Quantitative Analysis of Financial Markets GARCH, VaR and VECM

Benjamin Ee

Volatility Forecasting

[*this is a case study for ARIMA-X*]

Definition of Volatility and Variance Rate

- Define σ_n as the volatility of a market variable on day n, as estimated at the end of day $n-1$.

The square of the volatility σ_n^2 on day n is the **variance rate**.

Denote S_i as the value of a variable at the end of day

- Define the log return as $u_i := \ln \frac{S_i}{S_{i-1}}$.
- An unbiased estimate of the mean of u_i, using the most recent m observations with respect to "today" n is

$$\overline{u} = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}.$$

An unbiased estimate of the variance rate per day, σ_n^2 ,is

$$\sigma_n^2 = \frac{1}{m-1}\sum_{i=1}^{m}(u_{n-i} - \overline{u})^2.$$

Really, what is volatility?

- Intuitively, volatility is a measure of a financial asset's readiness to move from one price to another price, in such a way that leaves the market participants unsure about the next quantum of price change.

- The more ready the asset's price is to move, the more uncertain its price in the future will be. The resulting price uncertainty makes the asset risky to the investors, as the price may move in the direction contrary to what they expect, and they become exposed to the consequent losses.

Alternative Estimation

- For monitoring daily volatility, we define u_i as a percentage change in the market variable instead:

$$u_i = \frac{S_i - S_{i-1}}{S_{i-1}}.$$

- Assume u to be zero.
- Replace $m - 1$ by m to obtain a maximum likelihood estimate:

$$\sigma_n^2 = \frac{1}{m} \sum_{i=1}^{m} u_{n-i}^2.$$

Volatility modelling: EWMA

Weighting Scheme

- So far, we give an equal weight to $u^2_{n-1}, u^2_{n-2}, \ldots, u^2_{n-m}$.

To estimate the *current* level of volatility, give more weight α_i to recent data.

$$\sigma_n^2 = \sum_{i=1}^{m} \alpha_i u_{n-i}^2. \qquad (1)$$

- The weights $\alpha_i, i = 1, 2, \ldots, m$ must satisfy the following:
  - positive: $\alpha_i > 0$.
  - lesser weight for older observations: $\alpha_i < \alpha_j$ when $i > j$.
  - summed to unity: $\sum_{i=1}^{m} \alpha_i = 1.$

Model for Estimating the Variance Rate: EWMA

- The EMWA model is a special case of (1) where the weights α_i decrease exponentially as we move back through time.

- Specifically, with $0 < \lambda < 1$,

$$\alpha_{i+1} = \lambda \alpha_i,$$

the EMWA model is

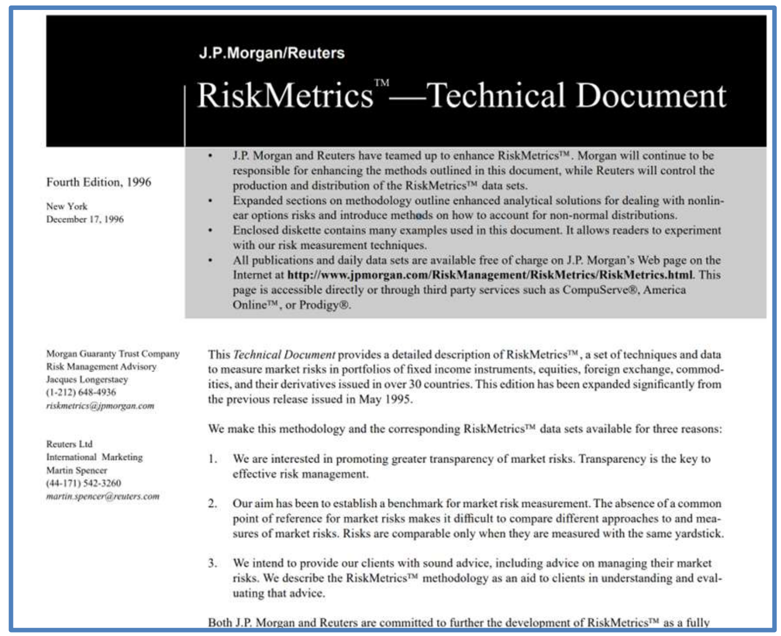
$$\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1 - \lambda) u_{n-1}^2.$$

- Only the current estimate of the variance rate and the most recent observation on the value of the market variable are needed.

- Risk Metrics database uses the EWMA model with $\lambda = 0.94$ for updating daily volatility estimates.

Attractions of EWMA

- Relatively little data needs to be stored
- We need only remember the current estimate of the variance rate and the most recent observation on the market variable
- Tracks volatility changes
- 0.94 is a popular choice for λ



J.P.Morgan/Reuters

RiskMetrics™—Technical Document

Fourth Edition, 1996

New York
December 17, 1996

- J.P. Morgan and Reuters have teamed up to enhance RiskMetrics™. Morgan will continue to be responsible for enhancing the methods outlined in this document, while Reuters will control the production and distribution of the RiskMetrics™ data sets.
- Expanded sections on methodology outline enhanced analytical solutions for dealing with nonlinear options risks and introduce methods on how to account for non-normal distributions.
- Enclosed diskette contains many examples used in this document. It allows readers to experiment with our risk measurement techniques.
- All publications and daily data sets are available free of charge on J.P. Morgan's Web page on the Internet at **http://www.jpmorgan.com/RiskManagement/RiskMetrics/RiskMetrics.html**. This page is accessible directly or through third party services such as CompuServe®, America Online™, or Prodigy®.

Morgan Guaranty Trust Company
Risk Management Advisory
Jacques Longerstaey
(1-212) 648-4936
riskmetrics@jpmorgan.com

Reuters Ltd
International Marketing
Martin Spencer
(44-171) 542-3260
martin.spencer@reuters.com

This *Technical Document* provides a detailed description of RiskMetrics™, a set of techniques and data to measure market risks in portfolios of fixed income instruments, equities, foreign exchange, commodities, and their derivatives issued in over 30 countries. This edition has been expanded significantly from the previous release issued in May 1995.

We make this methodology and the corresponding RiskMetrics™ data sets available for three reasons:

1. We are interested in promoting greater transparency of market risks. Transparency is the key to effective risk management.
2. Our aim has been to establish a benchmark for market risk measurement. The absence of a common point of reference for market risks makes it difficult to compare different approaches to and measures of market risks. Risks are comparable only when they are measured with the same yardstick.
3. We intend to provide our clients with sound advice, including advice on managing their market risks. We describe the RiskMetrics™ methodology as an aid to clients in understanding and evaluating that advice.

Both J.P. Morgan and Reuters are committed to further the development of RiskMetrics™ as a fully

Please read this document when you have time:

https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

Pros and Cons of EWMA

Advantage: Able to model volatility clustering

Disadvantage: Unable to directly account for observation that long run volatility tends to mean revert to a steady state

Volatility clustering and mean reversion [1/3]

1. Empirically, we can use the VIX index to **measure actual 'real world' volatility**

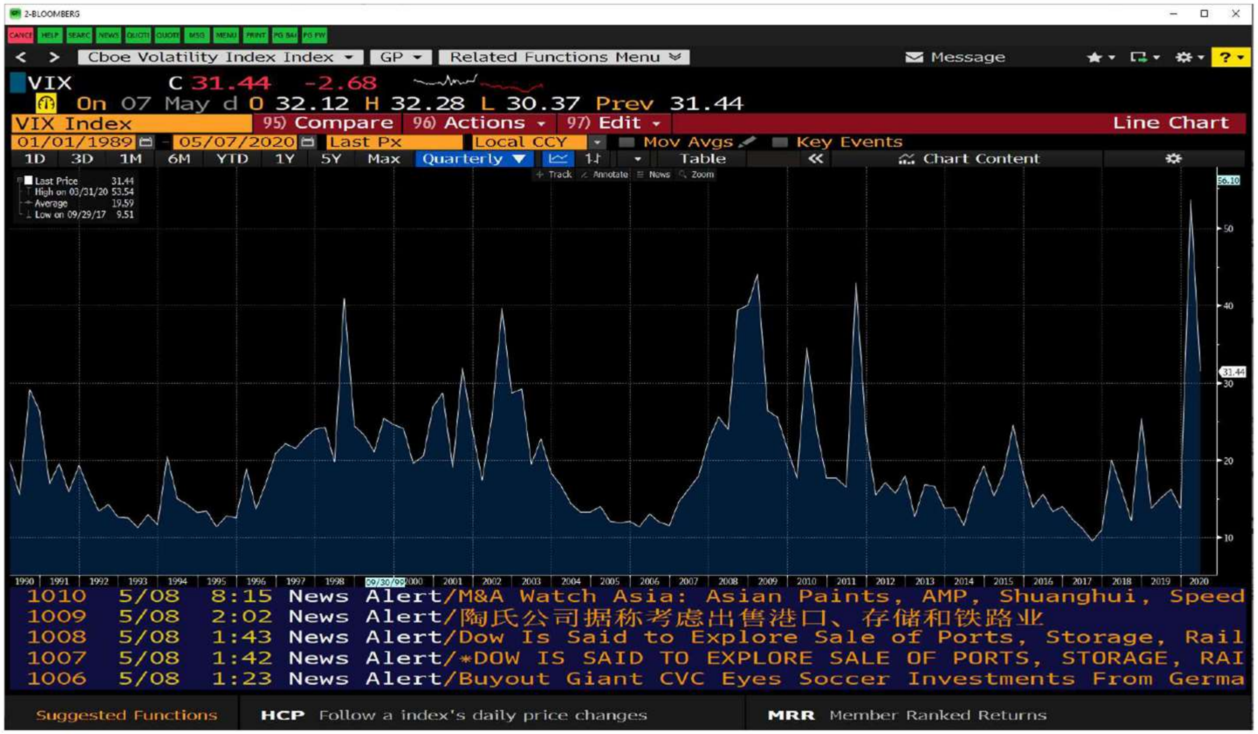
2. Volatility clustering means that if actual volatility is high 'recently', **it will likely remain elevated above long run mean in near future**

3. Mean reversion means that eventually, **volatility will fall to long run steady state levels**

WHY IS VOLATILITY MEAN REVERTING? [2/3]

- VIX is formally a measure of volatility in US equities market (S&P 500).
- Specifically it is a function of ATM implied volatility on near month options for the S&P 500
- Periods of high volatility tend to coincide with high *downside* volatility such as 4Q 2008 / 1Q 2009 / 1Q 2020
- This is because put option prices tend to increase sharply during times of financial stress. The effect is not symmetrical on call option prices. This leads to the "volatility skew"
- Why is downside volatility mean reverting?

Volatility clustering and mean reversion [3/3]



MEAN REVERTING VOLATILITY

Long run 30-year average value of VIX index is around 20

Intuitively, we can formulate trading strategies based on the assumption that whenever VIX increases above its long run average, it will eventually mean revert

We can directly model volatility clustering and mean reversion of volatility using GARCH model! [next slide]

Volatility modelling: GARCH

GARCH (1,1)

- Bollerslev in 1986 proposed the GARCH(1, 1) model.
- In GARCH(1,1), σ^2_n is calculated from a **long-run average variance rate**, V_L, as well as from σ^2_n and u^2_{n-1}
- The GARCH(1,1) model is

$$\sigma^2_n = \gamma V_L + \alpha u^2_{n-1} + \beta \sigma^2_{n-1}.$$

- Since the weights must sum to unity, it follows that

$$\gamma + \alpha + \beta = 1.$$

- The EWMA model is a particular case of GARCH(1,1) where $\gamma = 0$, $\alpha = 1 - \lambda$ and $\beta = \lambda$.

Parameters of GARCH (1,1)

- Let $\omega := \gamma V_L$. The GARCH (1,1) is rewritten as

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 .$$

- Once the parameters ω, α, and β have been estimated, we can calculate γ as $1 - \alpha - \beta$.

- The **long-term variance** V_L can then be calculated as

$$V_L = \frac{\omega}{\gamma} = \frac{\omega}{1 - \alpha - \beta}.$$

- For a stable GARCH(1,1) process, we require $\alpha + \beta < 1$, and (due to economic theory) $\omega > 0$ [**why?**]

Unconditional Stationarity

- GARCH processes are unconditionally stationary.

$$
\begin{aligned}
\sigma_n^2 &= \omega + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2 \\
&= \omega + \alpha u_{n-1}^2 + \beta\Big(\omega + \alpha u_{n-2}^2 + \beta\sigma_{n-2}^2\Big) \\
&= \omega(1+\beta) + \alpha\big(u_{n-2}^2 + \beta u_{n-2}\big) + \beta^2\Big(\omega + \alpha u_{n-3}^2 + \beta\sigma_{n-3}^2\Big) \\
&= \omega\big(1 + \beta + \beta^2 + \cdots\big) + \alpha\big(u_{n-1}^2 + \beta u_{n-2}^2 + \beta^2 u_{n-3}^2 + \cdots\big)
\end{aligned}
$$

- Taking unconditional expectation on both sides, so that

$$\sigma^2 := \mathbb{E}\left(u_n^2\right) = \mathbb{E}\left(u_{n-1}^2\right) = \mathbb{E}\left(u_{n-2}^2\right) = \cdots .$$

Unconditional Stationarity (cont'd)

- Then

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

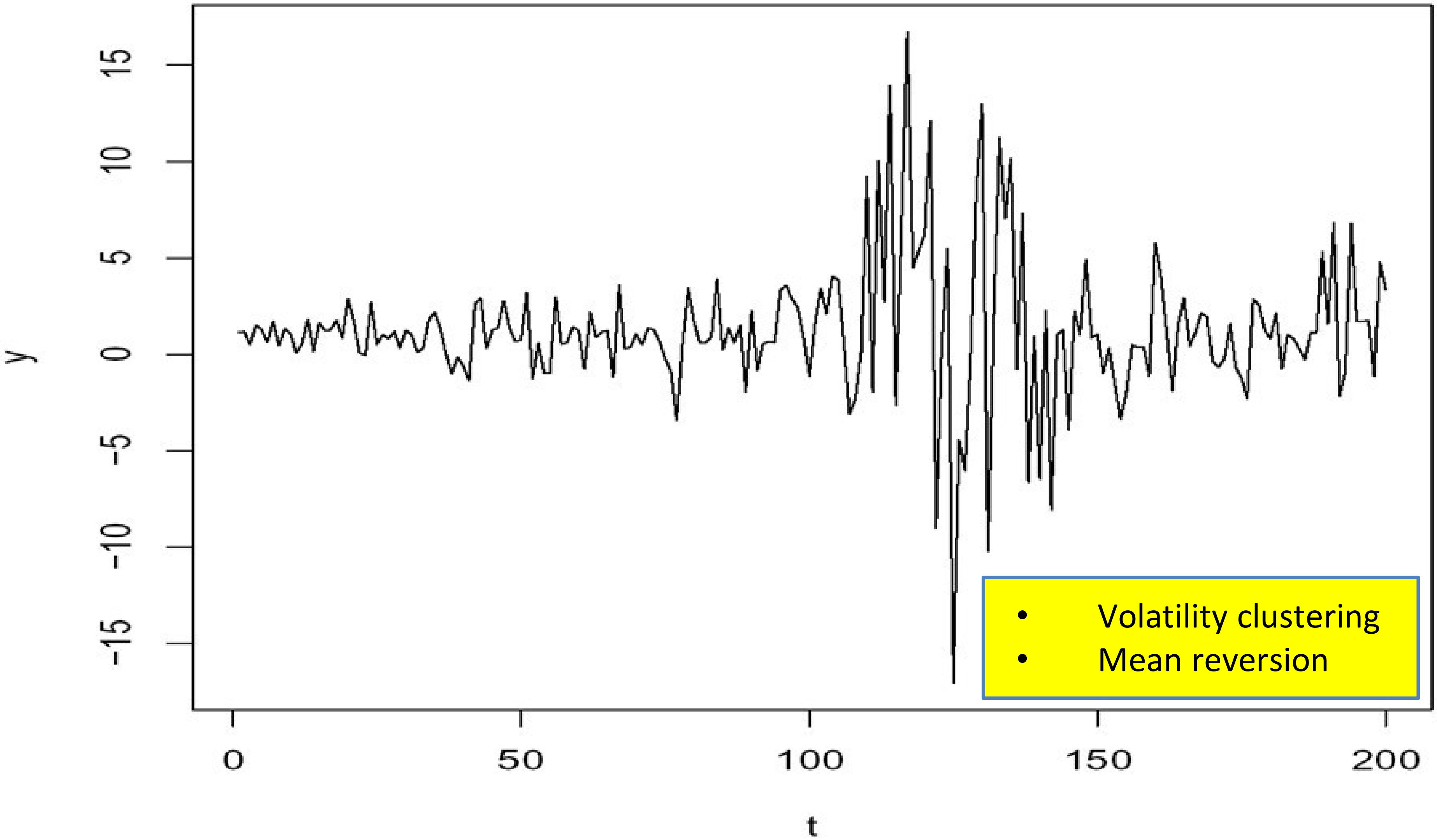
- The unconditional variance σ^2 is constant!

$$\sigma^2 = \frac{\omega}{1-\alpha-\beta} = V_L,$$

provided $\omega > 0$, and $|\alpha + \beta| < 1$.

Simulated GARCH Process

$$\sigma_n^2 = 0.50 + 0.25u_{n-1}^2 + 0.70\sigma_{n-1}{}^2$$



GARCH Example

- Suppose $\sigma_n^2 = 0.000002 + 0.13u_{n-1}^2 + 0.86\sigma_{n-1}^2$.
- The long-run variance rate is 0.0002 so that the long-run volatility per day is 1.41%.
- Suppose that the current estimate of the volatility is 1.6% per day and the most recent percentage change in the market variable is 1%.
- The new variance rate is

$$0.000002 + 0.13 \times 0.01^2 + 0.86 \times 0.016^2.$$

Consequently, the new volatility is 1.53% per day.

How Good Is the Model?

- If a GARCH model is working well, it should remove autocorrelation in u_i^2/σ_i^2.

- We can test whether it has done so by considering the autocorrelation structure for the variables u_i^2 If these show very little autocorrelation, our model for σ_i^2 has succeeded in explaining the autocorrelations in u_i^2.

- Ljung-Box statistic for the u_i^2/σ_i^2 time series.
  - Revision question: what distribution does the test statistic follow?

Variance Targeting

- One way of implementing GARCH(1,1) that increases stability is by using variance targeting.

- We set the long-run average volatility equal to the sample variance.

- Only two other parameters then have to be estimated.

Forecasting Future Volatility

- The variance rate estimated at the end of day $n-1$ for day n, when GARCH(1,1) is used, is

$$\sigma_n^2 = (1-\alpha-\beta)V_L + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2.$$

- It follows that

$$\sigma_n^2 - V_L = \alpha\left(u_{n-1}^2 - V_L\right) + \beta\left(\sigma_{n-1}^2 - V_L\right)$$

- On day $n+\ t$ in the future and with σ^2,

$$\sigma_{n+t}^2 - V_L = \alpha\left(u_{n+t-1}^2 - V_L\right) + \beta\left(\sigma_{n+t-1}^2 - V_L\right)$$

- The expected value of u_{n+t-1}^2 is σ_{n+t-1}^2. Hence

$$\mathbb{E}\left(\sigma_{n+t}^2 - V_L\right) = (\alpha+\beta)\,\mathbb{E}\left(\sigma_{n+t-1}^2 - V_L\right).$$

Forecasting Future Volatility (cont'd)

- Applying the above equation repeatedly yields

$$\mathbb{E}\left(\sigma_{n+t}^2 - V_L\right) = (\alpha + \beta)^t \left(\sigma_n^2 - V_L\right),$$

  which is

$$\mathbb{E}\left(\sigma_{n+t}^2\right) = V_L + (\alpha + \beta)^t (\sigma_n^2 - V_L).$$

- This equation allows you to forecast the volatility on day $n + t$ using the information available at the end of day n.

- **Margin**

  The maintenance margin x is set as, given today's settlement price of F_n,

$$x \geqq 1.645 \sqrt{\mathbb{E}\left(\sigma_{n+1}^2\right)}\, F_n$$

Volatility Term Structures

1 Suppose it is day n. Define

$$V(t) := \mathbb{E}\left(\sigma^2_{n+t}\right), \qquad \text{and} \qquad a := \ln \frac{1}{\alpha + \beta}.$$

2 The predictive equation becomes

$$V(t) = V_L + e^{-at}\big(V(0) - V_L\big).$$

Here, $V(t)$ is an estimate of the instantaneous variance rate in t days.

Volatility Term Structures

- The average variance rate per day between today and time T is given by

$$\frac{1}{T}\int_0^T V(t)\,dt = V_L + \frac{1-e^{-aT}}{aT}\big(V(0)-V_L\big).$$

- Then the volatility per annum for an option lasting T days is

$$\sigma(T) = \sqrt{252\left(V_L + \frac{1-e^{-aT}}{aT}(V(0)-V_L)\right)}.$$

- So from GARCH(1,1), we have a volatility term structure, which is the relationship between the forward-looking volatilities and the maturities.

S&P Volatility Term Structure Predicted from GARCH(1,1)

- Note that a is positive since $\alpha + \beta < 1$.
- ω= 0.0000013465, α = 0.083394, and β = 0.910116.

$$a = \ln\left(\frac{1}{0.083394 + 0.910116}\right) = 0.006511$$

Option Life (days)	10	30	50	100	500
Volatility (% per annum)	27.4	27.1	26.9	26.4	24.3

Impact of Volatility Changes

- We note that $V(0) = \sigma(0)^2/252$.

- When instantaneous volatility $\sigma(0)$ changes by $\Delta\sigma(0)$, volatility for T-day maturity changes by approximately

$$\Delta\sigma(T) \approx \frac{1-e^{-aT}}{aT}\frac{\sigma(0)}{\sigma(T)}\Delta\sigma(0).$$

- Impact of 1% change in the instantaneous volatility predicted from GARCH (1,1)

Option Life (days)	10	30	50	100	500
Volatility increase (%)	0.97	0.92	0.87	0.77	0.33

Summary

- In the EWMA and the GARCH(1,1) models, the weights assigned to observations decrease exponentially as the observations become older.
- The GARCH(1,1) model differs from the EWMA model in that some weight is also assigned to the long-run average variance rate. It has a structure that enables forecasts of the future level of variance rate to be produced relatively easily.
- Maximum likelihood methods are usually used to estimate parameters from historical data in the EWMA, GARCH(1,1), and similar models.
- Once its parameters have been determined, a GARCH(1,1) model can be judged by how well it removes autocorrelation from the u^2.

Volatility modelling applied to Value at Risk (VaR)

Risks and Risk Management

√ Major Risks

§ Market risks

§ Credit risks

§ Operational risks

§ Liquidity risks

§ Legal risks

§ Political risks

§ Model risks

√ Industry Practices

§ Regulatory capital adequacy

§ Bank's internal risk control

§ Corporations' investments

§ Firm's hedging of transactions

§ Exchanges' margining rules and practices

Introduction: Risk Measure

1. Risk Management is a procedure for shaping a loss distribution.

2. Despite **some serious shortcomings**, Value-at-Risk, or VaR, is the most popular portfolio risk measure used by risk management practitioners.

3. VaR is a number constructed on day t such that the portfolio losses on day $t + 1$ will only be larger than the VaR forecast with probability p, e.g. 5%.

4. The main objective of this discussion is to see how GARCH model is applied in forecasting VaR.

Value at Risk

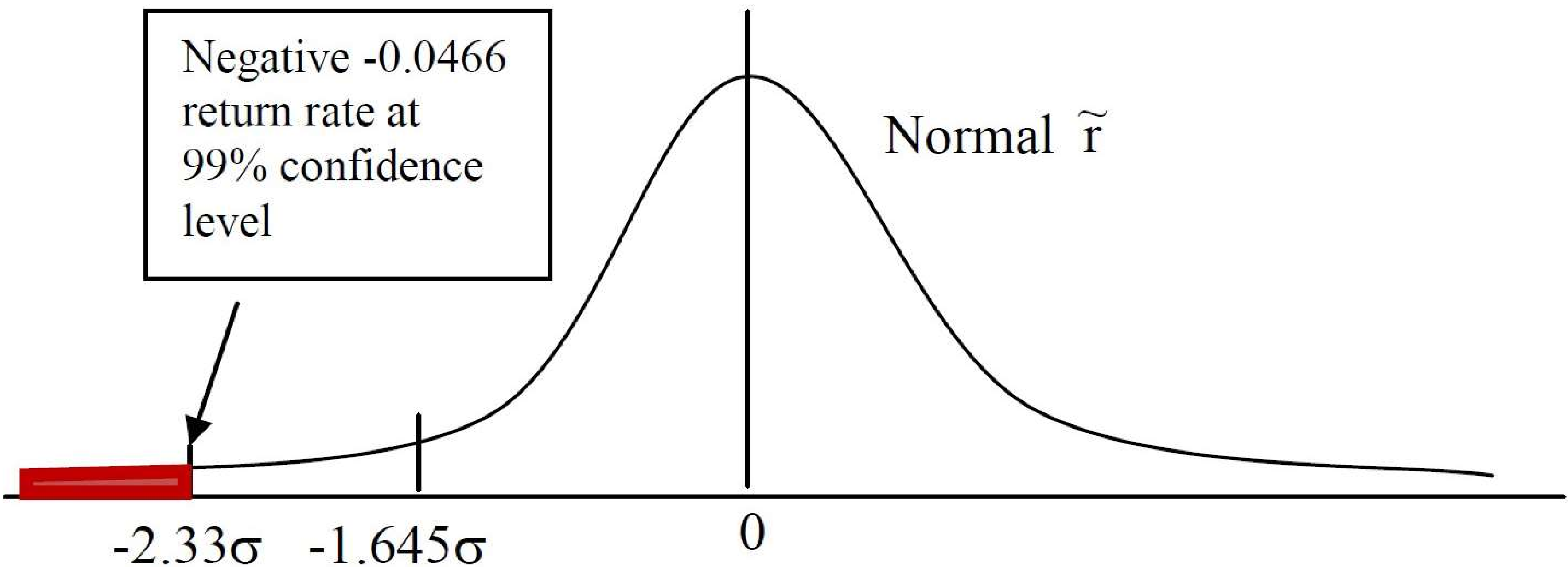
1. VaR was popularized by J.P. Morgan in the 1990s. The executives at J.P. Morgan wanted their risk managers to generate one number at the end of each day to summarize the risk of the firm's entire portfolio.
2. What they came up with was VaR.
3. If the 95% VaR of a portfolio is $400, then we expect the portfolio will lose $400 or less in 95% of the scenarios, and lose more than $400 in 5% of the scenarios.
4. We can define VaR for any confidence level, but 95% has become an extremely popular choice.
5. VaR is a one-tailed confidence interval.

Definition of Value at Risk and Example

Definition of VaR

VaR is the maximum loss over a specified horizon at a given confidence level (e.g. 95%).

Example: Suppose log return $\tilde{r} \overset{d}{\sim} N(\mu, \sigma^2)$ Suppose daily volatility σ = 2% and daily mean μ = 0.



Example

1. Suppose the portfolio value is $P_0 = \$100$ million.

2. The daily log return is

$$\tilde{r} = -\ln \frac{\tilde{P}_1}{P_0}$$

Since $\tilde{r} = -2.33 \times 0.02 = -0.0466$, we have

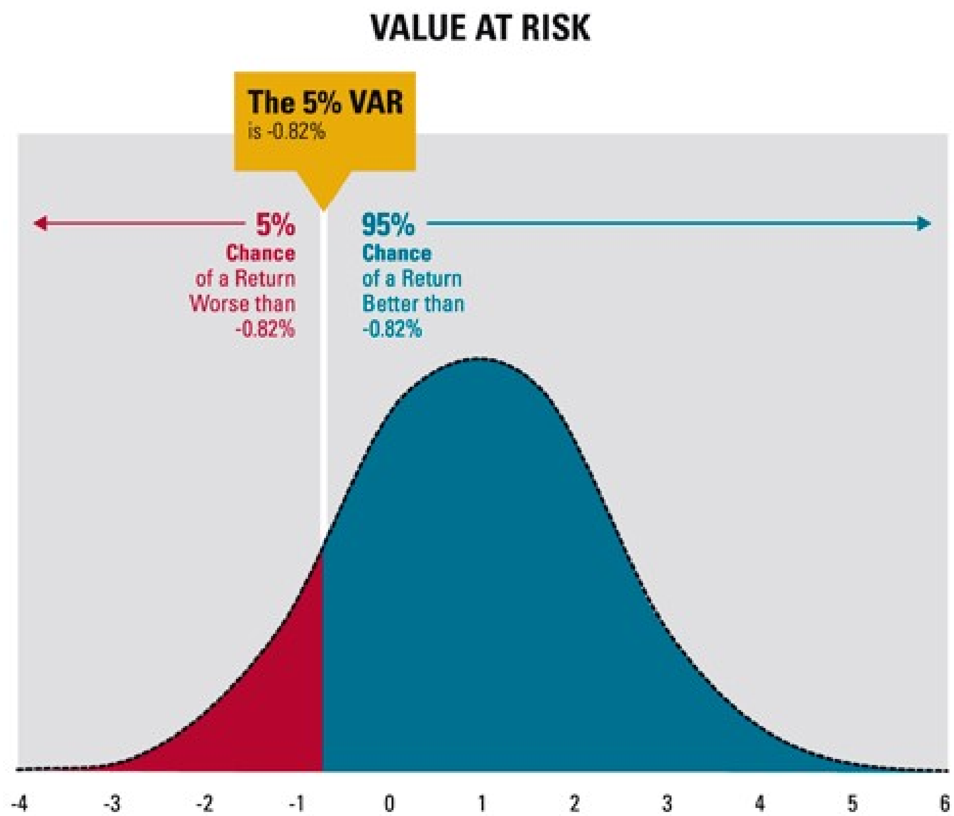
$$-0.0466 = \ln \frac{\tilde{P}_1}{\$100m} \qquad \text{or} \qquad \tilde{P}_1 = \$100m \times e^{-0.046}_{6}$$

3. The VaR is, at the 99% confidence level

$$\$100m \times (1 - e^{-0.0466}) = \$4.553m.$$

Value at Risk Time Horizon

1. Time horizon **needs** to be specified for VaR
2. On trading desks, with liquid portfolios, it is common to measure the one-day 95% VaR
3. In other settings, in which less liquid assets may be involved, time frames of up to one year are not uncommon



VaR Exceedance

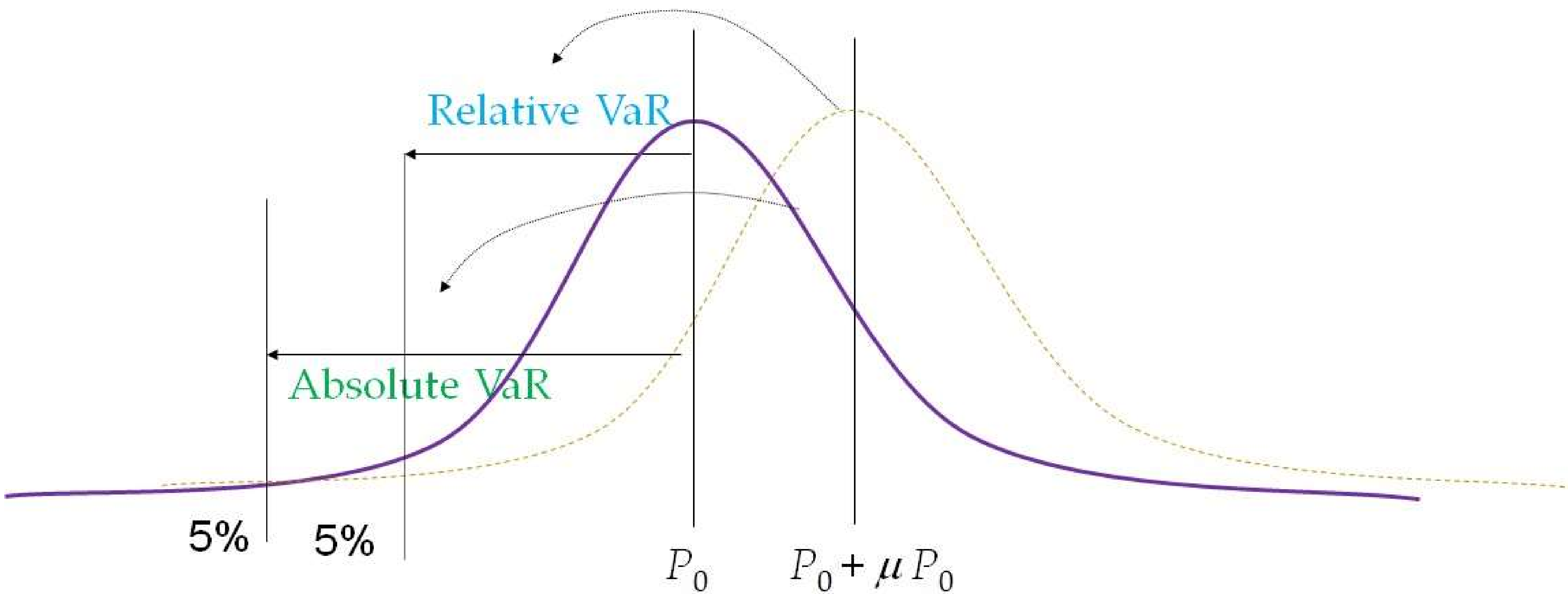
- If an actual loss equals or exceeds the predicted VaR, that event is known as an **exceedance**.
- For a one-day 95% VaR, the probability of an exceedance event on any given day is 5%.
- Let the random variable L represent the loss to your portfolio.
- For a given confidence level, α , then, we can define value at risk as

$$P\ (L \geq \text{VaR}_{\alpha}) = 1 - \alpha.$$

- If a risk manager says that the one-day 95% VaR of a portfolio is $400, it means that there is a 5% probability that the portfolio will lose $400 or more on any given day (that L will be more than $400).

Absolute versus relative VaR

[relative = add expected return to absolute VaR]



Recap on VAR and VECMs

Summary on VAR and VECMs [1/3]

1. For a Vector Auto Regression (VAR), we are using the **history of a set of variables to forecast the entire set into the future**

2. i.e. multiple dependent variables, **all of which depend on each other**

3. Each individual equation is **simply an AR model for the dependent variable** in question

4. We can also add **external independent variables [“exogeneous variables”]**

Summary on VAR and VECMs [2/3]

Relationship with ARIMA-X:

1. For ARIMA-X, we have **some trouble in forecasting the X variables**

2. We get around this with **ARIMA-X by using X_{t-1}** in the equation

3. This works for **1 period ahead forecast**

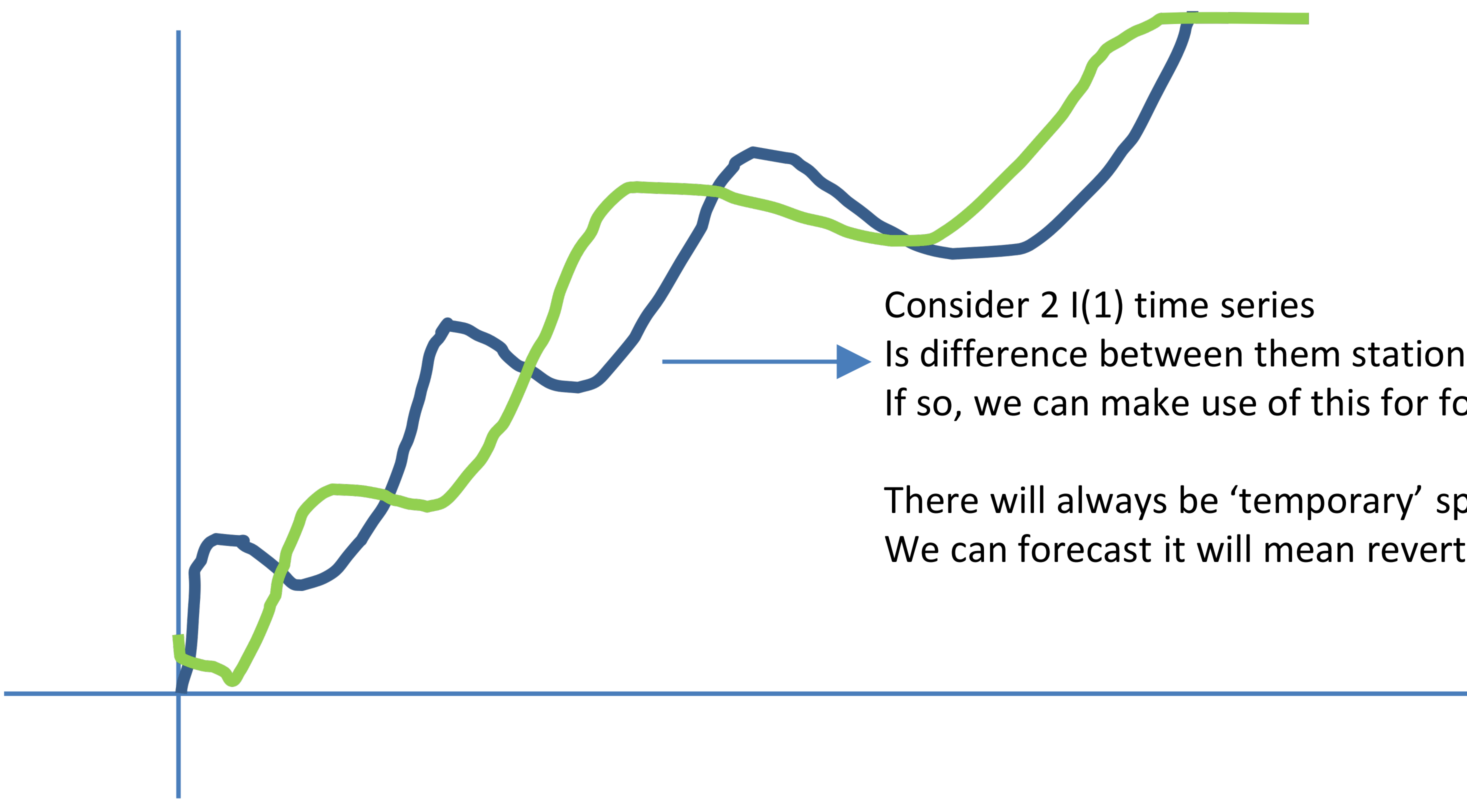
4. What if we wanted to forecast (say) 4 periods into the future? One option is to use X_{t-4} in the ARIMA-X model, but **this gets inflexible and cumbersome quickly**

5. A VAR model gets around this by enabling us to **integrate a forecast for all variables at the same time**

6. Mathematically, all the LHS variables in VAR are classified as dependent variables. Economically, there will **usually just be one out of those variables we are 'truly' interested in**, and the rest are playing a supporting role

Summary on VAR and VECMs [3/3]

Vector Error Correction Model [VECM] is a special case of VAR



Consider 2 I(1) time series
Is difference between them stationary?
If so, we can make use of this for forecasting

There will always be ‘temporary’ spikes in this difference.
We can forecast it will mean revert

Overview of VECM [just 2 equations as example]

1. **Eqn 1**: $Y_t =$ $f(Y_{t-1}, \ldots, Y_{t-k})$ +
$g(Z_{t-1}, \ldots, Z_{t-k})$ +
h(current difference between Y and Z, long term average difference) +
i(independent X variables)

2. **Eqn 2**: $Z_t =$ $a(Y_{t-1}, \ldots Y_{t-k})$ +
$b(Z_{t-1}, \ldots, Z_{t-k})$ +
c(current difference between Y and Z, long term average difference) +
d(independent X variables)

1. Highlighted in **yellow** is "**error correction term**"

2. We estimate **long term average difference using OLS**

3. Conclude long term average difference is stable **if residuals from OLS is stationary**

VAR & VECM
Exercises

Exercise 1: What order should we choose for this VAR model given that varselect() returned 1

```
>
>
>
>
> var1 <- VAR(uschange[,1:2], p=1, type="const")
> serial.test(var1, lags.pt=10, type="PT.asymptotic")

        Portmanteau Test (asymptotic)

data:  Residuals of VAR object var1
Chi-squared = 49.102, df = 36, p-value = 0.07144

> var2 <- VAR(uschange[,1:2], p=2, type="const")
> serial.test(var2, lags.pt=10, type="PT.asymptotic")

        Portmanteau Test (asymptotic)

data:  Residuals of VAR object var2
Chi-squared = 47.741, df = 32, p-value = 0.03633

>
>
> var3 <- VAR(uschange[,1:2], p=3, type="const")
> serial.test(var3, lags.pt=10, type="PT.asymptotic")

        Portmanteau Test (asymptotic)

data:  Residuals of VAR object var3
Chi-squared = 33.617, df = 28, p-value = 0.2138

>
>
>
>
> |
```

Ans: 3

Explanation:

1. We want no time series info in residuals
2. We do NOT want to reject the null hypothesis
3. Need a high p value
4. Order 1 is ambiguous; order 2 clearly has time series info in residuals, so go with order 3 (definitely do not reject the null hypothesis)

Exercise 2: Can we estimate this VAR model? [Yes/No]

Ans: No

1. Dataset length: 50 observations
2. Endogenous Variables: 3 variables
3. Proposed order of VAR model: 4

Yt = f(…)
Xt = g(...)
Zt = h(…)

Yt = c + a1 * Yt-1 + a2 * Yt-2 + a3 * Yt-3 + a4 * Yt-4 + b1 * Xt-1 + … + b4 * Xt-4 + c1 * Zt-1 + … + c4 * Zt-4
(total = 13 coefficients to estimate)

total = 39 hyperparameters
overfitting risk is extremely high

Exercise 3: Forecast Consumption 2 periods from now

```
VAR Estimation Results:
=======================

Estimated coefficients for equation Consumption:
================================================
Call:
Consumption = Consumption.l1 + Income.l1 + Consumption.l2 + Income.l2 + Consumption.l3 + Income.l3 + const

Consumption.l1      Income.l1 Consumption.l2      Income.l2 Consumption.l3      Income.l3          const
    0.19100120     0.07836635     0.15953548    -0.02706495     0.22645563    -0.01453688     0.29081124


Estimated coefficients for equation Income:
===========================================
Call:
Income = Consumption.l1 + Income.l1 + Consumption.l2 + Income.l2 + Consumption.l3 + Income.l3 + const

Consumption.l1      Income.l1 Consumption.l2      Income.l2 Consumption.l3      Income.l3          const
    0.45349152    -0.27302538     0.02166532    -0.09004735     0.35376691    -0.05375916     0.38749574
```

Exercise 4: VECM

Consider following facts:

a) Pepsi's stock price is I(1)

b) Coke's stock price is I(2)

c) An OLS regression of Pepsi's stock price against Coke's stock price:

   i. $PEP_t = c + B_1 KO_t + e_{t,}$

d) R-square from above estimation is >90%

e) T-statistic of B_1 is significant at 1% level of significance

f) e_t is not stationary according to ADF test

g) Based on the above, we form a pairs trading strategy, where we 'short the spread' between PEP and KO whenever it is very much above c, and we 'buy the spread' when it is very much below c

   a) "Short the spread" means trading on assumption that difference between PEP and KO will converge

   b) "Buy the spread" means trading on assumption that difference between PEP and KO will increase

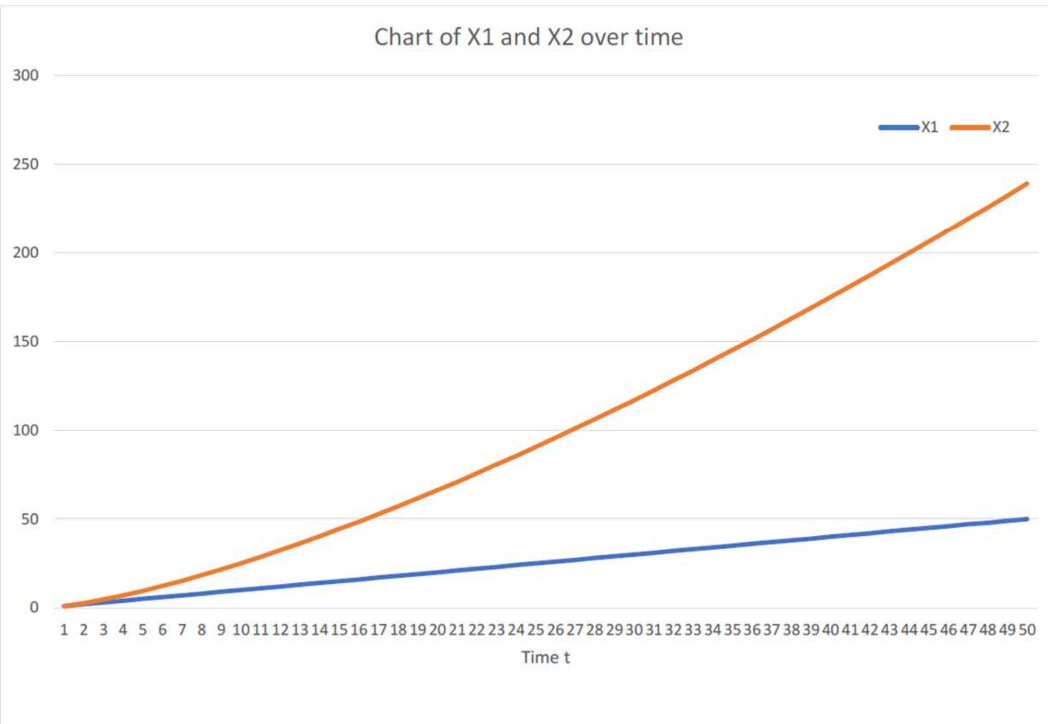
Question:

Explain **why we will probably lose money**

Error is not stationary

NEGATIVE EXAMPLE: NON STATIONARY SPREAD WITH HIGH R^2

Consider this **negative** example: $X_{2,t} = X_{1,t}^{1.4}$



What will be R^2 of OLS regression of $X_{2,t}$ against $X_{1,t}$?

NEGATIVE EXAMPLE: NON STATIONARY SPREAD WITH HIGH R^2

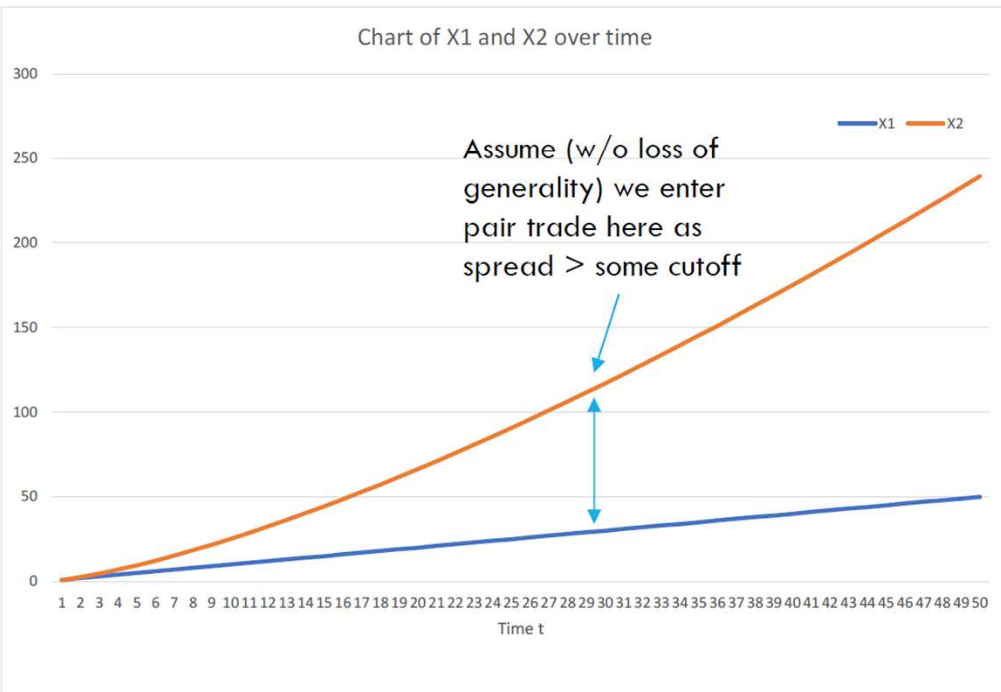
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.993452309							
R Square	0.98694749							
Adjusted R Square	0.986675563							
Standard Error	8.417708468							
Observations	50							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	257175.1646	257175.1646	3629.453738	6.8937E-47			
Residual	48	3401.175161	70.85781585					
Total	49	260576.3398						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-24.70792843	2.417054946	-10.22232799	1.22693E-13	-29.56774311	-19.84811374	-29.56774311	-19.84811374
X Variable 1	4.969778192	0.082492862	60.24494782	6.8937E-47	4.803915176	5.135641209	4.803915176	5.135641209

Very high R^2. > 0.98. Superficially this looks like good candidate for pair trading

This tells us that we should short around (1 / 4.97) shares of X2 for each share of X1 we buy

NEGATIVE EXAMPLE: NON STATIONARY SPREAD WITH HIGH R^2

Despite high R^2 and apparently reasonable hedge ratio, this pair will never converge.



Note that after trade entry (where we short X2, orange line and long X1, blue line), we will consistently lose money faster on the short leg than we make on the long leg. The spread keeps increasing without limit