CHARTA LEARNING · LESSON NOTES

Alpha and the Low-Risk Anomaly

FRM Part II

1 A story: calm stocks that kept up

- Kestrel buys the calmest, lowest-beta stocks
- Theory: less risk should mean less return
- Yet the fund kept pace with the market
- With much smaller swings along the way

IN THE LESSON

- Kestrel's founder had a simple idea.
- The Defensive Equity Fund buys the calmest stocks, those with the lowest beta.
- Beta measures how much a stock moves with the whole market: a beta of a half moves half as much.
- Theory says this should cost money.
- Less risk should mean less return, because investors are paid for bearing risk.
- Yet over many years the fund kept pace with the market, with much smaller swings along the way.
- It is a well known puzzle in finance, the low risk anomaly, and it is where our story about alpha begins.

2 The CAPM and the security market line

$E(R_i) = R_f + \beta_i \times (E(R_M) - R_f)$

Anomaly: the line found in the data is much flatter

Risk-free rate Rf	Paid for waiting, not for risk
Beta β	Sensitivity to the market
Market risk premium	Market return minus Rf

IN THE LESSON

- The theory comes from the capital asset pricing model, the C A P M.
- It says a stock's expected return equals the risk free rate plus its beta times the market risk premium.
- Plotted against beta, that is a straight line, called the security market line, rising at the market risk premium.
- The low risk anomaly is the finding that the line found in the data is much flatter.
- Low beta stocks have earned as much as, or more than, high beta stocks, so their risk adjusted performance has been better.
- Now the parts.
- The risk free rate is paid for waiting, not for risk.
- Beta is the sensitivity to the market.
- And the market risk premium is the market return minus the risk free rate.

3 Example 1: alpha along a flat line

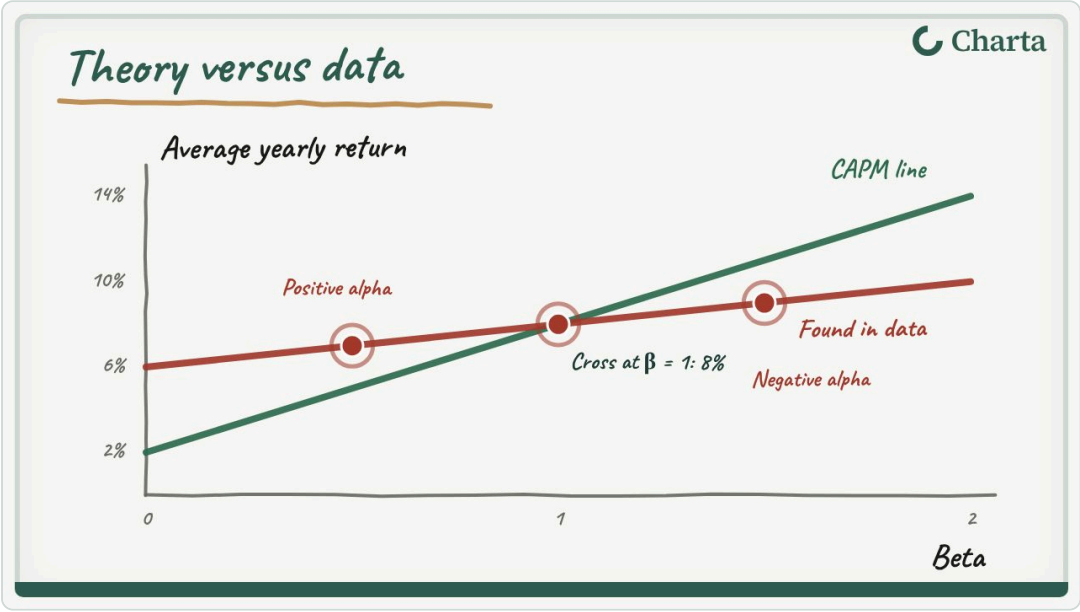
Rf 2%, market premium 6%. Realised: β 0.5 → 7%, β 1.0 → 8%, β 1.5 → 9%

CAPM return, β = 0.5	2 + 0.5 × 6	**5%**
CAPM return, β = 1.5	2 + 1.5 × 6	**11%**
Alpha, low beta	7 − 5	**+2%**
Alpha, high beta	9 − 11	**−2%**
Slope found in the data	(9 − 7) ÷ (1.5 − 0.5)	**2% per unit β**

IN THE LESSON

- The risk free rate is two per cent and the market risk premium six per cent.
- Stocks sorted into three made up groups by beta earned seven per cent at a beta of a half, eight per cent at a beta of one, and nine per cent at a beta of one and a half.
- The C A P M says a beta of a half should earn two plus a half times six, five per cent.
- A beta of one and a half should earn two plus one and a half times six, eleven per cent.
- So the low beta group's alpha is seven minus five, plus two per cent.
- The high beta group has nine minus eleven, minus two per cent.
- The slope found is nine minus seven, over one and a half minus a half: just two per cent per unit of beta, not six.

4 Theory versus data



IN THE LESSON

- Along the bottom is beta, from zero to two.
- Up the side is the average yearly return.
- The green C A P M line starts at the risk free rate of two per cent and climbs to fourteen per cent at a beta of two.
- The red line is the flatter line found in the data, starting at six per cent and rising only two per cent for each unit of beta.
- The lines cross at a beta of one, where both give eight per cent.
- To the left, the red line sits above the green, so low beta stocks show positive alpha.
- To the right it sits below, so high beta stocks show negative alpha.
- The dots are our three groups.

5 The case for, and the case against

- For: long samples, many countries, several asset classes
- For: survives controls for size, value and momentum
- So pure data mining is unlikely
- Against: part is other factor tilts
- Against: trading costs and limited capacity

IN THE LESSON

- Is the anomaly real, or a statistical fluke?
- The evidence covers long samples, many countries and several asset classes, not one market in one decade.
- It also survives controls for the size, value and momentum factors, so it is not a known effect in disguise.
- Pure data mining, searching the data until something turns up by chance, is therefore unlikely.
- The case against is that part of the measured alpha may be other factor tilts, such as a lean towards value stocks.
- Against it too are trading costs, which eat into paper returns, and limited capacity: the strategy may not survive large amounts of money chasing it.

6 Example 2: how long before alpha is believable?

t-statistic of alpha ≈ IR × √T, with T in years. About 2 is the usual bar

IR 0.5 over 16 years	0.5 × √16 = 0.5 × 4	**t = 2.0**
IR 0.5 over 4 years	0.5 × √4 = 0.5 × 2	**t = 1.0**
Years for t = 2 at IR 0.25	(2 ÷ 0.25)^2	**64 years**

Risk is not rewarded as the CAPM predicts; low risk does not win every year

IN THE LESSON

- The t statistic of an alpha is roughly the information ratio times the square root of the number of years, and about two is the usual bar.
- The information ratio, defined on the next slide, is alpha per unit of active risk.
- An information ratio of one half over sixteen years gives one half times four, a t statistic of two.
- Over four years it gives one half times two, just one.
- With a ratio of a quarter, the years needed are two divided by a quarter, squared: sixty four years.
- That is why long samples matter.
- And the exam point: the anomaly says risk is not rewarded as the C A P M predicts, not that low risk stocks win every year.

7 Four measures, defined

- Alpha: Rp − RB, or the regression intercept after beta
- Tracking error: SD of active return, Rp − RB
- Sharpe ratio: mean(Rp − Rf) ÷ σp
- Information ratio: alpha ÷ tracking error
- Annualise: mean × periods; SD × √periods

IN THE LESSON

- Now the four measures.
- Alpha is the return earned beyond what the chosen benchmark delivers: simply the portfolio return minus the benchmark return, or, more carefully, the intercept of a regression once beta is accounted for.
- Tracking error is the standard deviation of the active return, portfolio return minus benchmark return.
- It measures how far the portfolio strays from its benchmark.
- The Sharpe ratio is the mean of the portfolio return minus the risk free rate, divided by sigma p, the standard deviation of the portfolio's own return.
- The information ratio is alpha divided by tracking error, the reward per unit of active risk.
- To annualise, multiply the mean by the periods per year, and the standard deviation by the square root of that number.

8 Kestrel's four quarters of data

Quarterly returns, %. Benchmark: a low-volatility index. Rf 0.5% a quarter

Quarter	Fund	Benchmark	Active
Q1	−1.0	−2.5	+1.5
Q2	1.0	2.5	−1.5
Q3	5.0	3.5	+1.5
Q4	2.0	2.5	−0.5
Mean	1.75	1.50	+0.25

IN THE LESSON

- Here are quarterly returns for the Defensive fund.
- Its benchmark is a low volatility index, and the risk free rate is half a per cent a quarter.
- In quarter one the fund lost one per cent while the benchmark lost two and a half, an active return of plus one and a half.
- In quarter two the fund made one against two and a half, minus one and a half.
- In quarter three, five against three and a half, plus one and a half.
- In quarter four, two against two and a half, minus a half.
- On average the fund earned one point seven five a quarter and the benchmark one point five, so the active return averaged plus a quarter of a per cent.

9 Example 3: alpha, tracking error and IR

Active returns: +1.5, −1.5, +1.5, −0.5 (% a quarter). Mean +0.25

Sum of squared deviations	1.25^2+1.75^2+1.25^2+0.75^2	**6.75**
Sample SD, quarterly	√(6.75 ÷ 3) = √2.25	**1.5%**
Annual alpha	0.25 × 4	**1.0%**
Annual tracking error	1.5 × √4	**3.0%**
Information ratio	1.0 ÷ 3.0	**0.33**

IN THE LESSON

- The active returns are plus one and a half, minus one and a half, plus one and a half and minus a half, with a mean of plus a quarter.
- Their deviations from the mean are one point two five, minus one point seven five, one point two five and minus point seven five.
- Squared and added, they give six point seven five.
- For a sample standard deviation, divide by quarters minus one, three, giving two point two five, whose square root is one and a half per cent a quarter.
- Now annualise.
- Alpha is a quarter of a per cent times four, one per cent a year.
- Tracking error is one and a half times the square root of four, three per cent a year.
- The information ratio is one divided by three, about point three three.

10 Example 4: the Sharpe ratio, and two slips

Fund returns: −1, 1, 5, 2 (% a quarter). Mean 1.75. Rf 0.5% a quarter

Fund SD, quarterly	√(18.75 ÷ 3)	**2.5%**
Annual excess return	(1.75 − 0.5) × 4	**5.0%**
Annual fund SD	2.5 × √4	**5.0%**
Sharpe ratio	5.0 ÷ 5.0	**1.0**
Slips: α ÷ σp; no Rf	1.0 ÷ 5.0; (1.75×4) ÷ 5.0	**0.2; 1.4**

IN THE LESSON

- The Sharpe ratio uses the fund's own returns: minus one, one, five and two per cent, with a mean of one point seven five.
- The squared deviations from that mean add up to eighteen point seven five.
- Divide by three and take the square root: two and a half per cent a quarter.
- The annual excess return is one point seven five minus the risk free half a per cent, times four, five per cent.
- The annual standard deviation is two and a half times the square root of four, also five per cent.
- So the Sharpe ratio is five over five, exactly one.
- Now the two usual slips.
- Dividing alpha by the fund's own volatility gives one over five, point two, which is not an information ratio.
- Leaving out the risk free rate uses one point seven five times four, seven per cent, over five: one point four, overstating the Sharpe ratio.

11 Alpha depends on the yardstick

- Alpha is always measured against something
- Change the benchmark: alpha changes, portfolio does not
- Small-cap value fund judged against a large-cap index
- Size and value rewards then look like skill
- A style-matched benchmark strips them out

IN THE LESSON

- Alpha is always measured against something.
- Change the benchmark and the measured alpha changes, even though the portfolio is untouched.
- Consider a fund that holds small cap value stocks.
- Small cap means small companies; value stocks look cheap relative to earnings or book value.
- Suppose this fund is judged against a broad large cap index.
- When small value stocks do well, much of the fund's gain is simply the reward for size and value, which then looks like skill.
- A style matched benchmark, one built from small value stocks, strips those rewards out and leaves the manager's true contribution.

12 Example 5: one fund, two benchmarks

Kestrel Small Value Fund returned 12% last year

vs large-cap index (8%)	12 − 8	**+4%**
vs small-value index (11%)	12 − 11	**+1%**
Credited to style, not skill	4 − 1	**3 of 4 points**
Index picked after the fact	whichever flatters	**biased up**

IN THE LESSON

- Kestrel's Small Value Fund returned twelve per cent last year.
- Against a large cap index that returned eight per cent, its alpha looks like twelve minus eight, plus four per cent.
- Against a small value index that returned eleven per cent, its alpha is twelve minus eleven, just plus one per cent.
- Same portfolio, both times.
- The difference, four minus one, is three of the four points: the reward for size and value, credited to skill.
- A second trap: if the index is picked after the fact, whichever flatters the record, the measured alpha is biased up.
- The exam point: alpha is benchmark relative, and a mismatched benchmark credits factor returns to skill.

13 What makes a good benchmark

- Specified in advance; unambiguous and transparent
- Investable at low cost
- Measurable at the same frequency as the portfolio
- Matched to style and risk, beta close to the fund's
- Free of ex post selection

IN THE LESSON

- So what makes a good benchmark?
- It is specified in advance, before the results are known, and it is unambiguous and transparent, so everyone knows exactly what it holds.
- It is investable at low cost, meaning the client could simply buy it instead of hiring the manager.
- It is measurable at the same frequency as the portfolio.
- It is matched to the manager's style and risk, with a beta close to the portfolio's: a low beta fund judged against a full beta index looks poor in rising markets and brilliant in falling ones.
- Finally, it is free of ex post selection: nobody picks it after seeing the results.

14 The fundamental law of active management

IR = IC × √BR
Expected alpha = IR × TE
With constraints: IR ≈ TC × IC × √BR

Information coefficient IC	Skill: forecasts vs outcomes
Breadth BR	Independent bets per year
Tracking error TE	Active risk taken
Transfer coefficient TC	Below 1 with a long-only limit

IN THE LESSON

- Grinold's fundamental law of active management says where a good information ratio comes from, under ideal assumptions.
- The highest attainable information ratio equals the information coefficient times the square root of breadth.
- Expected alpha is then the information ratio times the tracking error.
- And when constraints get in the way, the ratio is scaled down by a transfer coefficient.
- Now the parts.
- The information coefficient, I C, is the correlation between a manager's forecasts and the returns that follow, a measure of skill.
- Breadth is the number of independent bets per year.
- Tracking error is the active risk the manager chooses to take.
- The transfer coefficient measures how fully forecasts reach the portfolio, and it is below one with a long only limit.

15 Example 6: Kestrel's ceiling on the IR

Kestrel: IC = 0.05, 100 independent bets a year, tracking error 3%

Highest IR	0.05 × √100 = 0.05 × 10	**0.50**
Expected alpha	0.50 × 3%	**1.5%**
Realised IR (Example 3)	only 4 quarters: too noisy	**0.33**
Double the IR, same skill	100 × 2^2	**400 bets**

Skill and breadth trade off, with diminishing returns

IN THE LESSON

- Kestrel's analysts have an information coefficient of point zero five.
- They make one hundred independent bets a year, and the fund runs a tracking error of three per cent.
- The highest attainable information ratio is point zero five times the square root of one hundred, so point zero five times ten, which is point five.
- Expected alpha is point five times three per cent, one and a half per cent a year.
- The realised ratio from Example three, one third, comes from only four quarters, too noisy to compare with this ceiling.
- To double the ratio at the same skill, Kestrel would need the square of two, four times the bets, so four hundred a year.
- Skill and breadth trade off, with diminishing returns.

16 Example 7: when the law overstates the IR

Same IC 0.05 and tracking error 3%. Two real-world frictions

25 independent of 100	0.05 × √25 = 0.05 × 5	**0.25: half ideal**
Long-only: TC = 0.6	0.6 × 0.05 × √100	**IR 0.30**
Alpha with TC = 0.6	0.30 × 3%	**0.9%**

IN THE LESSON

- Now two real world frictions, with the same skill and tracking error.
- First, Kestrel's one hundred positions are not all independent.
- Many are low beta stocks from a few industries that move together, so in effect there are only twenty five independent bets.
- The ratio becomes point zero five times the square root of twenty five, point zero five times five, which is point two five, half the ideal.
- Second, the fund is long only, so it cannot sell short the stocks it dislikes.
- Suppose that gives a transfer coefficient of point six.
- The ratio is roughly point six times point zero five times the square root of one hundred, point three.
- Expected alpha is then point three times three per cent, point nine per cent a year.

17 Assumptions and limits of the law

- Assumes: independent bets, same skill on every bet
- Assumes: forecasts used optimally, no constraints
- Limit: correlated signals shrink effective breadth
- Limit: skill (IC) is hard to estimate
- Limit: costs. Trap: correlated positions counted as bets

IN THE LESSON

- The law is an ideal.
- It assumes the bets are independent, and that skill is the same on every bet.
- It assumes forecasts are used optimally, with no constraints on the portfolio.
- Each of these can fail.
- Correlated signals shrink effective breadth: ten forecasts all driven by the same view on interest rates are one bet, not ten.
- Skill is hard to estimate, because an information coefficient of point zero five is tiny next to the noise in returns, so it takes many years to tell it from zero.
- And costs are ignored, though more bets mean more trading.
- The trap to remember is counting correlated positions as independent bets.

18 A factor regression

r = α + b1 × F1 + b2 × F2 + … + ε
Benchmark = b1 × F1 + b2 × F2 + …; α = r − benchmark
t-stat = α ÷ its standard error (about 2)

Alpha α	Return left unexplained
Loadings b1, b2, …	Sensitivity to each factor
Factor returns F1, F2, …	Market, size, value, …
Residual ε	Noise, averaging zero

IN THE LESSON

- A factor regression measures alpha against more than the market.
- The portfolio's excess return, r, its return minus the risk free rate, is regressed on the excess returns of a set of factors.
- The model says r equals alpha, plus b one times F one, plus b two times F two, and so on, plus an error.
- The multifactor benchmark is the factors weighted by the loadings, and alpha is r minus that benchmark.
- Its reliability is judged by the t statistic, alpha over its standard error, with about two as the usual bar.
- Now the parts.
- Alpha is the average return left unexplained.
- The loadings measure the portfolio's sensitivity to each factor.
- Factor returns include the market, size and value.
- The residual is noise that averages zero.

19 Example 8: alpha shrinks as factors are added

Fund 8%, Rf 2%, so r = 6%. Factor returns: market 5, size 2, value 3, low-beta 3

Market	0.6 × 5	**3.0%**
Size (small minus big)	−0.2 × 2	**−0.4%**
Value	0.5 × 3	**1.5%**
Low-beta factor	0.5 × 3	**1.5%**
Alpha	6.0 − (3.0 − 0.4 + 1.5 + 1.5)	**0.4%**

IN THE LESSON

- Last year the Defensive fund returned eight per cent and the risk free rate was two, so its excess return is six per cent.
- The factor excess returns were five per cent for the market, two for size, meaning small minus big companies, three for value, and three for a low beta factor, which buys low beta stocks and sells high beta ones.
- The market loading of point six contributes three per cent.
- The size loading is minus point two, because the fund leans to large companies, contributing minus point four.
- The value and low beta loadings, point five each, add one and a half apiece.
- The benchmark is five point six, so alpha is six minus five point six, just point four per cent, against three per cent with the market alone, keeping its loading at point six.
- Apparent skill was factor exposure.

20 Example 9: is it significant? Two common errors

Standard error of alpha: 1.2% a year

Market-only alpha (β 0.6): t	3.0 ÷ 1.2	**2.5**
Four-factor alpha: t	0.4 ÷ 1.2	**0.33**
Error: Rf not subtracted	8.0 − 5.6	**2.4% (too high)**
Error: size sign ignored	6.0 − (5.6 + 0.8)	**−0.4%**

IN THE LESSON

- The standard error of alpha is one point two per cent a year.
- Is either alpha real?
- The market only alpha, loading still point six, has a t statistic of three divided by one point two, two and a half, above the bar of two.
- The four factor alpha has a t statistic of point four divided by one point two, about point three three, nowhere near significant.
- Now two common errors.
- Forgetting to subtract the risk free rate from the fund's return gives eight minus five point six, an alpha of two point four per cent, far too high.
- Ignoring the sign of the size loading, adding point four instead of subtracting it, raises the benchmark by point eight.
- Alpha becomes six minus six point four, minus point four per cent.

21 Style analysis: exposures that move

- Long regression assumes fixed exposures; style drifts
- Style indices; weights ≥ 0, sum 1: the effective mix
- Residual = selection return, alpha vs the style benchmark
- Re-estimate over a rolling window to follow drift
- Short window: fast but noisy. Long: precise but slow

IN THE LESSON

- One long regression assumes the exposures stay constant, which fails for managers who change style.
- Style analysis, due to William Sharpe, regresses a fund's returns on the returns of a set of asset class or style indices, such as large growth, large value, small companies and cash.
- The weights are constrained to be non negative and to sum to one, so they can be taken as the fund's effective portfolio mix.
- The residual is the selection return, the alpha relative to this style benchmark.
- To cope with change, re-estimate the constrained regression over a rolling window, say three years, so the weights and the style benchmark move with the manager.
- A short window reacts faster but gives noisier weights.
- A long window is more precise but slow to detect a change.

22 Example 10: a manager who drifted

Last year: large growth 15%, large value 5%, small 8%, cash 2%. Fund 10.5%

Fixed benchmark (old mix)	0.6×5 + 0.3×8 + 0.1×2	**5.6%**
Rolling benchmark (new mix)	0.5×15+0.3×5+0.1×8+0.1×2	**10.0%**
Selection vs fixed mix	10.5 − 5.6	**4.9%**
Selection vs rolling mix	10.5 − 10.0	**0.5%**

Rolling constrained regression tracks exposures that change over time

IN THE LESSON

- Last year large growth stocks returned fifteen per cent, large value five, small companies eight and cash two.
- Kestrel's Core fund, which began as a value fund, returned ten and a half.
- Its early style mix was sixty per cent large value, thirty small and ten cash.
- A fixed benchmark from that old mix returned point six times five, plus point three times eight, plus point one times two, five point six per cent.
- The rolling window now finds fifty per cent large growth, thirty large value, ten small and ten cash, a benchmark return of ten per cent.
- Selection return looks like four point nine per cent against the fixed mix, but half a per cent against the rolling one.
- The fixed benchmark credited the drift into growth to skill; had growth lagged, it would have blamed the manager.
- Rolling constrained regression tracks exposures that change over time.

23 Nonlinear payoffs break the usual measures

- Alpha, Sharpe assume linear, roughly symmetric returns
- Options and dynamic trading break that assumption
- Selling puts: steady premium, rare large losses
- Negative skew and fat tails that a calm sample hides
- Convex (beta up in rallies): positive intercept, no skill

IN THE LESSON

- Standard alpha and Sharpe measures assume that a portfolio's return is linearly related to the benchmark or factor returns, and roughly symmetric.
- Options and dynamic trading break that assumption: such as selling options or replicating them.
- Take selling puts.
- A put pays its buyer if the market falls below a set level, so the seller is writing insurance.
- It collects a steady premium in calm markets, which shows up as high alpha and Sharpe ratio, yet it carries rare large losses.
- Its returns have negative skew, a long left tail, and fat tails, meaning extreme outcomes are more common than a normal curve suggests, and a calm sample hides them.
- A convex strategy, whose beta rises in up markets and falls in down markets, causes another problem: a positive intercept in a straight line regression, with no skill.

24 Example 11: the put writer's record

Put writer: 8% in each of 9 calm years, then −40% in a crash year

Nine calm years: mean	steady premium	**8.0%**
Crash year	the insurance pays out	**−40%**
Ten-year mean	(9 × 8 − 40) ÷ 10	**3.2%**
Excess over Rf 2%: calm years	8 − 2	**6.0%**
Excess over Rf 2%: all ten	3.2 − 2	**1.2%**

A high alpha from selling options may be pay for tail risk, not skill

IN THE LESSON

- Kestrel is reviewing a put writer for a client.
- In each of nine calm years the strategy earned eight per cent from its steady premium.
- Then came a crash year, when the insurance paid out and the strategy lost forty per cent.
- Over ten years the mean is nine times eight, minus forty, divided by ten, just three point two per cent.
- With a risk free rate of two per cent, the excess return was eight minus two, six per cent, in the calm years, but three point two minus two, only one point two per cent, over all ten.
- The exam point: a high alpha from selling options may be pay for tail risk, not skill.

25 Example 12: a convex strategy fools the regression

Market excess return −10%, 0 or +10%, equally likely. Strategy: −5%, 0, +15%

Down beta	−5 ÷ −10	**0.5**
Up beta	15 ÷ 10	**1.5**
Linear fit: beta	(50 + 0 + 150) ÷ 200	**1.0**
Linear fit: intercept	10/3 − 1.0 × 0	**+3.33%**
Add a (market)^2 factor	y = 1.0x + 0.05x^2, x in %	**alpha 0**

IN THE LESSON

- The market's excess return is minus ten, zero or plus ten per cent, equally likely, and a strategy returns minus five, zero or plus fifteen.
- It is convex: its beta rises in up markets and falls in down markets.
- Its down beta is minus five over minus ten, point five, and its up beta fifteen over ten, one point five.
- For a straight line fit, the market's mean is zero, so the slope is the sum of products, fifty plus zero plus one hundred and fifty, over the sum of squares, two hundred: a beta of one.
- The intercept is the strategy's mean, ten thirds, minus one times zero, about three point three three per cent: alpha with no skill.
- Add the squared market return as a factor and the curve fits exactly, one times the market plus point zero five times its square, with returns in percentage points.
- Alpha is zero.

26 Remedies for nonlinear strategies

- Add option-like factors, such as (market return)^2
- Use long samples that contain stress periods
- Check skewness and drawdowns, not just mean and SD
- Ask: tail-risk pay, or skill?

IN THE LESSON

- What can a risk manager do?
- First, add option like factors to the regression, such as the squared market return we just used, so that convexity is captured by a loading rather than by alpha.
- Second, use long samples that contain stress periods.
- The nine calm years told us nothing about the crash.
- Third, look at skewness and drawdowns, as well as the mean and the volatility.
- A drawdown is the fall from a peak to the low that follows it, and it shows pain that the standard deviation hides.
- And always ask: is it pay for tail risk, or skill?

27 Two anomalies, not one

- Beta anomaly: low beta, higher CAPM alpha; flat SML
- Volatility anomaly: low total or idiosyncratic vol wins
- Idiosyncratic: measured against a factor model
- Data: highest beta and volatility portfolios earned less
- Trap: treating beta and volatility as interchangeable

IN THE LESSON

- The low risk anomaly comes in two forms.
- The beta anomaly is the finding that low beta stocks have earned higher risk adjusted returns, measured as C A P M alpha, than high beta stocks.
- So the measured security market line is flatter than the C A P M predicts, and can even slope downward.
- The volatility anomaly is the separate finding that stocks with low volatility have earned higher returns than high volatility stocks.
- Volatility here can be total volatility, or idiosyncratic volatility, the stock specific part, which is measured against a factor model as the volatility of the residual.
- The data agree for both: the highest beta and highest volatility portfolios have earned less than the low risk ones.
- The usual error is treating beta and volatility as interchangeable.

28 Total risk splits in two

σ^2 = β^2 × σM^2 + σE^2
Total = market part + stock-specific part

Beta β	Sensitivity to the market
Market volatility σM	The same for every stock
Idiosyncratic volatility σE	Stock-specific, diversifiable

IN THE LESSON

- Here is why.
- A stock's total variance splits into a market part and a stock specific part.
- Total variance, sigma squared, equals beta squared times the market variance, plus the idiosyncratic variance.
- In words, the market part plus the stock specific part.
- Now the parts.
- Beta is the stock's sensitivity to the market.
- Sigma M is the market's volatility, the same for every stock.
- Sigma E is the idiosyncratic volatility, the stock specific part that diversification can remove.
- Beta drives only the first part.
- So a stock can have a low beta and still have high total volatility.
- A beta sort and a volatility sort can therefore put the same stocks in different orders.

29 Example 13: low beta, high volatility

Market volatility 20%. Stock A: β 0.6, σE 40%. Stock B: β 1.3, σE 10%

Stock A: variance	0.36 × 0.04 + 0.16	**0.1744**
Stock A: volatility	√0.1744	**41.8%**
Stock B: variance	1.69 × 0.04 + 0.01	**0.0776**
Stock B: volatility	√0.0776	**27.9%**
Ranking	β: A lower; volatility: B lower	**they differ**

IN THE LESSON

- Market volatility is twenty per cent, so its variance is point zero four.
- Stock A has a beta of point six and idiosyncratic volatility of forty per cent.
- Stock B has a beta of one point three and idiosyncratic volatility of ten per cent.
- Stock A's variance is point six squared, point three six, times point zero four, plus forty per cent squared, point one six: point one seven four four.
- Its volatility is the square root, about forty one point eight per cent.
- Stock B's variance is one point six nine times point zero four, plus point zero one: point zero seven seven six.
- Its volatility is about twenty seven point nine per cent.
- Ranked by beta, Stock A is lower risk.
- Ranked by volatility, Stock B is.
- They differ.

30 Example 14: a beta-neutral bet, and the data

Long $100m of low-beta stocks (β 0.6); short high-beta stocks (β 1.5)

Long side: market exposure	100 × 0.6	**$60m**
Short needed for zero beta	60 ÷ 1.5	**$40m**
Equal dollars: net exposure	100 × 0.6 − 100 × 1.5	**−$90m**

IN THE LESSON

- Kestrel goes long one hundred million dollars of low beta stocks, with a beta of point six, and sells short high beta stocks, with a beta of one and a half.
- It wants to bet on the anomaly without betting on the market.
- The long side carries one hundred times point six, sixty million dollars of market exposure.
- To cancel it, the short side needs sixty divided by one and a half, forty million dollars, a beta adjusted amount.
- Shorting equal dollars instead gives one hundred times point six, minus one hundred times one and a half, a net exposure of minus ninety million dollars: a large bet that the market falls.

31 Why might the anomaly exist?

Match each explanation to its mechanism

Explanation	Mechanism
Leverage constraints	Can't borrow, so buy high beta; bid it up
Benchmarking and agency	Beta below 1 creates tracking error
Behavioural	Lottery-like stocks; overconfidence
Risk-based	Riskier in bad times: weak support

IN THE LESSON

- Why might the anomaly exist?
- Match each explanation to its mechanism.
- First, leverage constraints.
- Investors who cannot or will not borrow, but want higher expected return, reach for high beta stocks.
- That bids high beta prices up and their future returns down, while low beta stocks are neglected.
- Second, benchmarking and agency.
- Managers with a tracking error budget against an index find low beta stocks costly to hold, since a beta below one creates tracking error, so they have little incentive to exploit the anomaly.
- Third, behavioural.
- Investors overpay for lottery like, positively skewed, high volatility stocks, and are overconfident in their forecasts.
- Last, the risk based story, that low risk stocks are riskier in bad times, has little empirical support.

32 Example 15: why a benchmarked manager holds back

Tracking error budget 3%. Market volatility 20%. Kestrel's portfolio β 0.7

TE from beta alone	(1 − 0.7) × 20%	**6%**
Versus the budget	6 ÷ 3	**2 times**
Lowest beta, no other risk	1 − 3 ÷ 20	**0.85 or more**

Pushed towards β = 1. Mechanisms: constraint, benchmark, preference

IN THE LESSON

- A pension client gives Kestrel a tracking error budget of three per cent against a market index whose volatility is twenty per cent, and Kestrel's portfolio has a beta of point seven.
- Here is the agency story in numbers.
- The beta gap alone creates tracking error of one minus point seven, times twenty per cent: six per cent.
- That is six over three, twice the budget.
- To stay inside it, with no other active risk, beta can be no lower than one minus three over twenty, point eight five, and in practice higher.
- So the manager is pushed back towards a beta of one, even if it believes in the anomaly.
- For the exam, name the mechanism of each story: a constraint, benchmarking, or a preference.

33 Try three quick questions

1. Quarterly active: mean 0.5%, SD 2%. Annual alpha, TE, IR?
   Answer: Alpha 2%, TE 2 × √4 = 4%, IR 0.5.
2. IC 0.04, 225 independent bets, TE 5%. Highest IR, alpha?
   Answer: IR = 0.04 × 15 = 0.6; alpha = 0.6 × 5% = 3%.
3. Excess 7%. Market 0.9 × 6%, value −0.3 × 2%. Alpha?
   Answer: Benchmark 5.4 − 0.6 = 4.8%; alpha 2.2%.

IN THE LESSON

- Now three quick questions.
- Pause, then listen to the answers.
- Question one.
- Quarterly active returns have a mean of half a per cent and a standard deviation of two per cent.
- What are the annual alpha, tracking error and information ratio?
- Alpha is two per cent, tracking error two times the square root of four, four per cent, and the ratio point five.
- Question two.
- An information coefficient of point zero four, two hundred and twenty five independent bets and a tracking error of five per cent.
- Highest information ratio and alpha?
- The ratio is point zero four times fifteen, point six, and alpha point six times five, three per cent.
- Question three.
- Excess return seven per cent, market loading point nine on a six per cent factor, value loading minus point three on a two per cent factor.
- Alpha?
- The benchmark is five point four minus point six, four point eight, so alpha is two point two per cent.