FINM 36700 - Midterm Exam (with Solutions)

Portfolio and Risk Management

Winter 2026

- This exam contains 100 multiple choice questions.
- Questions marked with † are [**Select All That Apply**] – they may have zero, one, or multiple correct answers.
- No calculators or computers are allowed.
- All calculations are designed for mental math.

Section 1: Mean-Variance Optimization

Questions 1–3: Three-Asset Snapshot

Reference Table: Expected returns and volatilities

Asset	$E[r]$	σ	Notes
A	6%	10%	$\text{Corr}(A, B) = 0$
B	8%	20%	$\text{Corr}(A, C) = 0.5$
C	10%	30%	$\text{Corr}(B, C) = 0.5$

1. A portfolio invests 50% in A and 50% in B. What is its expected return?

   x 6%
   ✓ 7%
   x 8%
   x 9%

   $0.5 \times 6\% + 0.5 \times 8\% = 7\%$.

2. Using $\text{Corr}(A, B) = 0$, what is the approximate volatility of the 50/50 A–B portfolio?

   x 10%
   ✓ 11%
   x 14%
   x 20%

   $\text{Var} = 0.5^2(0.1^2 + 0.2^2) = 0.0125$, so $\sigma \approx 11\%$.

3. Which pair offers the strongest diversification benefit based on correlations?

   ✓ A and B
   x A and C
   x B and C
   x All pairs are equally diversifying

   A–B has zero correlation, the lowest.

4. † [**Select All That Apply**] Which statements about the efficient frontier with a risk-free asset are correct?

   ✓ The tangency portfolio maximizes the Sharpe ratio among risky portfolios.
   ✓ Adding a new asset cannot reduce the maximum achievable Sharpe ratio.
   x The tangency portfolio must include every risky asset with positive weight.
   x The capital market line is unchanged if expected returns all rise by the same constant.

Tangency maximizes Sharpe; adding assets expands the set. If all expected returns shift up equally, Sharpe ratios change.

5. If the covariance matrix is assumed diagonal (all correlations set to zero), the MV optimal weights are proportional to:

   - x $1/\sigma$
   - ✓ μ/σ^2
   - x μ/σ
   - x σ/μ

   With diagonal covariance, $w \propto \Sigma^{-1}\mu$, so $w_i \propto \mu_i/\sigma_i^2$.

6. For a large N-asset portfolio with identical variances and average pairwise correlation ρ, the limiting portfolio variance as $N \to \infty$ is approximately:

   - x 0 for any $\rho < 1$
   - x σ^2/N
   - ✓ $\rho\,\sigma^2$
   - x σ^2 regardless of ρ

   Diversification removes idiosyncratic risk, leaving $\rho\sigma^2$.

7. If the correlation between two assets in a portfolio decreases while their individual volatilities and expected returns remain constant, the portfolio's Sharpe Ratio will generally:

   - ✓ Increase.
   - x Decrease.
   - x Remain exactly the same.
   - x Become negative.

   Lower correlation reduces portfolio variance while expected return stays constant, improving the Sharpe ratio.

8. In a MV-efficient world, what is the slope of the Capital Market Line?

   - ✓ Sharpe ratio of market portfolio
   - x Mean return of market portfolio
   - x Risk-free rate
   - x Treynor ratio of market portfolio

   The CML slope is (E[Rm] - Rf) / σ_m, which is the Sharpe ratio of the market (tangency) portfolio.

9. In portfolio optimization, measurement error is most problematic for:

   - x Transaction costs.
   - ✓ Expected returns.
   - x Covariances.

- x The risk-free rate.

Expected returns are notoriously difficult to estimate and have the largest impact on optimal weights.

10. Suppose all assets have identical Sharpe ratios and volatility but varying correlations. Which of the following is true of the tangency portfolio?

- x The tangency portfolio is any portfolio with positive weights summing to 1.
- x The tangency portfolio is undefined.
- x Any mean–variance efficient portfolio of risky assets becomes a tangency portfolio.
- ✓ The tangency portfolio reduces to the global minimum variance (GMV) portfolio.

When all assets have identical Sharpe ratios, the CML touches the frontier at the GMV point.

11. Consider the Two-Fund Separation Theorem. If the borrowing rate r_{borrow} is higher than the lending rate r_{lend}, how many tangency portfolios exist?

- x Still exactly one unique tangency portfolio.
- ✓ Two distinct tangency portfolios (one for each rate)
- x Infinitely many tangency portfolios
- x The efficient frontier ceases to be convex.

Different borrowing and lending rates create two CML lines, each tangent at a different portfolio.

Question 12: Gross Exposure Constraint

Reference Table: Unconstrained MVO weights

Asset	Weight	μ (Expected Return)	σ (Volatility)
AAPL	+240%	15%	32%
NVDA	+180%	18%	45%
SPY	−320%	10%	16%
Portfolio	100%	24.4%	28.1%

12. If you add a **gross exposure constraint of 200%** (weights sum to 100%), approximate new weights are:

- x AAPL = 120%, NVDA = 90%, SPY = −110%
- x Weights scale proportionally to original ratios
- x AAPL = 80%, NVDA = 60%, SPY = −40%
- ✓ AAPL = 80%, NVDA = 70%, SPY = −50% (sums to 100%, gross = 200%)

Gross 200% and net 100% implies longs sum to 150% and shorts sum to -50%. The keyed weights (80/70/−50) are the only choice satisfying both constraints. Note that constrained re-optimization does not simply scale the original unconstrained weights.

13. A constraint has a positive Lagrange multiplier at the optimum. This indicates that:

    x The optimization is infeasible

    ✓ Relaxing the constraint would improve the objective

    x The constraint is slack and irrelevant

    x The constraint lowers portfolio variance at no cost

    A positive multiplier means the constraint is binding and costly.

14. A long-only constraint forces an asset with negative unconstrained weight to:

    x An indeterminate weight

    ✓ Zero weight in the constrained solution

    x Positive weight in the constrained solution

    x Negative weight with reduced magnitude

    Binding long-only constraints set negative weights to zero.

15. † [**Select All That Apply**] Which inputs are most likely to make MV weights unstable?

    ✓ Highly correlated assets (ill-conditioned covariance matrix)

    ✓ Noisy expected-return estimates

    x Increasing sample size for estimating covariances

    x Imposing long-only constraints

    Ill-conditioning and noisy means produce extreme, unstable weights. Constraints actually help stabilize.

Section 2: Harvard Endowment Case (HW1)

Questions 16–25: HMC Asset Allocation

Background: Harvard Management Company (HMC) manages a multi-billion dollar endowment across 11 asset classes. They use a two-stage optimization process and place bounds on allocations rather than implementing unconstrained MV solutions.

Reference Table: Selected HMC Asset Class Statistics (Annualized)

Asset Class	Mean	Vol	Sharpe	Corr w/ Dom Eq	Weight (Bounded)
Domestic Equity	12%	16%	0.50	1.00	0%
Foreign Equity	10%	20%	0.35	0.60	15%
Private Equity	16%	24%	0.50	0.40	15%
Domestic Bonds	4%	8%	0.25	0.20	0%
TIPS	3%	6%	0.17	0.10	5%
Commodities	6%	20%	0.15	−0.10	10%
Real Estate	8%	12%	0.42	0.30	10%

16. † **[Select All That Apply]** Why does HMC use a two-stage optimization (first within asset classes, then across asset classes) rather than optimizing across all individual securities directly?

    ✓ Estimation error in the covariance matrix becomes severe with thousands of securities.

    ✓ The number of parameters to estimate grows as $N(N+1)/2$, creating dimensionality problems.

    x Regulatory requirements mandate asset-class-level reporting.

    x Two-stage optimization always produces higher Sharpe ratios than single-stage.

    With N securities, the covariance matrix has $N(N+1)/2$ parameters. For 1000 stocks, that's ∼500,000 parameters to estimate, leading to massive estimation error and unstable weights. Two-stage optimization reduces dimensionality at some cost to theoretical optimality.

17. In the HMC data, if we drop TIPS from the investment set entirely, the tangency portfolio's Sharpe ratio will:

    x Stay exactly the same, because TIPS have near-zero correlation with equities.

    x Become undefined, because TIPS are needed to span the efficient frontier.

    x Increase, because TIPS have the lowest Sharpe ratio.

    ✓ Decrease or stay the same, because removing an asset cannot expand the opportunity set.

    Removing any asset can only shrink (or leave unchanged) the efficient frontier. The tangency Sharpe cannot increase.

18. HMC focuses on **real** (inflation-adjusted) returns rather than nominal returns. This matters for MV optimization because:

    x Real returns are always higher than nominal returns.

- x Real returns have lower volatility than nominal returns.
- ✓ Inflation affects asset classes differently, changing both correlations and relative expected returns.
- x MV optimization requires returns to be expressed in real terms by construction.

Inflation hedges (like TIPS, commodities) look different in real vs nominal terms. Their correlations with other assets change when inflation is removed.

19. † **[Select All That Apply]** Commodities have a low Sharpe ratio (0.15) but receive 10% weight in the bounded portfolio. Which reasons could justify this allocation?

- ✓ Commodities have negative correlation with Domestic Equity, providing diversification.
- x Commodities have the highest expected return in the table.
- ✓ Adding a low-correlation asset can improve portfolio Sharpe even if its standalone Sharpe is low.
- x The bounded optimization requires at least 10% in each asset class.

Diversification benefit comes from low/negative correlation. An asset with Sharpe 0.15 and $\rho = -0.10$ can improve portfolio efficiency.

20. In the Constrained Optimization homework, weights were bounded between -10% and $+20\%$ per position. If the unconstrained weight for NVDA was $+180\%$, what would happen to NVDA's weight in the bounded solution?

- ✓ It would be exactly +20%, hitting the upper bound.
- x It would be +100%, the maximum allowed for any single position.
- x It would be 0%, because extreme weights are eliminated.
- x It would be +90%, half of the unconstrained value.

The upper bound of 20% becomes binding for any position the unconstrained solution wanted above 20%.

21. A fund has $R^2 = 0.85$ in a regression on market factors. What percentage of the fund's return variance is **idiosyncratic** (not explained by the factors)?

- x 0%
- ✓ 15%
- x 85%
- x Cannot determine from R^2 alone

Idiosyncratic variance fraction $= 1 - R^2 = 1 - 0.85 = 0.15 = 15\%$.

22. Based on the HMC table, which asset provides the strongest diversification benefit relative to Domestic Equity?

- x Private Equity
- ✓ Commodities
- x Foreign Equity

- x Real Estate

Commodities have correlation of −0.10 with Domestic Equity, the lowest (and only negative) in the table.

23. A constrained optimizer shows Lagrange multipliers for policy bounds below. Which bound is most costly?

Constraint	Lagrange Multiplier
Domestic Equity $\leq$ 30%	0.2
Private Equity $\leq$ 15%	1.1
Real Estate $\leq$ 10%	0.4

- ✓ Private Equity bound
- x Real Estate bound
- x All equal
- x Domestic Equity bound

Largest multiplier indicates most costly constraint. Relaxing the PE bound would most improve the objective.

24. Two portfolios have the following characteristics (risk-free rate 4%):

Portfolio	$E[r]$	σ
P1	10%	12%
P2	14%	20%

Which portfolio has the higher Sharpe ratio?

- x P1
- x P2
- ✓ Equal
- x Cannot be determined

Sharpe(P1) = (10-4)/12 = 0.50, Sharpe(P2) = (14-4)/20 = 0.50. Equal Sharpe ratios.

25. † **[Select All That Apply]** Why might an investment policy committee impose upper bounds on asset class weights even when the optimizer wants higher allocations?

- ✓ Liquidity constraints for large endowments
- ✓ Estimation error makes extreme optimal weights unreliable
- ✓ Governance and fiduciary responsibility concerns
- x Upper bounds always improve the portfolio Sharpe ratio

Constraints reflect real-world limitations (liquidity, governance) and protect against estimation error. Constraints cannot improve the theoretical optimum.

Section 3: Linear Factor Decomposition

Questions 26–36: Factor Regression

Reference Table: Monthly factor regressions

Fund	β_{Mkt}	β_{Bond}	α (ann.)	R^2	TE
Fund A	0.6	0.2	1.0%	0.80	4%
Fund B	1.0	-0.1	0.0%	0.40	8%
Fund C	0.3	0.5	2.0%	0.55	6%

26. Which fund has the highest tracking error?

    x Fund C

    x All equal

    x Fund A

    ✓ Fund B

    Fund B has TE = 8%, the highest.

27. In a Linear Factor Decomposition (LFD) regression $R_p = \alpha + \beta_1 F_1 + \beta_2 F_2 + \epsilon$, if we want to build a *tradeable* replication of R_p using ETFs for F_1 and F_2, we should probably:

    x Ignore the betas and just equal weight the ETFs.

    x Short the ETFs to capture the alpha.

    x Focus on maximizing the R^2 including the intercept.

    ✓ Force the intercept to zero.

    The intercept is not tradeable; forcing it to zero ensures the replication uses only investable positions.

28. In a replication context, basis risk refers to:

    x The risk that the hedge fund manager changes the strategy.

    x The risk that we cannot replicate the fund exactly due to differences in strategy.

    ✓ The volatility of the tracking error between the replication and the target.

    x The risk of interest rates rising.

    Basis risk is the residual risk from imperfect replication – the volatility of the tracking error.

29. † **[Select All That Apply]** Which statements about a "Time-Series Regression" (regressing one asset's returns on factor returns over time) are correct?

    ✓ It estimates the factor loadings (βs) of a specific asset.

    ✓ It can be used to test whether alpha is statistically different from zero.

    x It estimates the risk premiums (λ) of factors across many assets.

x It directly proves that mean-return estimates are accurate.

Time-series regression estimates an asset's betas and alpha. Testing alpha $\neq 0$ is a key CAPM test. Risk premiums (λ) come from cross-sectional regressions, not time-series.

30. † **[Select All That Apply]** Which statements about multicollinearity are correct?

    ✓ Betas can become unstable across samples

    ✓ Individual t-stats can be small even if R^2 is high

    x Multicollinearity guarantees low R^2

    x Multicollinearity eliminates tracking error

    Correlated regressors inflate standard errors and make betas unstable, but overall fit can still be good.

31. You estimate the regression $r_{NVDA} = \alpha + \beta\, r_{Mkt} + \epsilon$ and find $\alpha = 2\%$ (annualized). If you go long NVDA and short β units of the market, what is the expected excess return of this hedged position?

    x β times the market risk premium.

    x The tracking error of the regression.

    x Zero, because the market exposure is hedged out.

    ✓ Approximately α (2%), the portion of mean return not explained by the market.

    The hedged position removes market exposure, leaving $\alpha + \epsilon$. The expected value is α, the unexplained mean return.

32. † **[Select All That Apply]** A PM asks if EAFE Dividend (Div) offers any advantage over EAFE Value (Val) for understanding a fund's monthly excess returns. Which study designs constitute a rigorous test of Div's incremental value over Val?

    ✓ Regress the fund on EAFE Market, Val, and Div (all investable total-return series), including an intercept, assess whether Div's partial slope is significant and stable out-of-sample

    ✓ Before the LFD, orthogonalize Div with respect to Val within the EAFE and then test if the orthogonal Div reduces Tracking Error and improves stability.

    x Omit EAFE Market and compare in-sample R^2 between {Val} and {Val, Div}. Higher R^2 for the latter proves Div's advantage.

    x Whichever of Val or Div correlates more with the fund (pairwise) is the better descriptor.

    Rigorous tests require multivariate analysis and out-of-sample validation, not simple in-sample comparisons.

33. For a replication using only tradable assets of a target using liquid factors, which estimation choice is most consistent with the goal?

    ✓ Omit the intercept as it is not investible, accept lower R^2, and control tracking error.

    x Include the intercept to maximize in-sample R^2. The intercept can be purchased as a cash sleeve.

- x Include the intercept so the replication is guaranteed to have non-zero alpha out-of-sample.
- x Either way is identical for out-of-sample tracking error.

The intercept represents untradeable alpha; omitting it ensures the replication is fully investable.

34. A team argues: "Our replication corr = 0.90 to the target, so tracking error is negligible." The best response is:

- ✓ No, high correlation measures co-movement of changes, but level drift and volatility differences still create material tracking error.
- x Yes, corr $\geq$ 0.90 implies tracking error <10% by definition.
- x No, high corr guarantees negative alpha.
- x Yes, corr $\geq$ 0.90 implies we matched all moments.

Correlation measures co-movement, not level tracking. Tracking error measures the volatility of return differences.

35. † **[Select All That Apply]** Fund A trails its benchmark by exactly 1% every year (due to fees), while Fund B averages the same 0% excess return as the benchmark but with annual deviations (sometimes +2%, sometimes −2%). Which statements are true?

- x Fund A has higher tracking error than Fund B
- ✓ Fund B has higher tracking error than Fund A
- x Both funds have the same tracking error
- ✓ Fund A has tracking error near zero (constant underperformance is not volatility)

Tracking error is the volatility of excess returns. Fund A's constant -1% has zero volatility.

36. Fund X and Fund Y are managed identically, except Fund Y uses 2× leverage on every position compared to Fund X. Fund X's information ratio is 0.5. What will be Fund Y's information ratio?

- ✓ 0.5 (IR is scale-invariant: both alpha and tracking error double)
- x 1.0
- x 0.25
- x Cannot be determined.

IR = alpha / tracking error. Leverage scales both numerator and denominator equally.

Section 4: ProShares Replication Case (HW2)

Questions 37–44: Hedge Fund Replication

Background: ProShares HDG ETF attempts to replicate the HFRI hedge fund index using liquid factors. The replication uses a rolling 60-month regression to estimate factor weights.

Reference Table: Performance Comparison (Annualized, 2012–2023)

Series	Mean	Vol	Sharpe	Skew	Kurtosis	Max DD
HFRI (Target)	5.2%	5.8%	0.62	−0.8	4.5	−12%
HDG (Replication)	4.8%	6.2%	0.52	−0.6	3.8	−14%
SPY	12.4%	15.2%	0.68	−0.5	4.2	−34%

37. The term "Alternative Beta" in the ProShares case refers to:

    - x The beta of hedge funds relative to the S&P 500.
    - x Returns from investing in alternative assets like commodities and real estate.
    - ✓ Systematic risk factor exposures that explain hedge fund returns, captured via liquid instruments.
    - x The portion of hedge fund returns unexplained by any factor model.

    Alternative beta = hedge fund factor exposures (momentum, carry, value) that can be replicated with liquid ETFs, not alpha from manager skill.

38. † **[Select All That Apply]** Why does the HDG replication use a rolling 60-month window rather than the full sample for estimating factor weights?

    - ✓ Hedge fund exposures may change over time as strategies evolve.
    - x A rolling window guarantees higher R-squared than full-sample estimation.
    - ✓ Recent data may be more relevant for predicting near-term factor exposures.
    - x Regulatory requirements mandate 60-month lookback periods for ETFs.

    Rolling windows allow the replication to adapt to changing exposures. The tradeoff is more estimation noise from fewer observations.

39. In the LFD regression for replication, why might we choose to omit the intercept (force $\alpha = 0$)?

    - x Including an intercept always reduces out-of-sample R-squared.
    - ✓ The intercept represents alpha, which is not tradeable or investable.
    - x Omitting the intercept is required when using excess returns.
    - x The intercept has no economic interpretation in factor models.

    For a tradeable replication, we can only invest in the factors. The intercept (alpha) can't be purchased, so omitting it gives investable weights.

40. A fund has a positive alpha of 0.15% per month with t-stat = 1.88. At the 5% significance level (t-critical $\approx$ 2.0), we conclude:

- x The regression is misspecified.
- x We need more factors to determine significance.
- x The fund has statistically significant alpha; the manager adds value.
- ✓ The fund's alpha is not statistically significant; it could be zero.

t-stat 1.88 < 2.0 means we cannot reject $\alpha = 0$ at 5% level. The positive alpha may be due to chance.

41. † **[Select All That Apply]** The ProShares replication showed significant tracking error in 2008 and trailed HFRI in 2012–2013. Which mechanisms could explain this?

- ✓ The 60-month rolling window creates lag in adapting to regime changes.
- ✓ Factor exposures estimated from calm periods may not hold during crises.
- x High in-sample R-squared guarantees low out-of-sample tracking error.
- x The HFRI index methodology changed in 2012.

Rolling windows react slowly to structural breaks. 2008 was a regime change; estimated weights from 2003-2007 were inappropriate for the crisis.

42. † **[Select All That Apply]** Investors sometimes intentionally "leave a basis" (use a partial hedge). Which reasons could explain an intentional partial hedge?

- ✓ The asset has non-hedgeable risk factors or is illiquid.
- ✓ Shorting the exact asset or a perfect proxy is expensive or infeasible.
- ✓ The investor wants to retain some upside exposure to the asset's unique risks.
- x Leaving basis risk increases expected returns.

Partial hedges are intentional for practical reasons (cost, liquidity, desire for exposure). Basis risk itself does not increase expected returns.

43. † **[Select All That Apply]** The HFRI index and HDG replication both show negative skewness and excess kurtosis relative to a normal distribution. Which implications follow?

- ✓ Tail losses are more severe than a normal distribution would suggest.
- ✓ Normal VaR will underestimate the true downside risk.
- x The strategies must have higher Sharpe ratios than SPY.
- x VaR based on normal assumptions will overestimate risk.

Negative skew + high kurtosis = fat left tails. Normal VaR assumes thin, symmetric tails and would underestimate true downside risk. Sharpe ratio is unrelated to higher moments.

44. Comparing hedge fund fees (2% management + 20% performance) to HDG's 0.95% expense ratio: if both strategies have gross excess return of 8% and volatility of 10%, approximately how much higher is HDG's net Sharpe ratio?

- x Same; fees don't affect Sharpe ratio.
- x About 0.10 higher.
- ✓ About 0.25–0.30 higher.

x About 0.50 higher.

Hedge fund net: $(8\% - 2\% - 0.20 \times 6\%) = 4.8\%$. HDG net: $8\% - 0.95\% = 7.05\%$. Sharpe difference $\approx (7.05 - 4.8)/10 = 0.225$. NOTE: The performance-fee convention (20% applied to returns net of management fee) is not specified in the question. Under the alternative convention (20% of gross 8%), net = 4.4% and the difference is $\approx$0.265. Either way, "about 0.25–0.30" is the closest choice.

Section 5: Value-at-Risk and Risk Measures

Questions 45–51: Ordered Returns

Reference Table: 10 daily returns (worst to best)

-5%	-4%	-3%	-2%	-1%	0%	1%	2%	3%	4%

45. The 10% empirical VaR is:

    ✓ -5%
    x -4%
    x -3%
    x -2%

    10% VaR is the worst 1 of 10 observations. NOTE: Uses order-statistic convention (worst k of n). If students used a different empirical-quantile interpolation, consider leniency.

46. The 20% empirical VaR is:

    x -5%
    ✓ -4%
    x -3%
    x -2%

    20% VaR is the 2nd worst observation. Same order-statistic convention as Q45.

47. The 20% empirical CVaR is:

    ✓ -4.5%
    x -4.0%
    x -3.5%
    x -5.0%

    Average of worst two returns: (-5-4)/2 = -4.5%.

48. † **[Select All That Apply]** A VaR backtest shows 12% exceptions at the 5% level. This implies:

    ✓ The model is underestimating risk
    ✓ Losses are heavier-tailed than assumed
    x The model is conservative
    x The model has perfect calibration

    Too many exceptions implies underestimation of tail risk.

49. † **[Select All That Apply]** If a portfolio has a return distribution with negative skewness and high kurtosis, a standard Normal VaR will likely:

- ✓ Underestimate the true risk.
- x Overestimate the true risk.
- ✓ Fail to capture the fat left tail of the distribution.
- x Produce fewer VaR exceedances than predicted if backtested.

Normal VaR assumes thin, symmetric tails. Negative skew and high kurtosis mean fatter left tails, so true VaR is worse than normal VaR predicts. Backtests would show more exceedances than the stated confidence level.

50. † **[Select All That Apply]** When converting a 1-day normal VaR to a 10-day normal VaR using the $\sqrt{10}$ rule, we are implicitly assuming:

- ✓ Returns are serially uncorrelated (zero autocorrelation).
- ✓ Returns are identically distributed across days.
- x Returns are perfectly correlated across days (+1).
- x Volatility changes predictably over time.

The $\sqrt{T}$ scaling assumes i.i.d. returns: independent (zero autocorrelation) and identically distributed. If returns were perfectly correlated, VaR would scale linearly with T, not $\sqrt{T}$.

51. † **[Select All That Apply]** Which statements about "Parametric VaR" are correct?

- ✓ It assumes returns follow a specific distribution (e.g., normal).
- ✓ It uses estimated parameters (mean, volatility) to compute the VaR analytically.
- x It requires no distributional assumptions because it uses historical data directly.
- x It is always more accurate than empirical VaR for fat-tailed distributions.

Parametric VaR assumes a known distribution (typically normal) and uses estimated parameters. Empirical VaR uses historical data directly without distributional assumptions. Parametric VaR can underestimate risk for fat-tailed distributions.

Questions 52–56: VaR Methodology

Reference Table: Portfolio Risk Statistics (Weekly Data)

Measure	Portfolio A	Portfolio B	Equal-Weight	Market
Mean (weekly)	0.20%	0.15%	0.18%	0.18%
Vol (weekly)	4.0%	3.0%	2.5%	2.0%
Skewness	−0.3	−1.5	−0.6	−0.4
Kurtosis	4.0	8.0	5.0	3.5
Empirical VaR (5%)	−6.0%	−8.0%	−4.5%	−3.0%
Normal VaR (5%)	−6.4%	−4.8%	−3.9%	−3.1%

52. † **[Select All That Apply]** For Portfolio B, the empirical VaR (−8.0%) is much worse than normal VaR (−4.8%). Which factors explain this gap?

- ✓ Portfolio B has fat tails (high kurtosis = 8).
- ✓ Portfolio B has negative skewness (-1.5), pushing the left tail further out.
- x The normal VaR calculation used the wrong confidence level.
- x Empirical VaR is always more conservative than normal VaR regardless of distribution.

High kurtosis (fat tails) and negative skewness push the 5th percentile further left than a normal distribution predicts. Empirical VaR captures the actual tail; normal VaR assumes thin, symmetric tails.

53. Autocorrelation in squared returns implies:

- x Mean reversion.
- x Unconditional Heteroskedasticity.
- x Constant volatility.
- ✓ Volatility Clustering.

Autocorrelation in squared returns means high volatility periods tend to follow high volatility periods.

54. If we were to calculate empirical VaR for a portfolio of subprime assets (rare but severe losses), we would likely:

- ✓ Underestimate risk, since extreme losses are rare but severe when they do occur.
- x Overestimate risk, since extreme losses are rare but severe when they do occur.
- x Get an accurate estimate if we use enough historical data.
- x Get an estimate equal to normal VaR.

Historical data may not contain enough tail events, causing empirical VaR to understate true risk.

55. † **[Select All That Apply]** CVaR (Conditional VaR, or Expected Shortfall) is preferred over VaR because:

- ✓ CVaR is a coherent risk measure; VaR is not.
- ✓ CVaR tells you the average loss in the tail, not just the threshold.
- x CVaR is always smaller (less negative) than VaR.
- ✓ CVaR satisfies sub-additivity: diversification cannot increase CVaR.

CVaR is coherent (sub-additive); VaR can violate sub-additivity. CVaR = E[loss — loss > VaR], always $\geq$ VaR in magnitude.

56. † **[Select All That Apply]** A VaR model with 5% threshold is backtested over 200 days. Which statements are correct?

- ✓ If the model is correct, we expect approximately 10 exceedances.
- ✓ Observing 25 exceedances would suggest the model underestimates risk.
- x If the model is correct, we expect approximately 50 exceedances.
- x Observing 5 exceedances would suggest the model underestimates risk.

Expected exceedances = $0.05 \times 200 = 10$. Significantly more (e.g., 25) suggests underestimation; significantly fewer (e.g., 5) suggests overestimation (conservative model).

Section 6: Barnstable and Long-Run Risk (HW3)

Questions 57–61: Probability of Underperformance

Background: Barnstable College Endowment believes stocks are safer in the long run and maintains 100% equity allocation. The case examines the probability that stocks will underperform the risk-free rate over various horizons.

Reference Data: Log Return Statistics (Annualized, 1965–1999)

	Mean (μ)	Volatility (σ)
Market (log returns)	8%	16%
Risk-free rate (log)	2%	–
Excess log return	6%	16%

Note: Sharpe Ratio of log excess returns = 6%/16% = 0.375

57. The probability that stocks underperform the risk-free rate over horizon h years can be written as:

$$P(\text{underperform}) = \Phi(-\sqrt{h} \times \text{SR})$$

where Φ is the standard normal CDF and SR is the Sharpe ratio. For $h = 1$ year, using SR = 0.375, the probability is approximately:

- x 25%
- x 15%
- x 50%
- ✓ 35%

NOTE: Normal CDF table was not provided on this exam. $\Phi(-0.375) \approx \Phi(-0.4) \approx 0.35$. About 35% chance stocks underperform risk-free in any given year.

58. † **[Select All That Apply]** As the investment horizon h increases from 1 to 30 years, which statements about stocks vs the risk-free rate are correct (assuming i.i.d. returns and positive Sharpe ratio)?

- ✓ The probability of underperforming the risk-free rate decreases.
- x The probability of underperforming the risk-free rate increases.
- ✓ The formula $P = \Phi(-\sqrt{h} \times \text{SR})$ drives probability toward zero as h grows.
- x The volatility of cumulative returns decreases with horizon.

With positive SR and i.i.d. returns, the argument $-\sqrt{h} \times \text{SR}$ becomes very negative as h grows, so $\Phi \to 0$. Cumulative volatility grows as $\sigma\sqrt{h}$, not shrinks.

59. The volatility of **annualized** returns over an h-year horizon (assuming i.i.d. returns) is:

- ✓ $\sigma/\sqrt{h}$

- x $\sigma \times h$
- x σ / h
- x $\sigma \times \sqrt{h}$

Cumulative vol scales as $\sigma\sqrt{h}$; dividing by h to annualize gives $\sigma\sqrt{h}/h = \sigma/\sqrt{h}$.

60. † **[Select All That Apply]** Barnstable's belief that "stocks are safer in the long run" relies on which assumptions?

- ✓ Stock returns are approximately i.i.d. (independent across time).
- ✓ The equity risk premium is positive and stable over time.
- x Stock returns are negatively autocorrelated (mean-reverting).
- x Volatility decreases as the investment horizon increases.

The derivation assumes i.i.d. returns and positive premium. Total risk still grows with horizon; only the probability of underperformance decreases.

61. † **[Select All That Apply]** Using the 1965–1999 data, Barnstable calculated a 30-year underperformance probability of 3%, but actual returns from 2000–2024 were much lower than expected. Which explanations are valid?

- ✓ Parameter uncertainty: the true mean and Sharpe ratio may differ from historical estimates.
- ✓ Regime changes: the equity premium in 2000–2024 may have been structurally different from 1965–1999.
- x The formula for underperformance probability is mathematically incorrect.
- x A 30-year horizon is always too short for any probability formula to apply.

The formula is correct given parameters, but historical estimates may not reflect true future parameters. Structural breaks (dot-com crash, 2008 crisis) can cause regime changes that invalidate historical parameter estimates.

Section 7: The Capital Asset Pricing Model

Questions 62–67: CAPM Fundamentals

Assume: $r_f = 2\%$, Market risk premium $E[\tilde{r}_m] = 6\%$, $\sigma_m = 20\%$

Asset	β	σ_ϵ^2 (Idio. Var)	Notes
X	0.5	0.04	
Y	1.2	0.09	

62. The expected excess return for Asset X is:

    ✓ 3%

    x 6%

    x 8%

    x 2%

    $0.5 \times 6\% = 3\%$.

63. The expected total return for Asset Y is:

    x 6.0%

    x 8.0%

    ✓ 9.2%

    x 10.0%

    $r_f + \beta \times \text{premium} = 2\% + 1.2 \times 6\% = 9.2\%$.

64. The total variance of Asset X is:

    x 0.04

    ✓ 0.05

    x 0.08

    x 0.10

    $\sigma_m^2 = 0.2^2 = 0.04$. So $\beta^2\sigma_m^2 + \sigma_\epsilon^2 = 0.25 \times 0.04 + 0.04 = 0.05$.

65. Which asset has higher R^2 in a CAPM regression?

    ✓ Asset Y

    x Asset X

    x Equal

    x Cannot determine

    $R^2 = \beta^2\sigma_m^2/(\beta^2\sigma_m^2 + \sigma_\epsilon^2)$. X: $0.25 \times 0.04/(0.01 + 0.04) = 0.20$. Y: $1.44 \times 0.04/(0.0576 + 0.09) \approx 0.39$. Y is higher.

66. † **[Select All That Apply]** Under CAPM, which statements are correct?

 ✓ Only systematic risk earns a risk premium
 ✓ Assets with negative beta can have expected returns below r_f
 x Idiosyncratic risk increases expected return
 x A zero-beta asset must have zero expected return

 CAPM prices only systematic risk; zero beta implies expected return $= r_f$, not zero.

67. According to the CAPM, what should we expect for the intercept (α) in a time-series regression of any portfolio's excess return on the market's excess return?

 ✓ α equals zero.
 x α equals the portfolio's mean excess return.
 x α can be any value; CAPM makes no prediction about α.
 x α equals the risk-free rate.

 CAPM says $E[\tilde{r}^i] = \beta^i E[\tilde{r}^m]$. In regression form with intercept, this means $\alpha = 0$.

Questions 68–74: CAPM Regression Output

Reference Table: Time-series CAPM regressions

Portfolio	$\hat{\beta}$	Mean Excess	$\hat{\alpha}$	α t-stat	Vol	R^2
Small-Value	1.20	14%	4.0%	2.8	24%	0.60
Small-Growth	1.35	8%	−2.0%	−1.2	28%	0.55
Big-Value	0.90	11%	2.5%	2.1	18%	0.72
Big-Growth	1.05	10%	0.5%	0.4	16%	0.82
Market	1.00	9.5%	0.0%	–	16%	1.00

68. † **[Select All That Apply]** The Small-Value portfolio has $\alpha = 4.0\%$ with t-stat $= 2.8$. Which interpretations are valid?

 x The CAPM perfectly explains Small-Value returns.
 ✓ Small-Value earns returns beyond what its market beta would predict.
 ✓ This is evidence against the CAPM or suggests missing risk factors.
 x The high t-stat means Small-Value has lower total risk than average.

 Positive, significant alpha means returns exceed CAPM prediction. This is evidence against CAPM (or for missing factors like size/value). The t-stat measures statistical significance of alpha, not total risk.

69. The correlation between Big-Growth and the market is approximately:

 x 0.67
 x Cannot be determined without additional data.

x 0.82

✓ 0.91

For a single-factor regression, $\rho = \sqrt{R^2} = \sqrt{0.82} \approx 0.91$.

70. Small-Value has $\beta = 1.20$ and the market has expected excess return of 9.5%. Under the CAPM, Small-Value's expected excess return should be:

x 9.5%

x 10.0%

✓ 11.4%

x 14.0%

CAPM: $E[\tilde{r}] = \beta \times E[\tilde{r}^m] = 1.20 \times 9.5\% = 11.4\%$. Actual is 14%, hence positive alpha.

71. † **[Select All That Apply]** Which of the following reasons highlight the importance of the CAPM time-series regression?

✓ To estimate the asset's systematic risk exposure (β_i) and intercept (α_i).

x To estimate the price of risk directly from the slope of the regression line.

✓ To test the validity of the CAPM by examining if the intercept is statistically different from zero.

x To determine the optimal risk-free rate to use for the period.

Time-series regression estimates betas and tests CAPM by checking if alphas are zero.

72. If the CAPM holds strictly true, what values should we expect for the intercept (α) and slope (β) in a time series regression using the tangency portfolio's returns?

x $\alpha = R_f$, $\beta = 1$

x $\alpha = E[R_m]$, $\beta = 0$

x $\alpha = 0$, $\beta = 0$

✓ $\alpha = 0$, $\beta = 1$

The market/tangency portfolio regressed on itself has $\beta = 1$ and $\alpha = 0$.

73. Which portfolio has the highest proportion of variance from idiosyncratic factors?

x Big-Growth

x Big-Value

✓ Small-Growth (lowest $R^2 = 0.55$)

x Small-Value

Idiosyncratic proportion $= 1 - R^2$. Small-Growth has lowest R^2 (0.55), so highest idiosyncratic share (45%).

74. If asset returns are normally distributed such that portfolios are fully characterized by mean and variance, what is the CAPM significance of the market portfolio?

- x It serves as the Global Minimum Variance (GMV) portfolio.
- x It represents the portfolio with the highest possible absolute return.
- ✓ It is the tangency portfolio, maximizing the Sharpe Ratio.
- x It lies on the inefficient portion of the frontier.

Under CAPM assumptions, the market portfolio is mean-variance efficient (tangency portfolio).

Section 8: DFA and Factor Investing (HW4)

Questions 75–79: Testing the CAPM

Background: DFA believes in factor premiums (size, value) beyond the market. The CAPM predicts that alpha should be zero if the market is the only priced factor.

75. † [**Select All That Apply**] Which methods can be used to test whether the CAPM holds?

    ✓ Check if time-series alphas are statistically different from zero.

    ✓ Run a cross-sectional regression of mean returns on betas and check if intercept is zero.

    x Verify that all assets have the same Sharpe ratio as the market.

    x Confirm that all assets have the same expected return.

    Time-series alpha tests (checking $\alpha = 0$) and cross-sectional SML tests (two-stage regression) are the main CAPM testing approaches covered.

76. The cross-sectional regression of mean excess returns on betas gives an estimated "price of risk" $\hat{\lambda} = 7\%$. The market's actual mean excess return is 9.5%. This suggests:

    x The CAPM holds perfectly.

    ✓ The estimated compensation per unit of beta is less than the market premium implies.

    x All portfolios have positive alpha.

    x The risk-free rate is incorrectly measured.

    If CAPM holds, λ should equal the market premium. $\lambda < E[\tilde{r}^m]$ suggests the SML is flatter than predicted.

77. A portfolio has a CAPM regression alpha of 3% per year with a t-statistic of 2.5. Under the null hypothesis that CAPM holds, what should we conclude?

    x The CAPM is confirmed because alpha is positive.

    ✓ The positive alpha is statistically significant evidence against the CAPM for this portfolio.

    x The t-statistic is irrelevant; only the sign of alpha matters.

    x CAPM predicts alpha should be 3%, so this confirms the model.

    CAPM predicts $\alpha = 0$. A t-stat of 2.5 (above ∼2.0) means we reject the null at 5%, providing evidence against CAPM for this portfolio.

78. A time-series CAPM regression for Stock Z yields $R^2 = 0.25$. A colleague concludes: "The CAPM fails because it only explains 25% of Stock Z's return variation." The best response is:

    x Correct; $R^2 < 0.50$ means the CAPM is rejected.

    ✓ Incorrect; CAPM is a model of expected returns, not of return variation. Low R^2 does not reject CAPM.

    x Incorrect; CAPM predicts $R^2 = 1$ for all assets.

    x Correct; low R^2 implies positive alpha.

CAPM predicts $E[r^i] = \beta^i E[r^m]$, a statement about means. R^2 measures how much variance is explained by the market factor, which can be low due to idiosyncratic risk. The CAPM test is whether $\alpha = 0$, not whether R^2 is high.

79. A researcher finds that 3 out of 25 portfolios have statistically significant alphas (t-stat > 2.0). Another researcher argues this could be due to chance. The best response is:

    - x 3 out of 25 is definitely not due to chance.
    - ✓ With 25 tests at 5% level, we expect about 1-2 false positives by chance alone.
    - x Multiple testing never creates false positives.
    - x The t-statistic should be divided by 25 to account for multiple tests.

    With 25 independent tests at $\alpha = 0.05$, expected false positives $\approx 25 \times 0.05 = 1.25$. Finding 3 significant could partly reflect this. NOTE: Coverage sensitivity—multiple-testing intuition may not have been discussed in class. Consider lenient grading if needed.

Section 9: Additional Questions

Questions 80–100: Mixed Topics

80. What is the Global Minimum Variance (GMV) portfolio?
    - x The portfolio that maximizes expected return without regard to risk
    - x The portfolio with the minimum possible return
    - ✓ The portfolio with the lowest variance among all possible portfolios
    - x The portfolio consisting solely of the risk-free asset

    The GMV portfolio minimizes variance without regard to expected return.

81. Which statistical property leads to "fat tails" in return distributions?
    - x High Mean.
    - x High Variance.
    - ✓ Excess Kurtosis.
    - x Low Variance.

    Excess kurtosis (kurtosis > 3) indicates fatter tails than a normal distribution.

82. When optimizing over **excess returns** (returns in excess of the risk-free rate), how do the optimal risky-asset weights differ from the case with no risk-free asset?
    - x The weights of risky assets must all be positive
    - ✓ The weights of risky assets no longer must sum to one
    - x The Sharpe ratio of the tangency portfolio must increase
    - x The weights are identical to the no-risk-free-asset case

    With a risk-free asset (excess return framing), investors can borrow or lend at the risk-free rate, so the risky-asset weights need not sum to one—the remainder is implicitly allocated to the risk-free asset.

83. In MVO, why is the covariance matrix often more reliably estimated than expected returns?
    - x Covariances are constant over time, whereas expected returns change.
    - ✓ Second moments (variances, covariances) converge faster with sample size than first moments (means).
    - x Expected returns require daily data, while covariances can use monthly data.
    - x Covariances are directly observable, while expected returns must be estimated.

    Estimation error in means is proportional to $\sigma/\sqrt{T}$, which converges slowly. Covariance estimates converge faster and are more stable across samples.

84. A 24-month rolling window is used to estimate factor weights. After a regime shift, the replication will most likely:
    - ✓ Adjust slowly and lag the new regime

- x Immediately match the new regime
- x Eliminate tracking error entirely
- x Become unrelated to the benchmark

Rolling windows create lagged adaptation.

85. A hedge ratio greater than 1 in a cross-hedge most likely occurs when:

- ✓ The hedge instrument is less volatile than the asset
- x The asset and hedge are perfectly negatively correlated
- x The hedge instrument is more volatile than the asset
- x The hedge instrument is risk-free

Optimal hedge ratio $h = \rho\sigma_{asset}/\sigma_{hedge}$; if σ_{hedge} is smaller, $h > 1$.

86. Fund X has produced a very high Sharpe ratio over 3 years by grinding out small positive returns nearly every month. However, its return distribution shows significantly negative skewness and high kurtosis. What is the most likely interpretation?

- x Fund X is skillful at managing risk
- ✓ Fund X may be selling tail-risk (collecting premiums with rare large losses)
- x Fund X has a misreported Sharpe ratio
- x Fund X's strategy involves momentum trading

Negative skew + high kurtosis + steady small gains is the signature of selling tail risk—strategies that collect small premiums in exchange for exposure to rare but severe losses.

87. † [**Select All That Apply**] A replication shows high correlation but poor performance in a crisis. Which explanations are plausible?

- ✓ Model weights update with a lag
- ✓ Factors fail to capture tail behavior
- x High correlation guarantees no drawdowns
- x Correlation implies identical returns

High correlation does not prevent crisis underperformance.

88. A portfolio's 5% VaR is -3%. If the portfolio is doubled in size (2x leverage), the new 5% VaR is approximately:

- x -3%
- x -4.5%
- ✓ -6%
- x -9%

VaR scales linearly with position size. Doubling the portfolio doubles the VaR: $2 \times (-3\%) = -6\%$.

89. In an EWMA volatility model, increasing λ (e.g., 0.94 to 0.97) will:

- ✓ Increase the weight on older observations
- x Increase the weight on the most recent observation
- x Eliminate volatility clustering
- x Make volatility constant

Higher λ means slower decay, more weight on older data.

90. A portfolio has empirical 5% VaR = -2% and empirical 5% CVaR = -6%. This indicates:

- ✓ Severe tail losses relative to the VaR threshold
- x Near-normal tails
- x Low tail risk
- x Empirical VaR is overstated

Large gap between VaR (-2%) and CVaR (-6%) indicates fat tails—losses beyond VaR are much worse than VaR itself.

91. A portfolio manager argues: "The historical Sharpe ratio is 0.40, so over a 25-year horizon, the probability of underperforming the risk-free rate is negligible." What is the approximate probability using the formula $P = \Phi(-\sqrt{h} \times \text{SR})$?

- x 15%
- x 5%
- ✓ 2%
- x 0.5%

NOTE: Normal CDF table was not provided on this exam. $P = \Phi(-\sqrt{25} \times 0.40) = \Phi(-5 \times 0.40) = \Phi(-2.0) \approx 0.023 \approx 2\%$.

92. The volatility of **cumulative** returns over an h-year horizon (assuming i.i.d. returns) is:

- ✓ $\sigma \times \sqrt{h}$
- x $\sigma / \sqrt{h}$
- x $\sigma \times h$
- x σ / h

Under i.i.d., variance of sum = sum of variances. For h periods: $\text{Var} = h\sigma^2$, so $\text{SD} = \sigma\sqrt{h}$.

93. A portfolio has Sharpe ratio of 0.5 and Treynor ratio of 1.2. What can we infer?

- x The portfolio has negative alpha.
- ✓ The portfolio has substantial idiosyncratic (non-market) risk.
- x The portfolio's beta must be greater than 1.
- x The portfolio is perfectly correlated with the market.

Sharpe = excess return/total vol; Treynor = excess return/beta. Treynor/Sharpe = $\sigma_{\text{total}}/\beta$. Here $1.2/0.5 = 2.4$, meaning total vol is $2.4\times$ beta, implying substantial idiosyncratic volatility beyond systematic risk.

94. If a fund has a high R^2 to the S&P 500, but a low beta, the CAPM would suggest that:

- x The fund should have a higher expected return than the market.
- ✓ The fund should have a lower expected return than the market.
- x The fund should have the same expected return as the market.
- x The fund's expected return cannot be determined from this information.

Under CAPM, E[r] = rf + β(E[rm] - rf). Low beta implies lower expected return.

95. † [**Select All That Apply**] What should a CIO infer from 2013 performance (HDG 4.4% @ 4.8% vol vs S&P 32.4% @ 11% vol) for an alternatives allocation decision?

- ✓ HDG's "underperformance" in a bull year is irrelevant. Target role is diversification and drawdown mitigation, not beating pure equity beta in booms.
- ✓ The relevant yardstick is HFRI / hedge-fund beta behavior. Replication is about risk shape and regime robustness.
- ✓ Adding a sleeve like HDG can shift the portfolio efficient frontier even if it lags equities in a one-year bull.
- x Because 2013 was a bull, the only rational inference is that replication fails as an asset class.

Alternative allocations serve diversification purposes; single-year comparisons to equities are misleading.

96. What is the systematic volatility for an asset with $\beta = 0.6$ when market volatility is 15%?

- x 3%
- x 6%
- ✓ 9%
- x 15%

Systematic vol $= \beta \times \sigma_m = 0.6 \times 15\% = 9\%$.

97. Equity return is normal with mean 6% and vol 10%. A bond returns a constant 6%. What is the approximate probability that equity underperforms bonds in one year?

- x 10%
- x 25%
- ✓ 50%
- x 75%

Equity mean equals bond return, so underperformance is 50% under symmetry.

98. An investor uses the formula $P(\text{underperform}) = \Phi(-\sqrt{h} \times \text{SR})$ to calculate a 2% probability of stocks underperforming bonds over 25 years, based on a historical Sharpe ratio of 0.40. A colleague warns that the true Sharpe ratio might be 0.30. If the colleague is right, the actual probability of underperformance is approximately:

- x Still 2% (the formula is robust to parameter uncertainty)

- x 5%
- ✓ 7%
- x 15%

NOTE: Normal CDF table was not provided on this exam. With SR = 0.30: $P = \Phi(-\sqrt{25} \times 0.30) = \Phi(-5 \times 0.30) = \Phi(-1.5) \approx 0.07 = 7\%$. The probability more than triples when the true Sharpe ratio is lower than estimated. This illustrates why parameter uncertainty is a critical limitation of the "stocks are safer in the long run" argument.

99. † **[Select All That Apply]** In a two-stage CAPM test, the first stage estimates each asset's beta from time-series regressions. The second stage regresses average excess returns on betas across assets. If the CAPM holds, which statements about the second-stage regression are correct?

- ✓ The intercept should be zero (or close to zero).
- ✓ The slope should equal the market's average excess return.
- x The intercept should equal the risk-free rate.
- x The slope should equal zero if all assets are fairly priced.

Under CAPM: $E[\tilde{r}^i] = \beta^i \times E[\tilde{r}^m]$. The cross-sectional regression of mean excess returns on betas should have intercept = 0 and slope = market premium. If the intercept is positive, low-beta assets earn more than CAPM predicts (the "low-beta anomaly"). The risk-free rate is already subtracted when using excess returns.

100. A portfolio manager uses factor regression to decompose a fund's returns: $r_{fund} = \alpha + 0.8\, r_{Mkt} + 0.3\, r_{Bond} + \epsilon$. The fund has total volatility of 15%. If she hedges out both the market and bond exposures using short positions, what risk remains?

- x Zero risk—the hedge eliminates all volatility.
- x Only market risk remains.
- ✓ Idiosyncratic risk (the volatility of ϵ) remains.
- x The full 15% volatility remains unchanged.

Hedging removes systematic (factor) exposure but leaves idiosyncratic risk. The hedged position earns $\alpha + \epsilon$, with volatility equal to σ_ϵ. This is why factor hedging reduces but does not eliminate risk, and why the hedged position's expected return is α (the unexplained mean return).