
| | | 12 | 0.009 | 0.006 | 21.011 | 0.050 |



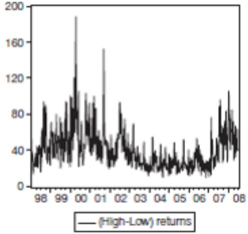

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.236 | 0.236 | 30 605 | 0.000 |
| | | 2 | 0.058 | 0.002 | 32.441 | 0.000 |
| | | 3 | 0.083 | 0.073 | 36.282 | 0.000 |
| | | 4 | 0.065 | 0.030 | 38.590 | 0.000 |
| | | 5 | 0.041 | 0.017 | 39.525 | 0.000 |
| | | 6 | 0.062 | 0.045 | 41.644 | 0.000 |
| | | 7 | 0.072 | 0.044 | 44.538 | 0.000 |
| | | 8 | 0.002 | -0.032 | 44.541 | 0.000 |
| | | 9 | 0.031 | 0.030 | 45.083 | 0.000 |
| | | 10 | 0.057 | 0.035 | 46.911 | 0.000 |
| | | 11 | 0.068 | 0.045 | 49.486 | 0.000 |
| | | 12 | 0.054 | 0.023 | 51.110 | 0.000 |

# 13.3 Is There Time Dependence in Volatility?

- absolute returns $|r_t|$ and the range $\max r_t - \min r_t$ for weekly S&P500 index




| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.187 | 0.187 | 19.217 | 0.000 |
| | | 2 | 0.109 | 0.077 | 25.738 | 0.000 |
| | | 3 | 0.153 | 0.125 | 38.715 | 0.000 |
| | | 4 | 0.138 | 0.088 | 49.233 | 0.000 |
| | | 5 | 0.120 | 0.067 | 57.270 | 0.000 |
| | | 6 | 0.071 | 0.011 | 60.066 | 0.000 |
| | | 7 | 0.085 | 0.037 | 64.116 | 0.000 |
| | | 8 | 0.067 | 0.012 | 66.581 | 0.000 |
| | | 9 | 0.078 | 0.036 | 69.985 | 0.000 |
| | | 10 | 0.111 | 0.068 | 76.830 | 0.000 |
| | | 11 | 0.109 | 0.058 | 83.446 | 0.000 |
| | | 12 | 0.098 | 0.042 | 88.885 | 0.000 |




| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.589 | 0.589 | 190.89 | 0.000 |
| | | 2 | 0.525 | 0.273 | 342.84 | 0.000 |
| | | 3 | 0.510 | 0.206 | 486.51 | 0.000 |
| | | 4 | 0.454 | 0.079 | 600.44 | 0.000 |
| | | 5 | 0.451 | 0.111 | 713.29 | 0.000 |
| | | 6 | 0.388 | -0.006 | 796.94 | 0.000 |
| | | 7 | 0.381 | 0.053 | 877.81 | 0.000 |
| | | 8 | 0.364 | 0.031 | 951.46 | 0.000 |
| | | 9 | 0.324 | -0.003 | 1010.2 | 0.000 |
| | | 10 | 0.383 | 0.137 | 1092.0 | 0.000 |
| | | 11 | 0.328 | -0.011 | 1152.5 | 0.000 |
| | | 12 | 0.304 | -0.007 | 1204.3 | 0.000 |

# 13.3 Is There Time Dependence in Volatility?

- similar pattern to that of the S&P 500 returns emerges for other two series
- daily returns to the yen/U.S. dollar exchange
  - raw returns $r_t$ do not exhibit any autocorrelation and they resemble a white noise process
  - squared and absolute returns, $r_t^2$ and $|r_t|$ do have significant autocorrelation resembling that of autoregressive processes
- daily returns to the 10-year Treasury note
  - some, but very little autocorrelation in the raw returns $r_t$
  - significant positive autocorrelation in the squared and absolute returns

# 13.3 Is There Time Dependence in Volatility?

so to summarize:

- volatility is not only time varying but also exhibits a time dependence
- we can use correlograms for $\{(Y_t - \mu_t)^2\}$ to determine what models to use to capture this time dependence

# 13.5 Simple Specifications for the Conditional Variance

- consider a general stochastic process for returns $\{r_t\}$
- our objective is to estimate and forecast the conditional volatility of this process
- we start with two simple specifications, MA and EWMA, that are very easy to implement because they will not require any estimation technique
- after that, we will construct ARCH and GARCH models that are more complex and require estimation and testing techniques

# 13.5 Simple Specifications for the Conditional Variance

- for the stochastic process $\{r_t\}$, we define the conditional variance at time $t$, that is $\sigma^2_{t|t-1}$, as the expectation of the squared deviations from the mean of the process, conditioning on the information up to time $t-1$:

$$\sigma^2_{t|t-1} = E[(r_t - \mu_{t|t-1})^2 | I_{t-1}]$$

- we will now analyze the following two models that are data-driven specifications
  1. Rolling window average or moving average (MA)
  2. Exponentially weighted moving average (EWMA) or RiskMetrics model
- both are considered smoothing algorithms because they are basically moving averages of past observations

## 13.5.1 Rolling Window Volatility

- simplest estimator of the conditional variance at time $t$, conditioning on the information up to time $t-1$, is a simple average of the squared returns (in deviation from the mean) over the last $n$ periods:

$$\sigma^2_{t|t-1} = \frac{1}{n}\sum_{i=1}^{n}(r_{t-i}-\mu)^2$$

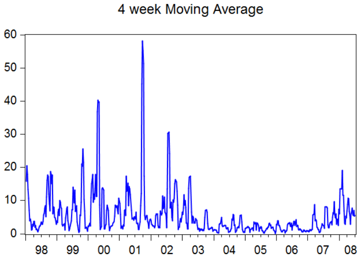
- information set is the collection of the past $n$ returns, so $I_{t-1} = \{r_{t-1}, r_{t-2}, \ldots, r_{t-n}\}$
- in this moving average all the components of the moving sum have the same weight $1/n$
- there is no rule to choose the optimal number of elements $n$
- a large number will produce a smoother and less noisy estimator as there are more components in the average

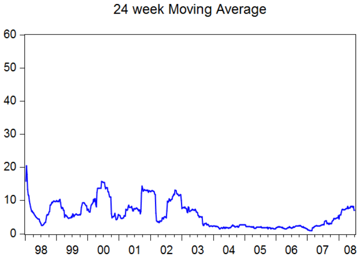
## 13.5.1 Rolling Window Volatility

- if for example $n = 4$ we can generate the time series of 1-step-ahead volatility forecast $\sigma^2_{t|t-1} = E[(r_t - \mu_{t|t-1})^2 | I_{t-1}]$ in EViews as follows
- we first generate the mean return $\mu$ by choosing **Object** $\rightarrow$ **Generate Series** and entering **r_mean = @mean(r)**
- next, we create squared demeaned returns $r_t^2 = (r_t - \mu)$ by choosing **Object** $\rightarrow$ **Generate Series** and entering **r2 = (r - r_mean)^2**
- finally, we create 4 period moving average $\sigma^2_{t|t-1}$ by choosing **Object** $\rightarrow$ **Generate Series** and entering **sigmasq_ma4 = 0.25*(r2(-1)+r2(-2)+r2(-3)+r2(-4))** or equivalently **sigmasq_ma4 = @mav(r2(-1),4)**
- note that even a 24 period moving average can be created easily using **sigmasq_ma24 = @mav(r2(-1),24)**

# 13.5.1 Rolling Window Volatility

- figure below plots 1-week-ahead volatility forecast for S&P 500 weekly returns for two cases: with a small rolling window (4 weeks) and a larger window (24 weeks)



# 13.5.1 Rolling Window Volatility

- market was highly volatile from 2000 to 2002, and much less from 2003 to 2006
- in 2007 and 2008, volatility started to increase as the economic conditions related to the subprime market of mortgages in the United States deteriorated greatly
- the 24-week window forecast is smoother and smaller in magnitude than the 4-week window forecast
- MA volatility estimator is not derived from a model explaining the dynamics of $\{r_t\}$
- the multistep volatility forecast is $\sigma^2_{t+h|t} = \sigma^2_{t+1|t}$

# 13.5.2 Exponentially Weighted Moving Average (EWMA) Volatility

- instead of equally weighting past squared realizations, so using weights $1/n$, we can assign more weight to the most recent realizations than to those in the far past
- we can thus define another estimator of the conditional variance as

$$\sigma^2_{t|t-1} = (1-\lambda)\sum_{i=1}^{n} \lambda^{i-1}(r_{t-i}-\mu)^2$$

  for some $\lambda \in (0,1)$
- the $r_{t-1}$ observation has a weight of $1-\lambda$, $r_{t-2}$ a weight of $(1-\lambda)\lambda$, $r_{t-3}$ a weight of $(1-\lambda)\lambda^2$, ...
- this weighting scheme is called **exponential weighting** or **exponential smoothing**
- note that the information set consists of all previous realizations, $I_{t-1} = \{r_{t-1}, r_{t-2}, \dots, r_1\}$
- because $\lambda < 1$, the observations in the far past have an almost negligible effect in the current estimator of the conditional variance

# 13.5.2 Exponentially Weighted Moving Average (EWMA) Volatility

- this estimator is very popular among the risk management professionals
- technical document of RiskMetrics (JP Morgan, 1996) uses the EWMA estimator with a value of $\lambda \in [0.94, 0.97]$ for financial time series
- when the sample size is large, the previous estimator is well approximated by the following recursive expression, which is very easy to implement once we choose the value of $\lambda$

$$\sigma^2_{t|t-1} = (1-\lambda)(r_t - \mu)^2 + \lambda \sigma^2_{t-1|t-2}$$

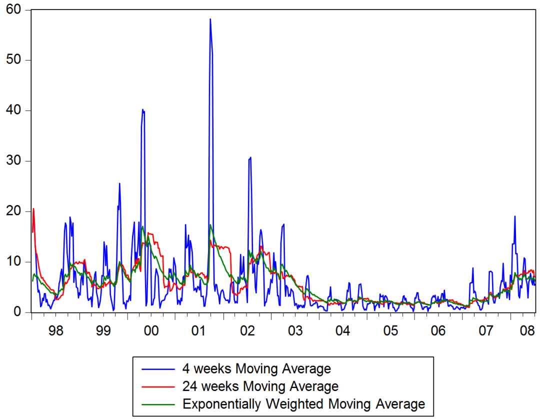
- this expression is a weighted average that assigns a weight of $\lambda$ to the previous estimator of the conditional variance and a weight of $1-\lambda$ to the most recent realization in the information set

## 13.5.2 Exponentially Weighted Moving Average (EWMA) Volatility

- we first generate the mean return $\mu$ by choosing **Object** $\rightarrow$ **Generate Series** and entering **r_mean = @mean(r)**
- next, we create squared demeaned returns $r_t^2 = (r_t - \mu)$ by choosing **Object** $\rightarrow$ **Generate Series** and entering **r2 = (r - r_mean)^2**
- we initialize the recursion using the unconditional variance, first to calculate it select the whole sample and generate series **r_var = @mean(r2)**
- next, set sample to the first observation, by entering **smpl @first @first** in the command window, and generate series **sigmasq_ewma = r_var**
- finally, we set the sample from the second observation 1/05/1998 to last observation 7/07/2008, using **smpl @first+1 @last** and provided that the smoothing parameter is $\lambda = 0.94$ we create the rest of the exponentially moving averages as **sigmasq_ewma = 0.06*r2(-1) + 0.94*sigmasq_ewma(-1)**

## 13.5.2 Exponentially Weighted Moving Average (EWMA) Volatility

- resulting 1-week-ahead volatility forecasts for the SP500 returns are plotted below
- note that time series for MA and EWMA are very similar, but EWMA forecast is much smoother than the MA(24) forecast



# 13.5.2 Exponentially Weighted Moving Average (EWMA) Volatility

summary:

- MA and EWMA specifications are not properly speaking “models” - their construction does not involve an analysis of the time dependence of the conditional variance
- we have not used the information about the time dependence structure observed in the correlograms
- MA and EWMA specifications should be understood as smoothing mechanisms - they average past squared returns
- MA and EWMA are appealing because of their simplicity: they are low cost, easy to implement, do not require any estimation, but are still valuable for detecting trends in volatility