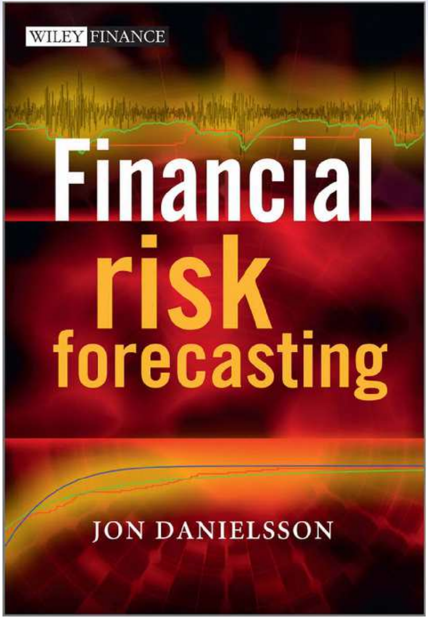
Financial Risk Forecasting
Chapter 2-b
Univariate Volatility Modelling — Part B

Jon Danielsson ©2025
London School of Economics

To accompany
Financial Risk Forecasting
FinancialRiskForecasting.com

Published by Wiley 2011

Version 10.0, August 2025


WILEY FINANCE
Financial
risk
forecasting
JON DANIELSSON

Univariate Volatility Modelling

Volatility

- Volatility is the main measure of risk (see Chapter 4)
- Investment decisions
- Portfolio construction
- Derivative pricing

Estimation and Forecasting

- In this Chapter we follow on from the previous Chapter (2-a) which focuses on the estimation and forecasting of volatility for a single asset (univariate)
- Here we look at estimation, analysis and diagnostics
- The next Chapter does multivariate volatility

Structure

- The theoretical specification of common volatility models
- The theory of estimating the models
- Practical implementation in estimation
- Diagnostics of estimated volatility models

Univariate Volatility Models

- Estimation of volatility models, maximum likelihood
- Diagnostics
- Implied volatilities, like the VIX volatility index
- Realised volatilities

Notation new to this Chapter

$\mathcal{L}$ likelihood function

Risk (and hence volatility) are latent variables

- It is not possible to measure volatility
- Instead, it has to be inferred from observed data
- Using some sort of model
- And that means there are multiple alternative ways to measure volatility
- And it can be very hard to discriminate between them

Learning outcomes

1. Understand and even derive the maximum likelihood estimator of ARCH and GARCH models
2. Implement the estimators of the various volatility models in R
3. Diagnose volatility models
4. Recognise how these models are different from realised volatility models and those obtained from implied volatility such as the VIX

Main challenges

- Understand the difference between conditional and unconditional
- For example, how can returns both be fat tailed and normal?
- What is a nested and un-nested model

Other reading

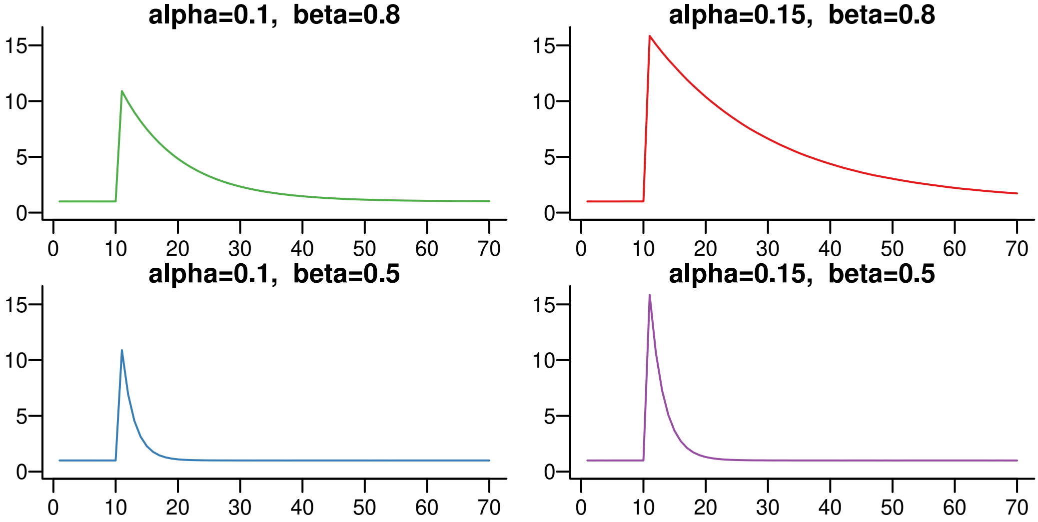
- Alexios Galanos, 2024, GARCH Models
  cran.r-project.org/web/packages/tsgarch/vignettes/garch_models.pdf
- Christian Francq and Jean-Michel Zakoian (2019), "GARCH Models: Structure, Statistical Inference and Financial Applications", 2nd Edition.

Meaning of The GARCH Parameters

Meaning of Parameters

- α is news – how volatility reacts to new information
- β is memory – how much volatility remembers from the past
- The size of $(\alpha + \beta)$ determines how quickly the predictability (memory) of the process dies out:
- If $(\alpha + \beta)$ is close to zero, predictability will die out very quickly
- If $(\alpha + \beta)$ is close to one, predictability will die out slowly

GARCH News and Memory



Half-Life

- In physics, half-life indicates how quickly a radioactive material decays to half the radiation
- The half-life of GARCH is how quickly a shock to volatility dies out to half the impact of the shock
- You can see that labelled on the next figure
- Start with GARCH

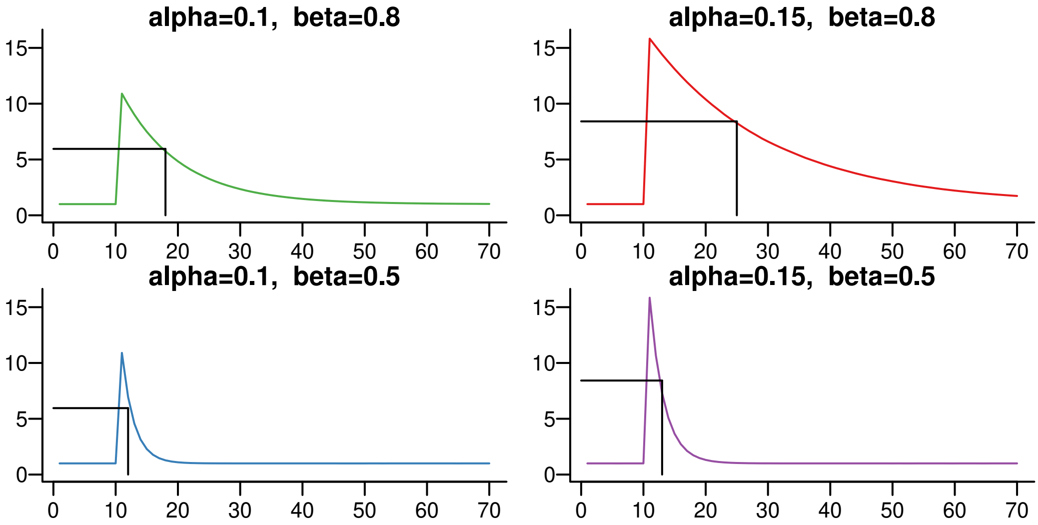
$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$

- We show in the Appendix that the halflife, n^* is

$$n^* = 1 + \frac{\log\left(\frac{1}{2}\right)}{\log(\alpha + \beta)}$$

- And as $(\alpha + \beta) \longrightarrow 1$, $n^* \rightarrow \infty$, memory is infinite

GARCH News, Memory and Halflife



S&P-500

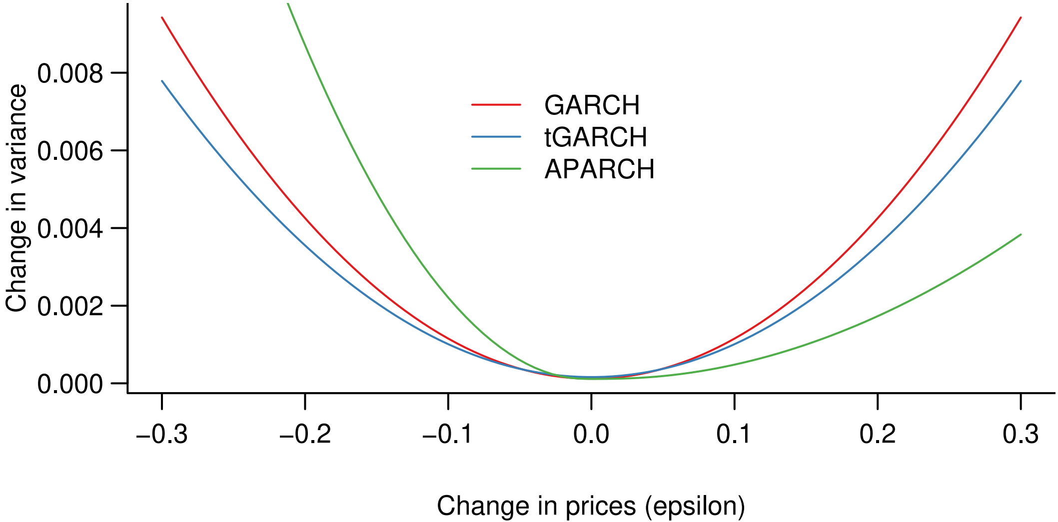
- 10,000 observations
- 12 November 1980 to 9 July 2020
- We discuss the significance a bit later

	ω	α	β	ζ	δ	ν	Half-life
1	0.0000018	0.09	0.89				48
2	0.0000010	0.08	0.91			5.9	110
3	0.0000001	0.06	0.89	0.35	2.6		47

News Impact Curve

- The news impact curve captures the relationship between shocks at time $t-1$ (ϵ_{t-1}) to the variance at time t, σ_t^2
- Recall the apARCH

GARCH News and Memory S&P-500



Maximum Likelihood

Maximum Likelihood Estimation

- A linear model is

$$y = a + bx + \epsilon$$

- Where ϵ relates linearly to the dependent variable y
- Why we can do ordinary least squares (OLS)
- But GARCH is *non-linear*
- Do maximum likelihood
- Consistent parameter estimates even if true density is non-normal

What Is Maximum Likelihood?

- We ask the question: Which parameters most likely generated the data?
- Suppose we have a sample of

$$(-0.2, 3, 4, -1, 0.5)$$

- Of the three possibilities, which is most likely?

	μ	σ
1	1	5
2	-2	2
3	1	2

Likelihood Function

- The likelihood function is *joint* density of all observation, given parameters θ

$$f(x_1, \ldots, x_t|\theta)$$

- Unlike a regular probability density, the likelihood treats the parameters as unknown and the data as fixed
- We use $\mathcal{L}$ to indicate the likelihood function

$$\mathcal{L}(\theta|\text{data})$$

- And the log likelihood by

$$\log \mathcal{L}(\theta|\text{data})$$

Maximum Likelihood

- We ask “what are the most likely parameters given the data?”

$$\hat{\theta}_{\text{ML}} = \max_{\theta} \ \log \mathcal{L}(\theta|\text{data})$$

- Why log?
- The shape of the log likelihood function at its maximum value gives the standard deviation of the parameter estimate
- In particular, the second derivative of the log likelihood
- We see this intuitively in a Figure a bit later

Curvature and Fisher Information

- The curvature of the log-likelihood near its maximum gives us information about parameter precision
- The steeper the curve, the smaller the standard error
- Formally, the *Fisher information* is defined as:

$$\mathcal{I}(\theta) = -\mathbb{E}\left[\frac{d^2}{d\theta^2} \log \mathcal{L}(\theta \mid \text{data})\right]$$

- The variance of the maximum likelihood estimator is approximately:

$$\text{Var}(\hat{\theta}_{\text{ML}}) \approx \mathcal{I}(\theta)^{-1}$$

- This is the foundation of asymptotic confidence intervals for MLEs

Normal and non-normal

- What follows, we only do the conditional normal densities
- It is straightforward to use others

Normal Density

- If:

$$x \sim \mathcal{N}(\mu, \sigma^2)$$

- Then its density is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)$$

Joint Normal Density

- If we have 2 x, i.e. x_1, x_2, uncorrelated and with the same mean and variance, the *joint* density is product of the two densities

$$
\begin{aligned}
f(x_1, x_2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x_1-\mu)^2}{\sigma^2}\right) \times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x_2-\mu)^2}{\sigma^2}\right) \\
&= \prod_{i=1}^{2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}\right)
\end{aligned}
$$

We will want this in logs

- Take logs

$$\log f(x_1, x_2) = \log f(x_1) + \log f(x_2)$$

- Then

$$\log f(x_1, x_2) = \underbrace{-\frac{2}{2}\log(2\pi)}_{\text{Constant}} - \frac{1}{2}\sum_{i=1}^{2}\left(\log\sigma^2 + \frac{(x_i-\mu)^2}{\sigma^2}\right)$$

If mean and variance are different

- If the two x remain uncorrelated, but with different mean and variance

$$\log f(x_1, x_2) = \underbrace{-\frac{2}{2}\log(2\pi)}_{\text{Constant}} - \frac{1}{2}\sum_{i=1}^{2}\left(\log \sigma_i^2 + \frac{(x_i - \mu_i)^2}{\sigma_i^2}\right)$$

Uncorrelated

- The only reason we can do this factorisation is because the two x are uncorrelated
- In what follows, we can do the same because
- Return on day t, *conditional* on return on day $t-1$, is uncorrelated with that return
- Because the conditional implies the previous returns is a known constant
- If it is constant, it is not random and hence there is no correlation

ARCH(1)

- ϵ_t in ARCH(1) is standard normally distributed

$$
\begin{aligned}
y_t =& \sigma_t \epsilon_t \\
\sigma_t^2 =& \omega + \alpha y_{t-1}^2 \\
\epsilon_t \sim& \mathcal{N}(0,1)
\end{aligned}
$$

- The day t density is:

$$
\begin{aligned}
f(y_t|y_{t-1}) =& \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{1}{2}\frac{y_t^2}{\sigma_t^2}\right) \\
=& \frac{1}{\sqrt{2\pi(\omega + \alpha y_{t-1}^2)}} \exp\left(-\frac{1}{2}\frac{y_t^2}{\omega + \alpha y_{t-1}^2}\right)
\end{aligned}
$$

The first day

- Return on day t depends on return on day $t-1$
- But what about day $t=1$?
- By definition, there is no day $t=0$
- So the first density is the $t=2$

$$f(y_2|y_1) = \frac{1}{\sqrt{2\pi(\omega+\alpha y_1^2)}} \exp\left(-\frac{1}{2}\frac{y_2^2}{\omega+\alpha y_1^2}\right)$$

Joint ARCH(1) Densities

- The joint density therefore starts on day 2

$$\prod_{t=2}^{T} f(y_t|y_{t-1}) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi(\omega + \alpha y_{t-1}^2)}} \exp\left(-\frac{1}{2}\frac{y_t^2}{\omega + \alpha y_{t-1}^2}\right)$$

ARCH(1) Likelihood Function

- The likelihood function is then the density, except we condition the parameters $\theta = (\omega, \alpha)$ on the data (returns)

$$\log \mathcal{L}(\theta | \text{data})$$

- The log likelihood function is then

$$\log \mathcal{L} = \underbrace{-\frac{T-1}{2}\log(2\pi)}_{\text{Constant}} - \frac{1}{2}\sum_{t=2}^{T}\left(\log(\omega + \alpha y_{t-1}^2) + \frac{y_t^2}{\omega + \alpha y_{t-1}^2}\right)$$

GARCH(1,1)

- ϵ_t in GARCH(1,1) is standard normally distributed

$$
\begin{aligned}
y_t =& \sigma_t \epsilon_t \\
\hat{\sigma}_t^2 =& \omega + \alpha y_{t-1}^2 + \beta \hat{\sigma}_{t-1}^2 \\
\epsilon_t \sim& \mathcal{N}(0,1)
\end{aligned}
$$

- The day t density is:

$$
\begin{aligned}
f(y_t|y_{t-1}) =& \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{1}{2}\frac{y_t^2}{\sigma_t^2}\right) \\
=& \frac{1}{\sqrt{2\pi(\omega + \alpha y_{t-1}^2 + \beta\hat{\sigma}_{t-1}^2)}} \exp\left(-\frac{1}{2}\frac{y_t^2}{\omega + \alpha y_{t-1}^2 + \beta\hat{\sigma}_{t-1}^2}\right)
\end{aligned}
$$

GARCH(1,1) Likelihood Function

- We also start on day 2
- And just as with the ARCH(1)

$$\log \mathcal{L} = \underbrace{-\frac{T-1}{2}\log(2\pi)}_{\text{Constant}} - \frac{1}{2}\sum_{t=2}^{T}\left(\log\left(\omega + \alpha y_{t-1}^2 + \beta\hat{\sigma}_{t-1}^2\right) + \frac{y_t^2}{\omega + \alpha y_{t-1}^2 + \beta\hat{\sigma}_{t-1}^2}\right)$$

Additional GARCH problem — σ_1

- When we did the ARCH likelihood, we started on day two, because if we had started on day 1, we would need to return for day 0, which does not exist
- It is more complicated with GARCH
- Because it also depends on previous volatility
- But, unlike y_1 in ARCH, we do not know σ_1
- We either need to estimate it or assume a value for it

Importance of σ_1

- Value of σ_1 can make a large difference
- Especially when the sample size is small
- Typically set $\sigma_1 = \hat{\sigma}$

Issues in Estimation

- Parameters are obtained by maximising the likelihood function

  θ The set of parameters

  Θ The space the parameters live in

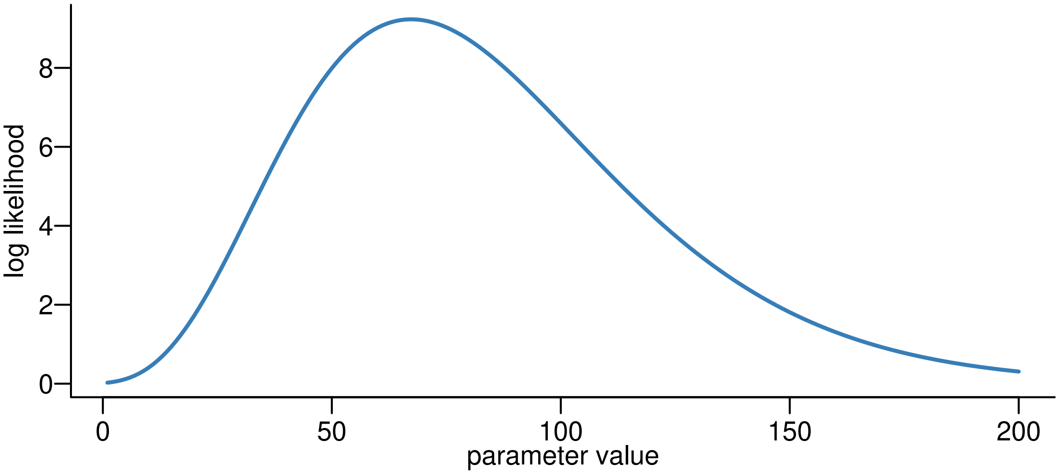
$$\max_{\theta \in \Theta} \log \mathcal{L}(\theta|y) = \hat{\theta}_{\mathrm{ML}}$$

- Done with an algorithm called *optimiser*
- There often are *numerical problems*

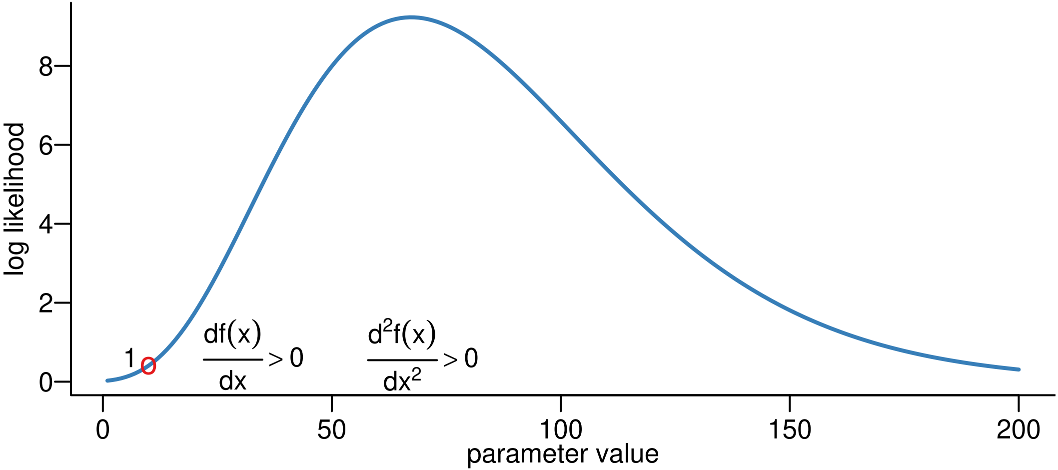
Optimising

- The computer uses an algorithm called optimiser to maximise the likelihood function
- There are many optimisers available
- A classical and not the best optimiser is *Newton-Raphson*

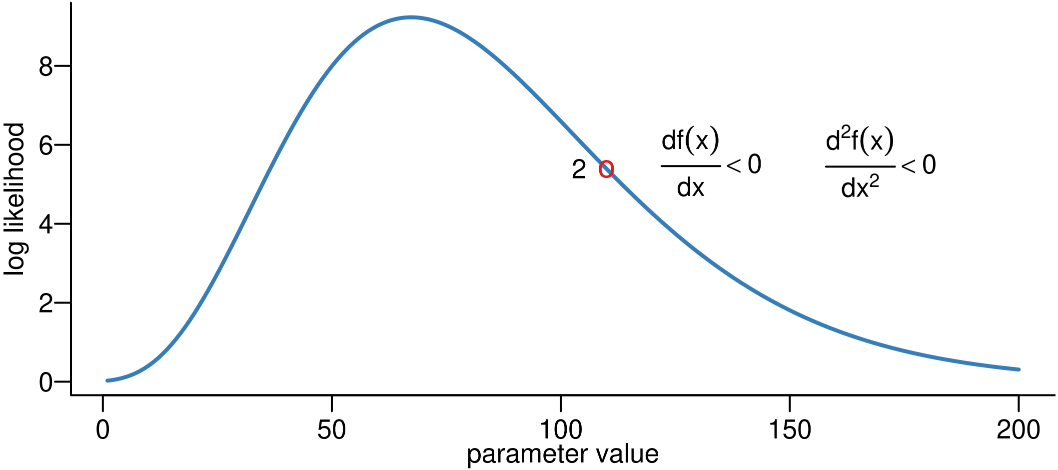
Newton-Raphson



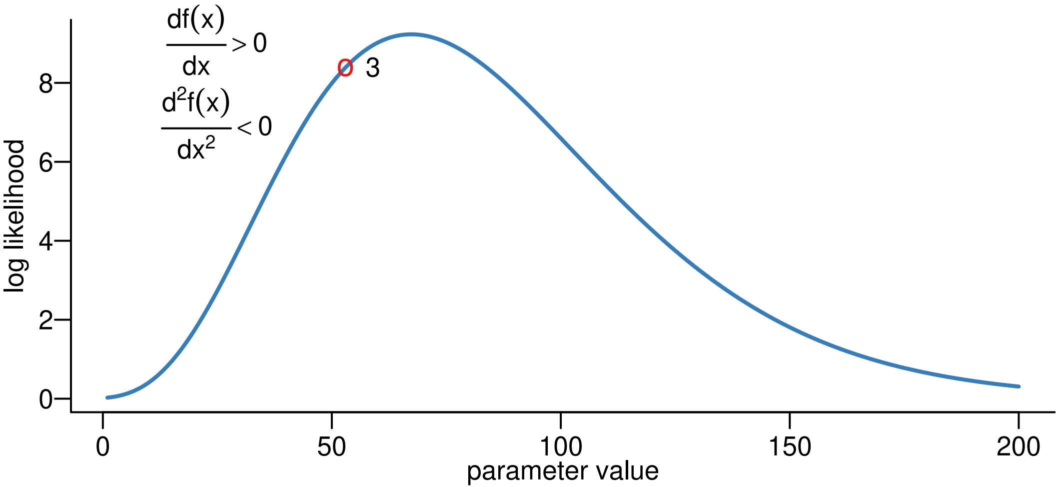
Newton-Raphson



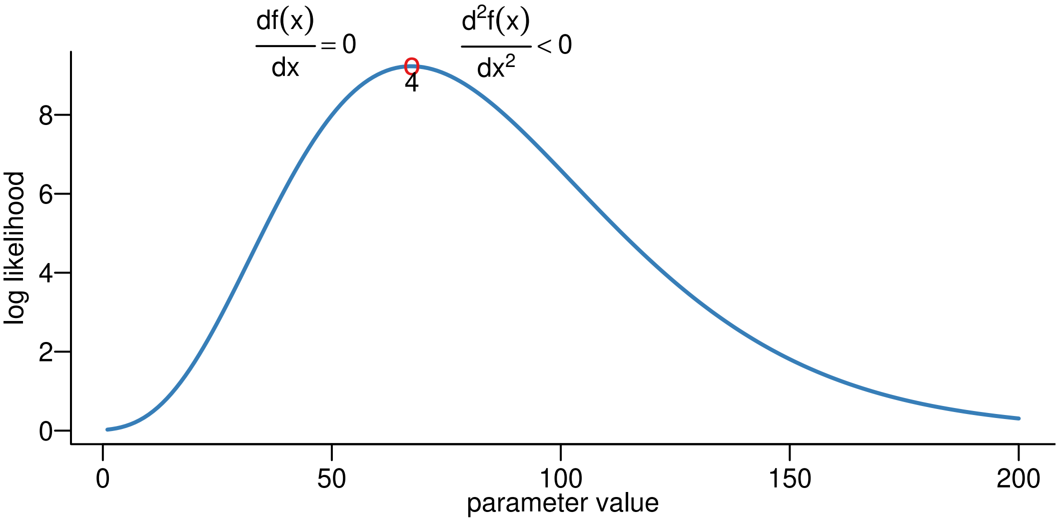
Newton-Raphson



Newton-Raphson



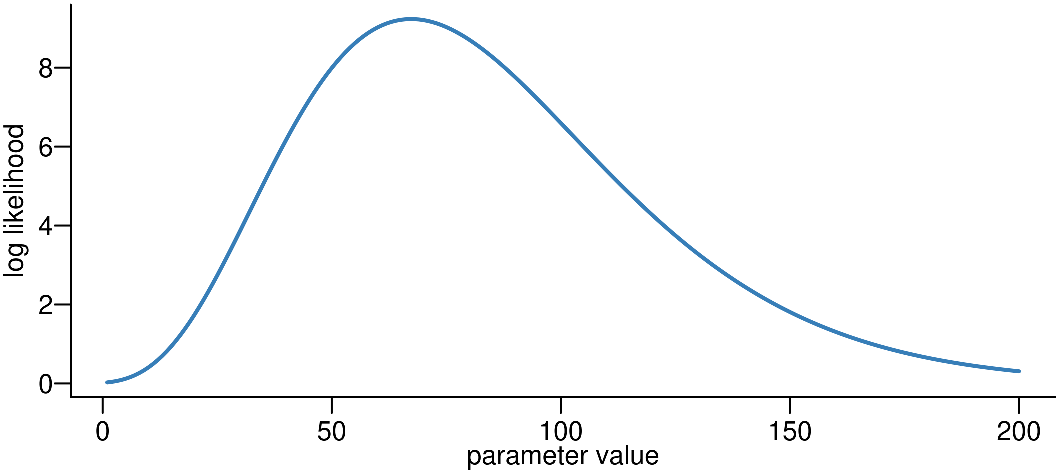
Newton-Raphson



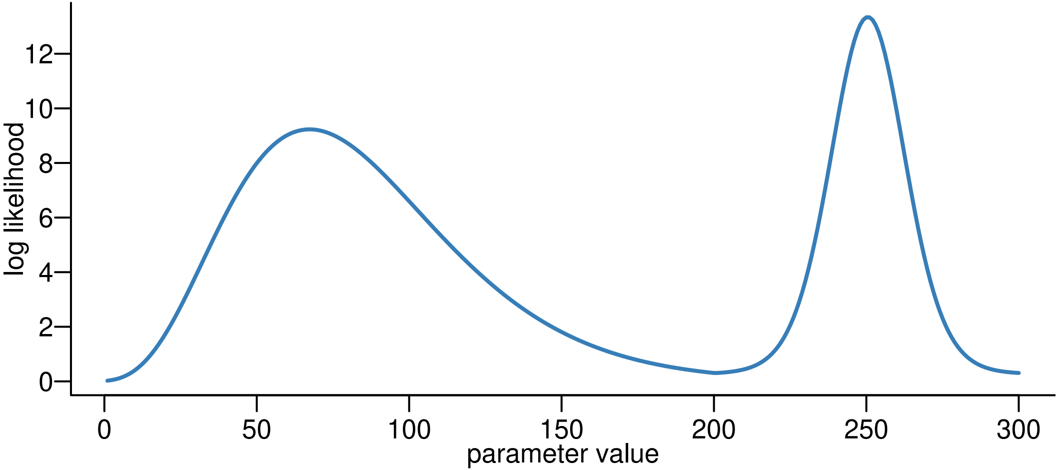
Pathologies

- It is straightforward to optimise the ARCH model
- And the GARCH
- For other models they can be a number of problems
- A common one is multiple local maxima

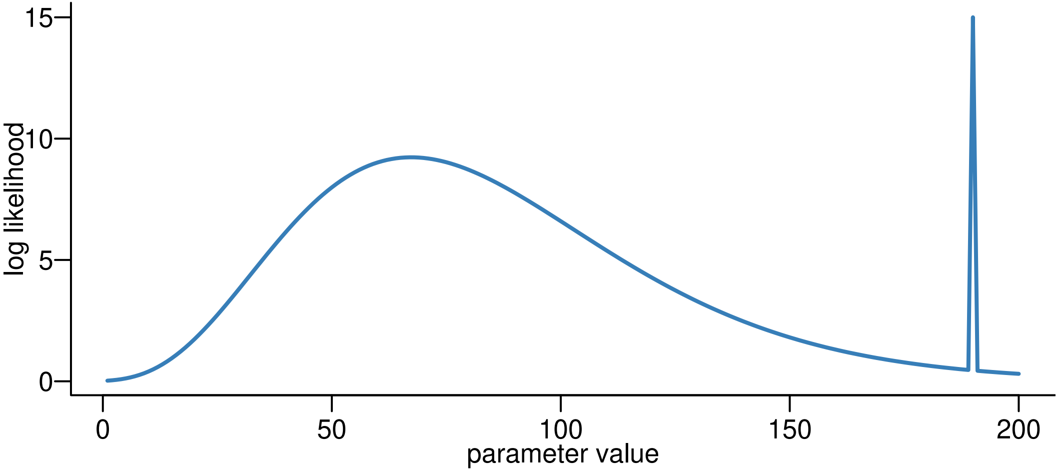
Pathologies



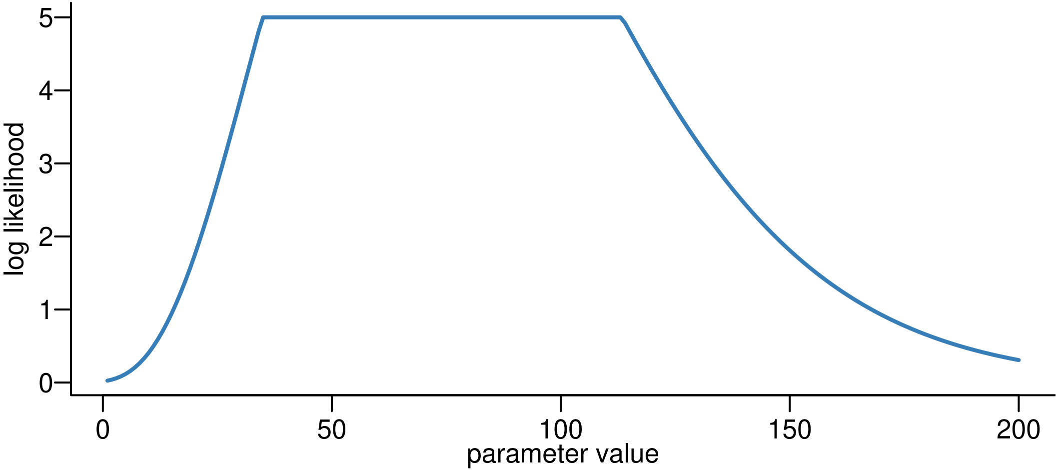
Pathologies



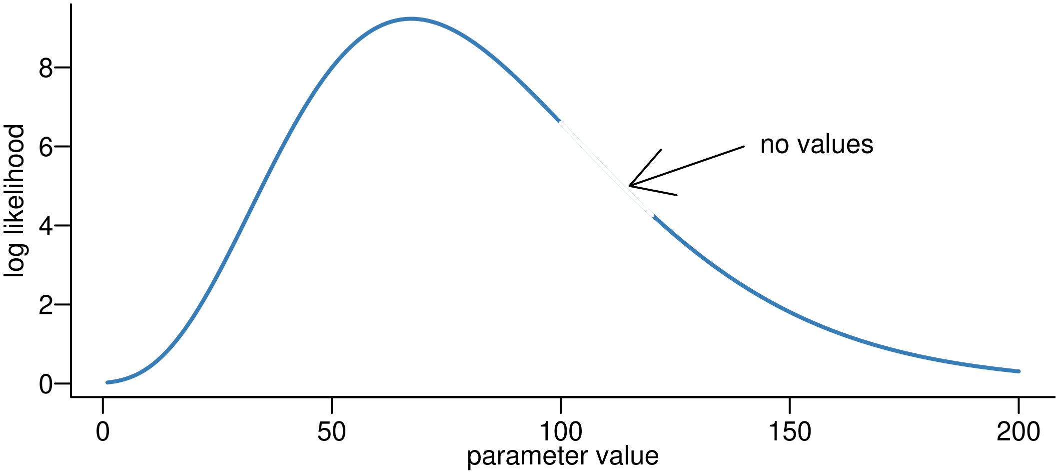
Pathologies



Pathologies



Pathologies



Issues

- Problems are rare for smaller models such as GARCH(1,1)
- The more parameters there are, the more likely problems are
- And the more data we need for estimation
- GARCH might need at least 500 observations, Student-t GARCH 3,000
- For the multivatiate models discussed in the next Chapter, estimation problems are common
- For example, multiple local minima

Estimation

Estimating volatility models

- Once we have data and specified the volatility model we want to estimate
- We estimate the model in R using a specialised volatility modelling package
- There are several alternatives, and we use `rugarch` by Alexios Ghalanos
- cran.r-project.org/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf
- cran.r-project.org/web/packages/rugarch
- The same author has a new library, but that is less mature, especially for the multivariate models in the next chapter
- See https://www.financialriskforecasting.com/notebook/Volatility/univariate.vol.html

Overview of what the estimation entails

- We have a vector of returns — y
- And pick a model — Normal GARCH(1,1) or any of the models we discussed earlier
- We have to inform the library what it is to do
- And then we get a wealth of output, both tables and plots

There are three steps

1. Specifying the model — with `ugarchspec`
2. Estimating the model — with `ugarchfit`
3. Analysing the output — estimated parameters, fitted volatility, log likelihood and other useful results

Specifying the model

- We can use default model specifications by
- `specification = ugarchspec()`
- Else
- We have to tell function that does the estimation the following things
  1. What model to estimate — ARCH(3), GARCH(1,1) apARCH or whatever. This is called the `variance.model` in the estimation
  2. How to model the conditional mean. Should we include one? What function for the mean do we want? This is called the `mean.model` in the estimation
  3. What is the conditional distribution? Normal? Student-t? skewed Student-t? This is called the `distribution.model` in the estimation

Model specification — `variance.model`

- `variance.model = list( )`
- This list contains the variance model specification
- The standard ARCH and GARCH is "`sGARCH`" (default). Another is "`apARCH`"
- `garchOrder`. The ARCH and GARCH orders (lags)
- `external.regressors` — Like geopolitics or climate risk
- Examples

GARCH(1,1) `variance.model = list()` — Its the default
GARCH(1,1) `variance.model = list( garchOrder = c(1, 1))`
ARCH(3) `variance.model = list( garchOrder = c(3,0))`
apARCH `variance.model = list(model = "apARCH")`

The Mean — `mean.model`

- `mean.model = list()` containing the mean model specification
- By default it estimates an ARMA(1,1) model
- We usually want to tell it not to estimate the mean
- `mean.model = list( armaOrder = c(0,0),include.mean = FALSE)`
- But if you are interested in forecasting returns, this would be one way to do it

Conditional Distribution — `distribution.model`

- The conditional density to use for the innovations
- The default is "`norm`" for the normal distribution
- We also use "`std`" for the Student-t
- But there are many others that could be used
- Examples

normal `distribution.model = "norm"`

Student-t `distribution.model = "std"`

Example Specifications

- Default

```
specification = ugarchspec()
```

- Normal GARCH(1,1), no mean

```
specification = ugarchspec(
          mean.model = list(
                    armaOrder = c(0,0),include.mean = FALSE))
```

- Student-t ARCH(2), no mean

```
specification = ugarchspec(
          variance.model = list( garchOrder = c(2, 0))
          mean.model =
                    list( armaOrder = c(0,0),include.mean = FALSE)
          distribution.model = "std"
)
```

Estimation

- Once we have specified the model we want to estimate
- We call ugarchfit() to estimate it
- Default model

```
library(rugarch)
specification = ugarchspec()
results = ugarchfit(spec = specification, data = y)
```

Analyse Results

- Once you have the `results` there are many ways to analyse the results
- You can print them on the screen by just typing `results` into the R console
- Or you can plot the results by `plot results`
- You can also extract specific information out of `results` such as the parameter estimates, the standard deviation of the parameters, log likelihood and the like
- Use functions like `coef(results)`, `sigma(results)`, `likelihood(results)` to extract key values
- When making a report, such as in your assignments, if you use `Quarto` you can directly put the estimation results into your report
- Which is both efficient and guards against mistakes
- Since, if you re-estimate the model, your results will always contain the latest results

Likelihood ratio tests

Alternatives

- Do not confuse the discussion here with backtesting as presented later

Likelihood Ratio Test

- Consider two models, where one is *nested* inside the other
- That is, one is a strict subset of the other
- Like ARCH(1) is a subset of GARCH(1,1)
- But ARCH(2) and GARCH(1,1) are not nested inside each other
- The likelihood ratio test helps us to ascertain whether the parameters specific to one model are statistically significant
- So, for example, is the GARCH(1,1) significantly better than ARCH(1)
- Or is the Student-t apARCH significantly better than the normal GARCH(1,1)
- Is akin to a standard t-test for one parameter
- But also allows testing of multiple parameters

Sample Tests

Unrestricted	Restricted	Test
ARCH(4)	ARCH(1)	$H_0 : \alpha_2 = \alpha_3 = \alpha_4 = 0$
GARCH(1,1)	ARCH(1)	$H_0 : \beta = 0$
GARCH(2,2)	GARCH(1,1)	$H_0 : \beta_2 = \alpha_2 = 0$
APARCH	GARCH	$H_0 : \delta = 2, \zeta = 0$

Each test sets some parameters to zero and checks if the simpler model suffices

Likelihood Ratio Test

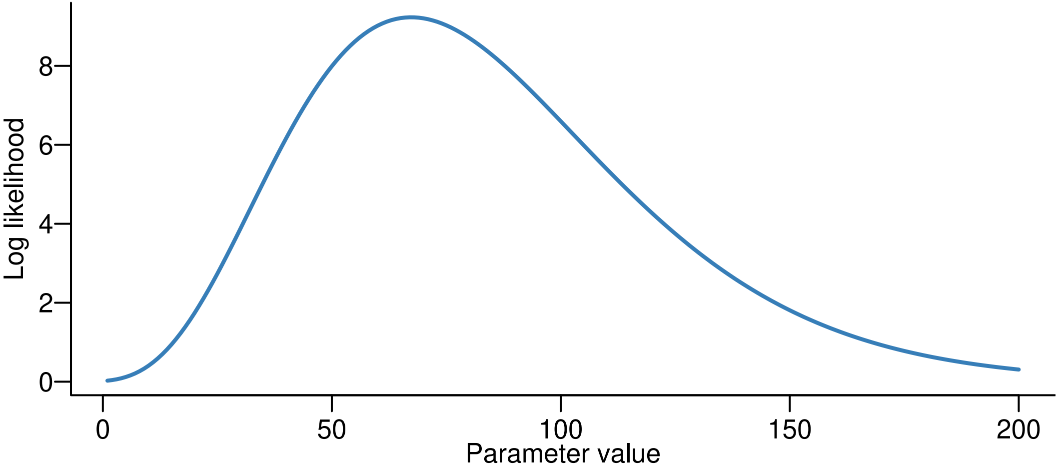
- Denote the log likelihood by $\log \mathcal{L}$
- Call the nested model *restricted* (R) and the other *unrestricted* (U)
- By definition

$$\log \mathcal{L}_R \leq \log \mathcal{L}_U$$

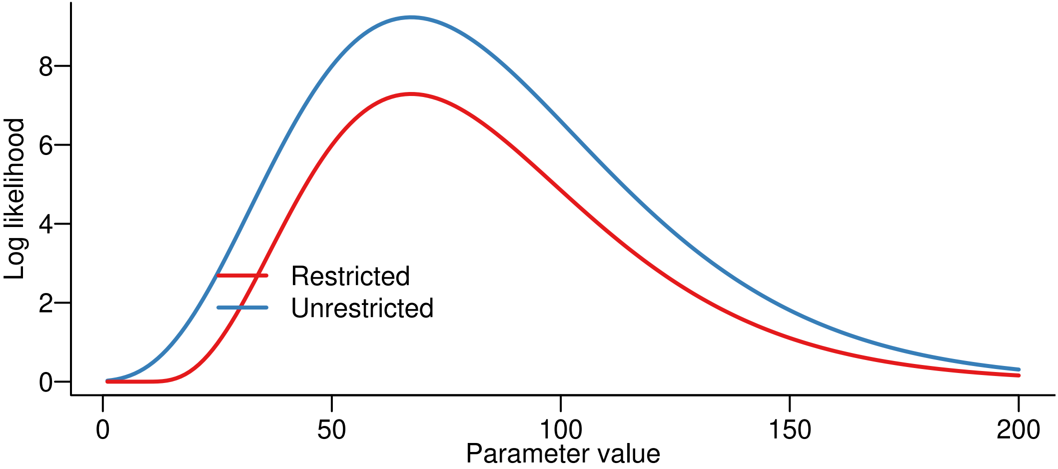
- We want to identify whether $\log \mathcal{L}_R - \log \mathcal{L}_U$ is statistically different from zero

$$H_0 : \mathcal{L}_R \text{ is sufficient (restricted model holds)}$$
$$H_A : \mathcal{L}_U \text{ is required (unrestricted model improves fit)}$$

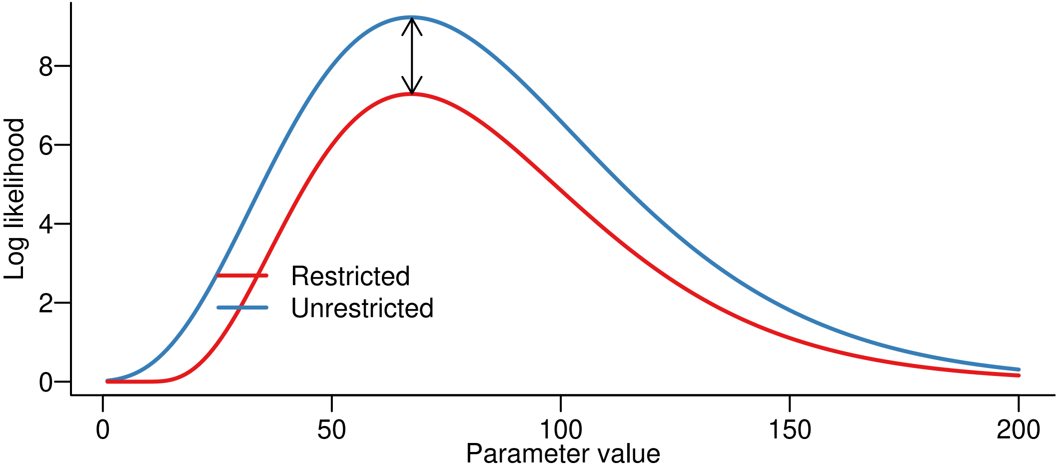
Likelihood Ratios



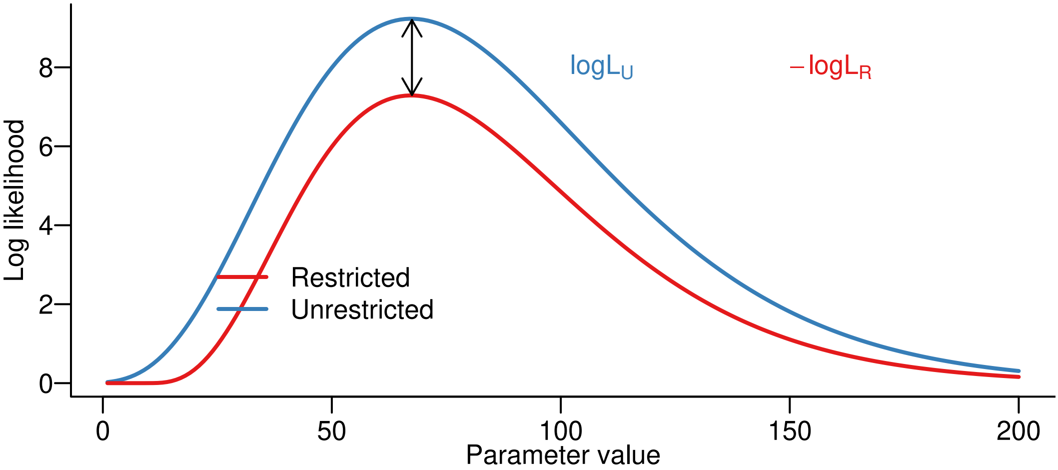
Likelihood Ratios



Likelihood Ratios



Likelihood Ratios



Likelihood Ratio Test

- Two times the difference between the unrestricted and restricted likelihoods is χ^2 distributed
- The degrees of freedom is equal to the number of restrictions

$$\text{LR} = 2(\log \mathcal{L}_U - \log \mathcal{L}_R) \sim \chi^2_{(\text{number of restrictions})}$$

```
qchisq(p=1-0.05,df=1)
 3.841459
```

χ^2 in R

```
x=seq(0.0,9,by=0.1)
plot(x,dchisq(x,3),type='l')
plot(x,pchisq(x,3),type='l')
plot(x,dchisq(x,1),type='l')
plot(x,pchisq(x,1),type='l')
```

Sign of Likelihood

- Most optimisers minimise
- And return the negative function value
- So, `res=garchFit(y)`
- Likelihood is the negative of `likelihood(res)=res@fit$llh`

Compare GARCH(1,1) to ARCH(1)

S&P-500 10,000 observations 12 November 1980 to 9 July 2020

```
spec0 = ugarchspec(variance.model = list( garchOrder = c(1, 0)),
  mean.model = list( armaOrder = c(0,0),include.mean = FALSE))
res0 = ugarchfit(spec = spec0, data = y,solver="hybrid")
likelihood(res0)

spec1 = ugarchspec(variance.model = list( garchOrder = c(1, 1)),
  mean.model = list( armaOrder = c(0,0),include.mean = FALSE))
res1 = ugarchfit(spec = spec1, data = y,solver="hybrid")
likelihood(res1)

LR=2 * (likelihood(res1)-likelihood(res0))
LR
740.9976
pvalue=1-pchisq(LR,1)
pvalue
0
```

Sample Results

S&P-500 10,000 observations 12 November 1980 to 9 July 2020

Model	ω	α	β	ζ	δ	ν	lik
ARCH(1)	0.0000861	0.325					31347
GARCH(1,1)	0.0000018	0.095	0.891				32693
tGARCH(1,1)	0.0000010	0.079	0.915			5.926	33012
APGARCH(1,1)	0.0000001	0.06	0.889	0.352	2.573		32794

The tGARCH has the highest likelihood, followed by the apGARCH. It might be worthwhile to estimate a Student-t apGARCH, which might be significantly better than the others

Diagnosing Volatility Models

Diagnosing Volatility Models

- Ideally we would want to know how a model performs operationally
- But usually we have to make do with an in-sample comparison

Residual Analysis

- Test if a single model is correctly specified
- Relies on checking whether the residuals (innovations) behave according to the model assumptions
- For example, when estimating a normal GARCH, are the standardised residuals approximately IID normal?

Residual Analysis of GARCH(1,1)

- Returns have a distribution (here assumed to be normal)

$$
\begin{aligned}
y_t \sim & \mathcal{N}(0, \sigma_t^2) \\
y_t = & \sigma_t \epsilon_t, \\
\epsilon_t \sim & \mathcal{N}(0, 1) \\
\hat{\sigma}_t^2 = & \hat{\omega} + \hat{\alpha} y_{t-1}^2 + \hat{\beta} \hat{\sigma}_{t-1}^2 \\
\hat{\epsilon}_t = & \frac{y_t}{\hat{\sigma}_t}
\end{aligned}
$$

- Can test the estimated residuals $\hat{\epsilon}$ with a Jarque-Bera test for normality and Ljung-Box test for autocorrelation of $\hat{\epsilon}$ and $\hat{\epsilon}^2$
- Can also do a QQ plot for $\hat{\epsilon}$

Residual Tests in R

- Suppose the returns are in vector `y`
- And the estimation results are in `results`
- We can extract the in sample estimated conditional standard deviation, $\hat{\sigma}_t$, by `sigma(res1)`
- We have one complication: the library we are using gives it to us as a time series, but we need to do the calculation

$$\hat{\epsilon}_t = \frac{y_t}{\hat{\sigma}_t}$$

- The R function `coredata()` removes time-series attributes and converts the object into a plain numeric vector
- `residuals=coredata(y)/coredata(sigma(results))`

Residual Tests in R on Vector `residuals`

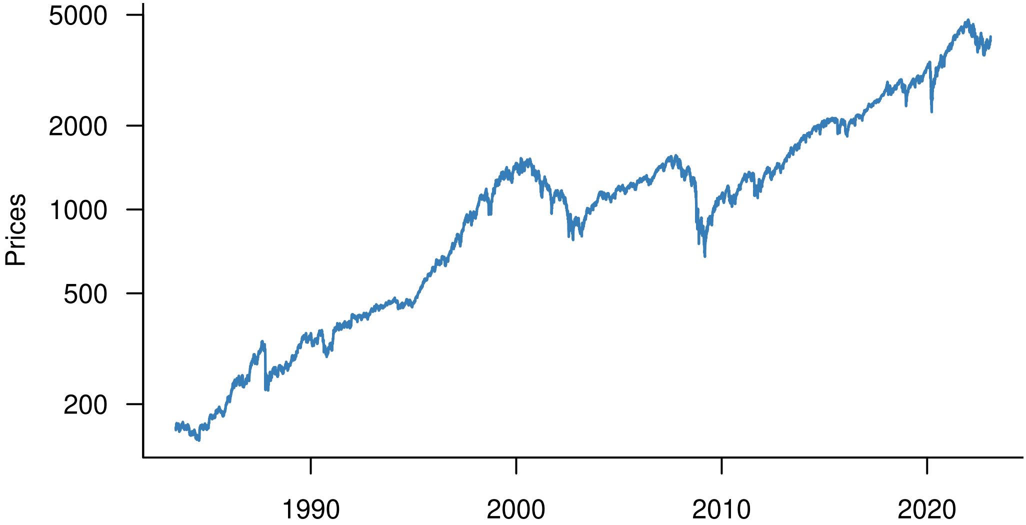
These calculations are familiar from Chapter 1

```
Box.test(residuals, lag = 20,type = c("Ljung-Box"))
library(tseries)
jarque.bera.test(residuals)
library(car)
qqPlot(residuals)
qqPlot(residuals,distribution="t",df=4)
```

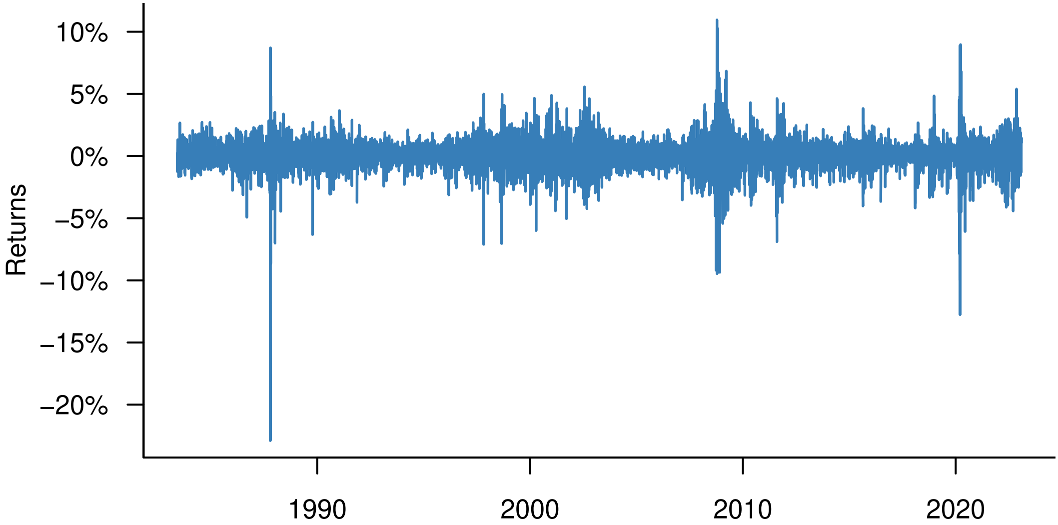
Estimation Results from the S&P-500

- We use a typical S&P-500 return series as example
- First issue prices and then returns
- After that the in sample conditional volatility
- Superimpose ± 2 conditional volatility bands on the return series
- And look at ACF and QQ plots
- We can do this with `plot(results)`

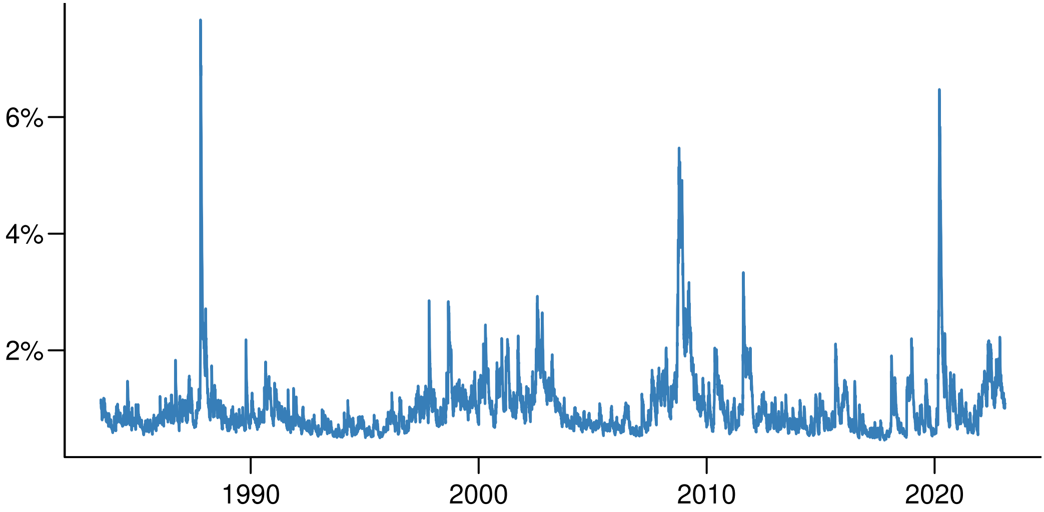
Price



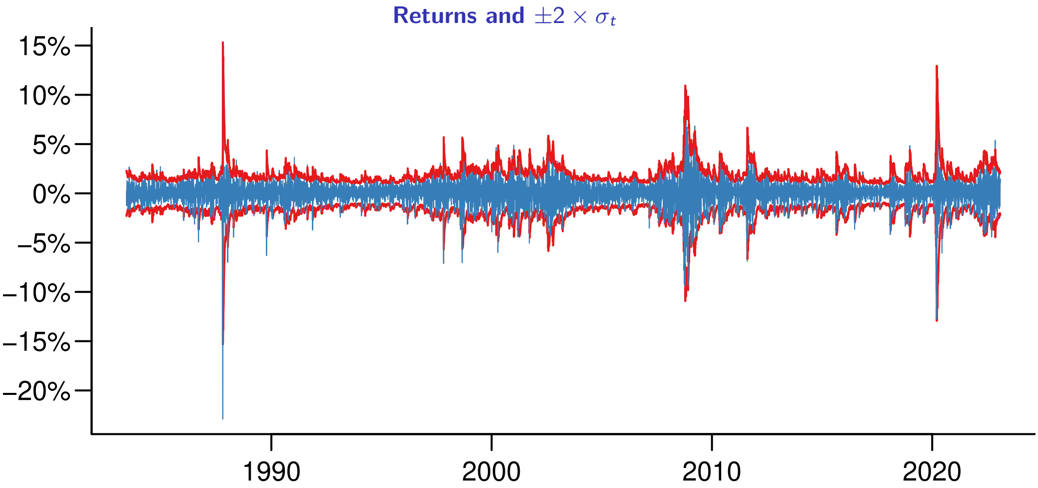
Return



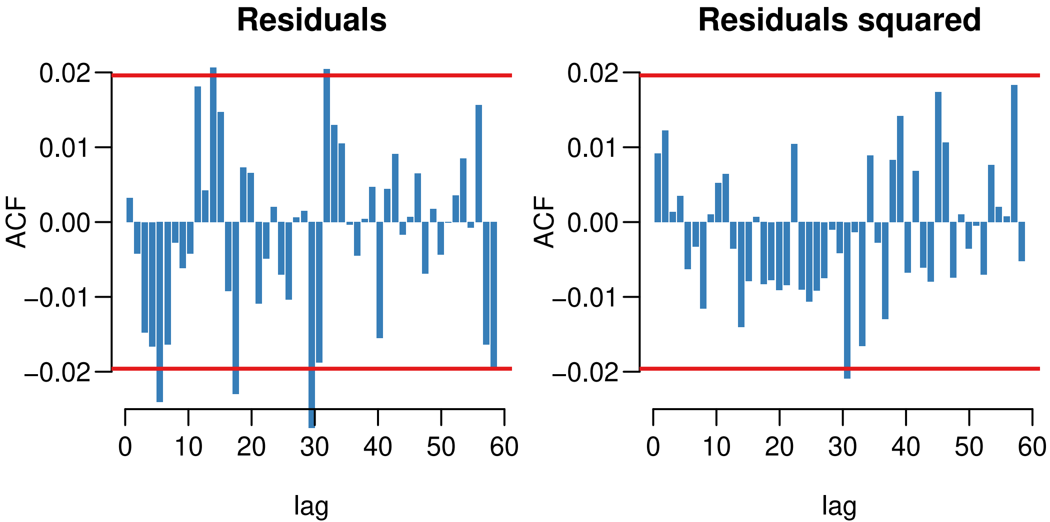
Conditional volatility (in-sample)



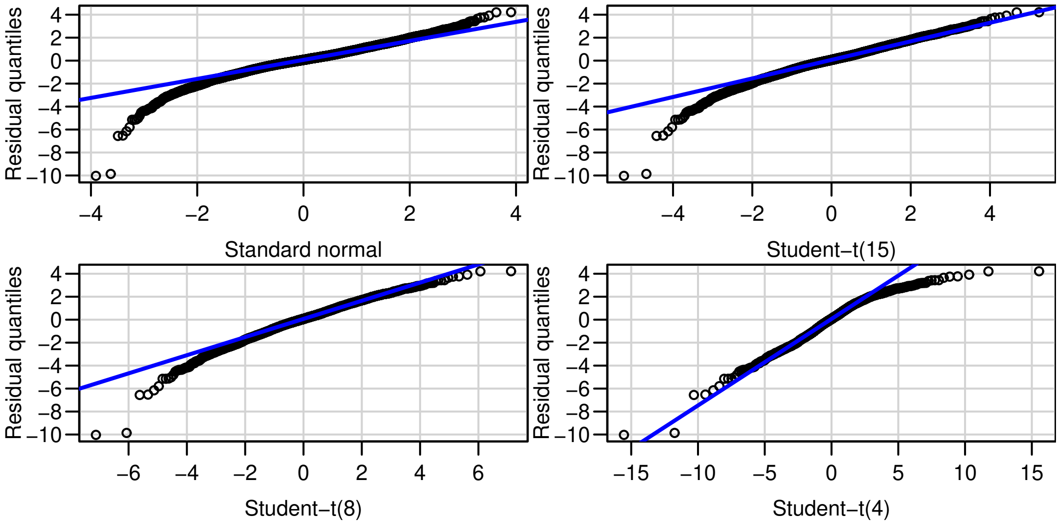
Conditional Volatility and Returns



ACF of Squared Residuals



QQ Plot of Residuals



What Do Good Residuals Look Like?

- No autocorrelation in residuals
- No autocorrelation in squared residuals
- Residuals approximately normal (or t if estimated as such)
- QQ plot follows 45-degree line

Alternative Volatility Models

Realised Volatility

- Calculated from intraday return data sampled at regular intervals to estimate volatility or the full covariance matrix
- Pros
  - Purely data driven and no reliance on parametric methods
- Cons
  - Intraday data must be available, data is difficult to obtain, hard to use, not very clean and expensive
  - Intraday data exhibit market microstructure noise, irregular trading and diurnal patterns that must be addressed

Implied Volatility

- Black-Scholes (BS) equation

$$\text{Option price} = BS(T, r, S, X, \sigma)$$

- Volatility implied by the price

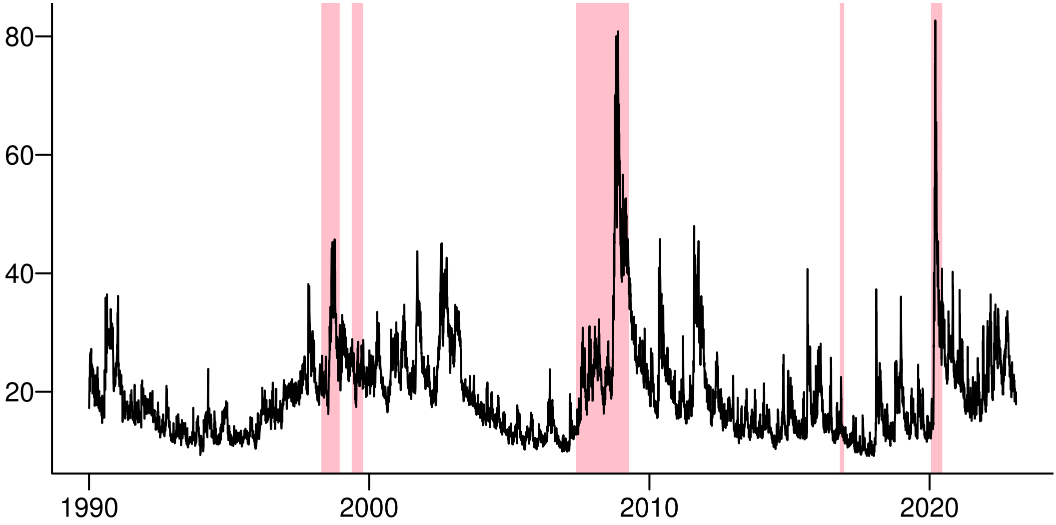
$$\text{Implied volatility} = BS^{-1}(T, r, S, X, \text{option price})$$

- Based on current market prices and not historical data
- Depends critically on accuracy of BS model (constant volatility and normal innovations)

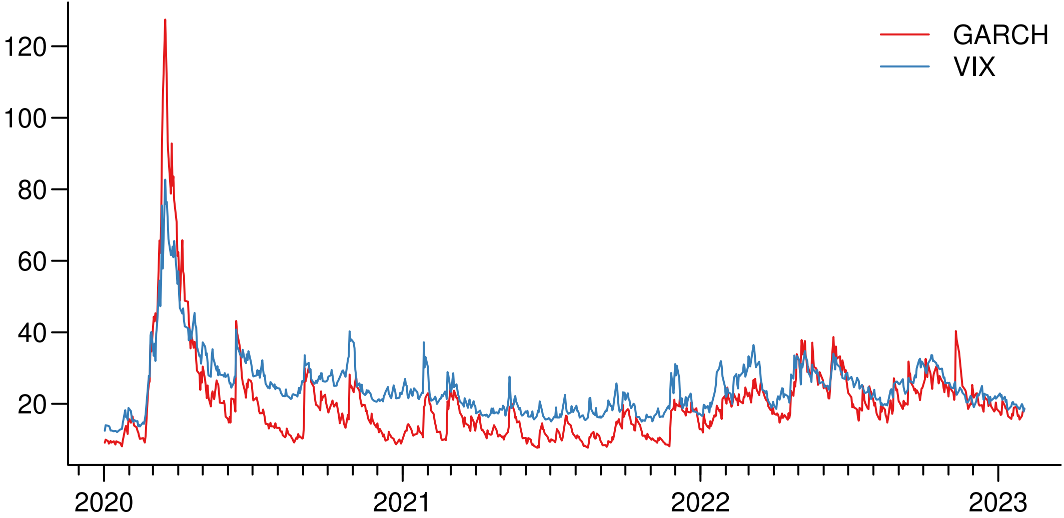
VIX

- From the Chicago Board Options Exchange (CBOE)
- Volatility index – VIX
- It reflects the market's expectation of annualised volatility over the next 30 days, derived from S&P-500 index option prices
- Similar to implied volatilities
- It is known as the *"fear index"* of the financial markets
- The VIX has entered popular culture, inspiring a thriller novel by Robert Harris titled *The Fear Index*
- The following figure shows the VIX along with key market events

VIX

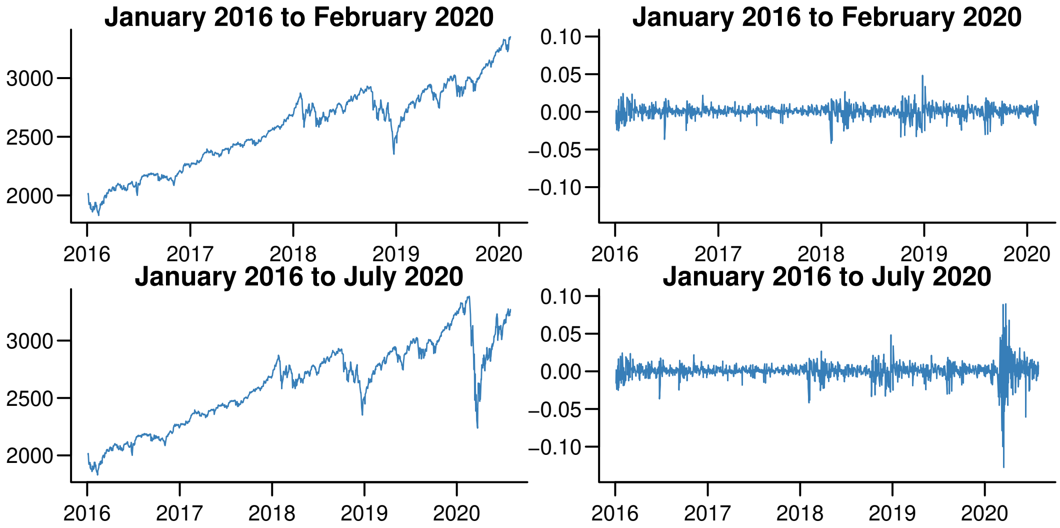


S&P-500 GARCH and VIX

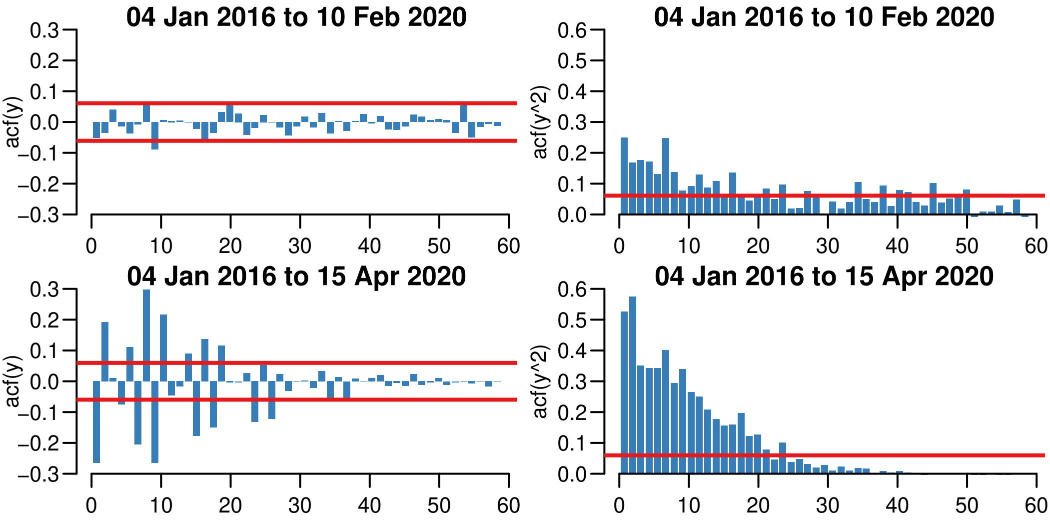


Case: Covid-19

S&P-500 With and Without 2020



S&P-500 ACF With and Without 2020



GARCH Analysis

- Consider the normal and student t GARCH
- When we include and exclude the crisis period
- The outcome is exactly what we would expect

S&P-500

2016-02-22 to 2020-02-10

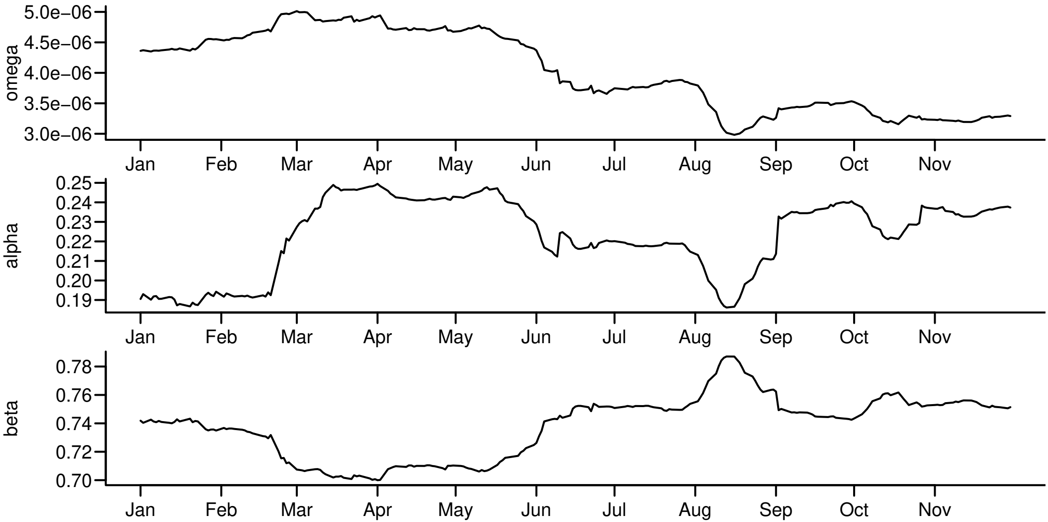
ω	α	β	ν	$\alpha + \beta$	Half
4.6e-06	0.19	0.73		0.93	9
2.7e-06	0.19	0.79	4.24	0.97	27

2016-07-14 to 2020-07-02

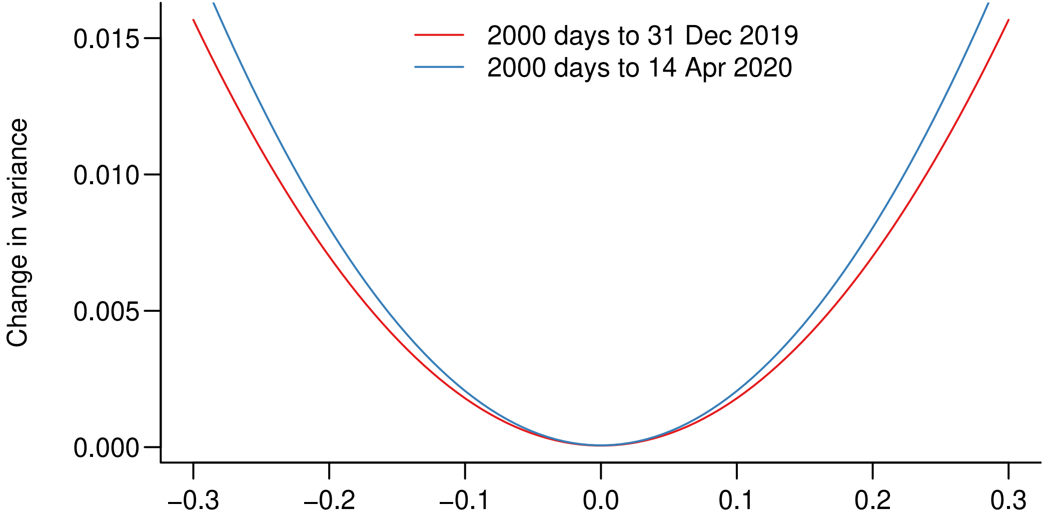
ω	α	β	ν	$\alpha + \beta$	Half
3.7e-06	0.22	0.75		0.97	23
2.1e-06	0.22	0.80	3.86	1.02	NA

Why do you think the half-life is NA for the Student-t?

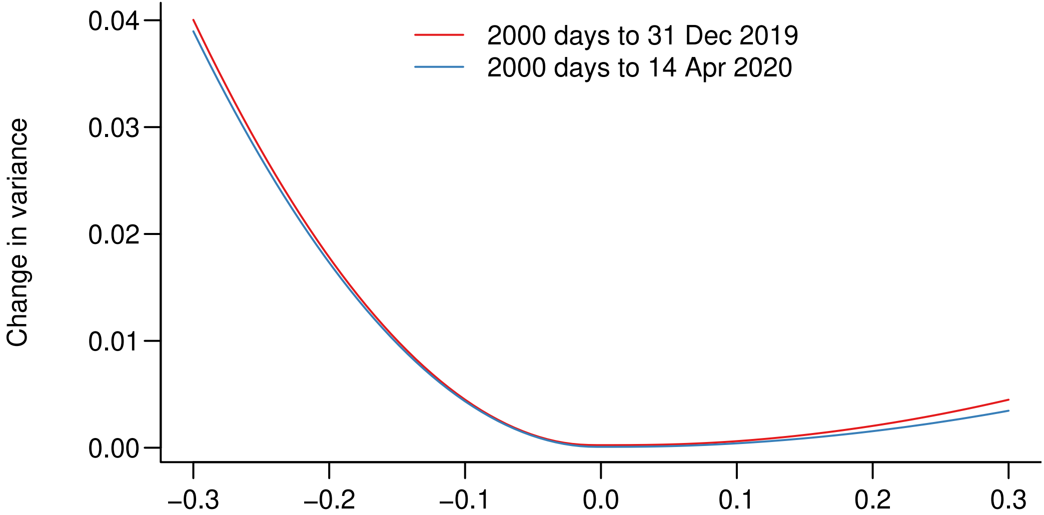
S&P-500 GARCH Coefficients in 2020



S&P-500 News Impact (GARCH)

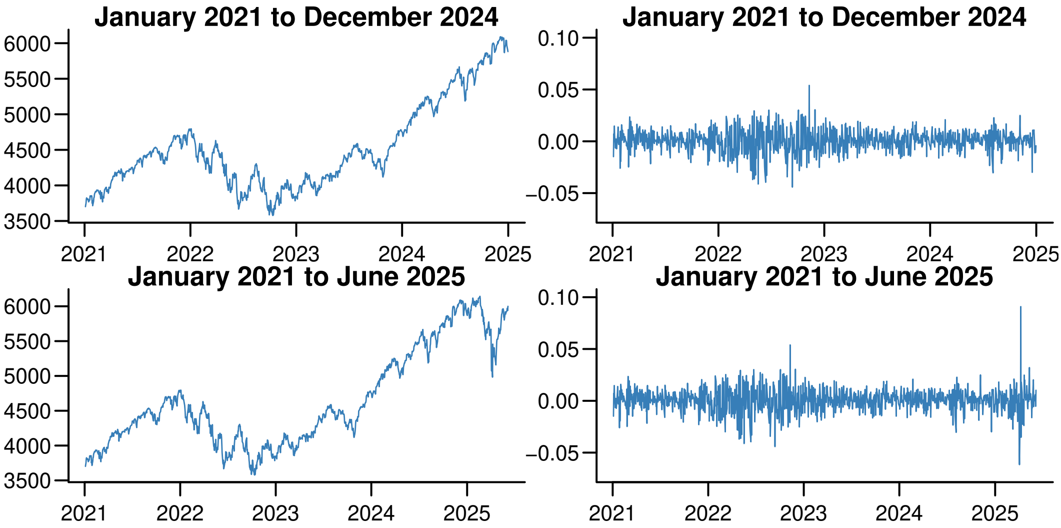


S&P-500 News Impact (APGARCH)

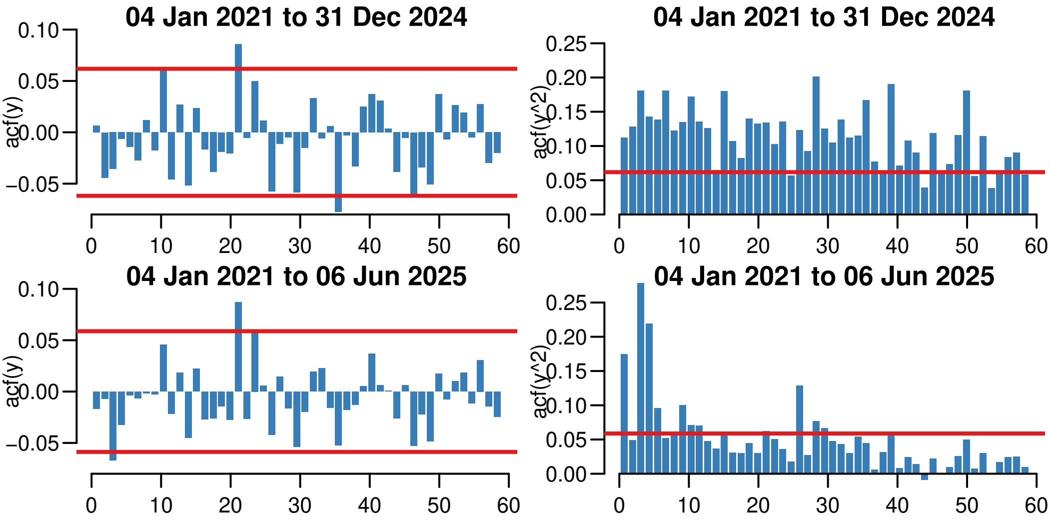


Case: Trump Tariffs

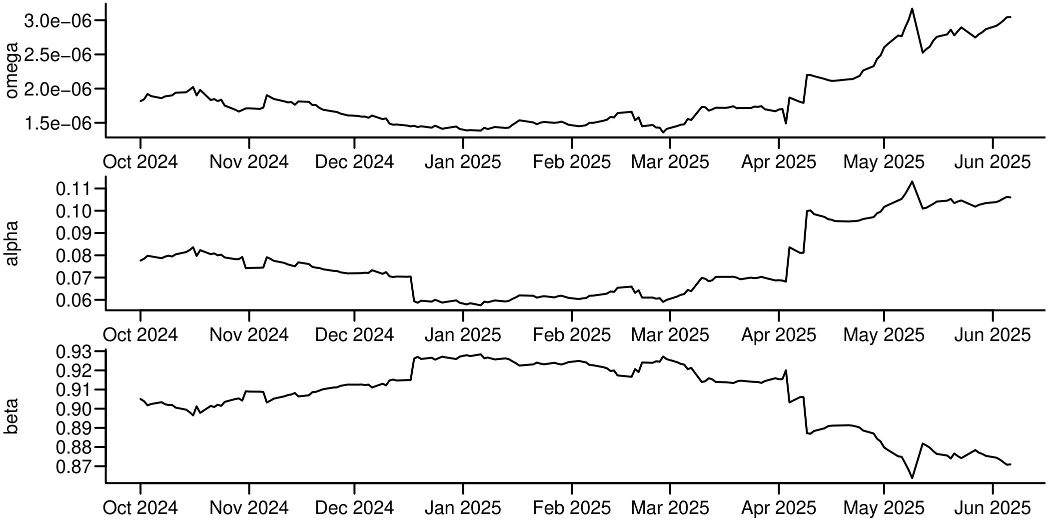
S&P-500 With and Without 2025



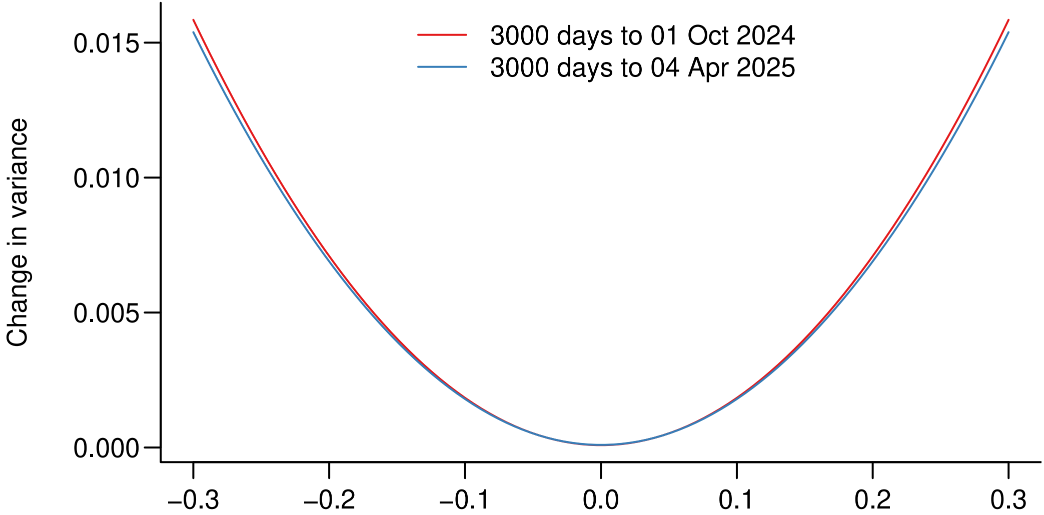
S&P-500 ACF With and Without 2025



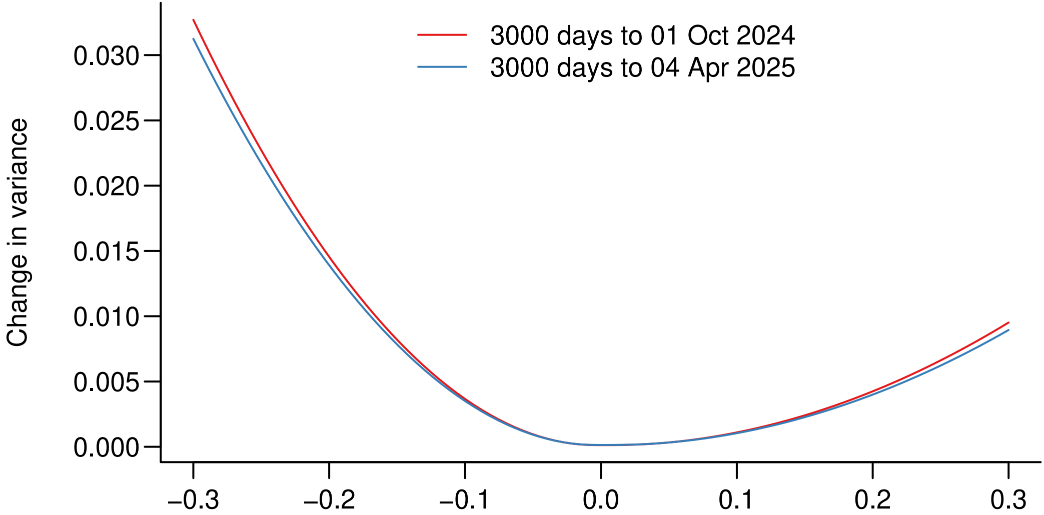
S&P-500 GARCH Coefficients in 2024 and 2025



S&P-500 News Impact (GARCH)



S&P-500 News Impact (APGARCH)



Appendix

GARCH Half-Life

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2$$

As

$$E_{t-1}[y_t^2] = \sigma_t^2$$

Then given $t-1$ and expectation

$$y_t^2 = \omega + (\alpha + \beta) y_{t-1}^2$$

Subtract $\sigma^2 = \omega/(1-\alpha-\beta)$ from both sides

$$\begin{aligned} y_t^2 - \sigma^2 =& \omega + (\alpha+\beta) y_{t-1}^2 - \frac{\omega}{1-\alpha-\beta} \\ =& (\alpha+\beta)(y_{t-1}^2 - \sigma^2) \end{aligned}$$

GARCH Half-Life (cont.)

- Suppose we are interested in σ_{t+n} as a function of y_{t+1}
- Then
$$y_{t+n}^2 - \sigma^2 = (\alpha + \beta)^{n-1}(y_{t+1}^2 - \sigma^2)$$
- The number of periods, n^*, it takes for conditional variance to revert back halfway towards unconditional variance
$$\sigma_{t+n^*}^2 - \sigma^2 = \frac{1}{2}(\sigma_{t+1}^2 - \sigma^2)$$
- So we want to solve for n^* as a function of the parameters

GARCH Half-Life (cont.)

- So

$$(\alpha+\beta)^{n^*-1}(\sigma_{t+1}^2-\sigma^2)=\frac{1}{2}(\sigma_{t+1}^2-\sigma^2)$$

- So

$$n^*=1+\frac{\log\left(\frac{1}{2}\right)}{\log(\alpha+\beta)}$$

- And as $(\alpha+\beta)\longrightarrow 1$, $n^*\to\infty$, memory is infinite

Back to main Memory slides