FINM 36700 - Final Exam (with Solutions)

Portfolio and Risk Management

Winter 2026

- This exam contains 100 multiple choice questions.
- Questions marked with † are **[Select All That Apply]** — choose all correct answers (there may be 0, 1, or multiple correct answers).
- No calculators or computers are allowed.
- All calculations are designed for mental math.

Questions

1. A student claims that diversification only works when assets are negatively correlated. This claim is:

   x Correct — portfolio variance can only decrease below the average when ρ ¡ 0

   ✓ Incorrect — diversification reduces portfolio volatility for any ρ ¡ 1

   x Correct — the equally-weighted portfolio formula shows variance increases with positive ρ

   x Incorrect — diversification works only when $\rho = 0$

   Diversification works whenever ρ ¡ 1. In class, the professor emphasized that typical equity correlations are ˜30-35% (positive) yet diversification still substantially reduces portfolio volatility. (Discussion 1.2, Video 2)

2. An equally-weighted portfolio of n stocks, each with volatility $\sigma = 20\%$ and average pairwise correlation $\rho = 0.04$, has portfolio volatility that converges to approximately:

   x 0% — all risk is diversified away

   x About 12% — only systematic risk remains

   x 20% — diversification has no effect

   ✓ About 4% — average covariance remains even in a very large portfolio

   As $n \to \infty$, $\sigma_p^2 \to \rho\sigma^2 = 0.04 \times 0.04 = 0.0016$, so $\sigma_p \approx \sqrt{0.0016} = 4\%$. Idiosyncratic risk vanishes but the average covariance term remains. (Discussion 1.2, diversification formula)

3. A stock has daily volatility of 0.5%. Its approximate annualized volatility (assuming 250 trading days) is closest to:

   x 500%

   x 32%

   x 2%

   ✓ 8%

   Annualized volatility is daily volatility times $\sqrt{250}$, so $0.5\% \times$ ˜$16 \approx 8\%$. Mean scales linearly with horizon, but volatility scales with the square root of time. (Discussion 1.1, annualization formulas)

4. A student ranks assets by individual Sharpe ratio and predicts the highest-Sharpe asset will receive the largest weight in the tangency portfolio. This prediction is:

   x Always correct — the tangency portfolio maximizes Sharpe by loading on high-Sharpe assets

   ✓ Incorrect — tangency weights depend on the full covariance structure, not just individual Sharpe ratios

   x Correct only if all pairwise correlations are equal

   x Correct only if the risk-free rate is zero

Tangency weights are $w^* \propto \Sigma^{-1}\mu$, which depends on the inverse covariance matrix. A high-Sharpe asset that is highly correlated with other holdings may receive less weight than a lower-Sharpe asset that provides diversification. This was demonstrated in the 442-stock exercise. (Discussion 1.3, Exercise 1.2)

5. In the S&P 500 constituent analysis, increasing the asset universe from 5 to 442 stocks caused the condition number of the covariance matrix to increase from ˜4 to ˜83,000. The practical consequence is:

   x The efficient frontier shifts inward, reducing the maximum achievable Sharpe ratio

   x The optimizer can no longer find a solution because the problem becomes non-convex

   ✓ Small estimation errors in covariances produce large, unstable swings in optimal weights

   x The covariance matrix becomes negative definite

   A high condition number means the matrix is ill-conditioned: small perturbations in inputs cause large changes in the solution. The frontier actually improves (more assets = higher Sharpe), but the weights become extremely sensitive to estimation error. The problem remains convex. (Discussion 1.3, Video 2)

6. In the HMC case, Private Equity appears to offer a large return premium over Domestic Equity with similar volatility and low correlation — a 'too good to be true' profile. The professor explained that the most likely cause of this apparent dominance is:

   x Private Equity managers generate genuine alpha through superior selection

   x Survivorship bias inflates average Private Equity returns

   ✓ Stale pricing of illiquid PE assets artificially smooths reported volatility and correlation

   x The sample period was unusually favorable for PE

   Private equity assets are priced infrequently (stale pricing), which smooths the return series, understating true volatility and correlation. As discussed in the managed funds lecture: 'these stale prices put in a lot of flatness that really isn't there.' This makes PE appear to have a better risk-return profile than it truly has. (HMC case C.1.0, Video 2, TA Review 9.3 on return smoothing)

7. An investor with access to a risk-free asset and the tangency portfolio wants to achieve a lower expected return than the tangency portfolio itself while remaining efficient. Under the Capital Market Line framework, the investor should:

   x Search for assets not on the efficient frontier

   x Lever up by borrowing at the risk-free rate and investing more than 100% in the tangency portfolio

   x Hold the global minimum variance portfolio instead

   ✓ Lend at the risk-free rate and invest less than 100% in the tangency portfolio

   With a risk-free asset, all efficient portfolios lie on the Capital Market Line and are combinations of the tangency portfolio and the risk-free asset. To target a lower return than the tangency portfolio, the investor lends at the risk-free rate and holds less than 100% in tangency. (Discussion 1.2, CML construction)

8. A bank calculates daily VaR and needs to report 100-day VaR to regulators. The standard convention is to scale daily VaR by:

   ✓ 10 (because $\sqrt{100} = 10$ under the square-root-of-time rule)

   x $\sqrt{10} \approx 3.2$ (square-root-of-time rule)

   x $10^2 = 100$ (variance scaling)

   x $\log(10) \approx 2.3$

   Under the usual square-root-of-time convention, τ-day VaR scales as $\sqrt{\tau}$ times daily VaR. For 100 days, that factor is $\sqrt{100} = 10$. The lecture emphasized that this is a convention under strong assumptions, not a law of nature. (Discussion 3.1, VaR horizon scaling)

9. A bank holds two trading desks. Each desk's standalone 5% VaR is \$10M. The combined desk's 5% VaR is \$22M. This violates which coherence property?

   x Positive homogeneity

   x Translation invariance

   ✓ Subadditivity

   x Monotonicity

   EXCLUDED - Not counted for grading. Disregard this question when evaluating performance. The canonical solution is retained below for reference only.
   Subadditivity requires $\rho(X+Y) \leq \rho(X)+\rho(Y)$. Here VaR(combined) = \$22M ¿ \$10M + \$10M = \$20M, violating subadditivity. This means VaR can penalize diversification — a known flaw. CVaR would not exhibit this behavior. (Discussion 3.X.1)

10. You backtest a 5% VaR model and observe that actual losses exceed the VaR threshold on 1% of trading days. This model is:

    ✓ Conservative — it overestimates risk

    x Anticonservative — it underestimates risk, which is dangerous for capital adequacy

    x Well-calibrated — fewer breaches just mean the realized sample happened to be calmer than usual

    x Uninformative — hit rates cannot evaluate VaR models

    A 5% VaR should be breached about 5% of the time. If actual breaches occur only 1% of the time, the threshold is too far into the tail and the model is conservative: it overstates risk relative to realized outcomes. (Discussion 3.1, hit test evaluation)

11. A portfolio has daily mean return of 0% and daily volatility of 0.1%. Using the normal distribution ($z_{0.05}$ = -1.645), the 5% daily VaR is approximately:

    ✓ -0.16%

    x -1.6%

    x -5%

    x -16%

Using the signed-return convention for VaR, the 5% return quantile is $\mu + z_q \times \sigma = 0 + (-1.645)(0.1\%) = -0.1645\% \approx -0.16\%$. If VaR were reported as a positive loss magnitude, the same tail size would be 0.16%; here the keyed answer follows the signed-return convention used in the course computation. (Discussion 3.1, normal VaR formula)

12. † **[Select All That Apply]** Which of the following are weaknesses of the expanding-window volatility estimator for conditional VaR?

- ✓ Each new observation receives vanishing weight (1/t), so it responds slowly to regime changes
- x It cannot be computed with fewer than 60 observations
- ✓ During a crisis, the estimate lags behind the true spike in volatility
- x It assigns negative weights to recent observations

The expanding window uses all historical data with equal weight (1/t per observation). As t grows, each new return barely moves the estimate — so it lags during crises when fast response matters most. There is no minimum observation requirement (B) and no negative weights (D). (Discussion 3.1, Video 5-6)

13. The EWMA volatility model assigns weights λ^i to past squared returns (with λ close to 1), so that recent observations receive more weight than distant ones. Compared to a 60-day rolling window, EWMA's primary advantage is:

- x It assigns equal weight to all observations in the window
- ✓ It produces a smoother volatility estimate because large past shocks decay geometrically rather than dropping off abruptly
- x It always produces lower volatility estimates
- x It requires fewer parameters to estimate via MLE

Rolling windows create discontinuities when a large return exits the window — as the TA review noted, 'one period of very high volatility' causes 'a gap up, and then a gap down.' EWMA weights decay geometrically, so old observations fade gradually — producing smoother transitions. The professor noted the rolling window problem is that 'one day the craziness just drops out.' (Discussion 3.1, TA Review 3.X.9, Video 5-6)

14. A stock has very high standalone volatility but zero correlation with the market. Under the CAPM, its expected excess return is:

- x High — investors demand compensation for bearing high volatility
- x Negative — uncorrelated assets are penalized
- x Equal to the market premium — all stocks earn the same expected return
- ✓ Zero — only systematic risk (market covariance) is compensated

$\beta = \rho(\sigma_i/\sigma_m)$. If $\rho = 0$, then $\beta = 0$, and $\mathrm{E}[\tilde{r}_i] = \beta \times \mathrm{E}[\tilde{r}_m] = 0$. The CAPM says only covariance with the market (systematic risk) is compensated; idiosyncratic volatility is diversifiable and earns no premium. This was shown in the Compensated Risk exercise (E.4.0): no relationship between mean return and standalone vol, skewness, kurtosis, or VaR. (Exercise 4.0, Video 4)

15. A time-series regression of stock excess returns on market excess returns yields $\alpha = 0.03\%$ per month (t-stat = 0.2). Under the CAPM framework, this alpha indicates:

    ✓ The CAPM is not rejected for this stock because alpha is economically tiny and statistically indistinguishable from zero

    x The stock is mispriced — it earns a return not explained by its market beta

    x The regression R-squared must be very high

    x The stock's beta is greater than 1

    Under CAPM, alpha should be zero. Here the estimated alpha is both economically tiny and statistically insignificant, so this stock does not provide evidence against the CAPM on its own. R-squared and beta level are separate issues. (Discussion 4.1-4.2, Video 4)

16. Under the CAPM, the Sharpe ratio of any individual asset i is related to the market Sharpe ratio by $SR_i = \rho_{i,m} \times SR_m$. This implies that no individual asset can have a Sharpe ratio exceeding the market's because:

    x Individual assets always have higher volatility than the market

    x The CAPM assumes all assets have the same expected return

    x Diversification always increases expected returns

    ✓ $—\rho_{i,m}— \leq 1$, so the scaling factor is at most 1

    Since correlation is bounded by $[-1, 1]$, the individual Sharpe ratio $SR_i = \rho_{i,m} \times SR_m$ cannot exceed SR_m in absolute value. Assets with low market correlation (high idiosyncratic risk) have proportionally lower Sharpe ratios. This elegant result follows directly from the CAPM and $\beta = \rho(\sigma_i/\sigma_m)$. (Discussion 4.1, TA Discussion 4.X.9)

17. A stock has volatility $\sigma_i = 30\%$, the market has volatility $\sigma_m = 15\%$, and their correlation is $\rho = 0.6$. The stock's market beta is:

    x 0.3

    x 0.6

    ✓ 1.2

    x 3.0

    $\beta = \rho \times (\sigma_i \ / \ \sigma_m) = 0.6 \times (30\%/15\%) = 0.6 \times 2 = 1.2$. Beta is correlation scaled by the ratio of volatilities. (Discussion 4.1, beta as scaled correlation)

18. In the DFA case, running the two-stage CAPM test on 25 size-value sorted portfolios produces a low cross-sectional R-squared and a positive intercept (γ ¿ 0). This result most directly motivates:

    x Rejecting all factor models as useless

    x Using a higher-frequency dataset to improve beta estimation

    x Replacing beta with volatility as the risk measure

    ✓ Adding size (SMB) and value (HML) factors to better explain cross-sectional return variation

The DFA case demonstrates that market beta alone cannot explain why small/value portfolios earn higher returns than large/growth portfolios. The low cross-sectional R-squared and significant residuals motivate the Fama-French extension (SMB, HML) — which is the direct bridge to Week 6 factor models. (DFA case C.4.0-C.4.1, Video 4)

19. The risk-free rate is 2% and the market excess return is 5%. A stock with $\beta = 0.2$ has a CAPM-predicted expected return of:

    x 9%

    x 11%

    x 6%

    ✓ 3%

    Under the CAPM, expected return equals the risk-free rate plus beta times the market excess return: $2\% + 0.2 \times 5\% = 3\%$. (Discussion 4.1, SML formula)

20. In the 49-industry CAPM regressions, the Software sector has a low R-squared but statistically significant alpha, while the Retail sector has a high R-squared but alpha near zero. The professor stated: 'Software had the worst R-squared... Retail looks great. Great R-squared. What does the CAPM say about that? Nothing.' Which statement best reflects this framing?

    x The CAPM works better for Software because its beta is higher

    x The CAPM works better for Retail because its R-squared is higher, showing a tighter fit

    ✓ The CAPM works better for Retail because its alpha is near zero — pricing accuracy, not R-squared, determines whether a factor model 'works'

    x Neither regression tells us whether the CAPM works

    The professor stated explicitly: 'Did you see anything in this theory that said anything about R-squared? ... The R-squared is a question of whether it's a good hedge or a good replicator.' R-squared measures hedging quality (Week 2 concept). Alpha measures pricing accuracy. Retail's near-zero α means the CAPM correctly prices it; Software's significant α means it doesn't — regardless of R-squared. (Video 4 lines 1930-1966, Discussion 4.2, connects to FACT-001)

21. The lecture says a momentum strategy built from the top decile has a much higher mean return than one built from the top half. What is the intended lesson?

    x Any broader portfolio must be mismeasured because means should be identical

    x Momentum only works if the long side is held without any short side

    x The top half construction is invalid because it uses too many stocks

    ✓ Extreme sorts can amplify a small underlying autocorrelation signal

    In the week 7 transcript the professor contrasts a top-decile momentum portfolio with a top-half construction and uses the difference to show how extreme sorting magnifies a small underlying persistence signal. (Video 7 - APT and Forecasting transcript around 1296-1300.)

22. The professor used a coin-flip analogy to explain momentum. A coin that comes up heads 50.5% of the time represents:

- x The small but positive alpha of a momentum strategy
- ✓ The tiny autocorrelation in individual stock returns, where diversification across many bets makes the edge reliable
- x The probability that a momentum strategy beats the market in any given year
- x The fraction of stocks that are winners minus losers

The professor asked: "50.5% coin... would you put your net worth on one flip? Would you rather play 10,000 times?" The point is that a tiny edge (autocorrelation ≈ 0.05) becomes reliable when you diversify across many stocks — just as a 50.5% coin becomes a near-certain winner over 10,000 flips.

23. The professor explained that naive monthly-rebalanced momentum strategies fail in practice because:

- ✓ The signal decays — big winners in January are rarely big winners in February, causing 100% monthly turnover
- x Short selling is prohibited for most stocks in the momentum universe
- x The momentum premium has been negative since 2000
- x Monthly data contains too much noise to identify momentum

The professor stated: "Big winners in January are rarely going to be big winners in February. That's the whole point." This leads to 100% turnover per month and enormous transaction costs: "If you try to implement momentum this way, you'll make money on paper and go out of business real fast."

24. The professor says that if you left the class knowing only three factors, they should be market, value, and momentum. If you knew a fourth, which one did he say to add?

- x SMB
- x CMA
- x The risk-free rate
- ✓ Profitability

In the week 6 lecture the professor says he would definitely want market, value, and momentum, and then would add profitability as the fourth factor. (Video 6 transcript around 1576-1579.)

25. A momentum strategy rebalances monthly with 100% turnover (the entire portfolio is replaced each month). If each trade costs approximately 50 basis points (0.50%), the annual transaction cost drag is approximately:

- x 0.5%
- x 3%
- ✓ 6%
- x 12%

The frozen item is using the lecture's one-way turnover-cost back-of-the-envelope: 100% turnover per month means one full-portfolio rebalance each month, so $12 \times 0.50\% = 6\%$ per year. If you instead count both sales and purchases separately, you would double that estimate to 12%; the keyed answer follows the one-way convention used in class when the professor said, 'Let's call it 6%.' (Video 6 ~L2144-2164.)

26. According to the professor, most practitioners diverge from the Fama-French 5-factor model by:

    x Dropping value and keeping investment

    ✓ Keeping profitability, using momentum, and dropping the investment factor

    x Replacing all FF factors with PCA-based statistical factors

    x Using only the market factor with leverage

    The professor stated: "Most people diverge from FF. They like profitability, use momentum, and most don't use investment." This reflects the practitioner consensus that differs from the academic FF5 specification.

27. In the week 6 lecture, the professor asks: what if a small stock behaves like a big stock? Under the Fama-French factor-pricing logic, what should NOT determine whether it earns a size premium?

    ✓ Its accounting label as a small-cap stock

    x Its beta or behavior with respect to the size factor

    x Its recent realized return over the last month

    x Whether it is in the Russell 2000 index

    The frozen stem is broader than the exact lecture contrast, but the intended point is still characteristics versus betas: Fama-French care about how the security behaves, not just how it is labeled. The keyed answer therefore targets the small-cap label itself as the thing that should not determine whether the stock earns the size premium. (Video 6 - Multifactor Models transcript, around the characteristics-versus-betas discussion.)

28. The practical solution to momentum's turnover problem, as described by the professor, is:

    x Use weekly instead of monthly rebalancing

    ✓ Use a 12-month ranking window instead of a 1-month window

    x Only trade the long side

    x Implement momentum with sector ETFs instead of individual stocks

    The professor stated that the solution to the turnover problem is using a "12-month ranking window instead of 1-month." A longer lookback window means the ranked portfolios change less frequently, reducing turnover and transaction costs.

29. Regarding the momentum premium's behavior in recent decades, the professor observed that:

    x Momentum has been consistently profitable every decade since publication

    ✓ Momentum performed well in the first ˜15 years after publication but poorly in the most recent ˜15 years

    x Momentum was only profitable among micro-cap stocks after publication

    x Momentum returns have become more volatile but maintained the same mean

The professor stated: "It does great in the first 15 years after published, then very poorly in the most recent 15 years." This raises questions about whether the premium is being arbitraged away, whether recent momentum crashes erased gains, or whether the pattern is within the range of normal variation for a volatile strategy. (Week 7 Lecture)

30. According to the professor, the fourth most important factor after market, value, and momentum is:

    x Size (SMB)
    x Investment (CMA)
    ✓ Profitability (RMW)
    x Low volatility (BAB)

    The professor stated his factor hierarchy: "If you left this class and you only knew three factors: market, value, and momentum. Fourth: profitability." Investment was explicitly described as less important.

31. According to the professor, what happened to the size factor premium after it became widely known in the 1980s?

    x It increased as more investors tried to exploit it
    x It remained stable but became more volatile
    ✓ Its Sharpe ratio dropped off a cliff
    x It reversed sign and became negative

    The professor stated: "1980s people know about the size factor, that starts getting published, and its Sharpe ratio drops off a cliff." This is a key example of factor decay after publication.

32. A researcher claims that momentum is entirely a small-stock phenomenon and does not exist among large-cap stocks. Based on the evidence discussed in class, this claim is:

    x Correct: momentum is only significant among small-cap stocks
    ✓ Incorrect: momentum is positive among both large and small stocks, though stronger in small stocks
    x Incorrect: momentum is actually stronger in large stocks
    x Correct: large-cap stocks are too efficiently priced for momentum to exist

    The professor stated: "If you restrict momentum to only the big stocks, you do still have a positive premium. It does show as being more powerful among small stocks, but it is still positive in the big stocks." Momentum is not solely a size effect. (Week 7 Lecture)

33. The professor described a company that sells products to small firms. Even though the company itself is large, it might still earn the size premium because:

    ✓ Its revenues are correlated with the small-cap sector, giving it a high SMB beta
    x Its book-to-market ratio is similar to small firms
    x It is included in the small-cap index by construction
    x Its dividend yield matches that of small firms

The professor's exact example: "A company that sells stuff to small firms — even though it's big, it might behave like small firms. FF says that's what matters." Under factor pricing, if the large company's returns co-move with small stocks (high SMB beta), it earns the small premium regardless of its own size.

34. For the same factor model applied to a set of test assets, the cross-sectional MAE is 1.2% and the time-series MAE is 2.5%. Is this result consistent with theory?

    x No: cross-sectional errors should always be larger than time-series errors

    ✓ Yes: the cross-sectional test has it easier, so its MAE must be less than or equal to the time-series MAE

    x No: the two error measures should be identical for any correctly specified model

    x Yes, but only if the model has at least three factors

    The professor stated this as a guaranteed exam fact: "The only thing I'm sure of in class or on a multiple choice exam is that the cross-sectional error has to be smaller than the time series error for the same model." The CS test fits factor premia to minimize cross-sectional pricing errors, while the TS test constrains premia to equal sample factor means. CS has more flexibility, so CS MAE ¡= TS MAE. (Week 7 Lecture)

35. A long-only momentum strategy that buys the top decile of past winners has a correlation to the overall market of approximately:

    x -30%

    x 10%

    x 50%

    ✓ 90%

    The professor stated: "The long only implementation has a correlation to the market of 90%. If you're an investor, it would take you years for this to feel very different than just being long the Russell 1000." Long-only momentum is dominated by market exposure, unlike the long-short version which has ˜-30% market correlation. (Week 7 Lecture)

36. A student claims: 'Alpha should always be zero in a good regression.' This statement is:

    ✓ Correct for factor pricing regressions only – in forecasting, alpha is just a re-leveling constant and irrelevant

    x Correct for both factor pricing and return forecasting regressions

    x Correct for forecasting regressions only – in pricing, alpha has no special meaning

    x Incorrect in all contexts – alpha is never exactly zero in practice

    The professor emphasized this dual meaning throughout the course. In pricing, $\alpha = 0$ is the test. In forecasting, 'we absolutely do not care about alpha ... alpha's just re-leveling.' A nonzero α in forecasting merely shifts the mean forecast without affecting whether the signal x_t has predictive power, which depends on β. (Video 7 ˜L2328-2360; QUESTION_GENERATION_PROMPT key distinctions.)

37. An analyst reports a forecasting R-squared of 0.01% for predicting monthly equity returns. The professor's response to this finding would be:

- x This is a useful result that should be traded aggressively
- ✓ This is too small to be economically meaningful — you would go bankrupt before the edge materializes
- x This is a strong result for monthly return forecasting
- x The R-squared is irrelevant; only the alpha matters for forecasting

The professor stated: "0.01% R-squared... Your rejoinder is, I'm going to go bankrupt before I ever see that edge." While any positive R-squared implies some predictability in theory, an R-squared this small means the signal is drowned out by noise. Transaction costs, estimation error, and finite capital make it economically unviable. (Week 7 Lecture)

38. Why does the course emphasize constructing factors as long-short portfolios rather than as long-only baskets?

- ✓ Long-short construction helps strip out broad market exposure so the factor is not just a disguised market bet
- x Long-only factors cannot be used in regressions
- x Long-short factors always have higher mean returns than long-only portfolios
- x Long-short construction guarantees zero volatility

The week 6 material repeatedly stresses that if you just grab a long-only basket of stocks, it will tend to be highly correlated with the market. Long-short construction is used to isolate the targeted exposure. (Video 6 transcript and discussions/6.0. Multifactor Pricing.md.)

39. In the AQR Momentum case, comparing momentum returns in small stocks versus large stocks (1994-2024), the evidence suggests:

- x Momentum is equally profitable in small and large stocks
- ✓ Small-stock momentum has a higher Sharpe (36.9%) than large-stock momentum (11.4%), but large-stock momentum still earns a positive mean return
- x Large-stock momentum has a higher Sharpe ratio than small-stock momentum
- x Momentum is only profitable in small stocks; large-stock momentum has a negative mean return

Small-stock momentum earns 6.2% with a Sharpe of 36.9%, while large-stock momentum earns 2.1% with a Sharpe of 11.4%. The all-stocks UMD earns 4.1% (Sharpe 24.9%). Large-stock momentum is weaker but still positive, which is relevant for AQR since their retail product focuses on larger, more liquid stocks for practical implementation.

40. † According to the professor, the two key ingredients that make momentum strategies profitable despite tiny individual stock autocorrelation are: (Select all that apply.)

- ✓ Trading only the extreme winners and losers (not the middle of the distribution)
- x Using leverage of at least 3:1
- ✓ Diversifying across many stocks to reduce noise
- x Holding positions for multiple years to let autocorrelation compound

The professor identified two ingredients: "(1) trade extreme winners/losers, (2) diversify across many stocks." Leverage and long holding periods were not part of the argument. The coin-flip analogy emphasized playing many times (diversification), not betting larger (leverage).

41. In the GMO/forecasting lecture, why is it a mistake to set a trading threshold using the full future sample rather than information available at time t?

    x It makes the strategy too conservative

    ✓ It introduces leakage and can generate phenomenal fake results

    x It eliminates the forecast intercept

    x It forces the strategy beta to equal one

    The lecture says even minor-looking leakage can create phenomenal fake results, and specifically warns that thresholds must be based only on information available at time t. (Video 8 - GMO, Forecasting, Managed Funds transcript around 1028-1038.)

42. GMO nearly went out of business in the late 1990s. According to lecture, what caused this?

    x GMO used too much leverage and suffered margin calls

    ✓ GMO's value-oriented strategy underperformed during the growth-dominated tech bubble

    x GMO's risk models failed to account for correlation spikes

    x GMO was caught in an accounting scandal

    "GMO almost goes out of business... because they're very much value investors. And the end of the 1990s are not a value time." Value strategies badly lagged growth/tech stocks during the late-1990s bubble. The professor noted AQR had the same experience, launching in the late 90s with a value tilt. (Lecture 8)

43. DFA's investment philosophy is best described as tilting toward which rewarded risk exposures?

    x Momentum beta and profitability beta

    ✓ Size beta and value beta

    x Dividend yield and earnings growth

    x Market timing and volatility targeting

    DFA's approach centers on loading on size beta and value beta — factors with equilibrium compensation. They are not forecasting returns; they are harvesting persistent factor premia. AQR is the firm that adds momentum and profitability to the value tilt. (Lecture 8)

44. What fundamentally distinguishes GMO's investment approach from DFA and AQR?

    x GMO uses leverage while DFA and AQR do not

    ✓ GMO forecasts where returns will go; DFA and AQR identify exposures with equilibrium high returns

    x GMO trades at higher frequency than DFA and AQR

    x GMO invests only in fixed income while DFA and AQR invest in equities

As stated in lecture: "DFA and AQR are NOT trying to predict where things will go. They predict what will have equilibrium high returns. GMO is the first firm in the course that is actually FORECASTING returns." DFA tilts toward size and value betas; AQR tilts toward value, momentum, and profitability betas. GMO uses fundamental signals (E/P, D/P) to forecast and holds positions for months. (Lecture 8)

45. The E/P tilt strategy stayed invested through the 2008 financial crisis but was largely out of the market from 2014 onward. What explains this pattern?

- x The strategy's risk model detected elevated volatility in 2014 and triggered a risk-off signal
- ✓ Earnings-price ratios remained below the expanding median after 2014, signaling poor valuations for equities
- x The strategy's momentum signal turned negative after 2014
- x Transaction costs made rebalancing prohibitively expensive after 2014

The professor noted: "Since 2014, this strategy has mostly been out of the market. Why? Because earning price ratios haven't been very good." The expanding median rose as E/P was historically moderate to low in the post-2014 bull market (prices high relative to earnings), keeping the signal below the threshold. During the 2008 crisis, E/P was still above the expanding median (prices had fallen, making valuations attractive), so the strategy stayed invested. (Lecture 8)

46. A student implements a forecasting strategy by computing $r_{\text{strategy},t} = w_t \times r_t$. What error has the student made?

- x The student should use excess returns instead of raw returns
- ✓ The student is multiplying a time-t weight by time-t returns instead of time-$(t+1)$ returns
- x The student forgot to subtract transaction costs
- x The student should be using log returns

The professor highlighted this as a common and critical mistake: "We're not used to multiplying a time t and a time t+1 variable, but that's absolutely what we need to do." The weight w_t is set based on time-t information, but the return it earns is r_{t+1}, the return realized in the next period. Multiplying by r_t introduces look-ahead bias because w_t was partly formed using information that includes r_t. (Lecture 8)

47. The professor used the analogy of a "voting machine" versus a "weighing machine." Which interpretation is correct?

- x In the short run, markets weigh fundamentals; in the long run, markets vote on sentiment
- ✓ In the short run, markets are a voting machine driven by sentiment; in the long run, they are a weighing machine driven by fundamentals
- x The voting machine refers to shareholder proxy votes affecting stock prices
- x The weighing machine refers to market-cap weighting in index construction

The professor invoked this classic Benjamin Graham distinction: "Short run: It's not just about whether you're right. It's about whether everyone else recognizes that you're right." In the short run, prices reflect collective sentiment (voting machine). Over the long run, fundamentals prevail (weighing machine). This is why GMO's value-based forecasts require patience. (Lecture 8)

48. The lecture notes that a forecast-target correlation around 24% in the classroom exercise looks suspiciously high. What is the professor's immediate concern?

    x The benchmark is too weak because it uses the expanding mean

    x The forecast is probably perfect and should be trusted without further checks

    ✓ There may be some data leakage or other backtest contamination

    x Correlation is mathematically impossible for forecasting returns

    In the lecture the professor says a 24% forecast correlation is suspicious and guesses there may be some leakage if one dug into it. (Video 8 transcript around 1522-1528.)

49. A forecaster correctly identifies that a stock is overvalued based on fundamentals. In the short run, the stock continues to rise. Which concept from lecture best explains this outcome?

    x The forecaster's model is wrong because prices are efficient

    ✓ Short-run prices reflect whether others recognize the same mispricing, not whether the forecast is correct

    x The stock must have a higher beta than the forecaster assumed

    x Survivorship bias is inflating the stock's apparent return

    The professor emphasized: "Short run: It's not just about whether you're right. It's about whether everyone else recognizes that you're right." A correct fundamental forecast can lose money in the short run because the market is a voting machine — consensus sentiment drives prices before fundamentals reassert. (Lecture 8)

50. A quant team's return forecasts have a negative OOS R-squared. Management considers firing the team. According to the lecture, why might this be a mistake?

    x Negative OOS R-squared always results from insufficient data and will improve with more observations

    x OOS R-squared is always negative for equity forecasts, so this is normal

    x The team should simply multiply their forecasts by -1 to flip the sign of R-squared

    ✓ The forecasts may still rank assets correctly — negative OOS R-squared reflects level/scale miscalibration, not sorting ability

    The professor shared a real-world anecdote: a quant team had negative OOS R-squared but their forecasts were "a good sorter" — when used to rank assets and build a simple long-short strategy, they generated "substantial positive alpha." The professor concluded: "You cannot decouple and assign a forecast score based on out-of-sample R-squared, then give it to another team." Correlation (sorting ability) matters more practically than R-squared (level accuracy). (Lecture 8)

51. When computing a threshold (e.g., a median) for a trading signal, why must the threshold be an expanding-window statistic rather than a full-sample statistic?

    x Expanding windows reduce estimation error

    x Expanding windows increase the strategy's Sharpe ratio

    x Full-sample statistics are computationally more expensive

    ✓ Full-sample statistics introduce look-ahead bias by using future data

    The professor was emphatic: "Even though that may seem like a very minor leakage of data, that kind of leakage will be enough to get phenomenal fake results." Using the full-sample median means the threshold at time t incorporates data from $t+1, t+2, \ldots, T$ — information the investor could not have known. (Lecture 8)

52. In the E/P tilt strategy discussed in lecture, when the E/P ratio is above its expanding median, the portfolio allocation is:

    x 50% equities, 50% risk-free

    x 100% equities

    ✓ 150% equities (borrowing 50% at the risk-free rate)

    x 200% equities (borrowing 100% at the risk-free rate)

    The professor described the strategy: "If E/P above expanding median: go 150% market, borrow 50% at risk-free. If E/P below expanding median: invest 50% equities, 50% risk-free." The strategy levers up when valuations are attractive and de-levers when they are not. (Lecture 8)

53. According to a formal theorem discussed in lecture, how should the variability of an optimal forecast compare to the variability of the quantity being forecast?

    x The forecast should be more variable than the target to capture tail events

    x The forecast should have the same variability as the target

    ✓ The forecast should be less variable than the target

    x There is no theoretical relationship between forecast and target variability

    The professor stated: "Your forecast should be less variable than what you're forecasting. There's a formal theorem about that." Intuitively, a forecast is a conditional expectation, which averages over residual uncertainty, making it smoother than the realized outcome. A forecast that is as volatile as the target is almost certainly overfitting. (Lecture 8)

54. When constructing a forecasting-based trading strategy, portfolio weights at time t must be a function of:

    x Returns realized at time t

    ✓ Information available at time t only

    x Returns realized at time $t+1$

    x The full-sample average of the signal

The professor emphasized: "When you build a strategy, weight at time t must be a function of what you could have known at time t." Using future information (including time $t+1$ returns or full-sample statistics) constitutes look-ahead bias. (Lecture 8)

55. The professor called this his 'trick question': In the forecasting regression $\tilde{r}_{t+1} = \alpha + \beta x_t + \epsilon_{t+1}$, the professor asks whether we want α to be zero. The answer is:

    x Yes – a nonzero alpha indicates model misspecification, just like in factor pricing

    x Yes – alpha should be zero for the forecast to be unbiased

    ✓ No – in forecasting, we only care whether beta is nonzero; alpha is just a re-leveling constant that does not affect forecast quality

    x No – we want alpha to be as large as possible to maximize expected returns

    The professor stated: 'This is my trick question ... we absolutely do not care. Alpha's just re-leveling them.' In factor pricing, $\alpha = 0$ is the key test. In forecasting, only β matters; a nonzero β means the signal x_t has predictive power for future returns. α just shifts the level and has no bearing on whether the forecast is informative. This is one of the professor's key 'alpha means opposite things' distinctions. (Video 7 ~L2328-2360.)

56. A fund forecasts returns for 500 individual stocks using separate models for each stock. According to lecture, what fundamental problem arises?

    x The fund will have too many positions to manage operationally

    ✓ Individual stock forecasts from separate models will be internally inconsistent, potentially creating near-arbitrage opportunities against the fund's own positions

    x Individual stock models always have lower R-squared than market models

    x Forecasting 500 stocks requires at least 500 factors

    The professor explained: "If you forecast 500 stocks individually with different models, forecasts will be internally INCONSISTENT." The example given was Diamondback Energy (FANG), which is 90% replicable by ConocoPhillips, Devon, and APA. If the model forecasts FANG at -10 bps but the replicating basket at flat, you have a near-arbitrage within your own portfolio. (Lecture 8)

57. How does the lecture suggest resolving the internal consistency problem when forecasting individual stock returns?

    x Use the same model specification for every stock

    x Average the forecasts across all stocks to remove inconsistencies

    ✓ Forecast a small number of factor returns and derive stock-level expected returns through factor loadings

    x Discard any stock that is highly correlated with other stocks

    The professor's solution: "Forecast a few factors, then make the premia attached to those factors" — using factor pricing as a no-arbitrage discipline. If stock returns are driven by a handful of factors, forecasting factor premia and mapping them to stocks via betas ensures internal consistency: stocks with similar factor exposures will automatically receive similar forecasts. (Lecture 8)

58. According to the classic Carhart finding on mutual fund persistence, a fund with strong recent performance:

    x Is very likely to continue outperforming

    ✓ Faces roughly a coin flip on whether it will outperform in the next period

    x Is guaranteed to mean-revert and underperform

    x Will outperform if it has positive momentum exposure

    The professor cited Carhart directly: "Funds with good returns face nearly a coin flip on whether they'll have good returns next period." Past winners do not reliably persist. However, past losers do tend to persist — making fund selection more about avoidance than selection. (Lecture 8)

59. The GMO lecture reports that GMWAX is about 80% correlated with SPY, about 86% with its benchmark, and has only a modest information ratio. What is the best allocator-level conclusion?

    x GMO is basically unrelated to broad equities

    x GMO's alpha is so strong that correlation information becomes irrelevant

    x The benchmark must be wrong because SPY correlation is not exactly 100%

    ✓ GMO behaves a lot like equities or its benchmark, with only a modest extra piece beyond that

    The lecture's point is that GMO has high correlation to broad equity benchmarks and only a modest residual component, so it should not be treated as if it were an entirely separate return stream. (Video 8 transcript around 584-650.)

60. In the forecasting SPY exercise, the DP (dividend-price ratio) forecast regression has an in-sample R-squared of approximately 0.68%. When this forecast is used out-of-sample with an expanding window, the most likely outcome for the OOS R-squared is:

    x Similar to in-sample, around 0.5-1.0%

    x Higher than in-sample, around 2-3%

    ✓ Approximately zero or negative

    x Around 5%, due to the longer estimation window

    EXCLUDED - Not counted for grading. Disregard this question when evaluating performance. The canonical solution is retained below for reference only.
    The assigned SPY forecasting exercise output actually reports a DP out-of-sample R-squared of about 1.6%, not zero or negative. That means the frozen item is flawed: the primary source does not support the keyed choice, and none of the printed answers matches the source value exactly. The broader lesson still holds that the out-of-sample gain is small, but this question should be treated as a problematic frozen item rather than as evidence that DP OOS R^2 was negative.

61. LTCM had approximately $8 billion in equity capital. Their convergence trades earned spreads of roughly 1% when they converged. At 12x leverage, the expected return on equity from a convergence trade was approximately:

- x 1%
- x 4%
- ✓ 12%
- x 50%

With 12x leverage, a 1% return on the traded notional becomes 12% on equity: 12 x 1% = 12%. The professor stated: 'Maybe these spreads are 1%. If the spreads converge, they make a percent. Multiplied by their leverage of 12.' LTCM wanted even higher leverage (20x) to generate 20% returns, which is why their actual reported gross returns of ˜24% annualized are consistent with high leverage on small spreads. (Video 9 ˜L1876-1888.)

62. Why did 'flight to quality' create a common-factor problem for LTCM's many spread trades?

- x Because it made all risky assets and low-quality instruments cheap at the same time
- ✓ Because many seemingly unrelated positions were effectively short lower-quality assets and long higher-quality ones
- x Because it only affected Russian bonds and nothing else in the portfolio
- x Because it eliminated all correlations across trades

The TA Review and lecture explain that many LTCM trades were effectively variations of quality spreads, so a broad flight-to-quality shock could hit many 'different' trades at once. (discussions/8.1. TA Review.ipynb and Video 9 transcript.)

63. The lecture's Royal Dutch/Shell example makes what asymmetry point about convergence trades?

- ✓ Potential profit from convergence is capped by the initial spread, but divergence can be much larger
- x Potential profit is unlimited if the spread keeps widening
- x Losses are capped at the initial spread because the firms are the same company
- x The trade cannot lose because the assets have the same underlying business

In the lecture, the professor explicitly says the spread can diverge a lot, but convergence gains are limited by the initial difference when the trade is entered. (Video 9 transcript around 1314-1318.)

64. The lecture explains that market timing a single security (SPY) using fundamental signals can produce a strategy with only 45% R-squared against passive SPY. This illustrates that:

- x The signals have no forecasting power
- ✓ Dynamic positioning in a single asset can create a return stream that behaves like an entirely different investment
- x SPY is not a valid benchmark for timing strategies
- x Regression R-squared is meaningless for timing strategies

The lecture states: "by adding market timing to a security, it becomes a totally different investment... my market timing of leaning in and out of Spy changes the investment stream to where it's really a totally new investment. That's why I can run a regression of my market timed spy. And when I run that regression, it says the R-squared is only 45%. Because the way I market time spy is so different from holding spy passively."

65. The lecture emphasizes that correlation between forecast and target may be superior to out-of-sample R-squared for evaluating forecasts because:

    x Correlation is always higher than R-squared

    ✓ Correlation is invariant to additive shifts in the forecast level

    x Correlation accounts for non-linear relationships

    x Correlation penalizes overfitting more harshly

    The lecture uses the analogy of a weather forecaster who is always off by 1000 degrees but perfectly tracks temperature movements. Out-of-sample R-squared would rate this forecaster terribly (huge squared errors), but correlation would be 100%. "My out-of-sample R-squared would be like -500%... But what would my correlation have been? 100%. That would have revealed you really do have something interesting here."

66. In the lecture, GMO is distinguished from DFA and AQR primarily because GMO:

    x Invests only in fixed income

    x Relies on equilibrium factor risk premia

    ✓ Actively forecasts returns using fundamentals like earnings and dividends

    x Uses high-frequency trading strategies

    The lecture states that DFA and AQR are "not trying to predict where things will go in the short term or the medium term" but instead predict "equilibrium high returns." GMO is described as "our first time where we're finally getting to forecast returns rather than take forecast as given," using signals like dividend-price ratio, earnings-price ratio, and interest rates.

67. The out-of-sample R-squared metric compares a model's forecast errors to:

    x The in-sample R-squared of the regression

    ✓ A benchmark forecast such as the expanding mean

    x The volatility of the target variable

    x The Sharpe ratio of the passive strategy

    The lecture defines out-of-sample R-squared as 1 minus the ratio of mean squared errors of the model to the mean squared errors of a benchmark (expanding mean). "A forecast is good if it has a small sum of squared errors... relative to the null." The DPI strategy achieved an OOS R-squared of 1.6%, and the EP strategy also 1.6%.

68. In the lecture, when discussing the earnings-price ratio tilt strategy, the instructor notes that since 2014 the strategy has mostly been out of the market. Why?

    x Earnings collapsed during this period

    ✓ Earnings-price ratios were below their expanding median, indicating expensive valuations

    x Interest rates were too high for equities to be attractive

    x The strategy's lookback window was too short

The lecture states: "Since 2014, this strategy has mostly been out of the market. Why? Because earning price ratios haven't been very good. What we mean is they've been expensive." The tilt strategy goes to 150% when E/P is above its expanding median and 50% when below, so persistently low E/P (expensive markets) keeps the strategy underweight.

69. In the lecture, the instructor describes a practical experience where a quant research team had negative out-of-sample R-squared but their forecasts still generated positive alpha. The key insight was:

    x Their forecasts had high R-squared in-sample

    ✓ Their forecasts did a good job of sorting bulls from bears, but were poorly scaled

    x They were using too few signals

    x Their forecasts were perfectly correlated with the market

    The lecture recounts: "Their forecast was doing a good job of sorting the bears and the bulls. But it was very bad at the actual scale." The out-of-sample R-squared penalized the poor scaling, but when used with smarter downstream positioning (ranking or Z-scoring rather than linear application), the forecasts worked well.

70. In the Forecasting with Fundamentals exercise, why does sector-hedging tend to reduce the apparent alpha of a dividend-yield carry strategy?

    ✓ Because much of the strategy's return is actually coming from sector tilts rather than pure stock-level carry

    x Because sector hedging mechanically forces the alpha to zero in every regression

    x Because dividend yield cannot be used as a forecasting variable

    x Because the strategy becomes a pure risk-free arbitrage after hedging

    The exercise logic is that high-dividend strategies often lean into defensive sectors, so once those sector exposures are hedged, much of the apparent alpha shrinks. The point is to separate stock-level signal from sector bet. (docs/exercises/S.8.1. Forecasting with Fundamentals_OUTPUT.ipynb.)

71. The lecture uses the analogy of a "voting machine vs weighing machine" in the context of GMO's investment philosophy. A convergence trader like LTCM faces the related risk that:

    x Fundamentals never determine prices in financial markets

    ✓ Even if you are right about fundamental value, the market may not recognize it before you run out of capital

    x Leverage is always beneficial for convergence trades

    x Flight-to-quality events are unpredictable and therefore unhedgeable

    The lecture states: "a firm can be right, history bears out their trade would have made money. But if other market participants don't update their views to your views and they just persist in what you would call the wrong view, then you're going to keep losing money until you're bankrupt and good for you. You were right. But in some sense, you were right too early." The TA review quotes Keynes: "markets can remain irrational longer than you can remain solvent."

72. In the LTCM case, extending the regression to include a quadratic market term (SPY-squared), the coefficient on SPY-squared was approximately -2.16. This indicates LTCM behaved as if it was:

    x Long market options (positive gamma)

    ✓ Short market options (negative gamma)

    x Delta-neutral with no volatility exposure

    x Long a straddle on the market

    The LTCM TA solution shows SPY-squared beta = −2.158. A negative quadratic coefficient means LTCM's returns were concave in market returns – it lost disproportionately in large market moves in either direction. This is the payoff profile of being short options (short gamma/short volatility), consistent with the case description that "LTCM is largely in the business of selling liquidity and volatility."

73. The professor argued that LTCM's convergence trades are 'a win-win – if I only used my own money.' If the spread converges, you profit. If the spread diverges, the trade looks even better. The professor then asked: what breaks this logic?

    x The counterparty might default, eliminating the position entirely

    x The spread can diverge without bound because the underlying assets are fundamentally different

    ✓ With leverage, a widening spread triggers margin calls that force you to close the position at the worst possible time – exactly when the trade looks best

    x Transaction costs on spread trades are too high to capture the small convergence premium

    The professor stated: 'The spread trades are a win-win if I only used my own money... But they just left out that one group – the funders.' With leverage, the trade's economics change fundamentally: 'When the trade has moved against you, that's exactly when you're going to get margin called. But that's exactly when you most want to stay in it.' The problem is not the trade's fundamentals – it is the interaction of leverage with mark-to-market margin requirements. (Video 9 ~L2524-2560.)

74. The LTCM case lists four funding risks. Which of the following is NOT one of them?

    x Collateral haircuts

    x Repo maturity

    x Equity redemption

    ✓ Dividend tax risk

    The LTCM case explicitly lists collateral haircuts, repo maturity, equity redemption, and loan access as funding risks. Dividend tax risk is not one of the four. (docs/case_studies/C.8.0. LTCM.ipynb.)

75. The Royal Dutch/Shell trade was an example of:

    x A momentum strategy

    ✓ An equity relative value trade

x A swap spread convergence trade

x A fixed-rate mortgage carry trade

The TA review states under "Equity relative value" that "Royal Dutch (listed in Amsterdam) and Shell (listed in London) are the same company. Royal Dutch trades at a premium. So enter into a total return swap where you short Royal Dutch and go long Shell."

76. In the TA review's collar example, a portfolio that is long SPY and buys a 3% OTM put while selling a 3% OTM call will have monthly returns bounded between:

✓ -3% and +3%

x -5% and +5%

x 0% and +3%

x -3% and unlimited upside

The collar caps upside at the call strike (+3%) and floors the downside at the put strike (-3%). The TA review shows the collar's Min = -0.03 and Max = 0.03, confirming this bounding.

77. In the LTCM case, the piecewise regression of LTCM net excess returns on SPY reveals that LTCM is:

x Long the call-like factor and long the put-like factor

x Short the call-like factor and short the put-like factor

✓ Short the call-like factor and long the put-like factor

x Long the call-like factor and short the put-like factor

This item is graded by reading the regression coefficients directly: SPY call beta -0.72 means LTCM loads negatively on the call-like factor, and SPY put beta $+1.04$ means LTCM loads positively on the put-like factor. So the keyed answer is 'short the call-like factor and long the put-like factor.' The explanation is intentionally limited to that sign pattern rather than to any broader options analogy.

78. In the TA review, LTCM's investment strategy is described as having a typical holding period of:

x Intraday to a few days

x A few weeks

✓ 6 months to 2 years

x 5 to 10 years

The TA review explicitly states LTCM had a "Long time horizon (6 months to 2 years)" for their trades. This is critical because convergence trades require patience for spreads to normalize.

79. † **[Select All That Apply]** LTCM's gross returns had annualized mean 24.4% and volatility 13.6%, yielding a Sharpe ratio of approximately 1.79. Over the same period, SPY had a Sharpe of approximately 1.36. Yet LTCM's net return skewness was -0.81 versus SPY's -0.41. Which combination correctly explains this pattern?

- ✓ LTCM's strategy was effectively short volatility, generating consistent small gains with occasional large losses
- ✓ LTCM's leverage amplified both returns and tail losses
- x LTCM's strategy had purely linear exposure to equity markets
- x LTCM's returns were positively skewed due to convergence trade payoffs

The LTCM TA solution shows net skewness = -0.81 and excess kurtosis = 2.93, compared to SPY skewness = -0.41 and kurtosis = -0.39. The TA review explains LTCM was "picking up pennies in front of a bulldozer" – consistent small gains (high Sharpe) with occasional large losses (negative skewness). The case states LTCM was "in the business of selling liquidity and volatility," and 50-100x leverage amplified both returns and tail events.

80. † **[Select All That Apply]** The TA review identifies the fundamental problem with all of LTCM's trades. Which of the following captures the core issue?

- ✓ The mispricings they exploited were typically very small
- ✓ Massive leverage was required to generate meaningful returns from small mispricings
- x Their trades had unlimited downside risk even without leverage
- x They could only trade in equity markets

The TA review states: "What is the problem with all of these trades? That typically these mispricings are very small. The only way to make a lot of money in them is to lever up in a serious way. On the repo market you can normally get between 50-100x levered." This is described as "picking up pennies in front of a bulldozer."

81. LTCM returned approximately $2 billion to outside investors before its collapse. According to the lecture discussion, why was this decision so consequential?

- x It violated the fund's lockup provisions and triggered regulatory action
- ✓ It increased the fund's leverage ratio dramatically (back to ˜20x), amplifying losses when spreads diverged
- x It forced the fund to close out its most profitable positions
- x It signaled weakness to counterparties, triggering margin calls

From lecture: "They returned this much money... their leverage goes from, well, back up to like 20x. Wildly. One year later, the firm has gone bust." The instructor notes: "If they hadn't increased their leverage in that way, I don't know if they would have made it or not... I think there is a chance they would have made it, because without the leverage... they might have been able to keep things from spiraling out of control."

82. What is Asness's main complaint in the lecture's lagged-beta example: that reported hedge-fund returns are wrong in level, or wrong in timing?

- ✓ Wrong in timing, because one month's reported return includes some of neighboring months
- x Wrong in level, because the average return must be negative
- x Wrong in both level and timing, and therefore unusable

x Neither; the point is only about taxes

The lecture says Asness is not even mainly complaining about the return level; he is saying the timing is wrong, with reported monthly returns containing pieces of earlier and later market moves. (Video 9 transcript around 2452-2456.)

83. The week 9 lecture jokes that if you held SPY in a private-equity vehicle and only reported returns with a lag, what would happen to the reported volatility and correlation?

✓ Both would look better than they really are

x Both would look worse than they really are

x Volatility would rise but correlation would stay unchanged

x Nothing would change because packaging cannot affect statistics

The lecture says that if you package liquid assets like SPY into a private-equity style reporting structure with lagged marks, the volatility will look incredible and correlations will be underreported. (Video 9 transcript around 2322-2332 and 2398-2406.)

84. Why does diversifying across many hedge fund managers potentially fail to add value, according to the lecture?

x Hedge fund returns are perfectly correlated with each other

✓ Combining many active long-short positions tends to reconstitute a passive market portfolio, but at much higher fees

x Diversification reduces alpha without reducing beta

x Hedge fund lockup periods prevent rebalancing

Slide 32 states: "Diversifying over investment managers just combines a lot of active positions back into a passive market position. But now you pay large fees to hold a passive position!" The long and short positions of different managers tend to cancel out, leaving market exposure with hedge fund fees.

85. Why do lock-up provisions in hedge funds potentially benefit remaining investors, according to the lecture?

x They allow the fund to charge higher management fees

✓ They prevent bank-run dynamics and fire-sale liquidation

x They guarantee a minimum return over the lock-up period

x They reduce the fund's tax obligations

Slide 40 explains that lock-ups "keep your fellow-investors locked in so that a run does not begin, with 'fire-sale' prices as a result." Without lock-ups, early redeemers get a higher price, forcing later withdrawals at fire-sale prices, creating bank-run dynamics (slide 39).

86. A hedge fund manages $100 million and charges 2% management fee plus 20% performance fee. In a year when the fund's gross return is 20% (above the high-water mark), the total fees paid to the manager are:

x $2 million

x $4 million

✓ $6 million

x $20 million

The keyed answer assumes the course's simplified 2-and-20 convention: 2% of beginning AUM plus 20% of gross profits above the high-water mark. Under that convention, the management fee is 2% x $100M = $2M and the performance fee is 20% x ($20M gross profit) = $4M, for a total of $6M. That is 6% of AUM, leaving investors with a 14% net return after fees in this example. (Discussion 9.1 slides; Video 9 ˜L2806-2828.)

87. A fund has iid annual returns with volatility 15% and a sample average return of 10% over T = 5 years. What is the approximate 95% confidence interval for the true mean return?

x 10% plus or minus 3%

x 10% plus or minus 7%

✓ 10% plus or minus 13%

x 10% plus or minus 20%

From slide 8: $CI(r, 95\%) = \bar{r} \pm 2 \times (15\%/\sqrt{5}) = 10\% \pm 2 \times 6.7\% \approx 10\% \pm 13\%$. So the true mean could plausibly be anywhere from -3% to 23%. This illustrates the enormous uncertainty in estimating mean returns with short samples.

88. † Which observations from the week 8 and week 9 materials support the course theme that labels can mislead about behavior? (Select all that apply.)

✓ GMO can be marketed as a distinct forecast-driven product and still end up highly correlated with broad equity benchmarks

✓ A private-equity reporting wrapper can make even liquid exposures look artificially smooth

✓ Return-based decomposition can reveal behavior even when the manager's narrative is vague

x A fund's stated label is enough to infer its true factor exposures without return data

Week 8 shows GMO behaving much more like broad equities than a casual label might suggest, and week 9 shows how private-equity style reporting can make ordinary exposures look smoother and less correlated. The lecture repeatedly says decomposition is how to test what a fund really behaves like. (Video 8 and Video 9 transcripts.)

89. Over 2000-2011, the GMO fund (GMWAX) and SPY show:

Fund	Cumulative Return	Ann. Sharpe
GMWAX	+25%	Positive
SPY	˜0%	Near zero

Over 2012-2025, these results reverse: SPY dominates with a Sharpe near 90%, while GMWAX lags. The professor's key concern about GMO's contrarian approach is:

x Contrarian strategies always underperform in the long run

- x GMO's models are fundamentally wrong because they missed the tech rally
- x All value-oriented strategies should be avoided because the value premium has disappeared
- ✓ If the periods of underperformance come first, the manager may lose capital (and clients) before the strategy pays off – this is career risk

The professor stated: 'It's hard to be a contrarian when others are making great returns and you're not. You may not survive to see the opposite... if you're immediately right, you'll be like a hero. But if it happens in the other order, you're probably going to have your capital pulled before you get a chance to redeem yourself.' GMO survived the late-1990s tech bubble, but career risk is the key structural concern with contrarian strategies. (Video 8 ˜L428-492.)

90. What strike price does STCM use for its put options relative to the current market price?

- x At-the-money (K = S)
- x 5% out of the money (K = 0.95S)
- ✓ 20% out of the money (K = 0.80S)
- x 30% out of the money (K = 0.70S)

The STCM notebook appendix specifies: "we sell puts with strike price of K = 0.80S, where S is the current market price." This 20% OTM strike means the puts only become costly when the market drops substantially.

91. A hedge fund reports a Sharpe ratio of 1.0 based on 25 years of annual data. Using $SE(SR) = \sqrt{(1 + SR^2/2)/T}$, the standard error of this Sharpe ratio estimate is approximately:

- x 0.04
- x 0.10
- ✓ 0.25
- x 1.00

$SE(SR) = \sqrt{(1 + 1^2/2)/25} = \sqrt{1.5/25} = \sqrt{0.06} \approx 0.245 \approx 0.25$. This means the 95% confidence interval for the true Sharpe ratio is approximately 1.0 ± 0.5, i.e., $[0.5, 1.5]$. Even with 25 years of data and an impressive reported Sharpe of 1.0, the true Sharpe could plausibly be as low as 0.5. (Discussion 9.1 slides; final_coverage.md.)

92. Two hedge funds each report a single +20% year. Fund A launched 8 months ago. Fund B has a 10-year history of mostly ordinary results. According to the week 9 lecture's Bayesian flow discussion, which fund should see the larger investor inflow from that one good year?

- ✓ Fund A, because one strong year updates beliefs more when the manager's skill is still highly uncertain
- x Fund B, because a long track record makes the new return more informative than for a young fund
- x Both should attract roughly the same inflow, because investors care only about the most recent return
- x Fund B, because older funds face less redemption risk and therefore stronger return chasing

In the week 9 Panopto lecture, the professor says that with a young hedge fund, "no one's sure if it has skill or not," so a good return causes a much larger Bayesian update and much more new money. For an old fund, a single great year is interpreted much more as luck than as new evidence of skill. (Video 9 - LTCM and Managed Funds, around the "young hedge fund" / "Bayesians" discussion at 2768-2780 in the transcript.)

93. According to Carhart (1997), as discussed in lecture, which type of fund performance is persistent?

- x Top-decile outperformance persists strongly
- x Middle-decile performance is the most stable
- ✓ Poor performance (bottom decile) is persistent
- x Both top and bottom decile performance persist equally

The lecture slides state that Carhart (1997) finds "no persistent outperformance, but persistent underperformance." The TA review reinforces: "The only significant persistence is by the worst-return performers, suggesting that poor performance is more persistent than good performance (hence Mark talks about avoiding disaster)."

94. What is the standard error formula for the Sharpe ratio, as presented in the TA review notebook?

- x $SE(SR) = 1/\sqrt{T}$
- x $SE(SR) = SR/\sqrt{T}$
- ✓ $SE(SR) = \sqrt{(1 + SR^2/2)/T}$
- x $SE(SR) = \sqrt{(1 + SR)/T}$

The TA review notebook states: $SE(SR) = \sqrt{(1 + SR^2/2)/T}$. This comes from the delta method, accounting for uncertainty in both the numerator (mean) and denominator (volatility) of the Sharpe ratio.

95. The professor's STCM (Short-Term Capital Management) simulation showed that conditional on surviving months, the fund had a Sharpe ratio of 150% and mean of 25%/year. But when 3 bust months were added back, the mean dropped to -27% and the Sharpe fell to -13%. The key lesson is:

- x Tail-risk strategies should be evaluated using only the months when they are operational, since bust months are by definition extreme outliers
- x The Sharpe ratio is robust to outliers and should not change this dramatically from three data points
- x The STCM strategy is profitable as long as it avoids the three specific crisis months
- ✓ Any investor being pitched a hedge fund is seeing conditional-on-survival statistics; the unconditional distribution – which includes the busts – tells a fundamentally different story

The professor stated: 'The return stats that you saw – that's conditional on being months I survived. I excluded three months from that sample. Because if you're getting pitched a hedge fund, is their sample including the month they go bust? No, they haven't gone bust

yet... you are staring at the left column, not the right.' Only 3 data points out of ˜420 months fundamentally alter the picture because tail-risk strategies have extreme payoffs conditional on rare events. (Video 9 ˜L884-944; STCM notebook bust dates.)

96. LTCM's convergence trades on bond spreads were leveraged approximately 30-50x. If a spread trade earns 30 basis points unleveraged and is leveraged 10x, what is the approximate leveraged return?

    ✓ 3%

    x 9%

    x 15%

    x 30%

    The same lecture logic applies at lower leverage: 0.30% times 10 is about 3%. This illustrates why tiny spread trades only become economically meaningful once leverage is applied. (Week 9 LTCM discussion.)

97. A hedge fund has a 5-year track record with an observed annualized mean return of 10% and annualized volatility of 15%. Using the CLT approximation, the 95% confidence interval for the fund's true mean annual return is approximately:

    x 7% to 13%

    x 4% to 16%

    x 0% to 20%

    ✓ -3% to 23%

    $SE(\text{mean}) = \sigma/\sqrt{T} = 15\%/\sqrt{5} \approx 15\%/2.24 \approx 6.7\%$. The 95% CI is approximately $10\% \pm 2 \times 6.7\% = 10\% \pm 13.4\%$, giving roughly -3% to 23%. The professor gave this exact example: 'Your 95% confidence interval would say that hedge fund might have a mean of 10%, but I really couldn't distinguish ... negative three to positive 23. And note again, that's after five years.' (Video 8 ˜L2646-2660; Discussion 9.1 slides.)

98. An ETF markets itself as a value strategy, but a return-based factor decomposition shows little loading on value and much larger exposure to other risks. According to the course's week 9 takeaway, what should an allocator conclude?

    x The label is decisive; if the prospectus says value, the fund should still be treated as value

    x Without seeing every holding, no inference about behavior is possible

    ✓ Behavior matters more than marketing labels, and decomposition is the right tool for checking what the fund really does

    x Weak value loading proves the fund has generated true alpha

    Near the end of week 9, the professor emphasizes that "what something's labeled is not how it behaves" and immediately ties that lesson to linear factor decompositions. Earlier in the lecture he also notes that, with enough data, decomposition lets you say "I don't know what you do, but I know what you behave like." The examable point is that allocators should trust behavior revealed in the return series more than the label. (Video 9 transcript around 2360-2363 and 2940-2947.)

99. What does the phrase "you can't eat Sharpe ratio" mean in the context of hedge fund evaluation?

    x The Sharpe ratio is unreliable for non-normal distributions

    ✓ A high Sharpe ratio on a small asset base does not generate enough PnL to matter

    x Sharpe ratios cannot be compared across different funds

    x The Sharpe ratio ignores transaction costs and taxes

    In lecture, the instructor explains: "You could have a great Sharpe ratio, but if it's on $1 million, it's just not going to generate enough PnL to pay anybody... to pay staff or anything." Scale matters beyond risk-adjusted returns.

100. According to lecture, what was the Royal Dutch / Shell trade that LTCM engaged in?

    x Betting that Shell's oil production would exceed Royal Dutch's

    ✓ Exploiting a price discrepancy between two share classes of the same company that should converge

    x A currency arbitrage between Dutch guilder and British pound denominated shares

    x A merger arbitrage bet that Royal Dutch would acquire Shell

    In lecture, Royal Dutch and Shell are described as "the same company, same corporate assets, just different wrappings" — two share classes trading at different prices due to a tax issue. LTCM bet the spread would converge. The instructor uses the analogy of Berkshire Class A and Class B trading differently.