

# Risk and Reward in Momentum Trading

Richard Martin

Founding Partner, Longwood Credit Partners LLP  
Imperial College London  
University College London

# What momentum trading is

---

- Flippantly: Buy high and sell low
- What I actually mean by this is: Buy high, sell higher; Sell low, buy (back) lower
- Or: Buy what is going up, sell what is going down
  
- “It works because ...”
  - anchoring bias
  - information leakage
  - in some asset classes (short-term govt bonds, STIRs) Government policy

# Is it really that simple?

---

- Buy what is going up; sell what is going down
- “Momentum investing shouldn’t work. Markets are supposed to be efficient ... But it does work, prompting Eugene Fama to describe it as ‘the premier anomaly’. The team at London Business School mined their database of long-term stock returns all the way back to 1900 and found that buying the 20 best performers from the 100 biggest UK shares generated an average annual return of 14.1%” (Jonathan Eley, FT, 8 Mar 2013)
- “If momentum investing is so simple, why don’t more people do it? One reason might be that its crudeness offends our intellectual sensibilities. Investing is supposed to involve a degree of judgment and intelligence, whereas momentum is a mechanical process involving little conscious thought. Anyone could do it.” (ibid)

# Specific conclusions of this talk

---

- The first moment (expected return) is a matter of empiricism/belief
- But the second and third moments of momentum trading returns are of technical interest
- Momentum trading returns are known to have positive skewness
  - thus they hold onto profits, which is good
  - also implies that the proportion of winning trades may be less than one-half
  - this is true even if expected return is zero
  - intriguing and useful that we can design skewness into a trading strategy
- How does this help us design systems?

# Design principles

---

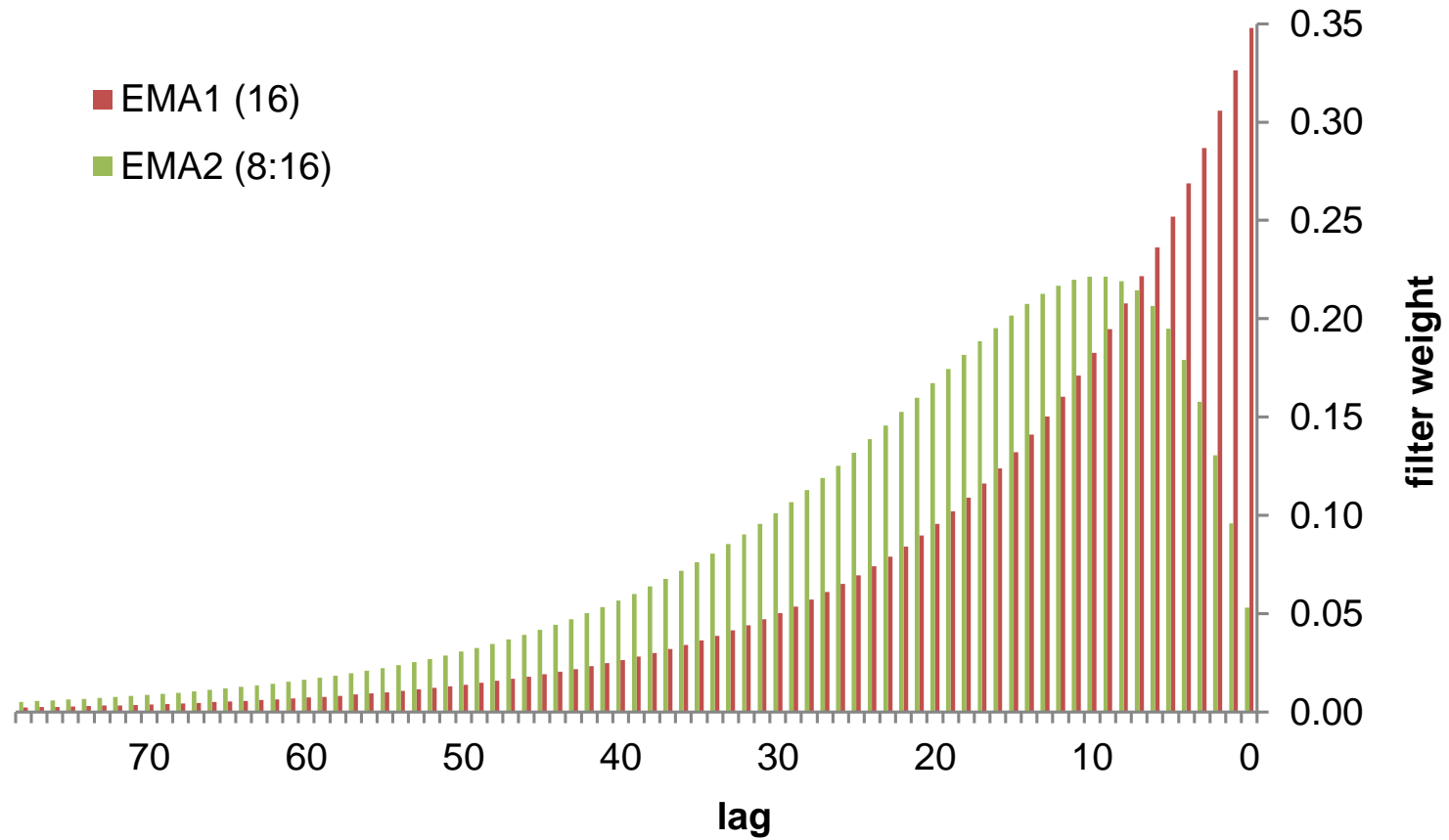
- General trading principle (not just momentum): Scale positions inversely to volatility
    - thereby keeping bets evenly-sized
    - or, strategy (trading) returns have roughly constant volatility as a result
  - so, position  $\propto$  “signal”  $\div$  vol :  $\theta_t = \mu_t G / \sigma_t^2 = \phi_t G / \sigma_t$
  - “Signal” will be related to previous price history:  $>0$  in uptrend,  $<0$  in downtrend
    - essentially, all these operate by taking weighted average of previous returns
    - we will do this using exponentially-weighted moving averages, though we don’t have to
  - Transaction costs
    - as each time we change position we incur costs
    - not for this seminar, but I have done quite a bit of work in this area
-

# Moving averages

- A common device in momentum / technical analysis is the moving average
- Use exponentially-weighted (EMA) and compare either of:
  - EMA1: spot minus moving average. When spot > EMA, uptrend
  - or EMA2: fast minus slow
  - example: USDJPY

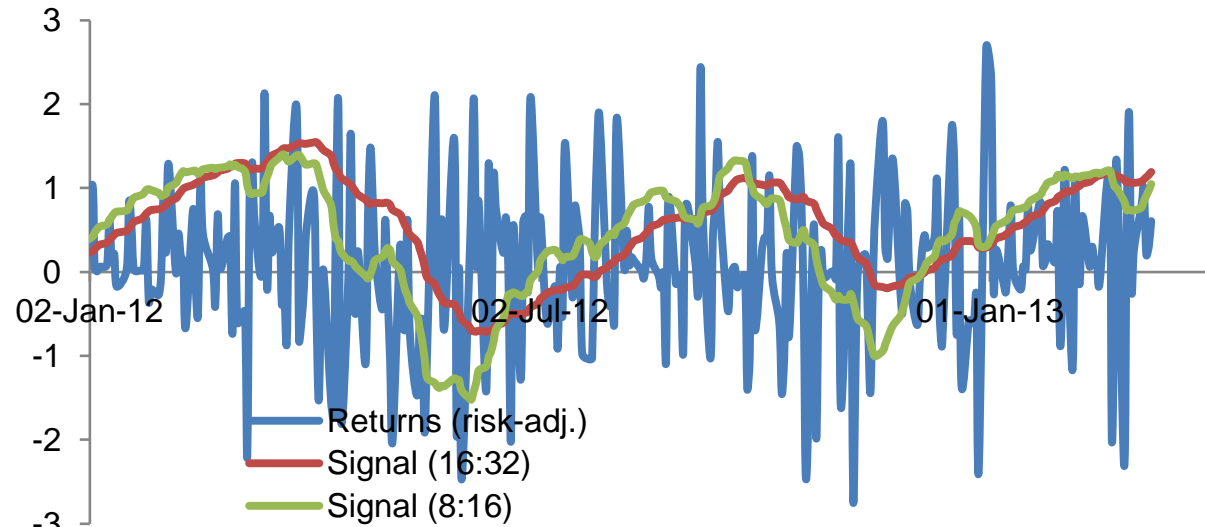


# Filter weights for EMA1 and EMA2

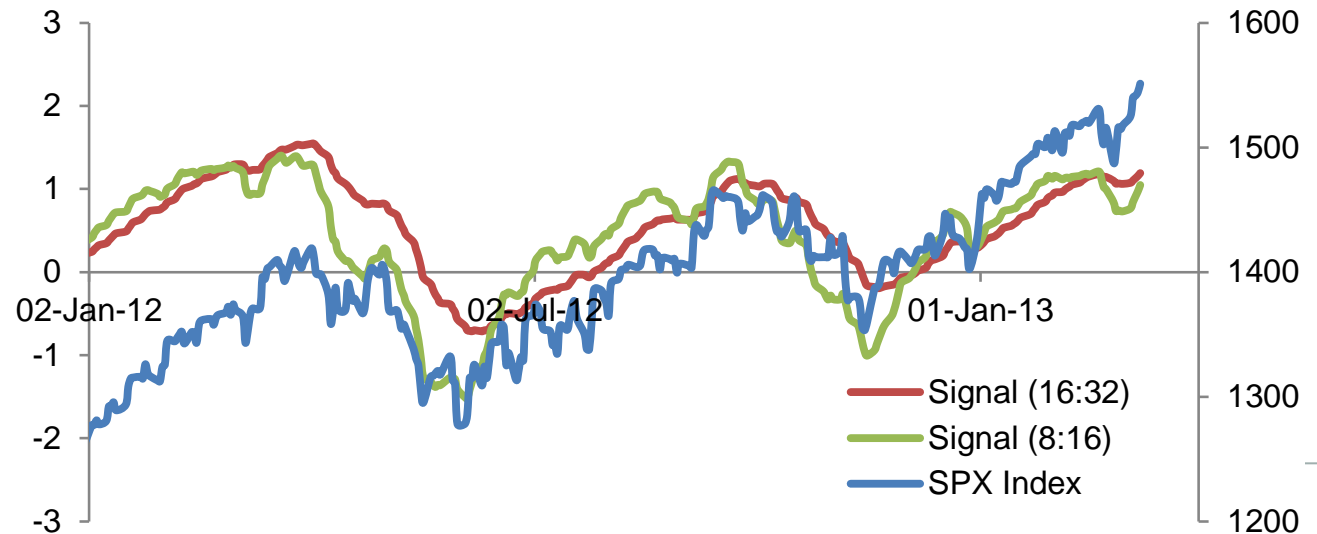


# SP1 momentum signals

Momentum signals are smoothed return series (above). They also follow the price series (below).



(NB : SPX Index is the index level and is shown as a convenient guide. The implementation uses SP1, the front futures.)





# Some technicalities

---

- Two simple things were done that I've not mentioned, both pertinent to standardisation.

1. Adjusted the returns for volatility

- obtained by averaging squared returns and imposing a floor (a pretty standard idea)

2. Divided by a normalising constant so that the momentum signal has unit std deviation

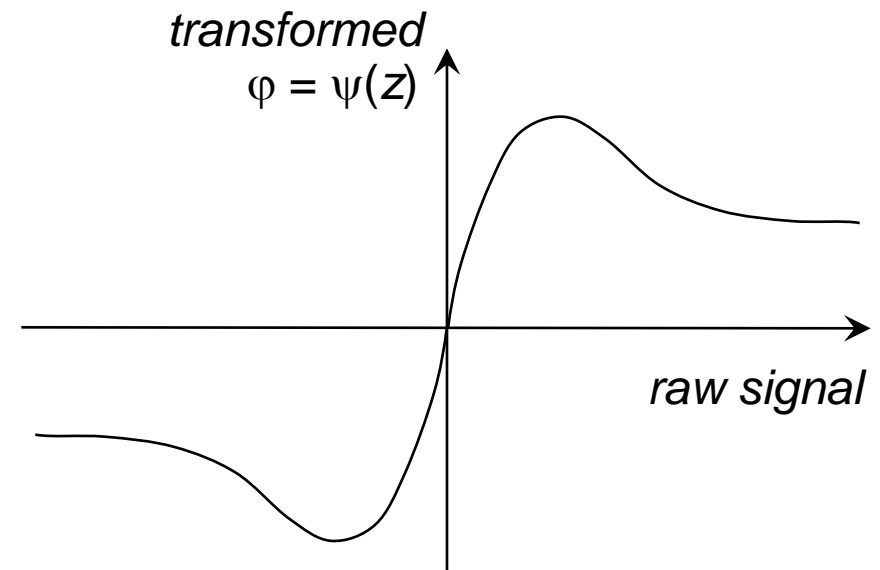
$$\text{EMA2} = \frac{\sum_{j=0}^{\infty} \beta^{j+1} (X_{n-j} - X_{n-j-1}) - \sum_{j=0}^{\infty} \alpha^{j+1} (X_{n-j} - X_{n-j-1})}{(\beta - \alpha) \left( \frac{1 + \alpha\beta}{(1 - \alpha\beta)(1 - \alpha^2)(1 - \beta^2)} \right)^{1/2}}$$

- Now, any asset class will give momentum signals that “look the same”
    - a bit like a “common currency” for momentum
-

# Turning a momentum signal into a position

---

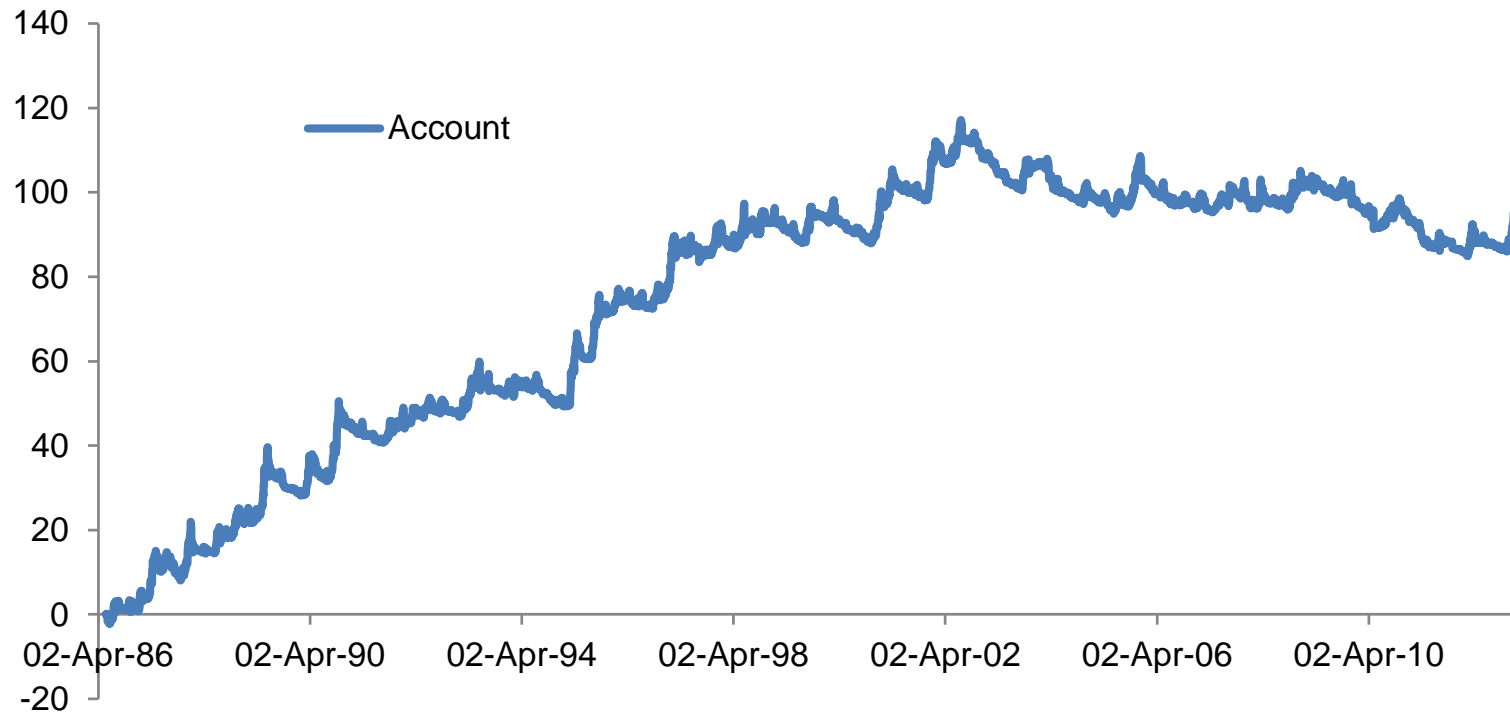
- Binary rule (already covered):  $\phi = \text{sgn}(\text{mom})$ 
  - Pro: simple, positions bounded
  - Con: sudden change in position at the origin.
  - Con: no positive skew in trading returns (▶▶)
- Linear rule:  $\phi = \text{mom}$ 
  - Pro: simple
  - Con: unbounded position
- “Optimal” shape shown :  $\phi = \psi(\text{mom})$ 
  - Bounded position
  - Smooth behaviour
  - Takes risk off when trend v strong
- Linear combination of different speeds



# And here is a typical result

---

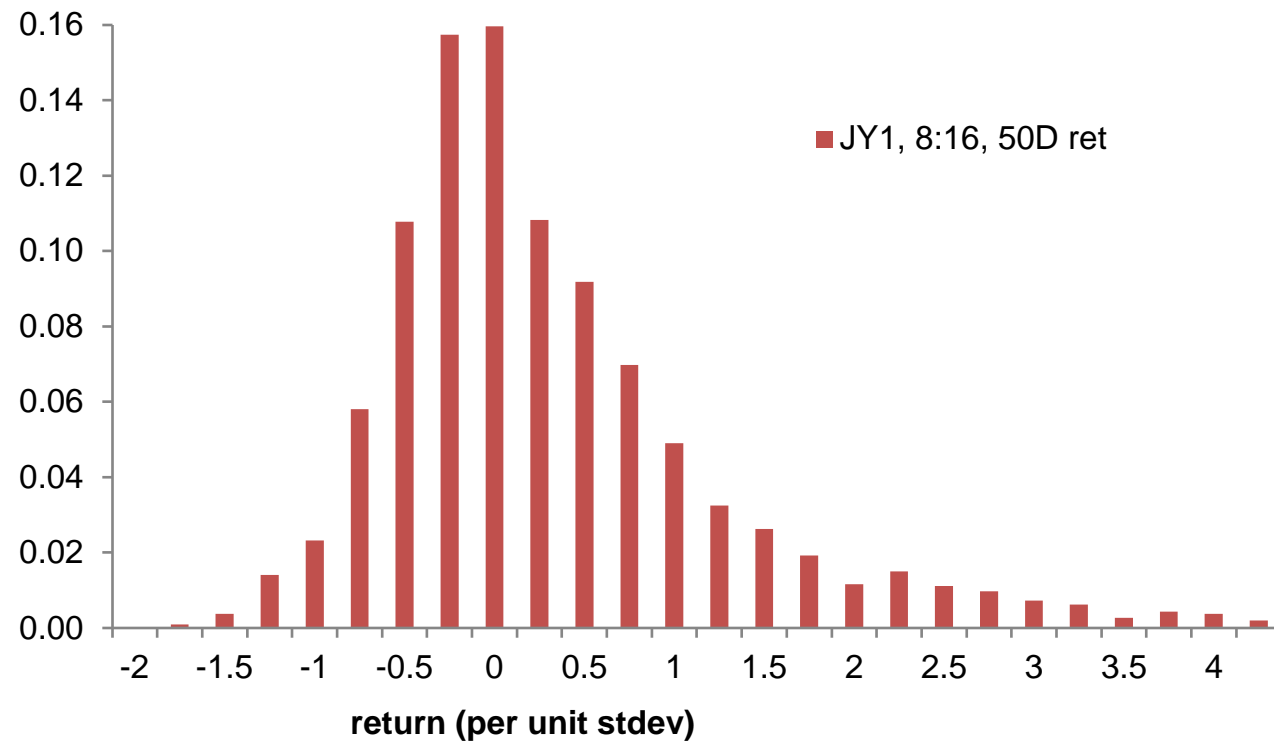
- USDJPY using 8:16
- Notice characteristic pattern of the trading returns. Up fast, down slow.



# Skewness seen empirically

---

- USDJPY using 8:16
- SR ~ +0.50 but success rate 48% (!)



# Statistical characteristics of momentum systems

---

- Skewness of trading returns is positive, under general assumptions
    - This is quite robust to market environment
    - Correct even during ‘fallow periods’ where money is not being made
    - R J Martin and A Bana, “Nonlinear momentum strategies”, RISK, Nov 2012
    - R J Martin and D Zou, “Momentum trading: ‘skews me””, RISK, Aug 2012
    - M Potters and J-P Bouchaud, “Trend follows lose more often than they gain”, Wilmott, Nov 2005
    - E Acar and S Satchell, “Advanced trading rules”, Butterworth-Heinemann, 2002
  - Derivation of second and third moments of trading returns
  - Calculation in special cases
  - Linear and nonlinear systems, i.e. incorporate the response function  $\psi$  we saw earlier
  - Hybrid systems which combine positive and negative momentum bets
-

# Trading returns: moments

---

- Period- $M$  trading return, with  $X$  the adjusted price of the instrument:

$$Y_n^{(M)} = \sum_{k=0}^{M-1} \varphi_{n+k} U_{n+k+1} \quad U_{n+1} = \frac{X_{n+1} - X_n}{\hat{\sigma}_n}$$

- First moment is zero (assuming independent returns)
- Second and third moments

$$\langle (\varphi_0 U_1 + \dots + \varphi_{M-1} U_M)^2 \rangle \quad \langle (\varphi_0 U_1 + \dots + \varphi_{M-1} U_M)^3 \rangle$$

- which evaluate to

$$M \langle \varphi^2 \rangle \quad 3 \sum_{0 \leq m < n < M} \langle \varphi_n^2 \varphi_m U_{m+1} \rangle$$

- returns are uncorrelated but not independent
-

# Linear systems

---

- Signal is a weighted (linear) combination of previous returns

$$\varphi_n = \sum_{j=0}^{\infty} a_j U_{n-j}$$

- Expression for third moment in terms of autocorrelation function

$$3 \sum_{0 \leq m < n < M} \langle \varphi_n^2 \varphi_m U_{m+1} \rangle = 6 \sum_{k=1}^{M-1} (M-k) a_{k-1} R_k^a.$$

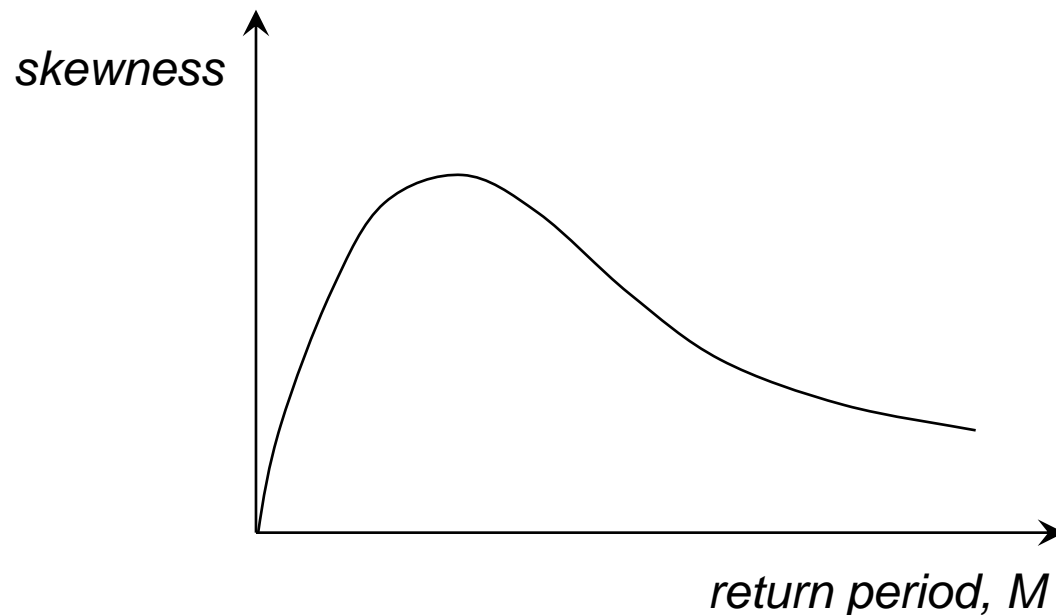
- where autocorrelation function  $R$  is, in terms of the filter weights:

$$R_k^a = \sum_{j=0}^{\infty} a_j a_{j+k}, \quad k \geq 0$$

# Trading returns: basic characteristics of skewness

---

- Second moment is directly proportional to  $M$
- Third moment is proportional to  $M^2$  for small  $M$ , but to  $M$  for large  $M$
- So skewness is proportional to  $M^{+1/2}$  for small  $M$ , but to  $M^{-1/2}$  for large  $M$
- Sketch for a pure-momentum system (all filter weights positive):



$$\kappa_3^{(M)} = \mu_3^{(M)} / (\mu_2^{(M)})^{3/2}$$



# Linear systems

---

- EMA1, weighting is an exponential

- where  $\alpha = 1 - 1/N$  and  $N =$  'period' or 'lookback', e.g.  $N=20$  gives  $\alpha=0.95$

$$a_j = \alpha^{j+1}$$

- EMA2, weighting is a difference of exponentials

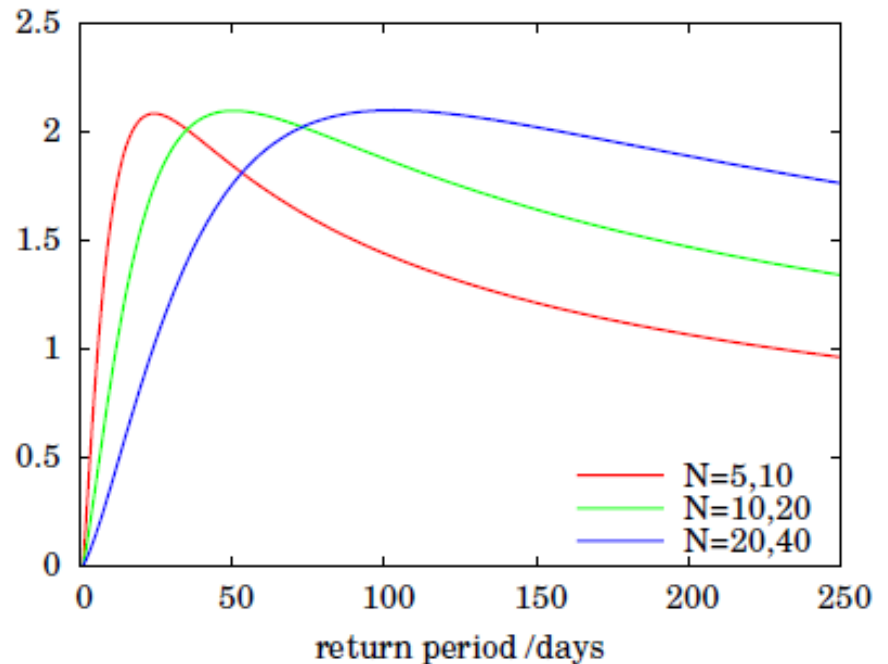
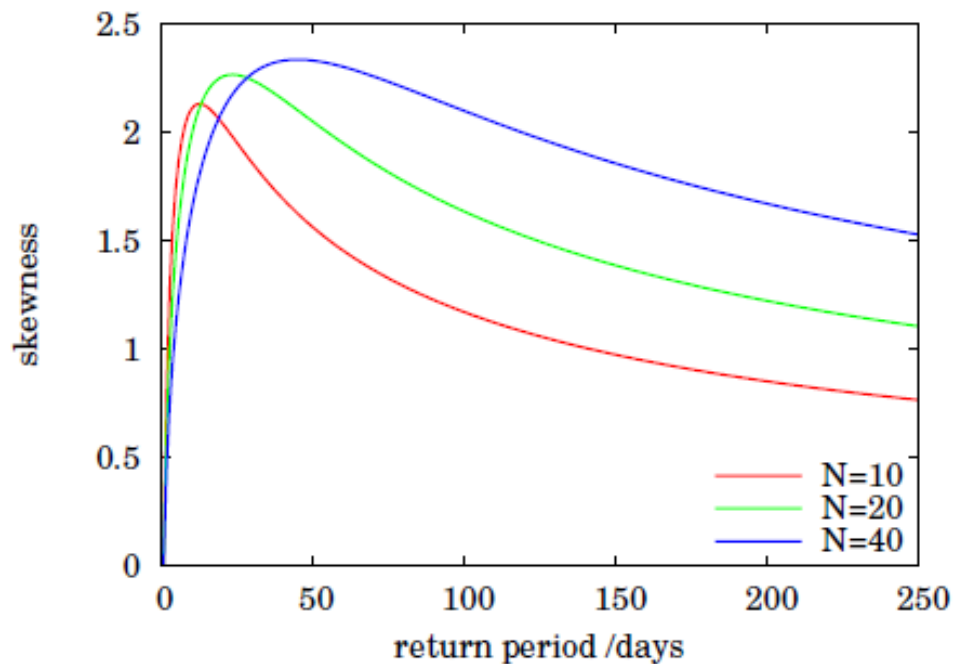
$$a_j = \frac{\alpha^{j+1} - \beta^{j+1}}{\alpha - \beta}$$

- where  $\alpha = 1 - 1/N_\alpha$  and  $\beta = 1 - 1/N_\beta$

- From these I can calculate  $R$ , & then second and third moments, & then the skewness
-

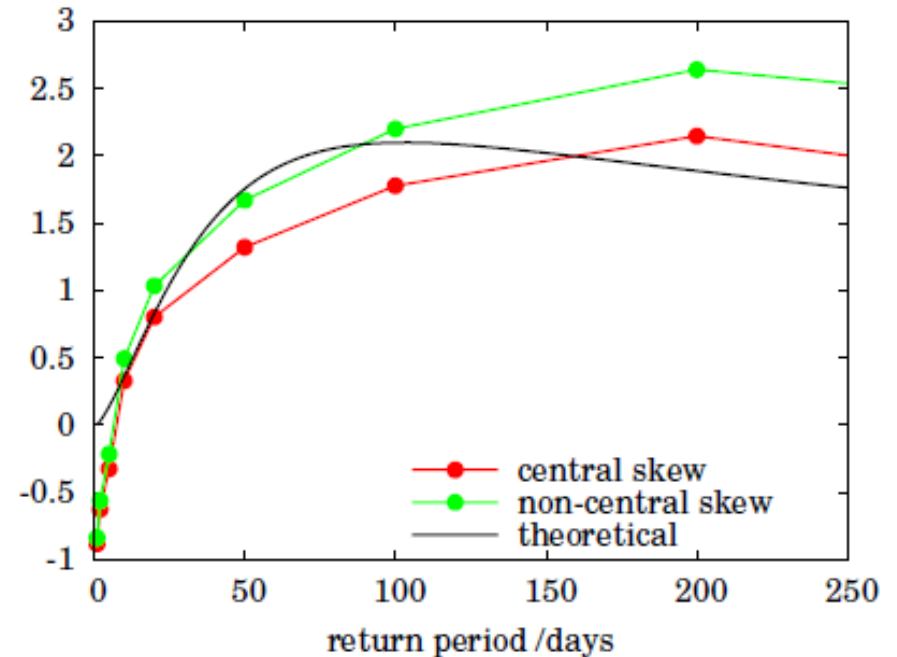
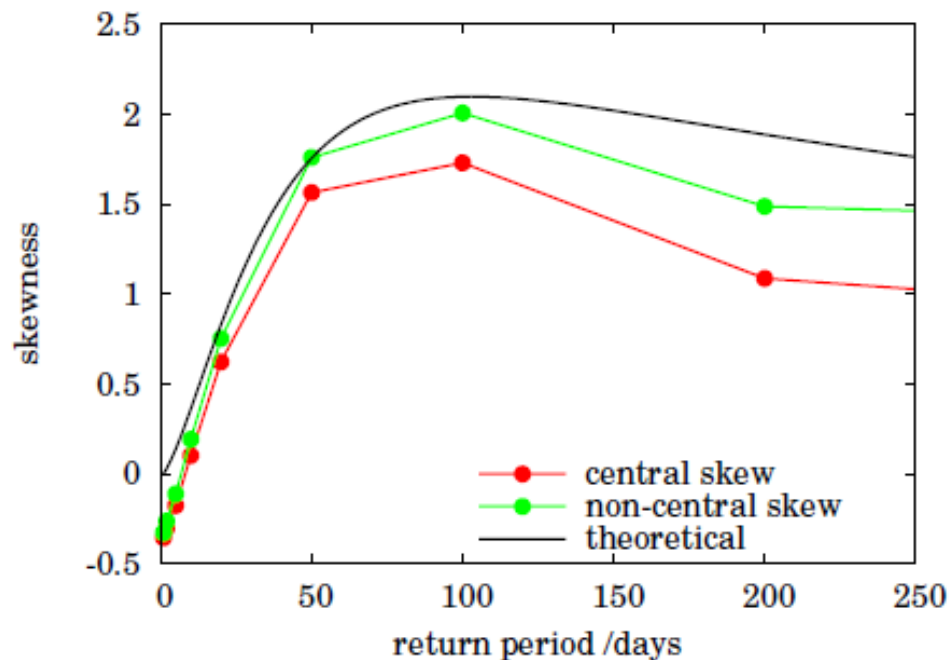
# Skewness term structure

- EMA1 and EMA2
- Slower oscillators basically stretch in horizontal direction



# Skewness term structure, empirically

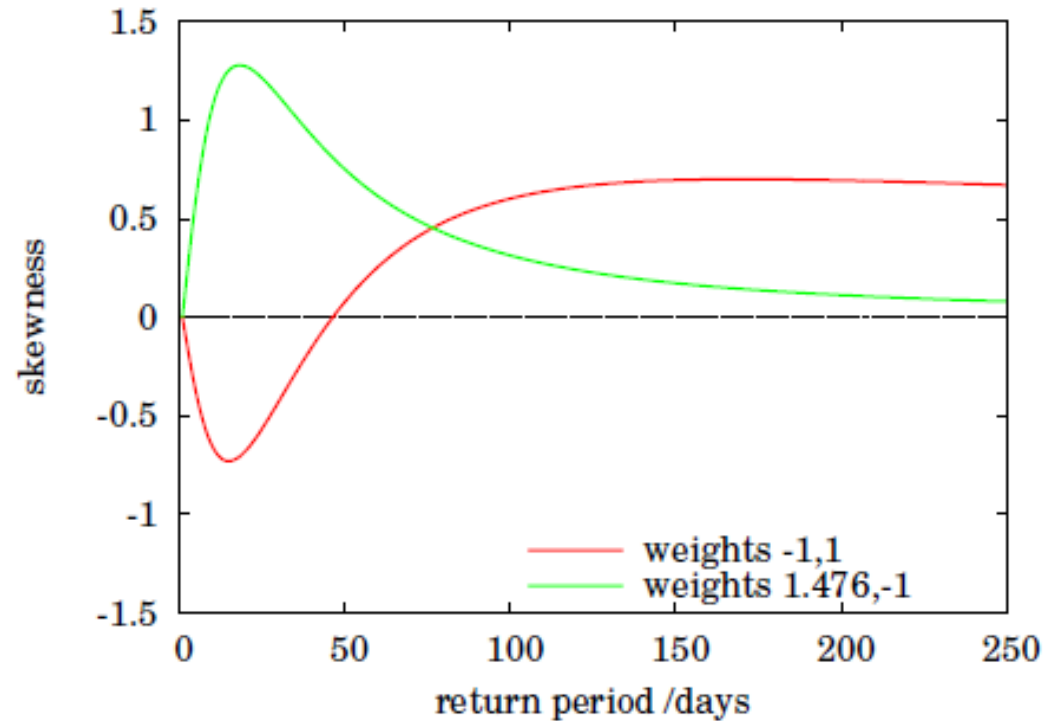
- We have done analysis on basis that mean trading return = 0
- However, trend following does work! So mean won't be zero
- So plot central skew (moments about mean) & non-central skew (moments about 0)
- CHFUSD and SPX. NB: empirical estimation of skew is very sensitive to outliers



# Hybrid systems

---

- 5:10 (short-term) and 20:40 (longer-term) oscillators
- Red: negative bet on short-term momentum, positive bet on longer-term
- Green: positive bet on short-term momentum, “small” negative bet on longer-term



# Nonlinear systems

---

- As before, momentum factor is going to be weighted combination of risk-adjusted returns and position  $\theta = \varphi/\sigma$  (= signal  $\div$  vol):

$$V_n = \sum_{j=0}^{\infty} a_j U_{n-j} \quad \varphi_n = \psi(V_n)$$

- Third moment

$$\langle (Y_n^{(M)})^3 \rangle = 3 \sum_{k=1}^{M-1} (M-k) H_k, \quad H_{n-m} = \langle \varphi_n^2 \varphi_m U_{m+1} \rangle.$$

- and this can be evaluated via

$$\begin{aligned} H_k &= 2a_{k-1} \left\langle \psi(Z_1)^2 \psi(Z_2) \frac{Z_1 - \rho Z_2}{2(1 - \rho^2)} \right\rangle \Big|_{\rho=R_k^a} \\ &= 2a_{k-1} \langle \psi(Z_1) \psi'(Z_1) \psi(Z_2) \rangle \Big|_{\rho=R_k^a}, \end{aligned}$$

- where  $(Z_1, Z_2) \sim$  standard Bivariate Normal with correl  $\rho$
-

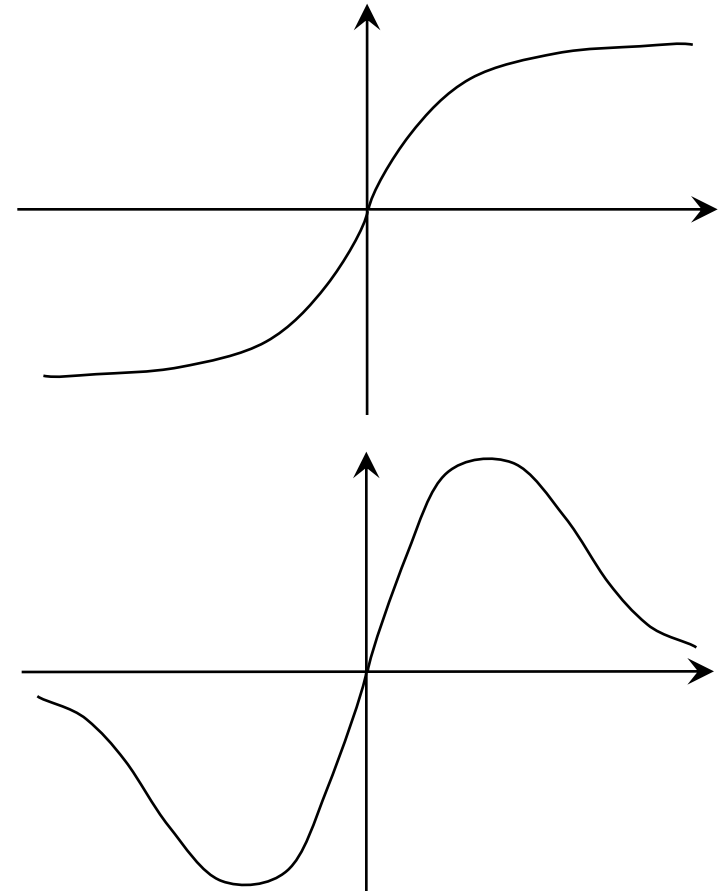
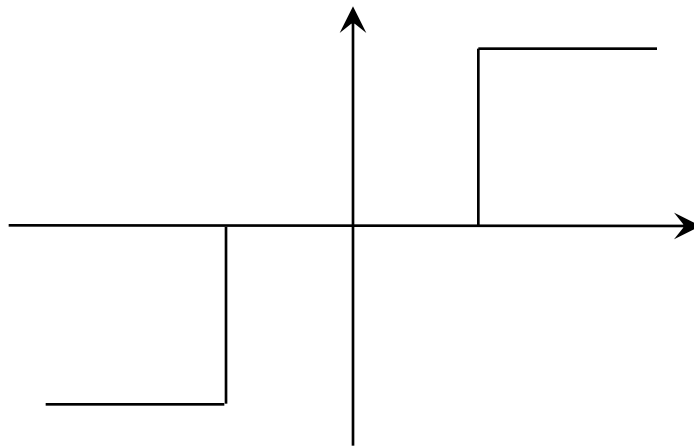
# Choice of $\psi$

- For some choices of  $\psi$ , the expectation (double-integral) can be evaluated in closed form
- Two such:

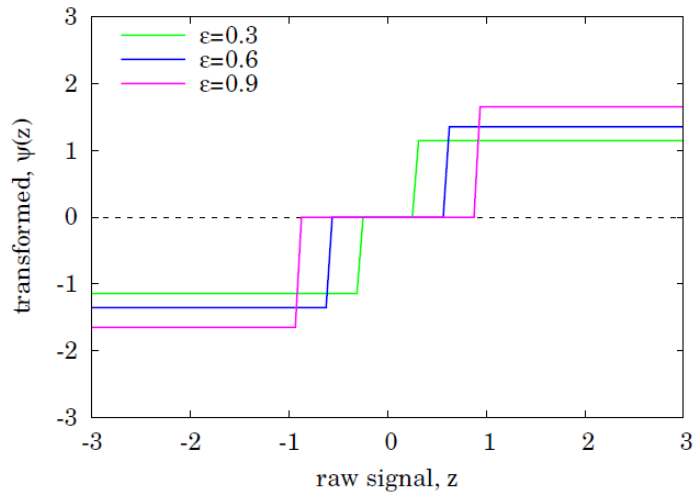
**Simple sigmoid**,  $\psi(z) = c_\lambda \cdot (2\Phi(\lambda z) - 1)$

**Reverting sigmoid**,  $\psi(z) = c_\lambda \cdot ze^{-\lambda^2 z^2/2}$

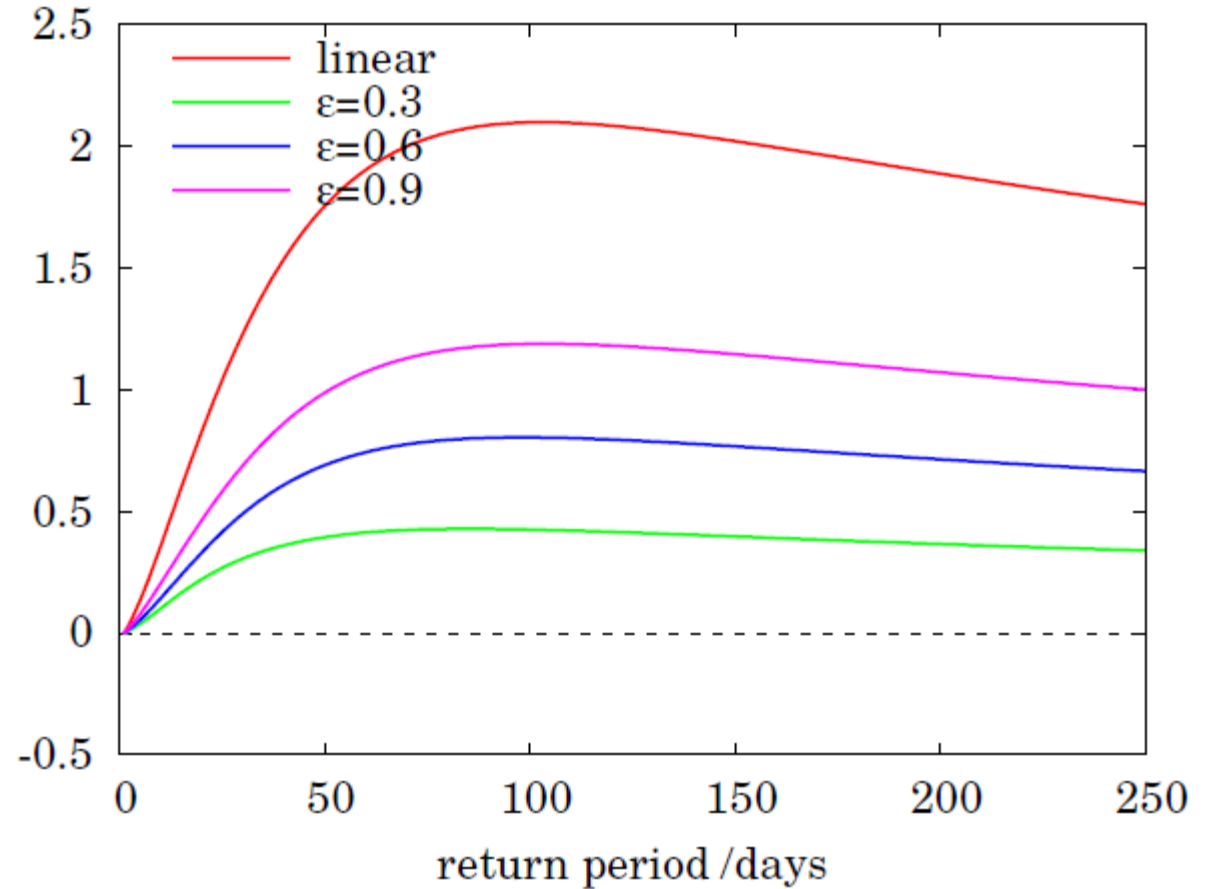
**Double-step**,  $\psi(z) = c_\lambda \cdot (\mathbf{1}_{z>\varepsilon} - \mathbf{1}_{z<-\varepsilon})$



# Results for double-step

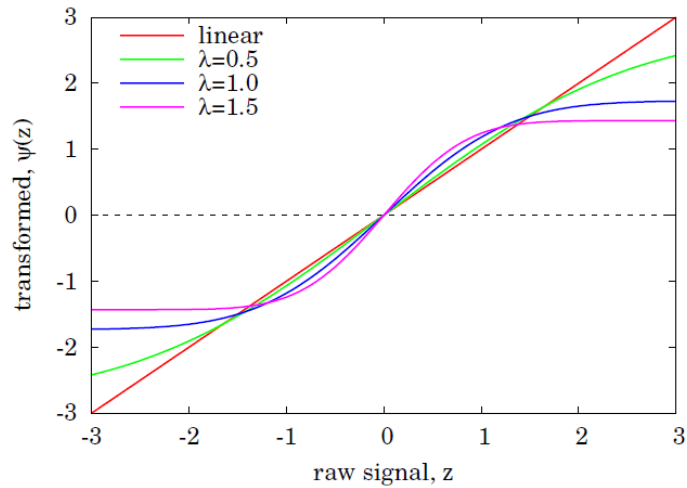


Narrower separation ►  
Lower skewness.  
Zero separation ►  
No skewness (Why?)

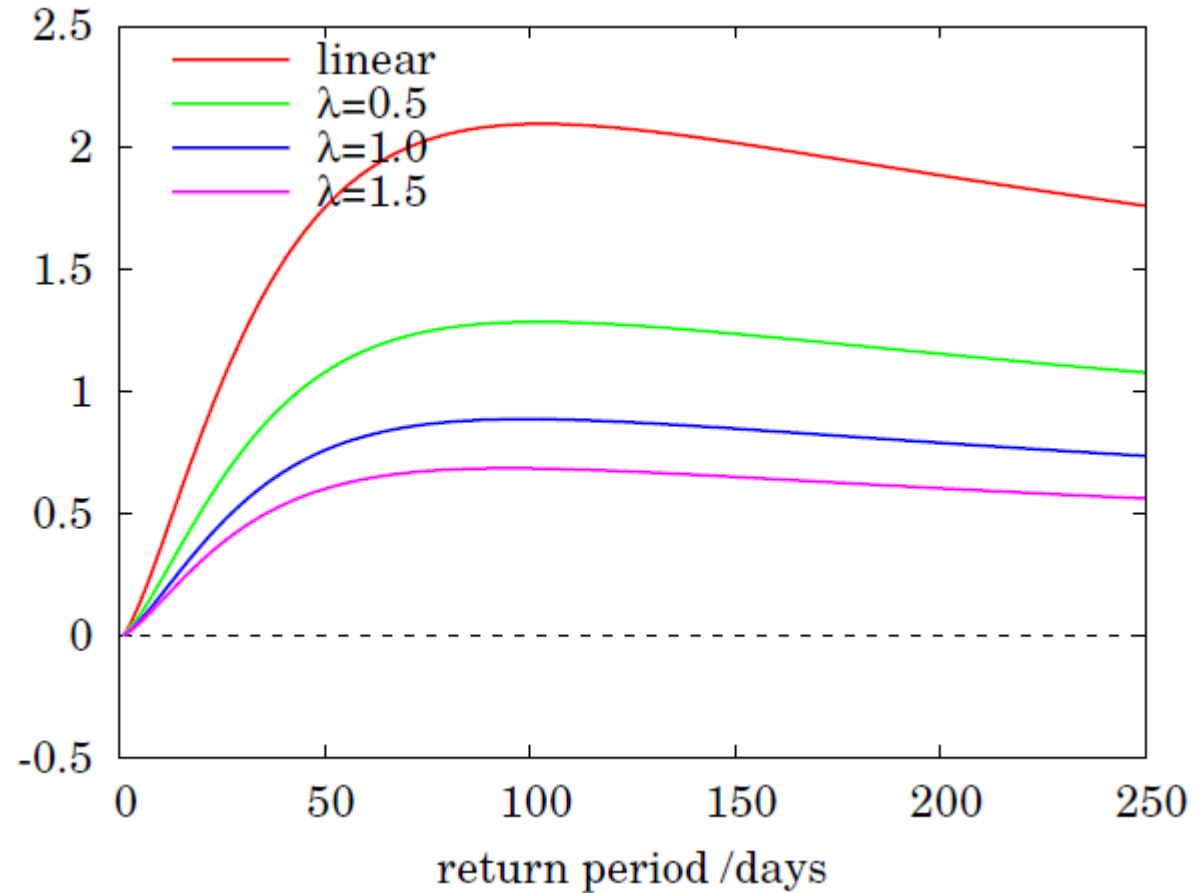


$N=20,40$  throughout

# Results for sigmoid

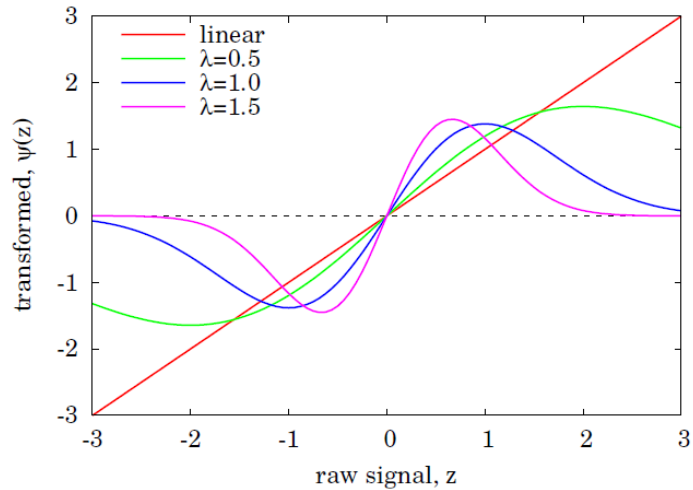


Position-limiting  
reduces skewness

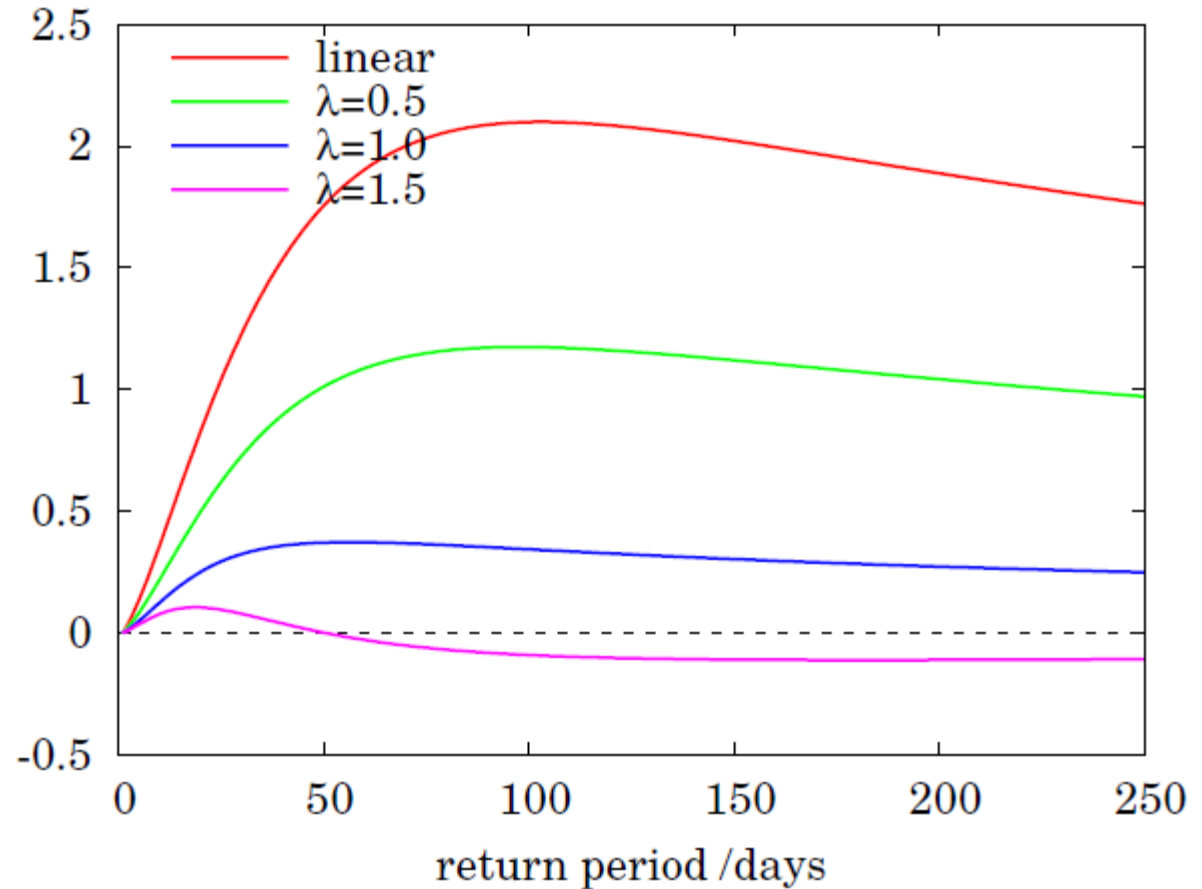




# Results for reverting sigmoid



Fast reversion causes negative skewness.  
Critical  $\lambda \sim 1.3$



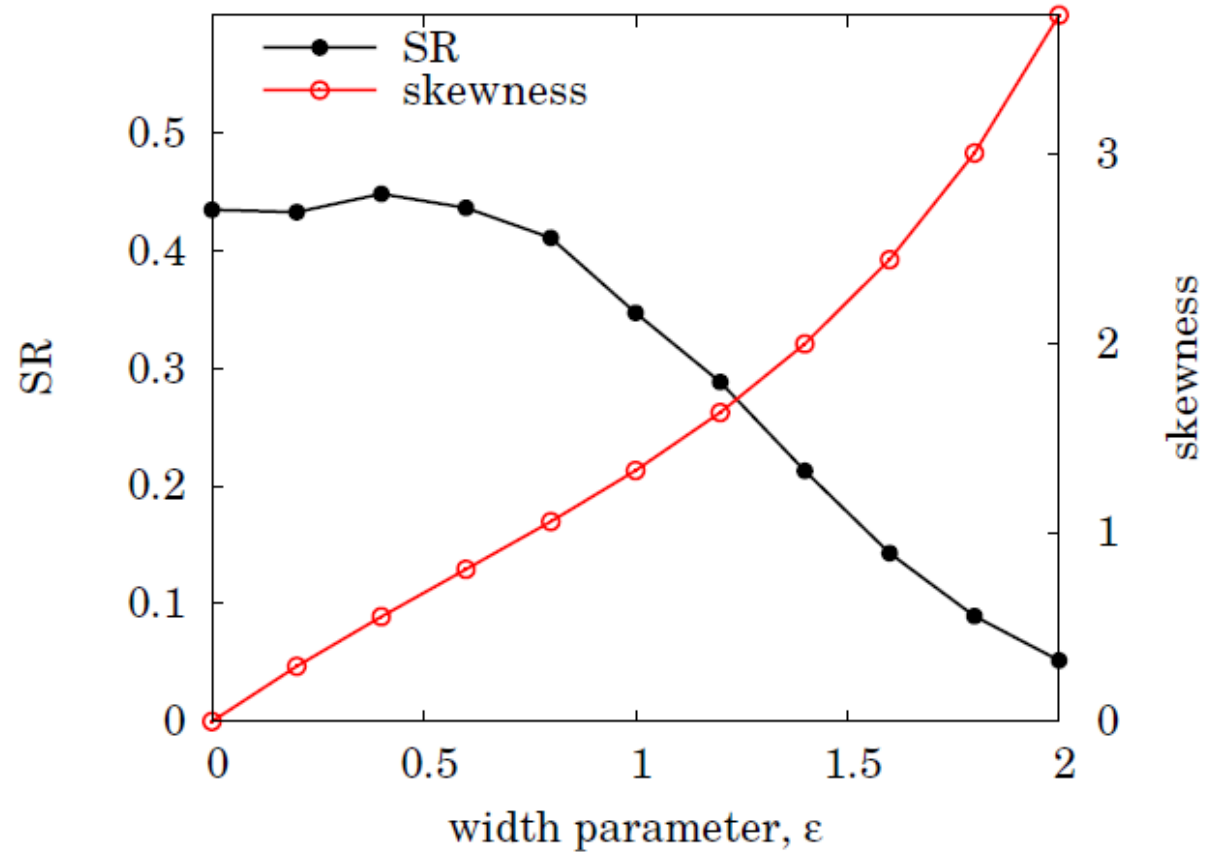
# Optimisation of $\psi$

---

- Can now consider two criteria
    1. The classical Sharpe ratio
      - take a selection of futures contracts
      - backtest each one individually and take average SR as the objective function
      - try different shapes
    2. Skewness
      - no backtesting of this required! Just compute the skewness analytically
      - and go for highest skewness you can
  - Of course, these two may well conflict
-

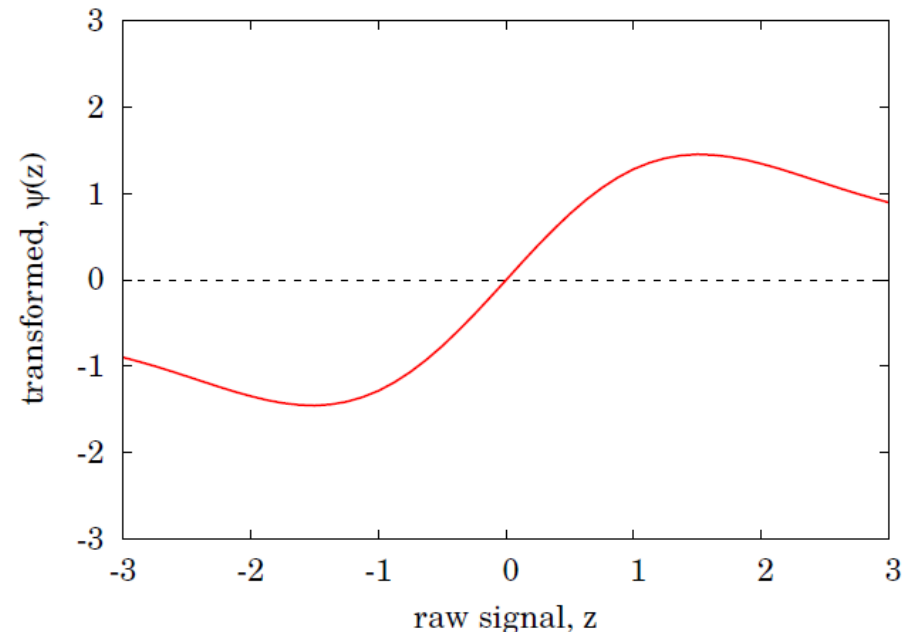
# Optimising the double-step

- Optimal  $\varepsilon \sim 0.6$ : SR not affected, but skewness higher



# Optimisation of a smooth function

- We can take a linear combination of the two sigmoids  $2(\Phi(\lambda z) - 1)$ ,  $ze^{-\lambda^2 z^2 / 2}$ 
  - if we choose the same  $\lambda$  for each, then we have two parameters,  $\lambda$  and the weight-ratio
- Use on a universe of different futures contracts and trading speeds
- Two criteria, as we wish to maximise Sharpe ratio and Skewness
  - SR: the performance surface rather flat, i.e. hard to be confident about “best” soln.
  - SR prefers reverting sigmoid  $ze^{-\lambda^2 z^2 / 2}$
  - but skewness doesn't
  - so we get a compromise like this:



# Optimisation of a smooth function

← Linear      Sharpe      Reverting      ↑

Skew

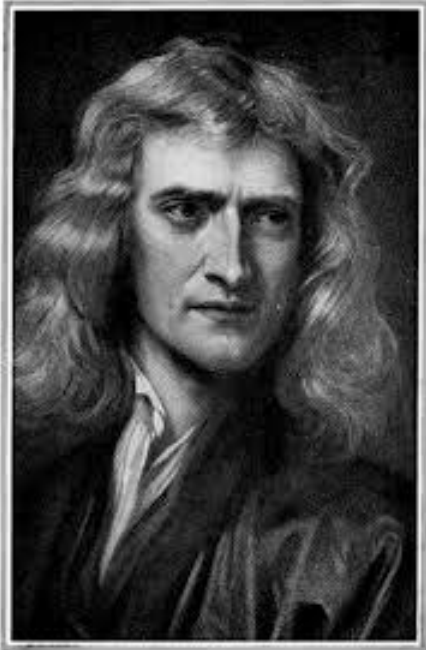
$\frac{w_R}{w_S} \setminus \lambda$	0.25	0.50	0.75	1.00	1.25	1.50	1.75
$\infty$	0.39; 1.77	0.45; 1.17	0.48; 0.69	0.48; 0.34	0.45; 0.09	0.40; -0.09	0.35; -0.23
5.0	0.38; 1.74	0.44; 1.19	0.48; 0.75	0.48; 0.43	0.46; 0.21	0.43; 0.04	0.39; -0.09
2.4	0.38; 1.73	0.43; 1.21	0.47; 0.80	0.48; 0.50	0.47; 0.29	0.45; 0.14	0.42; 0.02
1.5	0.38; 1.71	0.43; 1.22	0.46; 0.83	0.48; 0.56	0.47; 0.37	0.46; 0.22	0.43; 0.11
1	0.38; 1.70	0.42; 1.23	0.46; 0.87	0.47; 0.62	0.47; 0.43	0.46; 0.30	0.45; 0.19
0.67	0.38; 1.69	0.41; 1.24	0.45; 0.91	0.47; 0.67	0.47; 0.50	0.46; 0.37	0.45; 0.28
0.4	0.37; 1.67	0.41; 1.25	0.44; 0.94	0.46; 0.73	0.46; 0.57	0.46; 0.46	0.46; 0.37
0.2	0.37; 1.65	0.40; 1.27	0.43; 0.99	0.45; 0.80	0.46; 0.66	0.46; 0.56	0.46; 0.48
0	0.37; 1.63	0.39; 1.29	0.41; 1.05	0.43; 0.89	0.44; 0.77	0.45; 0.69	0.45; 0.62

← Linear      Saturating      ↓

# In the words of Alexander Pope

---

- The famous triplet (not couplet, as is often supposed):



Nature and nature's laws lay hid in night;  
God said, "Let Newton be!" and all was light;  
Ride good, cut bad, for this we find is right.

