Boston University

Questrom School of Business

MF 793 - Fall 2021

Eric Jacquier

The Distribution of Financial Returns:

(Log)Normality and other aspects

Outline

- Why we like normality in financial returns
- Normality through aggregation: across time: multi-period returns
  across assets: portfolios
- What is normal: returns or logarithms of returns?
  When does normality vs **log-normality** matter?
- Log-normality: multi-period implementation
- Using (log)normality to compute VaR (Value at Risk) .. Why is **VaR** important?
- So … are returns (log-)normally distributed?
- Arithmetic or geometric average to estimate mean returns?

Why we like Normality

- We can concentrate on the mean-variance decision framework
  Markowitz portfolio theory only cares about mean and variance
  Skewness and Kurtosis irrelevant to decision making

- Intertemporal asset pricing (modeling) is more conveniently done with (often continuous-time) processes based on normal shocks, e.g., Brownian motions.

- VaR (Value at risk), **shortfall** probabilities, etc.., are easy to compute
  If $R \sim N(\mu, \sigma)$ then $VAR_{5\%} = \mu - 1.644\ \sigma$
  $Prob(R<0) = Prob(Z < -\mu/\sigma)$

- Result: **If (X,Y) are jointly normally distributed, E(Y|X) is a linear function of X.**
  Take two stocks or portfolio returns: $E(R_i|R_M) = \alpha_i + \beta_i R_M$ [1]
  If [1], then $E(R_i) = \alpha_i + \beta_i E(R_M)$ [2]
  There is a well-specified linear regression: $R_{i,t} = \alpha_i + \beta_i R_{M,t} + \varepsilon_{i,t}$

  What is the difference between [1] and [2]

- All this is very nice … but is it realistic?
  - Sometimes … after some adaptation ! Such as time varying volatility, adding jumps..

Why is VaR important?

- Financial institutions must compute and report their VaR (trading desks, etc..)

- Institutions choose their own statistical model (normal, log-normal, fat-tail, etc…)

- VaR is used to compute capital reserves mandated to cover ***market risk***,
  i.e., the risk due to movements in the market prices of trading positions.

- Required ***Capital Reserve*** *is* set to be the larger of:
  - Average 10 day VAR reported in the last 60 days
    times a multiplier (3 at least)
  - Last-reported 10 day VAR

- Regulator keeps track of ***exceptions***, days with losses larger than the VAR:
  Too many exceptions -> regulator increases the multiplier

Normality: through time or cross-sectional aggregation

Central Limit Theorems: Under mild conditions, distribution of a sum of N non-normal random variables converges to normality.

Portfolios: weighted sum of stock returns

Monthly, quarterly returns: sums (? in fact compounding) of daily returns.

One random variable: hist(runif(10000,0,10),freq=F,nclass=40)

Many random variables

```
nvar <- 5
many <- matrix(runif(10000*nvar,0,10),ncol=nvar)
hist(apply(many,1,mean),freq=F,nclass=40)
```

uncorrelated

- Does it work for stocks ?
  - Time aggregation does seem to bring normality.
  - Are the daily Nasdaq, NYSE, S&P500, FTSE, returns normally distributed? **No !**

- Are we talking about returns or log of returns being normal?

Normality or Log-normality ?

What is best described as Normal, returns or log-returns?

- Return: $R_t = P_t / P_{t-1} - 1$
  larger than -1 by construction (limited liability).
  If it is bounded, it can't be normal

- Log return: $r_t = \log(1+R_t)$ is unbounded: $r_t \to -\infty$ as $P_t \to 0$.
  Better

- T-period return is **not** the sum of T one-period returns

  Let $V_T = 1 + R_{1,T}$ value at time T of \$1 invested at time 0

  $$V_T = 1(1+R_1)(1+R_2)\ldots(1+R_T)$$

- T-period log-return: **is** the simple sum of T one-period log-returns:

  $$Log V_T = \log(1+R_1)(1+R_2)\ldots(1+R_T) = \sum_t \log(1+R_t) = \sum_t r_t$$

Log-normality is preserved with time-aggregation

Log-returns have convenient aggregation formulas for mean and variance

$$LogV_T = \log(1+R_1)(1+R_2)\ldots(1+R_T) = \sum_t \log(1+R_t) = \sum_t r_t$$

- Then $E(LogV_T) = E\sum_t r_t = T\mu$
- If $r_t \sim N(\mu, \sigma)$: then $Log(V_T)$ is also normal, as sum of normals.
- If $r_t \sim$ **iid**: $V(LogV_T) = V(r_1 + \ldots + r_T) = (\sigma^2 + \ldots + \sigma^2 + 0 + \ldots 0) = T\sigma^2$

  Zero (or low) **autocorrelation** of returns is an important aspect of stockreturns

- Multi-period VaR and Shortfall probabilities follow:

  $$\text{Prob}[\, V_{12} < (1+R_f)^{12} \,] = \text{Prob}[\text{Log}(V_{12}) < \text{Log}(1+R_f)^{12} \,]$$

- In this setup, we compute log returns, estimate mean and variance **of log returns, not** of the discrete holding period returns.

Log-normality: Mean, Variance and Time-aggregation

Time-aggregation results for the parameters of multi-period returns

$$LogV_T = \log(1+R_1)(1+R_2)\ldots(1+R_T) = \sum_t \log(1+R_t) = \sum_t r_t$$

- If r_t ~ i.i.d. N(μ, σ), the **N**-period return is log-normal.

$$\boxed{\begin{aligned} E(LogV_N) &\equiv \mu_N = N\mu \\ V(LogV_N) &\equiv \sigma_N^2 = N\sigma^2 \end{aligned}}$$

$Log(V_N) \sim N\ (N\mu, N\sigma^2)$

- The *aggregation formulas* are used to annualize estimates of mean and variance obtained from higher frequency data.

- **We don't multiply the data by N!**

  We aggregate the higher frequency parameter estimates

Normal or log-normal returns: when does it matter?

- It matters for longer horizons.

- <u>Short horizon:</u> daily returns are small typically, e.g., 0.01

  for x small, $x \approx \log(1+x)$

  => At short horizons, the return R_t is small and very close to the log-return $\log(1+R_t)$.

- <u>Long horizon:</u> a year and above:

  – The two distributions start differing.

  – Approximating long term returns as normal rather than log-normal results in measurably different probabilities of shortfall, VAR, etc…

Daily (log) returns: Mean, Variance, Skewness, Kurtosis

Panel A: Daily Stock Returns

Company	$\bar{r}$	σ	Skewness	Kurtosis	$\rho_1(r)$
Merck	0.239	0.241	-0.02	3.0	0.02
Boeing	0.170	0.298	0.03	5.2	0.03
Dole Food	0.125	0.349	0.15	11.8	0.02
H. P.	0.229	0.349	-0.10	4.9	0.01
FEDEX	0.225	0.347	0.26	3.1	0.09
Ford Motor	0.242	0.304	0.16	3.2	0.01
Sony	0.176	0.329	0.72	5.5	-0.02
Fleet Bank	0.242	0.281	0.72	9.8	0.05
Exxon	0.206	0.219	-0.47	24.1	-0.06
Merrill Lynch	0.279	0.382	0.01	6.5	0.00
Equal weight	0.213	0.179	-0.77	16.4	0.04
Value weight	0.239	0.179	-1.01	21.5	-0.04
S&P500	0.140	0.154	-2.04	48.7	0.00
Small (d10)	0.169	0.147	-1.78	40.0	0.00

Mean and standard deviations are annualized.

5057 daily returns from 1979 to 1998. About 253 days per year.

Kurtosis computed in excess of 3.

Mean and standard deviation of short vs long term returns

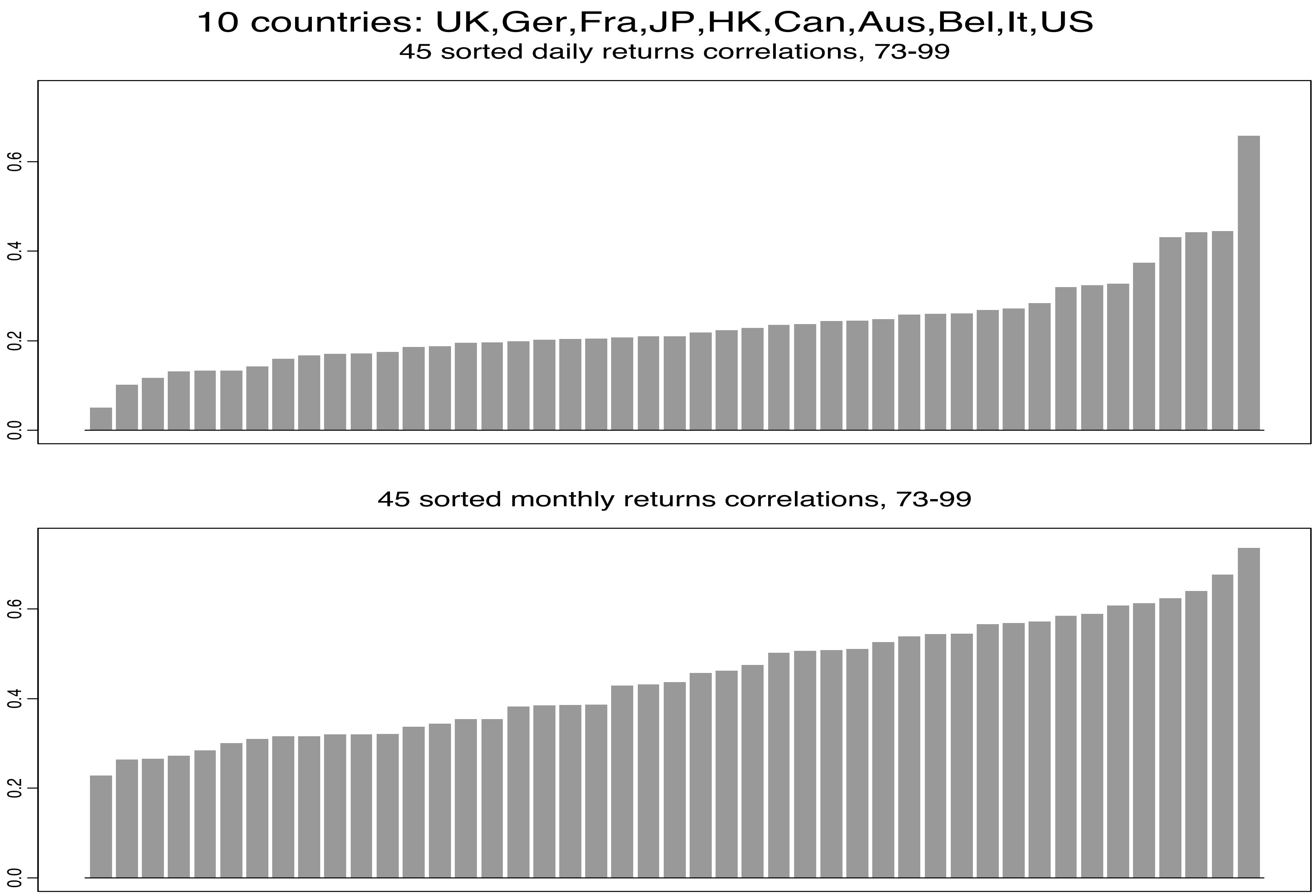
Daily log-return on VW US Market index, 1979 – 1998

	Direct Estimate		By aggregation				
	μ	σ	μ_{ann}	σ_{ann}	Sk	Ku	$\rho(1)$
Log(1+R) daily	0.00063	0.00892	0.159	0.142	-2.41	52	0.12
Log(1+R) weekly							
Log(1+R) monthly							
Log(1+R) annual							

Daily Return: $\sigma_D \approx 14\ \mu_D$ Annually aggregated values: $\sigma_A = 1.1\ \mu_A$

Higher frequency stock returns are tremendously variable relative to their mean

Country stock markets cross-correlations



US sectors cross-correlations (quarterly from daily returns)