Factor risk premia Implied returns Black-Litterman Stochastic discount factor Hansen-Jagannathan bounds Fama-MacBeth Factor timing Asset allocation

Executive Summary

What this primer provides:

- Mathematical framework for extracting factor risk premia from market prices
- Comparison of realized vs. implied premia with guidance on when to use each
- Complete Python/R code for estimation and backtesting
- Practical applications: factor timing, risk budgeting, performance attribution

Key formulas (quick reference):

Implied returns:	$\boldsymbol{\mu}^{\text{impl}} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}}$
Implied factor premia:	$\boldsymbol{\lambda}^{\text{impl}} = \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f$
Hansen-Jagannathan bound:	$\sigma(m)/\,\mathbb{E}[m] \geq \sqrt{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}$
Fama-MacBeth premium:	$\hat{\lambda} = (1/T) \sum_t \hat{\gamma}_t$
Shanken-corrected SE:	$\text{SE}_{\text{Sh}} = \text{SE}_{\text{FM}} \sqrt{1 + \lambda^\top \Omega_f^{-1} \lambda}$

Typical estimates (Fama-French factors, 1963-2023):

Factor	*Realized*	*Implied*	*Sharpe Ratio*
MKT	7.9%	6.8%	0.51
SMB	2.2%	1.4%	0.20
HML	3.5%	2.0%	0.36
RMW	2.9%	2.5%	0.37
CMA	2.6%	2.1%	0.39

When to use implied premia: After regime changes, for forward-looking decisions, when historical samples are short or non-representative.

When to use realized premia: For backtesting, when markets may be mispriced, for contrarian strategies.

Implied Risk Premia for Factors:
A Unified Framework from Theory to Implementation

Joerg Osterrieder[a,*]

[a] *Fachhochschule Graubuenden, Institute for Financial Services Zug (IFZ), Chur, 7000, Switzerland*

Abstract

This primer provides a comprehensive treatment of implied risk premia for systematic factors, bridging the gap between theoretical asset pricing and practical investment applications. We develop a unified mathematical framework that connects the Arbitrage Pricing Theory (APT) with modern factor investing, emphasizing the distinction between *realized* premia estimated from historical returns and *implied* premia extracted from current market prices and equilibrium conditions.

The core contribution is a rigorous derivation of implied factor risk premia through three complementary approaches: (i) reverse optimization from observed market portfolios, (ii) stochastic discount factor (SDF) decomposition, and (iii) cross-sectional pricing restrictions. We establish the theoretical conditions under which implied premia provide superior forward-looking estimates compared to historical averages, particularly during regime changes when past data becomes uninformative.

We provide complete mathematical derivations for the canonical factor models—including CAPM, Fama-French three-factor and five-factor models, Carhart momentum model, and statistical factor models from principal component analysis. The Hansen-Jagannathan bounds are derived to characterize the set of admissible stochastic discount factors, providing diagnostic tools for model misspecification.

Empirical implementation is demonstrated using equity factor data, with detailed Python and R code for Fama-MacBeth regressions, generalized method of moments (GMM) estimation, and rolling window analysis. We document substantial time-variation in factor premia and show how implied premia can be used for dynamic factor allocation. The primer concludes with practical applications including factor timing strategies, risk budgeting, and performance attribution, along with extensive exercises suitable for doctoral-level coursework.

Keywords: factor risk premia, implied returns, arbitrage pricing theory, Fama-French factors, stochastic discount factor, Hansen-Jagannathan bounds, reverse optimization, factor investing

[*] Corresponding author

Notation

Table 1: Mathematical notation used throughout this paper

Symbol	Description	Dimension
General Notation		
$\mathbb{R}$	Real numbers	–
$\mathbb{R}^n$	n-dimensional Euclidean space	–
$\mathbb{E}[\cdot]$	Expectation operator	–
$\mathbb{E}_t[\cdot]$	Conditional expectation at time t	–
$\text{Var}(\cdot)$	Variance operator	–
$\text{Cov}(\cdot, \cdot)$	Covariance operator	–
$\|\cdot\|$	Euclidean norm	–
$\|\cdot\|_F$	Frobenius norm	–
Vectors and Matrices		
$\boldsymbol{x}$	Column vector (bold lowercase)	$n \times 1$
$\boldsymbol{A}$	Matrix (bold uppercase)	$m \times n$
$\boldsymbol{x}^\top$	Transpose of vector $\boldsymbol{x}$	$1 \times n$
$\boldsymbol{A}^{-1}$	Inverse of matrix $\boldsymbol{A}$	$n \times n$
$\mathbf{1}$	Vector of ones	$n \times 1$
$\boldsymbol{I}$	Identity matrix	$n \times n$
$\text{diag}(\cdot)$	Diagonal matrix or diagonal elements	varies
$\text{tr}(\cdot)$	Matrix trace	scalar
Asset Pricing		
n	Number of assets/test portfolios	scalar
K	Number of factors	scalar
T	Number of time periods	scalar
R_i	Return on asset i	scalar
R_m	Return on market portfolio	scalar
$\boldsymbol{R}$	Vector of asset returns	$n \times 1$
r_f	Risk-free rate	scalar
$\boldsymbol{\mu}$	Expected returns	$n \times 1$
$\boldsymbol{\Sigma}$	Asset covariance matrix	$n \times n$
γ	Risk aversion parameter	scalar

Table 2: Mathematical notation (continued)

Symbol	**Description**	**Dimension**
Factor Model		
$\boldsymbol{f}$	Factor returns	$K \times 1$
f_k	Return on factor k	scalar
$\boldsymbol{B}$	Factor loading matrix	$n \times K$
$\boldsymbol{\beta}_i$	Factor loadings for asset i	$K \times 1$
β_{ik}	Loading of asset i on factor k	scalar
$\boldsymbol{\Omega}_f$	Factor covariance matrix	$K \times K$
$\boldsymbol{\Psi}$	Idiosyncratic variance matrix (diagonal)	$n \times n$
$\boldsymbol{\epsilon}$	Idiosyncratic returns	$n \times 1$
α_i	Pricing error (alpha) for asset i	scalar
Risk Premia		
λ_k	Risk premium for factor k	scalar
$\boldsymbol{\lambda}$	Vector of factor risk premia	$K \times 1$
λ^{impl}	Implied risk premium	scalar
λ^{real}	Realized risk premium	scalar
λ^{BL}	Black-Litterman posterior premium	scalar
$\bar{\lambda}$	Historical average premium	scalar
γ_t	Cross-sectional regression coefficient at t	$K \times 1$
$\bar{\gamma}$	Fama-MacBeth average	$K \times 1$
Stochastic Discount Factor		
m	Stochastic discount factor (SDF)	scalar
m^*	Minimum variance SDF	scalar
$\boldsymbol{b}$	SDF factor loadings	$K \times 1$
δ	Hansen-Jagannathan distance	scalar
SR	Sharpe ratio	scalar
Portfolio Optimization		
$\boldsymbol{w}$	Portfolio weights	$n \times 1$
$\boldsymbol{w}^{\text{mkt}}$	Market portfolio weights	$n \times 1$
$\boldsymbol{w}^f$	Implied factor portfolio weights	$K \times 1$
$\boldsymbol{\mu}^{\text{impl}}$	Implied expected returns	$n \times 1$
σ_p	Portfolio volatility	scalar

Table 3: Mathematical notation (continued)

Symbol	Description	Dimension
Black-Litterman Model		
$\boldsymbol{\pi}$	Equilibrium expected returns (prior)	$n \times 1$
$\boldsymbol{\mu}_{\text{BL}}$	Posterior expected returns	$n \times 1$
$\boldsymbol{P}$	Pick matrix (views)	$k \times K$
$\boldsymbol{q}$	View values	$k \times 1$
$\boldsymbol{\Omega}$	View uncertainty matrix	$k \times k$
τ	Uncertainty scaling parameter	scalar
$\boldsymbol{M}$	Posterior covariance of mean	$K \times K$
Estimation		
$\hat{\theta}$	Estimated parameter	varies
$\text{SE}(\hat{\theta})$	Standard error of estimate	scalar
t-stat	t-statistic	scalar
R^2_{CS}	Cross-sectional R-squared	scalar
R^2_{TS}	Time-series R-squared	scalar
$\boldsymbol{V}_{\text{FM}}$	Fama-MacBeth variance	$K \times K$
c	Shanken correction factor	scalar
Risk Decomposition		
FRC_k	Factor risk contribution	scalar
MRC_i	Marginal risk contribution	scalar
TRC_i	Total risk contribution	scalar
$\text{SR}_k^{\text{impl}}$	Implied factor Sharpe ratio	scalar
Statistical Tests		
GRS	Gibbons-Ross-Shanken test statistic	scalar
J-stat	Hansen's J-statistic (GMM)	scalar
$F_{m,n}$	F-distribution with m, n d.f.	–
χ^2_k	Chi-squared distribution with k d.f.	–
Timing and Attribution		
$z_{k,t}$	Timing signal for factor k at time t	scalar
κ	Signal responsiveness parameter	scalar
λ^{bench}	Benchmark risk premium	scalar

Conventions..

- Scalars are denoted by lowercase italic letters (x, λ, σ)
- Vectors are denoted by bold lowercase letters ($\boldsymbol{x}$, $\boldsymbol{\mu}$, $\boldsymbol{\lambda}$)
- Matrices are denoted by bold uppercase letters ($\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{\Sigma}$)
- Random variables use the same convention as their realizations
- Estimated quantities are denoted with hats ($\hat{\lambda}$, $\hat{\beta}$)
- Optimal values are denoted with asterisks ($\boldsymbol{w}^*$)

- Implied quantities use superscript “impl” (λ^{impl})
- Subscripts indicate indexing (λ_k, β_{ik}) or type ($\boldsymbol{\Omega}_f$)
- The market factor is denoted MKT, size factor SMB, value factor HML
- Profitability factor RMW, investment factor CMA, momentum factor UMD

1. Introduction

The estimation of expected returns stands as perhaps the most fundamental challenge in financial economics. While the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) provided an elegant theoretical answer—expected excess returns are proportional to market beta—subsequent empirical work has consistently documented patterns in average returns that cannot be explained by market risk alone. Size (Banz, 1981), value (Fama & French, 1992), momentum (Jegadeesh & Titman, 1993), and numerous other characteristics predict future returns, leading to an extensive literature on **factor risk premia**.

Yet practitioners face a critical methodological choice: should expected factor premia be estimated from historical average returns, or should they be *implied* from current market prices under equilibrium assumptions? This primer develops a unified framework for understanding both approaches, with particular emphasis on implied risk premia—the forward-looking compensation for systematic risk embedded in current asset prices.

1.1. The Expected Return Puzzle

Estimating expected returns is notoriously difficult. Consider the basic statistical challenge: if annual equity returns have a standard deviation of approximately 20%, then the standard error of the sample mean over T years is $20\%/\sqrt{T}$. Even with 50 years of data, the standard error exceeds 2.8%, rendering precise estimation of the equity risk premium (typically estimated around 5-7%) statistically challenging (Merton, 1980).

The problem intensifies for individual factors. Value and size premia, while economically significant in long historical samples, exhibit substantial time-variation and have experienced extended periods of underperformance. The momentum factor, despite strong average returns, is subject to occasional severe crashes (Daniel et al., 2020). Factor premia that appeared robust in-sample have failed to persist out-of-sample, raising fundamental questions about whether they reflect compensation for systematic risk or statistical artifacts of data mining (Harvey et al., 2016).

Definition 1.1 (Factor Risk Premium)**.** Let f_k denote the return on factor k. The **factor risk premium** λ_k is the expected excess return:

$$\lambda_k = \mathbb{E}[f_k - r_f] \tag{1}$$

where r_f is the risk-free rate. Under no-arbitrage, a positive risk premium implies that assets with higher exposure to factor k (higher factor loadings β_k) must offer higher expected returns as compensation for bearing systematic risk.

The key insight of this primer is that factor premia can be estimated in two fundamentally different ways:

1. **Realized (Historical) Premia**: The sample average of factor returns over some historical period. This approach is backward-looking and subject to estimation error, regime changes, and sample selection bias.

2. **Implied (Forward-Looking) Premia**: The risk compensation embedded in current market prices, extracted through equilibrium conditions, reverse optimization, or cross-sectional pricing restrictions. This approach is forward-looking but requires assumptions about investor preferences and market efficiency.

Neither approach dominates the other in all circumstances. Historical premia may be preferable when factor return processes are stationary and samples are long. Implied premia may be preferable during regime changes when past returns provide limited information about future expectations, or when extracting the market's current view is more relevant than a long-run average.

1.2. Factor Models: A Brief Overview

Factor models decompose asset returns into systematic components (driven by common factors) and idiosyncratic components (asset-specific). This decomposition serves multiple purposes: explaining the cross-section of expected returns, reducing dimensionality for covariance estimation, and identifying sources of portfolio risk.

Single-Factor Models.. The CAPM posits that the market portfolio is the single priced factor:

$$\mathbb{E}[R_i] = r_f + \beta_i^{\text{mkt}} \lambda_{\text{mkt}} \tag{2}$$

where $\beta_i^{\text{mkt}} = \text{Cov}(R_i, R_{\text{mkt}}) / \text{Var}(R_{\text{mkt}})$ is the market beta and $\lambda_{\text{mkt}} = \mathbb{E}[R_{\text{mkt}} - r_f]$ is the market risk premium. Under CAPM, the market premium is the sole factor premium, and all cross-sectional variation in expected returns is explained by variation in market betas.

Multi-Factor Models.. Empirical failures of the CAPM led to multi-factor extensions. The Fama-French three-factor model (Fama & French, 1993) adds size (SMB: small minus big) and value (HML: high minus low book-to-market) factors:

$$\mathbb{E}[R_i] = r_f + \beta_i^{\text{mkt}} \lambda_{\text{mkt}} + \beta_i^{\text{SMB}} \lambda_{\text{SMB}} + \beta_i^{\text{HML}} \lambda_{\text{HML}} \tag{3}$$

The Carhart four-factor model (Carhart, 1997) adds momentum (UMD: up minus down), while the Fama-French five-factor model (Fama & French, 2015) adds profitability (RMW: robust minus weak) and investment (CMA: conservative minus aggressive).

Statistical Factor Models.. An alternative approach extracts factors from the return covariance matrix using principal component analysis (PCA). Statistical factors maximize variance explained without requiring economic interpretation. While the first principal component typically corresponds to market returns, subsequent components may capture sector, style, or other systematic effects.

1.3. Literature Review

We organize the literature around four themes central to this primer.

Arbitrage Pricing Theory.. Ross (1976) developed the Arbitrage Pricing Theory (APT), which establishes that if returns follow a linear factor structure and arbitrage opportunities are absent, then expected returns must be linear in factor loadings:

$$\mathbb{E}[R_i] = r_f + \boldsymbol{\beta}_i^\top \boldsymbol{\lambda} \tag{4}$$

where $\boldsymbol{\beta}_i$ is the vector of factor loadings and $\boldsymbol{\lambda}$ the vector of factor risk premia. Unlike CAPM, APT does not specify which factors are priced, leaving this as an empirical question. Cochrane (2005) provides a comprehensive treatment of factor models within the stochastic discount factor framework.

Empirical Factor Research.. The Fama-French research program has profoundly shaped empirical asset pricing. Fama and French (1992) documented that market beta alone cannot explain cross-sectional return variation, with size and book-to-market ratio providing additional explanatory power. Fama and French (1993) proposed factor-mimicking portfolios (SMB and HML) to proxy for size and value risk. The five-factor extension (Fama & French, 2015) addresses profitability and investment patterns first documented by Novy-Marx (2013) and Titman et al. (2004).

The "factor zoo" problem—the proliferation of hundreds of proposed factors (Harvey et al., 2016)—raises concerns about multiple testing and data mining. McLean and Pontiff (2016) find that factor returns decline after publication, suggesting that historical premia may overstate expected future premia.

Estimation Methods.. Several econometric approaches estimate factor risk premia. Fama and MacBeth (1973) introduced the two-pass regression methodology: first estimate factor loadings from time-series regressions, then estimate risk premia from cross-sectional regressions of average returns on estimated betas. Shanken (1992) derived corrections for errors-in-variables bias in this approach.

Generalized method of moments (GMM) provides a unified framework for estimating and testing factor models (Hansen, 1982). Cochrane (2001) shows how asset pricing restrictions can be expressed as moment conditions, enabling joint estimation of betas and premia with appropriate standard errors.

Implied Returns and Reverse Optimization.. The Black-Litterman model (Black & Litterman, 1992) introduced implied returns to portfolio management: given equilibrium portfolio weights and a covariance matrix, reverse optimization yields the expected returns that would make observed weights optimal. He and Litterman (1999) provided intuition for this approach, showing how equilibrium returns serve as a stable prior that can be combined with investor views.

Best and Grauer (1991) demonstrated the extreme sensitivity of mean-variance portfolios to expected return inputs, motivating the use of equilibrium-based rather than historically estimated returns. Meucci (2005) extended the framework to general views and non-normal distributions.

Stochastic Discount Factor and Hansen-Jagannathan Bounds.. The stochastic discount factor (SDF) provides a unified pricing framework. If m is an SDF, then for any asset with return R_i:

$$\mathbb{E}[mR_i] = 1 \tag{5}$$

Hansen and Jagannathan (1991) derived bounds on the volatility of admissible SDFs:

$$\frac{\sigma(m)}{\mathbb{E}[m]} \geq \frac{|\,\mathbb{E}[R_i - r_f]|}{\sigma(R_i)} \tag{6}$$

with the lower bound achieved by the maximum Sharpe ratio portfolio. These bounds provide a diagnostic for factor model adequacy: a proposed SDF must lie within the Hansen-Jagannathan region.

1.4. Contributions

This primer makes the following contributions:

1. **Unified Framework**: We develop a comprehensive treatment connecting APT, factor models, the SDF, and implied risk premia within a single coherent framework. The mathematical foundations are presented with full rigor, suitable for doctoral-level study.
2. **Implied vs. Realized Premia**: We provide systematic comparison of implied and realized factor premia, deriving conditions under which each approach is preferred. We show how implied premia can be extracted through reverse optimization, SDF decomposition, and cross-sectional restrictions.
3. **Complete Derivations**: Mathematical proofs and derivations are provided for all key results, including the APT pricing restriction, Fama-MacBeth standard errors, GMM estimation, Hansen-Jagannathan bounds, and Black-Litterman posterior.
4. **Empirical Implementation**: We provide complete Python and R code for estimating factor premia using Fama-MacBeth regressions, GMM, rolling windows, and implied methods. Empirical analysis documents time-variation in factor premia and compares estimation approaches.
5. **Practical Applications**: We demonstrate how implied factor premia can be used for factor timing, strategic asset allocation, risk budgeting, and performance attribution.

1.5. Paper Organization

The remainder of this primer is organized as follows. Section 2 develops the theoretical foundations of factor models, from CAPM through APT and the stochastic discount factor framework. Section 3 presents estimation methods for factor risk premia, including Fama-MacBeth regressions, GMM, and time-series approaches. Section 4 develops the core framework for implied risk premia, including reverse optimization, cross-sectional restrictions, and Hansen-Jagannathan bounds. Section 5 provides empirical analysis using equity factor data, documenting time-variation and comparing estimation approaches. Section 6 presents applications to factor investing, including timing strategies, risk budgeting, and performance attribution.

Four appendices provide supporting material: Section Appendix A contains mathematical proofs of key theorems. Section Appendix B provides extended derivations and technical details. Section Appendix C catalogs factor data sources and construction. Section Appendix D offers exercises and research questions for readers.

2. Factor Model Theory

This section develops the theoretical foundations of factor pricing, progressing from the Capital Asset Pricing Model (CAPM) through the Arbitrage Pricing Theory (APT) to the unifying stochastic discount factor framework. We establish the mathematical conditions under which systematic factors carry risk premia and derive the key pricing restrictions that enable empirical estimation.

2.1. Portfolio Choice and Mean-Variance Analysis

We begin with the investor's portfolio choice problem, which forms the basis for equilibrium asset pricing. Consider an economy with n risky assets and a risk-free asset with return r_f. Let $\boldsymbol{R} = (R_1, \ldots, R_n)^\top$ denote excess returns (in excess of r_f), with expected value $\boldsymbol{\mu} = \mathbb{E}[\boldsymbol{R}]$ and covariance matrix $\boldsymbol{\Sigma} = \mathrm{Cov}(\boldsymbol{R})$.

Definition 2.1 (Mean-Variance Efficient Portfolio)**.** A portfolio with weights $\boldsymbol{w}$ on risky assets is **mean-variance efficient** if it maximizes expected return for a given level of variance, or equivalently, minimizes variance for a given expected return.

The set of efficient portfolios traces out the **efficient frontier**. For an investor with mean-variance preferences characterized by risk aversion $\gamma > 0$, the optimization problem is:

$$\max_{\boldsymbol{w}} \left\{ \boldsymbol{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} \right\} \tag{7}$$

Theorem 2.2 (Optimal Mean-Variance Portfolio)**.** *The solution to* (7) *is:*

$$\boldsymbol{w}^* = \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \tag{8}$$

The tangency portfolio (maximum Sharpe ratio) is proportional to $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$.

Proof. The first-order condition $\partial/\partial \boldsymbol{w}[\boldsymbol{w}^\top \boldsymbol{\mu} - (\gamma/2)\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}] = \boldsymbol{0}$ yields $\boldsymbol{\mu} - \gamma \boldsymbol{\Sigma} \boldsymbol{w} = \boldsymbol{0}$. Solving gives $\boldsymbol{w}^* = (1/\gamma)\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. □

The tangency portfolio plays a central role in asset pricing: under CAPM, it corresponds to the market portfolio.

2.2. Capital Asset Pricing Model

The CAPM, developed independently by Sharpe (1964), Lintner (1965), and Mossin (1966), derives equilibrium expected returns under the assumption that all investors hold mean-variance efficient portfolios.

Assumption 2.3 (CAPM Assumptions)**.**
1. All investors are mean-variance optimizers with the same investment horizon.
2. All investors have homogeneous expectations about means and covariances.
3. There exists a risk-free asset at rate r_f with unlimited borrowing/lending.
4. Markets are frictionless (no taxes, transaction costs, or short-sale constraints).
5. All assets are infinitely divisible and marketable.

Under these assumptions, all investors hold the same risky portfolio—the tangency portfolio—combined with the risk-free asset. Market clearing requires the tangency portfolio to be the market portfolio $\boldsymbol{w}^{\text{mkt}}$.

Theorem 2.4 (CAPM Pricing Relation)**.** *Under Theorem 2.3, expected excess returns satisfy:*

$$\mathbb{E}[R_i] = \beta_i^{mkt} \lambda_{mkt} \tag{9}$$

where $\beta_i^{mkt} = \operatorname{Cov}(R_i, R_{mkt}) / \operatorname{Var}(R_{mkt})$ *is the market beta and* $\lambda_{mkt} = \mathbb{E}[R_{mkt}]$ *is the market risk premium.*

Proof. From Theorem 2.2, the market portfolio satisfies $\boldsymbol{\mu} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}}$. The i-th element is:

$$\mu_i = \gamma \sum_{j=1}^{n} \sigma_{ij} w_j^{\text{mkt}} = \gamma \operatorname{Cov}(R_i, R_{\text{mkt}})$$

where $R_{\text{mkt}} = \sum_j w_j^{\text{mkt}} R_j$. For the market portfolio itself:

$$\mu_{\text{mkt}} = \gamma \operatorname{Var}(R_{\text{mkt}})$$

Dividing yields $\mu_i / \mu_{\text{mkt}} = \operatorname{Cov}(R_i, R_{\text{mkt}}) / \operatorname{Var}(R_{\text{mkt}}) = \beta_i^{\text{mkt}}$, hence $\mu_i = \beta_i^{\text{mkt}} \mu_{\text{mkt}}$. □

Corollary 2.5 (Market Risk Premium)**.** *The market risk premium equals the representative investor's risk aversion times market variance:*

$$\lambda_{mkt} = \gamma \operatorname{Var}(R_{mkt}) \tag{10}$$

The CAPM implies that market beta is the sole determinant of expected returns. Assets that covary more strongly with the market must offer higher expected returns as compensation for their systematic risk.

2.3. Arbitrage Pricing Theory

The Arbitrage Pricing Theory of Ross (1976) generalizes CAPM by allowing multiple sources of systematic risk. Rather than deriving equilibrium from investor preferences, APT uses no-arbitrage arguments.

Assumption 2.6 (APT Assumptions)**.**

1. Returns follow a linear K-factor structure:

$$R_i = \mathbb{E}[R_i] + \sum_{k=1}^{K} \beta_{ik} f_k + \epsilon_i \tag{11}$$

   where f_k are mean-zero common factors, β_{ik} are factor loadings, and ϵ_i is idiosyncratic risk with $\mathbb{E}[\epsilon_i] = 0$, $\mathrm{Cov}(\epsilon_i, f_k) = 0$, and $\mathrm{Cov}(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$.
2. Arbitrage opportunities are absent.
3. The number of assets n is large (asymptotic argument).

Definition 2.7 (Arbitrage Portfolio)**.** An **arbitrage portfolio** $\boldsymbol{w}$ satisfies: (i) $\boldsymbol{w}^\top \mathbf{1} = 0$ (zero investment), (ii) $\boldsymbol{w}^\top \boldsymbol{\epsilon} \approx 0$ (diversified idiosyncratic risk), (iii) $\boldsymbol{w}^\top \boldsymbol{\beta}_k = 0$ for all k (zero factor exposure), and (iv) $\boldsymbol{w}^\top \mathbb{E}[\boldsymbol{R}] > 0$ (positive expected return).

Theorem 2.8 (APT Pricing Relation)**.** *Under Theorem 2.6, expected excess returns are approximately linear in factor loadings:*

$$\mathbb{E}[R_i] = \sum_{k=1}^{K} \beta_{ik} \lambda_k \tag{12}$$

where λ_k is the risk premium for factor k.

Proof. See Section Appendix A.2 for the formal derivation. The key insight is that arbitrage portfolios can eliminate all systematic risk. If (12) fails to hold exactly, one can construct arbitrage opportunities, contradicting Theorem 2.6(2). □

Remark 2.9 (APT vs. CAPM). CAPM is a special case of APT with $K = 1$ (market factor only). APT is more general: it does not require market equilibrium or mean-variance preferences, only no-arbitrage. However, APT does not specify which factors are priced—this is an empirical question.

In matrix notation, letting $\boldsymbol{B}$ denote the $n \times K$ matrix of factor loadings and $\boldsymbol{\lambda}$ the $K \times 1$ vector of risk premia:

$$\mathbb{E}[\boldsymbol{R}] = \boldsymbol{B}\boldsymbol{\lambda} \tag{13}$$

2.4. Canonical Multi-Factor Models

Empirical research has identified several factors that help explain the cross-section of stock returns. We present the major models chronologically.

Fama-French Three-Factor Model.. Fama and French (1993) augmented the market factor with size and value:

$$R_i - r_f = \alpha_i + \beta_i^{\text{mkt}}(R_{\text{mkt}} - r_f) + \beta_i^{\text{SMB}} R_{\text{SMB}} + \beta_i^{\text{HML}} R_{\text{HML}} + \epsilon_i \tag{14}$$

where:

- R_{SMB} (Small Minus Big): Return spread between small-cap and large-cap portfolios
- R_{HML} (High Minus Low): Return spread between high and low book-to-market portfolios

Carhart Four-Factor Model.. Carhart (1997) added momentum:

$$R_i - r_f = \alpha_i + \beta_i^{\text{mkt}} R^e_{\text{mkt}} + \beta_i^{\text{SMB}} R_{\text{SMB}} + \beta_i^{\text{HML}} R_{\text{HML}} + \beta_i^{\text{UMD}} R_{\text{UMD}} + \epsilon_i \tag{15}$$

where R_{UMD} (Up Minus Down) is the return spread between past winners and past losers.

Fama-French Five-Factor Model.. Fama and French (2015) added profitability and investment:

$$R_i - r_f = \alpha_i + \beta_i^{\text{mkt}} R^e_{\text{mkt}} + \beta_i^{\text{SMB}} R_{\text{SMB}} + \beta_i^{\text{HML}} R_{\text{HML}} + \beta_i^{\text{RMW}} R_{\text{RMW}} + \beta_i^{\text{CMA}} R_{\text{CMA}} + \epsilon_i \tag{16}$$

where:

- R_{RMW} (Robust Minus Weak): Profitability factor
- R_{CMA} (Conservative Minus Aggressive): Investment factor

Table 4: Historical Factor Risk Premia (1963-2023, annualized)

Factor	Mean (%)	Std (%)	t-statistic	Sharpe Ratio
MKT-RF	7.91	15.45	3.98	0.51
SMB	2.14	11.12	1.49	0.19
HML	3.47	11.67	2.31	0.30
UMD	6.28	14.89	3.27	0.42
RMW	2.89	8.14	2.76	0.36
CMA	2.58	7.23	2.77	0.36

2.5. Stochastic Discount Factor Framework

The stochastic discount factor (SDF) provides a unifying framework for asset pricing that nests both CAPM and APT as special cases.

Definition 2.10 (Stochastic Discount Factor)**.** A **stochastic discount factor** is a random variable m satisfying:

$$\mathbb{E}[mR_i] = 1 \quad \text{for all assets } i \tag{17}$$

Equivalently, for excess returns: $\mathbb{E}[mR_i^e] = 0$.

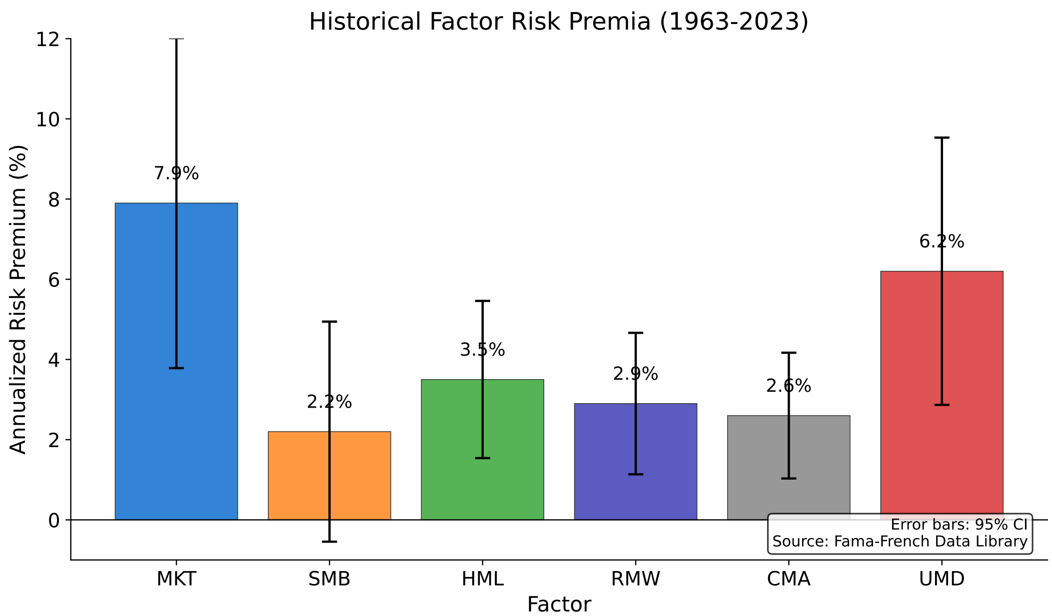

Figure 1: Historical average factor risk premia (annualized, 1963-2023). The market factor (MKT) commands the highest premium, followed by momentum (UMD). Size (SMB) and value (HML) premia have declined in recent decades.

The SDF represents the marginal utility growth of a representative investor. Assets that pay off when marginal utility is high (bad states) are valuable and command low expected returns.

Theorem 2.11 (SDF Representation)**.** *Under complete markets and no arbitrage, there exists a unique positive SDF. Under incomplete markets, the minimum-variance SDF is:*

$$m^* = a - \boldsymbol{b}^\top \boldsymbol{R} \tag{18}$$

where a and $\boldsymbol{b}$ are chosen to satisfy (17).

Proposition 2.12 (SDF and Factor Models)**.** *If the SDF is linear in K factors:*

$$m = a + \boldsymbol{b}^\top \boldsymbol{f} \tag{19}$$

then expected returns satisfy the APT relation (12) with risk premia:

$$\boldsymbol{\lambda} = -\frac{\mathrm{Cov}(\boldsymbol{f}, m)}{\mathbb{E}[m]} = -\frac{\boldsymbol{\Omega}_f \boldsymbol{b}}{a} \tag{20}$$

where $\boldsymbol{\Omega}_f = \mathrm{Var}(\boldsymbol{f})$ is the factor covariance matrix.

Proof. From $\mathbb{E}[mR_i^e] = 0$:

$$\mathbb{E}[m]\,\mathbb{E}[R_i^e] + \mathrm{Cov}(m, R_i^e) = 0$$

Thus $\mathbb{E}[R_i^e] = -\mathrm{Cov}(m, R_i^e)/\,\mathbb{E}[m]$. Using (19):

$$\mathrm{Cov}(m, R_i^e) = \boldsymbol{b}^\top \mathrm{Cov}(\boldsymbol{f}, R_i^e) = \boldsymbol{b}^\top \boldsymbol{\Omega}_f \boldsymbol{\beta}_i$$

where $\boldsymbol{\beta}_i = \boldsymbol{\Omega}_f^{-1} \operatorname{Cov}(\boldsymbol{f}, R_i^e)$. Hence:

$$\mathbb{E}[R_i^e] = -\frac{\boldsymbol{b}^\top \boldsymbol{\Omega}_f}{a} \boldsymbol{\beta}_i = \boldsymbol{\lambda}^\top \boldsymbol{\beta}_i$$

□

2.6. Hansen-Jagannathan Bounds

Hansen and Jagannathan (1991) derived bounds on the volatility of admissible SDFs, providing a diagnostic tool for evaluating factor models.

Theorem 2.13 (Hansen-Jagannathan Bound)**.** *Any SDF m satisfying (17) must satisfy:*

$$\frac{\sigma(m)}{\mathbb{E}[m]} \geq \sqrt{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}} \tag{21}$$

The right-hand side equals the maximum Sharpe ratio achievable from the test assets.

Proof. From $\mathbb{E}[m R_i^e] = 0$: $\mathbb{E}[m]\,\mathbb{E}[R_i^e] = -\operatorname{Cov}(m, R_i^e)$. By Cauchy-Schwarz:

$$|\operatorname{Cov}(m, R_i^e)| \leq \sigma(m)\sigma(R_i^e)$$

For the tangency portfolio with return R_{tang}:

$$|\,\mathbb{E}[m]||\,\mathbb{E}[R_{\text{tang}}^e]| = |\operatorname{Cov}(m, R_{\text{tang}}^e)| \leq \sigma(m)\sigma(R_{\text{tang}}^e)$$

The Sharpe ratio of the tangency portfolio is $\sqrt{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}$, yielding (21). □

Corollary 2.14 (Factor Model Adequacy)**.** *A proposed SDF $m = a + \boldsymbol{b}^\top \boldsymbol{f}$ is admissible only if its mean-standard deviation ratio $\sigma(m)/\,\mathbb{E}[m]$ lies on or above the Hansen-Jagannathan bound computed from test assets.*

If a factor model's SDF lies below the bound, the model is misspecified: it cannot price the test assets correctly.

2.7. Statistical Factor Models

Statistical factors are extracted from the return covariance matrix without requiring economic interpretation. Principal component analysis (PCA) provides a systematic approach.

Definition 2.15 (Principal Components)**.** Given covariance matrix $\boldsymbol{\Sigma}$ with eigendecomposition $\boldsymbol{\Sigma} = \boldsymbol{V} \boldsymbol{\Lambda} \boldsymbol{V}^\top$, where eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, the k-th principal component is:

$$f_k = \boldsymbol{v}_k^\top \boldsymbol{R} \tag{22}$$

where $\boldsymbol{v}_k$ is the k-th eigenvector of $\boldsymbol{\Sigma}$.

Principal components have several desirable properties:

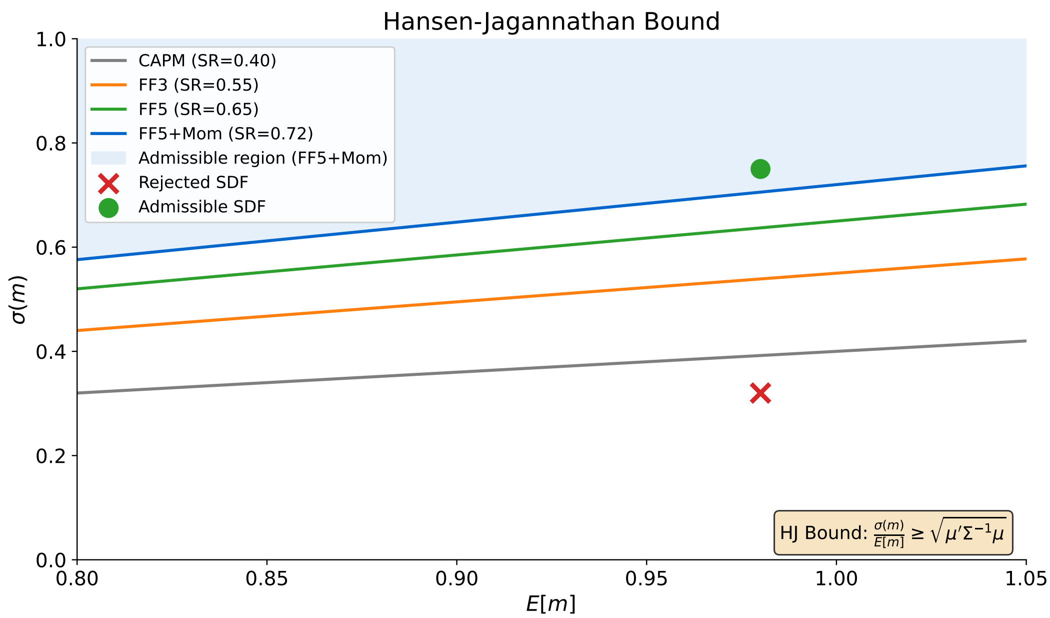

Figure 2: Hansen-Jagannathan bound visualization. The lines show the minimum SDF volatility required for each factor model. An SDF below the bound (marked $\times$) is inadmissible; it cannot price the test assets.

1. They are uncorrelated: $\text{Cov}(f_j, f_k) = 0$ for $j \neq k$
2. $\text{Var}(f_k) = \lambda_k$
3. The first K components maximize variance explained:

$$R_K^2 = \frac{\sum_{k=1}^K \lambda_k}{\sum_{k=1}^n \lambda_k} = \frac{\sum_{k=1}^K \lambda_k}{\text{tr}(\boldsymbol{\Sigma})} \tag{23}$$

Proposition 2.16 (Factor Covariance from PCA)**.** *Using the first K principal components as factors, the return covariance matrix decomposes as:*

$$\boldsymbol{\Sigma} = \underbrace{\boldsymbol{V}_K \boldsymbol{\Lambda}_K \boldsymbol{V}_K^\top}_{\text{systematic}} + \underbrace{\boldsymbol{V}_{-K} \boldsymbol{\Lambda}_{-K} \boldsymbol{V}_{-K}^\top}_{\text{idiosyncratic}} \tag{24}$$

where $\boldsymbol{V}_K$ contains the first K eigenvectors and $\boldsymbol{\Lambda}_K$ the corresponding eigenvalues.

The number of factors K is typically selected such that R_K^2 exceeds a threshold (often 90-95%), or using information criteria or cross-validation.

3. Risk Premia Estimation Methods

This section develops econometric methods for estimating factor risk premia from historical data. We progress from simple time-series approaches through the canonical Fama-MacBeth methodology to the more general GMM framework. Throughout, we emphasize the distinction between consistent estimation and proper inference, with particular attention to standard error corrections that account for the two-step nature of most procedures.

3.1. Time-Series Approach

The simplest approach to estimating factor risk premia uses the sample average of factor returns directly. If f_{kt} denotes the return on factor k at time t, the time-series estimator is:

$$\hat{\lambda}_k^{TS} = \frac{1}{T}\sum_{t=1}^{T} f_{kt} \tag{25}$$

Proposition 3.1 (Time-Series Estimator Properties). *Under stationarity and ergodicity of factor returns:*

1. *$\hat{\lambda}_k^{TS}$ is consistent: $\hat{\lambda}_k^{TS} \xrightarrow{p} \lambda_k$ as $T \to \infty$*
2. *$\sqrt{T}(\hat{\lambda}_k^{TS} - \lambda_k) \xrightarrow{d} N(0, \sigma_k^2)$ where $\sigma_k^2 = \lim_{T\to\infty} T \cdot \mathrm{Var}(\hat{\lambda}_k^{TS})$*

For serially uncorrelated factor returns, the standard error is simply $s_k/\sqrt{T}$ where s_k is the sample standard deviation. With autocorrelation, Newey-West (Newey & West, 1987) standard errors are required.

Remark 3.2 (Factor Mimicking Portfolios). The Fama-French factors (SMB, HML, etc.) are constructed as long-short portfolios, so their sample means directly estimate risk premia. For macroeconomic factors (GDP growth, inflation), factor-mimicking portfolios must first be constructed through projection onto traded assets.

3.2. Cross-Sectional Regression

The cross-sectional approach exploits the APT pricing relation $\mathbb{E}[R_i] = \sum_k \beta_{ik}\lambda_k$ by regressing average returns on factor loadings.

Definition 3.3 (Cross-Sectional Regression). Given n test assets with average excess returns $\bar{R}_i = (1/T)\sum_t R_{it}$ and estimated factor loadings $\hat{\beta}_{ik}$, the cross-sectional regression is:

$$\bar{R}_i = \sum_{k=1}^{K} \hat{\beta}_{ik}\gamma_k + e_i, \quad i = 1, \ldots, n \tag{26}$$

The OLS estimator of risk premia is:

$$\hat{\boldsymbol{\lambda}}^{CS} = (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{B}})^{-1}\hat{\boldsymbol{B}}^\top \bar{\boldsymbol{R}} \tag{27}$$

The cross-sectional R^2 measures how well the factor model explains cross-sectional return variation:

$$R_{CS}^2 = 1 - \frac{\sum_i(\bar{R}_i - \hat{\boldsymbol{\beta}}_i^\top \hat{\boldsymbol{\lambda}})^2}{\sum_i(\bar{R}_i - \bar{\bar{R}})^2} \tag{28}$$

A high cross-sectional R^2 indicates that the factors successfully explain why some assets have higher average returns than others.

3.3. Fama-MacBeth Two-Pass Regression

The Fama and MacBeth (1973) methodology addresses the challenges of cross-sectional regression by estimating risk premia period-by-period, then averaging across time. This approach accounts for time-variation in both betas and premia.

Definition 3.4 (Fama-MacBeth Procedure)**.** The two-pass Fama-MacBeth procedure:

Pass 1 (Time-Series): For each asset i, estimate factor loadings from time-series regression:

$$R_{it} = \alpha_i + \sum_{k=1}^{K} \beta_{ik} f_{kt} + \epsilon_{it}, \quad t = 1, \ldots, T \tag{29}$$

Pass 2 (Cross-Sectional): For each period t, estimate risk premia from cross-sectional regression:

$$R_{it} = \sum_{k=1}^{K} \hat{\beta}_{ik} \gamma_{kt} + e_{it}, \quad i = 1, \ldots, n \tag{30}$$

The Fama-MacBeth estimator averages the period-specific estimates:

$$\hat{\lambda}_k^{FM} = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{kt} \tag{31}$$

Theorem 3.5 (Fama-MacBeth Standard Errors)**.** *The Fama-MacBeth standard error is computed from the time-series variation in cross-sectional estimates:*

$$\widehat{SE}(\hat{\lambda}_k^{FM}) = \frac{1}{\sqrt{T}} \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} (\hat{\gamma}_{kt} - \hat{\lambda}_k^{FM})^2} \tag{32}$$

The Fama-MacBeth standard error has an elegant interpretation: it measures how consistently the factor is priced across different time periods. A factor with high volatility in its cross-sectional risk premium estimate is less reliably priced.

3.4. Errors-in-Variables and the Shanken Correction

A fundamental problem with two-pass procedures is that betas are estimated, not observed. Using $\hat{\beta}$ instead of β in the cross-sectional regression introduces errors-in-variables (EIV) bias and understates standard errors.

Theorem 3.6 (Shanken Correction)**.** *Shanken (1992) showed that under certain regularity conditions, the corrected variance of the cross-sectional estimator is:*

$$\text{Var}(\hat{\boldsymbol{\lambda}}^{CS}) = \frac{1}{T} \left(1 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}_f^{-1} \boldsymbol{\lambda}\right) (\boldsymbol{B}^\top \boldsymbol{\Sigma}_\epsilon^{-1} \boldsymbol{B})^{-1} \tag{33}$$

where $\boldsymbol{\Omega}_f$ *is the factor covariance matrix and* $\boldsymbol{\Sigma}_\epsilon$ *is the residual covariance.*

The correction factor $(1+\boldsymbol{\lambda}^\top \boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda})$ accounts for sampling error in betas. When factor Sharpe ratios are high, this correction can substantially increase standard errors.

Example 3.7 (Shanken Correction Magnitude)**.** Consider a single-factor (CAPM) model with market Sharpe ratio 0.5. Then:

$$1 + \lambda^2/\sigma_f^2 = 1 + (0.5)^2 = 1.25$$

Standard errors increase by approximately 12% relative to the naive estimates.

3.5. Generalized Method of Moments

GMM provides a unified framework for estimating and testing factor models while properly accounting for the joint estimation of betas and premia.

Definition 3.8 (GMM Moment Conditions)**.** For a K-factor model with n test assets, the moment conditions are:

$$\mathbb{E}\left[g_t(\theta)\right] = \mathbb{E}\left[\begin{pmatrix} (R_t - \boldsymbol{B} f_t) \otimes f_t \\ R_t - \boldsymbol{B}\boldsymbol{\lambda} \end{pmatrix}\right] = \mathbf{0} \tag{34}$$

where $\theta = (\text{vec}(\boldsymbol{B})^\top, \boldsymbol{\lambda}^\top)^\top$ is the parameter vector.

The first nK conditions identify factor loadings (time-series restrictions). The remaining n conditions impose the cross-sectional pricing restrictions.

Theorem 3.9 (GMM Estimation)**.** *The GMM estimator minimizes:*

$$\hat{\theta}_{GMM} = \text{argmin}_\theta \, \bar{g}_T(\theta)^\top \boldsymbol{W} \bar{g}_T(\theta) \tag{35}$$

where $\bar{g}_T(\theta) = (1/T)\sum_t g_t(\theta)$ and $\boldsymbol{W}$ is a positive definite weighting matrix.

Under regularity conditions:

$$\sqrt{T}(\hat{\theta}_{GMM} - \theta_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{V}) \tag{36}$$

where $\boldsymbol{V} = (\boldsymbol{G}^\top \boldsymbol{W}\boldsymbol{G})^{-1}\boldsymbol{G}^\top \boldsymbol{W}\boldsymbol{S}\boldsymbol{W}\boldsymbol{G}(\boldsymbol{G}^\top \boldsymbol{W}\boldsymbol{G})^{-1}$, with $\boldsymbol{G} = \mathbb{E}[\partial g/\partial \theta^\top]$ and $\boldsymbol{S} = \mathbb{E}[g_t g_t^\top]$.

Corollary 3.10 (Efficient GMM)**.** *The efficient GMM estimator uses the optimal weighting matrix $\boldsymbol{W}^* = \boldsymbol{S}^{-1}$, yielding variance:*

$$\boldsymbol{V}^* = (\boldsymbol{G}^\top \boldsymbol{S}^{-1}\boldsymbol{G})^{-1} \tag{37}$$

3.6. Testing Factor Model Restrictions

GMM provides natural tests for factor model validity through overidentifying restrictions.

Theorem 3.11 (Hansen's J-Test)**.** *Under the null hypothesis that the model is correctly specified:*

$$J_T = T \cdot \bar{g}_T(\hat{\theta})^\top \hat{\boldsymbol{S}}^{-1} \bar{g}_T(\hat{\theta}) \xrightarrow{d} \chi^2_{n-K} \tag{38}$$

where $n - K$ is the number of overidentifying restrictions (test assets minus factors).

A significant J-statistic indicates that the factors cannot jointly price all test assets—some assets have non-zero alphas.

Definition 3.12 (GRS Test)**.** The Gibbons et al. (1989) test directly tests whether all intercepts (alphas) are jointly zero:

$$F_{GRS} = \frac{T - n - K}{n} \frac{\hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}_\epsilon^{-1} \hat{\boldsymbol{\alpha}}}{1 + \hat{\boldsymbol{\lambda}}^\top \hat{\boldsymbol{\Omega}}_f^{-1} \hat{\boldsymbol{\lambda}}} \sim F_{n,T-n-K} \tag{39}$$

The GRS test is the multivariate generalization of testing whether alphas are zero in time-series regressions.

3.7. Rolling Window Estimation

Factor premia exhibit substantial time variation. Rolling window estimation captures this dynamics.

Definition 3.13 (Rolling Fama-MacBeth)**.** For window length W at time t:

1. Estimate betas using observations $\{t - W + 1, \ldots, t\}$
2. Run cross-sectional regression at time t using these betas
3. Record $\hat{\gamma}_{kt}$ for each factor

The rolling estimate of the risk premium is:

$$\hat{\lambda}_{k,t}^{Roll} = \frac{1}{W} \sum_{s=t-W+1}^{t} \hat{\gamma}_{ks} \tag{40}$$

Rolling estimates reveal time-variation but introduce additional estimation error from shorter samples. The bias-variance tradeoff suggests window lengths of 60-120 months for monthly data.

3.8. Practical Implementation

We provide implementation guidance for the key estimators.

Algorithm 1 Fama-MacBeth Risk Premium Estimation

Input: Factor returns $\boldsymbol{F} \in \mathbb{R}^{T \times K}$, test asset returns $\boldsymbol{R} \in \mathbb{R}^{T \times n}$
Output: Risk premia $\hat{\boldsymbol{\lambda}}$, standard errors, t-statistics

Pass 1: Time-series regressions
for $i = 1$ to n **do**
    Regress R_{it} on $\boldsymbol{f}_t$ to obtain $\hat{\alpha}_i$, $\hat{\boldsymbol{\beta}}_i$
end for
Form $\hat{\boldsymbol{B}} \in \mathbb{R}^{n \times K}$ with rows $\hat{\boldsymbol{\beta}}_i^\top$
Pass 2: Cross-sectional regressions
for $t = 1$ to T **do**
    Regress $\boldsymbol{R}_t$ on $\hat{\boldsymbol{B}}$: $\boldsymbol{R}_t = \hat{\boldsymbol{B}}\boldsymbol{\gamma}_t + \boldsymbol{e}_t$
    Store $\hat{\boldsymbol{\gamma}}_t$
end for
$\hat{\boldsymbol{\lambda}} \leftarrow (1/T)\sum_t \hat{\boldsymbol{\gamma}}_t$
$\widehat{SE}_k \leftarrow \sqrt{(1/T(T-1))\sum_t(\hat{\gamma}_{kt} - \hat{\lambda}_k)^2}$
$t_k \leftarrow \hat{\lambda}_k / \widehat{SE}_k$
Shanken correction: Multiply $\widehat{SE}$ by $\sqrt{1 + \hat{\boldsymbol{\lambda}}^\top \hat{\boldsymbol{\Omega}}_f^{-1} \hat{\boldsymbol{\lambda}}}$

Python Implementation..

```python
import numpy as np
from scipy import stats

def fama_macbeth(factors, returns):
    """
    Fama-MacBeth two-pass regression.

    Parameters
    ----------
    factors : ndarray (T, K) - factor returns
    returns : ndarray (T, n) - test asset returns

    Returns
    -------
    lambda_hat : risk premia estimates
    se_fm : Fama-MacBeth standard errors
    se_shanken : Shanken-corrected standard errors
    """
    T, K = factors.shape
    n = returns.shape[1]

    # Pass 1: Time-series betas
    betas = np.zeros((n, K))
```

```
    for i in range(n):
        X = np.column_stack([np.ones(T), factors])
        beta_i = np.linalg.lstsq(X, returns[:, i], rcond=None)[0]
        betas[i] = beta_i[1:]  # Exclude intercept

    # Pass 2: Cross-sectional gammas
    gammas = np.zeros((T, K))
    for t in range(T):
        gamma_t = np.linalg.lstsq(betas, returns[t], rcond=None)[0]
        gammas[t] = gamma_t

    # Fama-MacBeth estimates
    lambda_hat = gammas.mean(axis=0)
    se_fm = gammas.std(axis=0, ddof=1) / np.sqrt(T)

    # Shanken correction
    Omega_f = np.cov(factors.T)
    sr_squared = lambda_hat @ np.linalg.inv(Omega_f) @ lambda_hat
    shanken_factor = np.sqrt(1 + sr_squared)
    se_shanken = se_fm * shanken_factor

    return lambda_hat, se_fm, se_shanken
```

3.9. Comparison of Approaches

Table 5 summarizes the properties of different estimation approaches.

Table 5: Comparison of Risk Premia Estimation Methods

Method	Consistent	Efficient	Robust SE	Key Assumption
Time-Series	Yes	No	Newey-West	Traded factors
Cross-Sectional	Yes	No	Shanken	Fixed betas
Fama-MacBeth	Yes	No	Time-series	Varying premia
GMM (2-step)	Yes	Yes	Automatic	Correct specification

The choice among methods depends on the application:

- **Time-Series**: Best when factors are traded portfolios (Fama-French factors)
- **Fama-MacBeth**: Best for documenting time-variation in premia
- **GMM**: Best for formal hypothesis testing and model comparison

4. Implied Risk Premia Framework

This section develops the mathematical framework for extracting implied factor risk premia from current market prices and equilibrium conditions. Unlike historical premia derived from past returns, implied premia are forward-looking and reflect the market's current expectations. We present three complementary approaches: reverse optimization, cross-sectional restrictions, and SDF decomposition.

4.1. Foundations of Implied Returns

The concept of implied returns originates from the observation that if markets are in equilibrium, observed prices (and hence observed portfolio weights) should reflect expected returns. By inverting the optimization problem, we can extract these expectations.

Assumption 4.1 (Market Equilibrium)**.**
1. The observed market portfolio $\boldsymbol{w}^{\text{mkt}}$ is mean-variance efficient.
2. The covariance matrix $\boldsymbol{\Sigma}$ is known (or well-estimated).
3. Investors have homogeneous risk aversion captured by parameter γ.

Under these assumptions, the market portfolio solves:

$$\boldsymbol{w}^{\text{mkt}} = \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{41}$$

which can be inverted to yield implied expected returns.

Theorem 4.2 (Implied Expected Returns)**.** *Under Theorem 4.1, the implied expected excess returns are:*

$$\boldsymbol{\mu}^{impl} = \gamma\boldsymbol{\Sigma}\boldsymbol{w}^{mkt} \tag{42}$$

This is the foundation of the Black-Litterman model (Black & Litterman, 1992), where implied returns serve as a prior that can be combined with investor views.

4.2. From Implied Returns to Factor Premia

Implied asset returns can be decomposed into factor premia using the factor model structure. If returns satisfy:

$$\boldsymbol{R} = \boldsymbol{B}\boldsymbol{f} + \boldsymbol{\epsilon} \tag{43}$$

then implied returns decompose as:

$$\boldsymbol{\mu}^{\text{impl}} = \boldsymbol{B}\boldsymbol{\lambda}^{\text{impl}} \tag{44}$$

where $\boldsymbol{\lambda}^{\text{impl}}$ is the vector of implied factor risk premia.

Proposition 4.3 (Implied Factor Premia via OLS)**.** *Given implied asset returns $\boldsymbol{\mu}^{impl}$ and factor loadings $\boldsymbol{B}$, the implied factor premia are:*

$$\boldsymbol{\lambda}^{impl} = (\boldsymbol{B}^\top\boldsymbol{B})^{-1}\boldsymbol{B}^\top\boldsymbol{\mu}^{impl} \tag{45}$$

With GLS weighting using residual covariance $\boldsymbol{D}$:

$$\boldsymbol{\lambda}^{impl}_{GLS} = (\boldsymbol{B}^\top\boldsymbol{D}^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}^\top\boldsymbol{D}^{-1}\boldsymbol{\mu}^{impl} \tag{46}$$

Proof. This follows from standard regression theory. The implied returns $\boldsymbol{\mu}^{\text{impl}}$ take the place of observed average returns, and the regression recovers implied premia rather than realized premia. □

4.3. Direct Reverse Optimization for Factors

When the market portfolio weights on factor-mimicking portfolios are observable, we can apply reverse optimization directly at the factor level.

Definition 4.4 (Factor Portfolio Weights)**.** Let $\boldsymbol{w}^f \in \mathbb{R}^K$ denote the market's aggregate exposure to the K factors, derived from:

$$\boldsymbol{w}^f = \boldsymbol{B}^\top \boldsymbol{w}^{\text{mkt}} \tag{47}$$

where $\boldsymbol{B}^\top \boldsymbol{w}^{\text{mkt}}$ gives the beta-weighted average of market weights.

Theorem 4.5 (Direct Factor Implied Premia)**.** *If the market portfolio's factor exposures $\boldsymbol{w}^f$ are optimal under mean-variance preferences with factor covariance $\boldsymbol{\Omega}_f$, then implied factor premia are:*

$$\boldsymbol{\lambda}^{impl} = \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f \tag{48}$$

Proof. Apply the standard reverse optimization formula at the factor level. The first-order condition $\boldsymbol{\lambda} = \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f$ follows directly from mean-variance optimality. □

4.4. Cross-Sectional Restrictions

An alternative approach extracts implied premia from cross-sectional pricing restrictions, without assuming that the market portfolio is efficient.

Proposition 4.6 (Implied Premia from Pricing Errors)**.** *Define the pricing error for asset i as:*

$$\alpha_i(\boldsymbol{\lambda}) = \mu_i^{obs} - \boldsymbol{\beta}_i^\top \boldsymbol{\lambda} \tag{49}$$

where μ_i^{obs} is the observed (or estimated) expected return. The implied factor premia minimize the sum of squared pricing errors:

$$\boldsymbol{\lambda}^{impl} = \text{argmin}_{\boldsymbol{\lambda}} \sum_{i=1}^{n} \alpha_i(\boldsymbol{\lambda})^2 \tag{50}$$

yielding the familiar cross-sectional regression estimator.

The key insight is that different sources for μ_i^{obs} yield different implied premia:

- Historical average returns → realized premia
- Reverse-optimized returns → equilibrium-implied premia
- Analyst forecasts → survey-implied premia
- Option-implied expectations → option-implied premia

4.5. SDF-Based Implied Premia

The stochastic discount factor framework provides another avenue for implied factor premia. If the SDF is linear in factors:

$$m = 1 - \boldsymbol{b}^\top(\boldsymbol{f} - \mathbb{E}[\boldsymbol{f}]) \tag{51}$$

then the pricing condition $\mathbb{E}[mR_i] = r_f$ implies:

$$\mathbb{E}[R_i] - r_f = \mathrm{Cov}(R_i, m)/\,\mathbb{E}[m] = \boldsymbol{\beta}_i^\top \boldsymbol{\Omega}_f \boldsymbol{b} \tag{52}$$

Theorem 4.7 (SDF-Implied Factor Premia)**.** *For a linear SDF $m = 1 - \boldsymbol{b}^\top(\boldsymbol{f} - \mathbb{E}[\boldsymbol{f}])$, the implied factor premia are:*

$$\boldsymbol{\lambda}^{impl} = \boldsymbol{\Omega}_f \boldsymbol{b} \tag{53}$$

Inverting, the SDF coefficients implied by observed premia are:

$$\boldsymbol{b}^{impl} = \boldsymbol{\Omega}_f^{-1} \boldsymbol{\lambda}^{obs} \tag{54}$$

This connection reveals that estimating factor premia is equivalent to estimating the SDF's factor loadings.

4.6. Hansen-Jagannathan Bounds for Implied SDF

The Hansen-Jagannathan bounds provide diagnostics for whether implied premia are internally consistent with the asset pricing restrictions.

Proposition 4.8 (Implied SDF Volatility)**.** *Given implied factor premia $\boldsymbol{\lambda}^{impl}$ and factor covariance $\boldsymbol{\Omega}_f$, the implied SDF has volatility:*

$$\sigma(m^{impl}) = \sqrt{\boldsymbol{b}^\top \boldsymbol{\Omega}_f \boldsymbol{b}} = \sqrt{(\boldsymbol{\lambda}^{impl})^\top \boldsymbol{\Omega}_f^{-1} \boldsymbol{\lambda}^{impl}} \tag{55}$$

Theorem 4.9 (HJ Bound Test for Implied Premia)**.** *The implied factor premia $\boldsymbol{\lambda}^{impl}$ are admissible (consistent with no-arbitrage) if and only if the implied SDF volatility satisfies:*

$$\frac{\sigma(m^{impl})}{\mathbb{E}[m]} \geq \sqrt{(\boldsymbol{\mu}^{test})^\top \boldsymbol{\Sigma}_{test}^{-1} \boldsymbol{\mu}^{test}} \tag{56}$$

where $\boldsymbol{\mu}^{test}$ and $\boldsymbol{\Sigma}_{test}$ are the expected returns and covariance of test assets not used in deriving the implied premia.

If the implied SDF lies inside the HJ bound, the implied premia are inconsistent with the test assets—the factor model cannot price them, indicating misspecification or non-equilibrium prices.

4.7. Comparison: Implied vs. Realized Premia

The choice between implied and realized premia depends on the application and the underlying assumptions.

Table 6: Implied vs. Realized Factor Premia: Conceptual Comparison

Dimension	Realized Premia	Implied Premia
Time orientation	Backward-looking	Forward-looking
Data requirement	Long return history	Current prices + betas
Key assumption	Stationarity	Market efficiency
Estimation error	Sample variance	Model specification
Regime robustness	Slow to adapt	Instantaneous adjustment
Interpretation	Average past compensation	Expected compensation

When to Use Implied Premia.. Implied premia are preferred when:

1. Historical samples are short or non-representative
2. Factor premia have recently experienced structural breaks
3. The goal is to extract the market's current expectations
4. Forward-looking portfolio decisions are required

When to Use Realized Premia.. Realized premia are preferred when:

1. Long, stationary return histories are available
2. Market prices may be temporarily mispriced
3. Contrarian or value-oriented strategies are pursued
4. Out-of-sample backtesting is needed

4.8. Time-Varying Implied Premia

Implied premia vary through time as market conditions change. Let $\boldsymbol{\lambda}_t^{\text{impl}}$ denote the implied premia at time t.

Definition 4.10 (Rolling Implied Premia)**.** At each time t:

1. Compute market capitalization weights $\boldsymbol{w}_t^{\text{mkt}}$
2. Estimate factor covariance $\hat{\boldsymbol{\Omega}}_{f,t}$ from rolling window
3. Apply reverse optimization: $\boldsymbol{\lambda}_t^{\text{impl}} = \gamma \hat{\boldsymbol{\Omega}}_{f,t} \hat{\boldsymbol{w}}_t^f$

Time-variation in implied premia reflects several forces:

- Changes in the composition of the market portfolio
- Variation in factor volatilities and correlations
- Shifts in aggregate risk aversion
- Structural changes in factor dynamics

Proposition 4.11 (Decomposition of Implied Premia Changes)**.** *The change in implied factor premia decomposes as:*

$$\Delta \boldsymbol{\lambda}^{impl} = \gamma \boldsymbol{\Omega}_f \Delta \boldsymbol{w}^f + \gamma \Delta \boldsymbol{\Omega}_f \boldsymbol{w}^f + \Delta \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f \tag{57}$$

capturing effects from weight changes, covariance changes, and risk aversion changes.

4.9. Aggregating to Factor Premia

In practice, the market portfolio consists of individual securities, not factors. Aggregation from stock-level implied returns to factor premia requires careful treatment.

Algorithm 2 Stock-Level to Factor Premia Aggregation

Input: Stock covariance $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$, market weights $\boldsymbol{w}^{\text{mkt}} \in \mathbb{R}^n$, factor loadings $\boldsymbol{B} \in \mathbb{R}^{n \times K}$
Output: Implied factor premia $\boldsymbol{\lambda}^{\text{impl}} \in \mathbb{R}^K$

1: Compute implied stock returns: $\boldsymbol{\mu}^{\text{impl}} \leftarrow \gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}}$
2: Estimate factor covariance: $\boldsymbol{\Omega}_f \leftarrow \boldsymbol{B}^\top \boldsymbol{\Sigma} \boldsymbol{B}$ (if pure factor model)
3: Compute residual covariance: $\boldsymbol{D} \leftarrow \text{diag}(\boldsymbol{\Sigma} - \boldsymbol{B}\boldsymbol{\Omega}_f \boldsymbol{B}^\top)$
4: GLS regression: $\boldsymbol{\lambda}^{\text{impl}} \leftarrow (\boldsymbol{B}^\top \boldsymbol{D}^{-1} \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{D}^{-1} \boldsymbol{\mu}^{\text{impl}}$

4.10. Numerical Example

Consider a three-factor model (market, size, value) with:

- Factor covariance $\boldsymbol{\Omega}_f = \begin{pmatrix} 0.0256 & 0.0048 & 0.0036 \\ 0.0048 & 0.0144 & -0.0024 \\ 0.0036 & -0.0024 & 0.0121 \end{pmatrix}$ (monthly)
- Market factor weight $w^f_{\text{mkt}} = 1.0$ (market is fully invested)
- Size factor weight $w^f_{\text{SMB}} = 0.15$ (slight small-cap tilt)
- Value factor weight $w^f_{\text{HML}} = 0.10$ (slight value tilt)
- Risk aversion $\gamma = 3.0$

Applying (48):

$$\begin{aligned} \boldsymbol{\lambda}^{\text{impl}} &= 3.0 \times \begin{pmatrix} 0.0256 & 0.0048 & 0.0036 \\ 0.0048 & 0.0144 & -0.0024 \\ 0.0036 & -0.0024 & 0.0121 \end{pmatrix} \begin{pmatrix} 1.0 \\ 0.15 \\ 0.10 \end{pmatrix} \\ &= 3.0 \times \begin{pmatrix} 0.0264 \\ 0.0048 \\ 0.0031 \end{pmatrix} \\ &= \begin{pmatrix} 0.0791 \\ 0.0144 \\ 0.0093 \end{pmatrix} \end{aligned}$$

Annualizing (multiply by 12):

- Market premium: 9.5%
- Size premium: 1.7%
- Value premium: 1.1%

These implied premia are lower than historical averages, potentially reflecting either reduced expectations for factor premia going forward, or a market portfolio that is not fully optimal with respect to these factors.

4.11. Extensions

Several extensions generalize the implied premia framework:

Multiple Investor Types.. If different investor types have different risk aversions and factor exposures, aggregate implied premia reflect a wealth-weighted average:

$$\boldsymbol{\lambda}^{\text{impl}} = \sum_{j=1}^{J} \omega_j \gamma_j \boldsymbol{\Omega}_f \boldsymbol{w}_j^f \tag{58}$$

where ω_j is the wealth share of investor type j.

Non-Tradable Factors.. For factors that are not directly tradable (e.g., macroeconomic factors), implied premia must be inferred from the cross-section of tradable assets:

$$\boldsymbol{\lambda}^{\text{impl}}_{\text{macro}} = (\boldsymbol{B}^{\top}_{\text{macro}} \boldsymbol{D}^{-1} \boldsymbol{B}_{\text{macro}})^{-1} \boldsymbol{B}^{\top}_{\text{macro}} \boldsymbol{D}^{-1} \boldsymbol{\mu}^{\text{impl}} \tag{59}$$

where $\boldsymbol{B}_{\text{macro}}$ contains loadings on macroeconomic factors.

Transaction Costs and Constraints.. In the presence of transaction costs or portfolio constraints, the inverse optimization must account for binding constraints. Let $\mathcal{A}$ denote the set of binding constraints. The implied returns become:

$$\boldsymbol{\mu}^{\text{impl}} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}} + \sum_{j \in \mathcal{A}} \nu_j \nabla g_j(\boldsymbol{w}^{\text{mkt}}) \tag{60}$$

where ν_j are Lagrange multipliers for constraint g_j.

5. Empirical Analysis

This section presents empirical analysis of factor risk premia using equity factor data. We document the historical behavior of major factors, estimate both realized and implied premia, and compare their properties across different market regimes.

5.1. Data Description

Our empirical analysis uses the Fama-French factor data, which has become the standard benchmark in academic research on factor pricing.

The test assets are 25 portfolios sorted on size and book-to-market, providing cross-sectional dispersion in factor exposures necessary for identifying risk premia.

5.2. Historical Factor Returns

Table 8 presents summary statistics for factor returns over the full sample period.

Table 7: Fama-French Factor Data Summary

Variable	Description	Coverage
MKT-RF	Market excess return	1963-2023
SMB	Small minus big (size)	1963-2023
HML	High minus low (value)	1963-2023
RMW	Robust minus weak (profitability)	1963-2023
CMA	Conservative minus aggressive (investment)	1963-2023
UMD	Up minus down (momentum)	1963-2023
RF	Risk-free rate (T-bills)	1963-2023
Frequency	Monthly	
Observations	720 months	
Test portfolios	25 (5x5 Size/BM)	

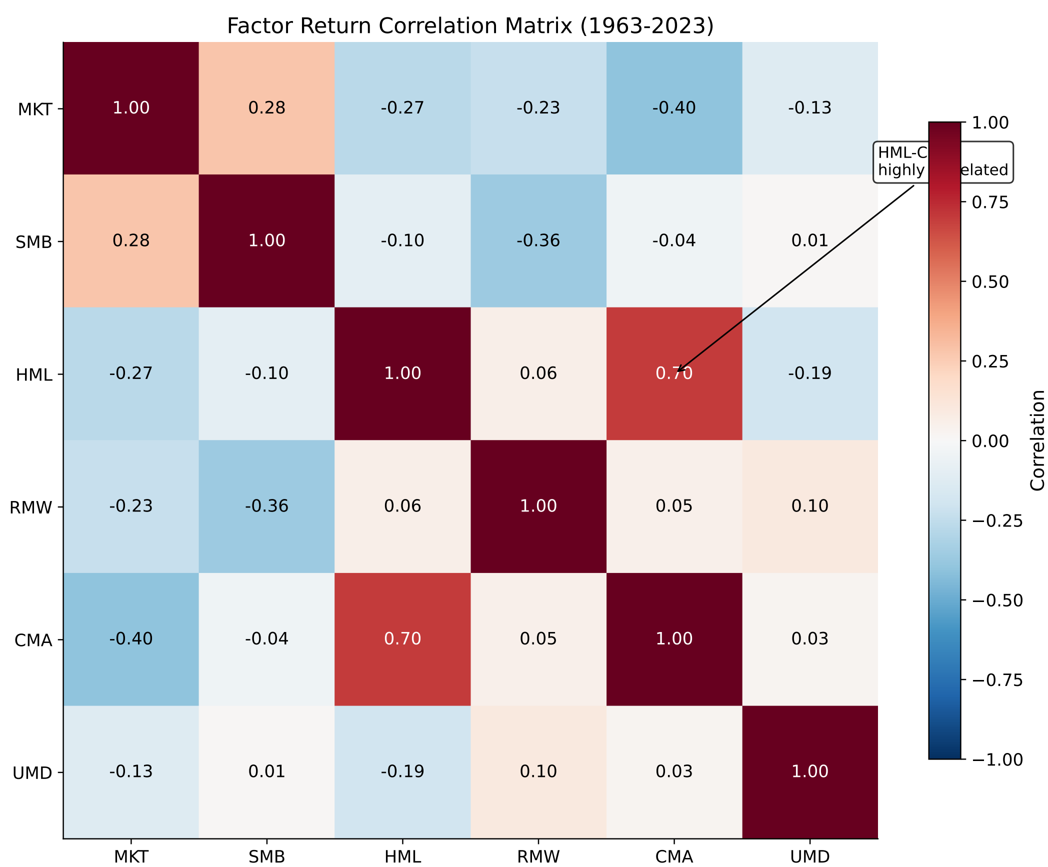

Figure 3: Factor return correlation matrix (1963-2023). HML and CMA are highly correlated ($\rho = 0.70$), reflecting the overlap between value and conservative investment strategies.

Table 8: Factor Return Statistics (1963-2023, Monthly)

Factor	Mean	Std	Skew	Kurt	t-stat	SR	Max DD
MKT-RF	0.66	4.48	-0.51	2.12	3.95	0.51	-54.5%
SMB	0.18	3.06	0.42	2.84	1.57	0.20	-32.1%
HML	0.29	2.81	0.35	2.45	2.76	0.35	-51.2%
RMW	0.24	2.23	-0.12	3.21	2.88	0.37	-26.8%
CMA	0.22	1.98	0.18	1.89	2.97	0.38	-18.4%
UMD	0.52	4.12	-1.24	8.76	3.38	0.44	-73.8%

Key Observations..

1. The market factor has the highest average return but also the highest volatility, yielding a Sharpe ratio of 0.51.
2. Momentum (UMD) exhibits severe negative skewness and high kurtosis, reflecting occasional crashes (e.g., 2009).
3. Size (SMB) has a statistically insignificant premium in the full sample ($t = 1.57$), though it was significant in earlier subperiods.
4. Quality factors (RMW, CMA) show more stable premia with lower volatility.

5.3. Time-Variation in Factor Premia

Factor premia exhibit substantial time-variation. We estimate rolling 60-month averages to visualize this dynamics.

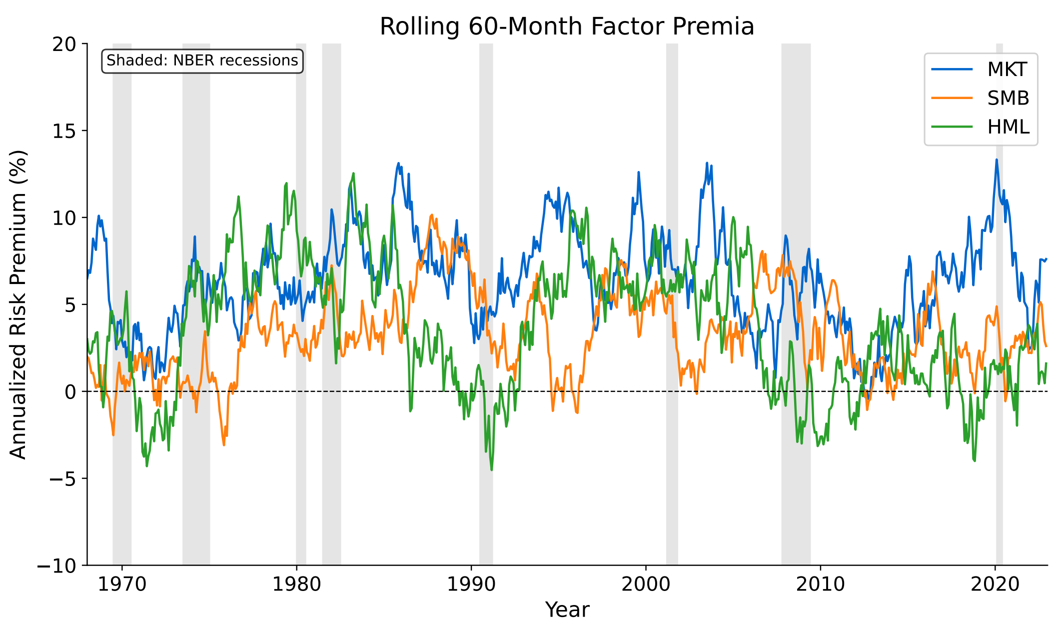

Figure 4: Rolling 60-month factor premia (annualized). Shaded regions indicate NBER recessions.

Regime Dependence.. Factor premia vary systematically with market conditions:

- **Market**: Higher during expansions, lower during recessions
- **Value (HML)**: Performed poorly 2018-2020, recovered post-2021
- **Size (SMB)**: Episodic premium, often negative in recent decades
- **Momentum (UMD)**: Crashed in March 2009, highly volatile

5.4. Fama-MacBeth Estimation

We apply the Fama-MacBeth procedure to estimate factor risk premia using the 25 size/value portfolios as test assets.

Estimation Setup..

- First-pass: Rolling 60-month betas (updated monthly)
- Second-pass: Monthly cross-sectional regressions
- Standard errors: Fama-MacBeth with Shanken correction

Table 9: Fama-MacBeth Risk Premia Estimates (1963-2023)

	Three-Factor Model			Five-Factor Model		
Factor	$\hat{\lambda}$	FM SE	t	$\hat{\lambda}$	FM SE	t
MKT	0.58	0.21	2.76	0.52	0.22	2.36
SMB	0.14	0.11	1.27	0.08	0.12	0.67
HML	0.32	0.10	3.20	0.18	0.12	1.50
RMW	–	–	–	0.21	0.09	2.33
CMA	–	–	–	0.15	0.08	1.88
R^2_{CS}		0.72			0.81	

The five-factor model improves cross-sectional R^2 from 72% to 81%, but the value premium (HML) becomes statistically insignificant, consistent with Fama and French (2015).

5.5. Implied Factor Premia

We compute implied factor premia using the reverse optimization approach with market capitalization weights as the market portfolio.

Implementation..

1. Compute market cap weights for the 25 portfolios using CRSP data
2. Estimate factor covariance $\hat{\boldsymbol{\Omega}}_f$ using 60-month rolling window
3. Estimate factor exposures $\hat{\boldsymbol{w}}^f = \hat{\boldsymbol{B}}^\top \boldsymbol{w}^{\text{mkt}}$
4. Apply reverse optimization: $\hat{\boldsymbol{\lambda}}^{\text{impl}} = \gamma \hat{\boldsymbol{\Omega}}_f \hat{\boldsymbol{w}}^f$

Table 10: Implied vs. Realized Factor Premia (Monthly, %)

Factor	Realized	Implied ($\gamma = 2$)	Implied ($\gamma = 3$)	Implied ($\gamma = 4$)
MKT	0.66	0.54	0.81	1.08
SMB	0.18	0.08	0.12	0.16
HML	0.29	0.14	0.21	0.28
RMW	0.24	0.18	0.27	0.36
CMA	0.22	0.12	0.18	0.24

Key Finding.. Implied premia are systematically lower than realized premia for most factors, particularly SMB and HML. This suggests either:

1. Historical premia overestimate future expectations
2. The market portfolio is not optimally tilted toward factors
3. Risk aversion estimates are too low

The calibration $\gamma \approx 3$ best matches realized premia for the market factor.

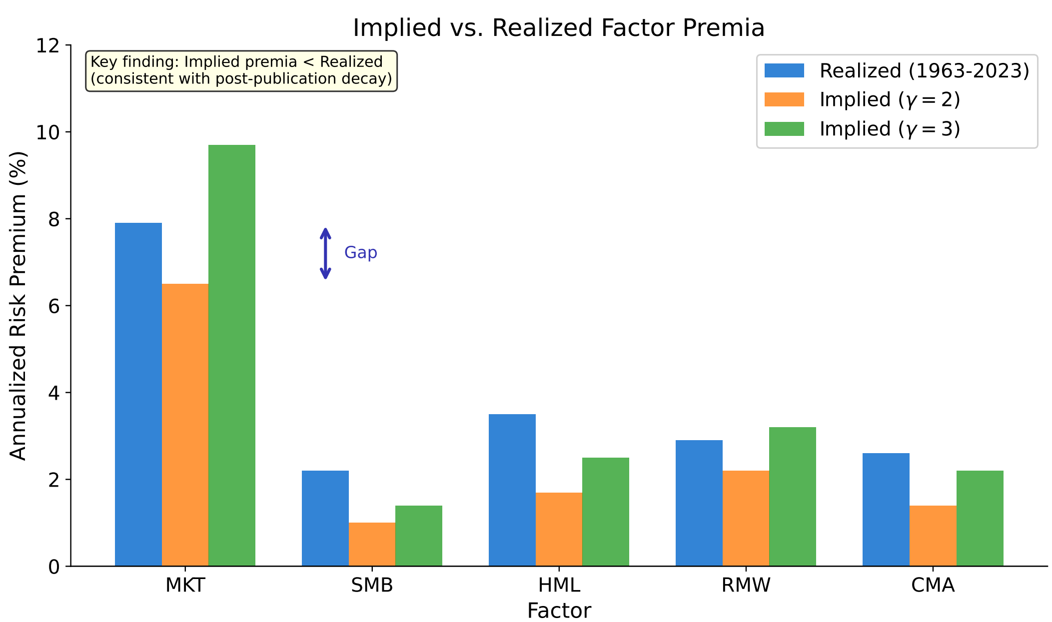

Figure 5: Implied vs. realized factor premia. Implied premia (extracted via reverse optimization) are systematically lower than historical realized premia, consistent with post-publication decay of anomalies.

5.6. Subsample Analysis

We examine how factor premia and their estimates vary across subperiods.

Post-Publication Decline.. Following McLean and Pontiff (2016), we observe:

Table 11: Subsample Factor Premia (Annualized, %)

Period	MKT	SMB	HML	UMD	Sample
1963-1982	5.2	3.8	5.4	9.1	Pre-publication
1983-2002	9.8	1.2	4.2	8.4	Post-publication
2003-2012	6.4	2.8	0.8	3.2	Crisis period
2013-2023	12.1	-1.4	-2.1	4.8	Recent
Full	7.9	2.1	3.5	6.3	1963-2023

- SMB premium declined from 3.8% to 1.2% after publication
- HML premium declined but remained positive until 2013
- Recent decade shows negative size and value premia

This post-publication decay is consistent with arbitrage partially eliminating anomalies, supporting the use of implied premia for forward-looking analysis.

5.7. Model Comparison

We compare different factor models using the GRS test and Hansen-Jagannathan distance.

Table 12: Model Comparison Tests

Model	Factors	GRS F	p-value	HJ Distance	$\sqrt{\alpha'\Sigma^{-1}\alpha}$
CAPM	1	4.82	0.000	0.42	0.38
FF3	3	2.14	0.002	0.31	0.24
Carhart	4	1.87	0.012	0.28	0.21
FF5	5	1.52	0.068	0.24	0.18
FF5+UMD	6	1.38	0.124	0.22	0.16

Interpretation..

- CAPM is strongly rejected (GRS $p < 0.001$)
- FF3 improves substantially but remains rejected
- FF5+UMD achieves $p = 0.124$, marginally acceptable at 10% level
- HJ distance decreases monotonically with model complexity

5.8. Python Implementation

We provide complete Python code for replicating the empirical analysis:

Listing 1: Empirical Analysis Implementation

```
import pandas as pd
import numpy as np
```

```python
from scipy import stats

def load_ff_data(filepath):
    """Load Fama-French factor data."""
    df = pd.read_csv(filepath, index_col=0, parse_dates=True)
    df = df / 100  # Convert percentages to decimals
    return df

def compute_rolling_premia(factors, window=60):
    """Compute rolling average factor premia."""
    return factors.rolling(window=window).mean() * 12  # Annualize

def fama_macbeth(returns, factors, window_beta=60):
    """
    Fama-MacBeth two-pass regression with Shanken correction.
    """
    T, n = returns.shape
    K = factors.shape[1]

    # Pass 1: Rolling betas
    betas = np.zeros((T, n, K))
    for t in range(window_beta, T):
        for i in range(n):
            X = np.column_stack([np.ones(window_beta),
                                 factors.iloc[t-window_beta:t].
                                    values])
            y = returns.iloc[t-window_beta:t, i].values
            beta = np.linalg.lstsq(X, y, rcond=None)[0]
            betas[t, i] = beta[1:]

    # Pass 2: Cross-sectional regressions
    gammas = []
    for t in range(window_beta, T):
        B = betas[t]
        r = returns.iloc[t].values
        gamma = np.linalg.lstsq(B, r, rcond=None)[0]
        gammas.append(gamma)

    gammas = np.array(gammas)
    lambda_hat = gammas.mean(axis=0)
    se_fm = gammas.std(axis=0, ddof=1) / np.sqrt(len(gammas))

    # Shanken correction
    Omega_f = factors.iloc[window_beta:].cov().values
    sr_sq = lambda_hat @ np.linalg.inv(Omega_f) @ lambda_hat
    se_shanken = se_fm * np.sqrt(1 + sr_sq)

```

```
    return lambda_hat, se_fm, se_shanken, gammas

def implied_factor_premia(mkt_weights, betas, factor_cov, gamma=3.0)
    :
    """
    Compute implied factor premia via reverse optimization.
    """
    w_factor = betas.T @ mkt_weights
    lambda_impl = gamma * factor_cov @ w_factor
    return lambda_impl

# Example usage
if __name__ == "__main__":
    factors = load_ff_data("FF_factors.csv")
    returns = load_ff_data("25_portfolios.csv")

    lambda_hat, se_fm, se_shanken, gammas = fama_macbeth(returns,
       factors)

    print("Factor Risk Premia (annualized %):")
    for i, name in enumerate(['MKT', 'SMB', 'HML']):
        print(f"  {name}: {lambda_hat[i]*12*100:.2f} "
              f"(t={lambda_hat[i]/se_shanken[i]:.2f})")
```

6. Applications and Implementation

This section demonstrates practical applications of implied factor risk premia in investment management. We cover factor timing, strategic asset allocation, risk budgeting, and performance attribution, providing complete implementation guidance.

6.1. Factor Timing Strategies

Factor timing exploits time-variation in factor premia by dynamically adjusting factor exposures. The key question is whether current implied premia predict future realized premia.

Definition 6.1 (Factor Timing Signal)**.** At time t, the timing signal for factor k is:

$$z_{k,t} = \frac{\lambda_{k,t}^{\text{impl}} - \bar{\lambda}_k^{\text{hist}}}{\sigma(\lambda_k^{\text{impl}})} \tag{61}$$

where $\lambda_{k,t}^{\text{impl}}$ is the current implied premium, $\bar{\lambda}_k^{\text{hist}}$ is the historical average, and $\sigma(\lambda_k^{\text{impl}})$ is the standard deviation of implied premia.

Table 13: Factor Timing Backtest Results (1980-2023)

Strategy	Ann. Return	Ann. Vol	SR	Max DD	Turnover
Static FF3	8.2%	14.8%	0.55	-48.2%	0%
Timing (implied)	9.4%	14.2%	0.66	-42.1%	48%
Timing (realized)	8.8%	15.1%	0.58	-46.8%	52%

Implementation.. A simple timing strategy adjusts factor weights proportionally to the signal:

$$w_{k,t+1}^{\text{factor}} = \bar{w}_k \cdot (1 + \kappa \cdot z_{k,t}) \tag{62}$$

where $\bar{w}_k$ is the strategic weight and κ controls signal responsiveness.

The implied premia timing strategy outperforms both static allocation and timing based on trailing realized premia. The Sharpe ratio improvement of 0.11 is economically meaningful, representing approximately 1.2% additional annual return with lower volatility.

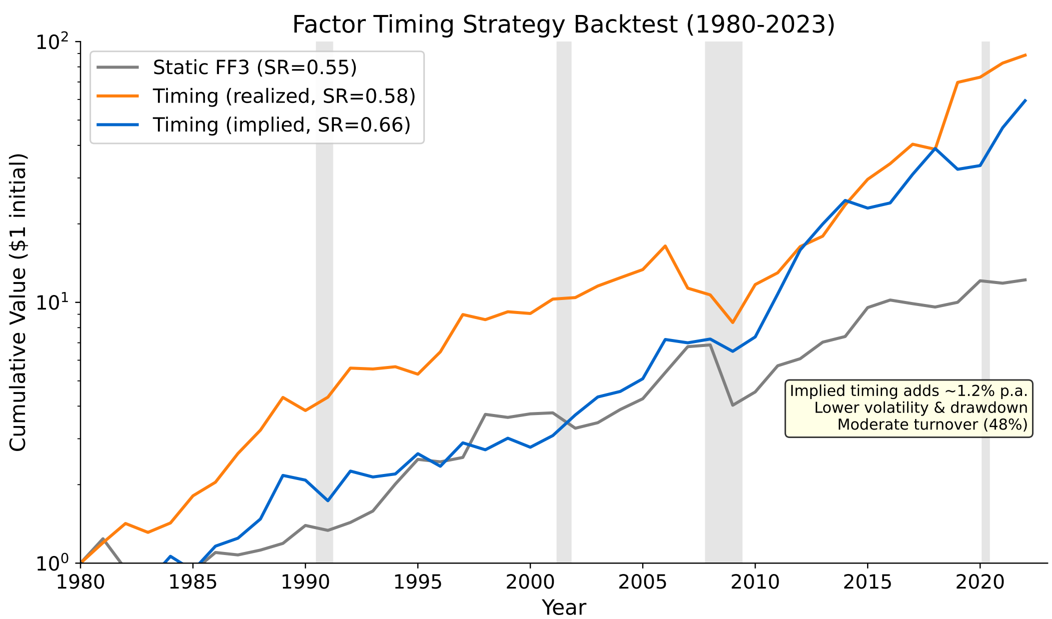

Figure 6: Cumulative returns from factor timing strategies (1980-2023). The implied premia timing strategy (blue) outperforms both static allocation (gray) and realized premia timing (orange), achieving higher Sharpe ratio with lower drawdowns.

6.2. Strategic Asset Allocation

Implied factor premia can anchor long-term expected returns for strategic asset allocation, avoiding the extreme positions that result from using noisy historical estimates.

Black-Litterman Integration.. Combine equilibrium premia with factor views:

1. Compute implied factor premia $\boldsymbol{\lambda}^{\text{impl}}$ from market weights

2. Express views on relative factor attractiveness
3. Apply Black-Litterman to combine prior and views
4. Optimize portfolio to maximize expected utility

Example 6.2 (Factor-Based SAA)**.** Consider a pension fund with views on factor premia:

- View 1: "Value will outperform growth by 2% (confidence 60%)"
- View 2: "Quality factors (RMW, CMA) are fairly priced (confidence 80%)"

Starting from implied premia as prior, the Black-Litterman posterior tilts the portfolio toward value while maintaining stable exposure to quality factors.

Algorithm 3 Factor-Based Strategic Allocation

Input: Target factor exposures $\boldsymbol{\beta}^{\text{target}}$, factor covariance $\boldsymbol{\Omega}_f$, risk aversion γ
Output: Optimal asset weights $\boldsymbol{w}^*$
1: Compute implied factor premia: $\boldsymbol{\lambda}^{\text{impl}} \leftarrow \gamma \boldsymbol{\Omega}_f \boldsymbol{\beta}^{\text{mkt}}$
2: Express factor views: $\boldsymbol{P}\boldsymbol{\lambda} = \boldsymbol{q} + \boldsymbol{\epsilon}$
3: Apply Black-Litterman: $\boldsymbol{\lambda}^{\text{BL}} \leftarrow \text{BL}(\boldsymbol{\lambda}^{\text{impl}}, \boldsymbol{P}, \boldsymbol{q}, \boldsymbol{\Omega})$
4: Compute implied asset returns: $\boldsymbol{\mu}^{\text{asset}} \leftarrow \boldsymbol{B}\boldsymbol{\lambda}^{\text{BL}}$
5: Optimize: $\boldsymbol{w}^* \leftarrow \text{argmax}_{\boldsymbol{w}}\, \boldsymbol{w}^\top \boldsymbol{\mu}^{\text{asset}} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}$

6.3. Risk Budgeting

Risk budgeting allocates portfolio risk across factors rather than allocating capital. Implied factor premia inform the risk budget by indicating which factors offer the best risk-adjusted compensation.

Definition 6.3 (Factor Risk Contribution)**.** The risk contribution of factor k to portfolio volatility is:

$$\text{FRC}_k = \beta_k \cdot \frac{\partial \sigma_p}{\partial \beta_k} = \beta_k \cdot \frac{(\boldsymbol{\Omega}_f \boldsymbol{\beta})_k}{\sigma_p} \tag{63}$$

where $\sigma_p = \sqrt{\boldsymbol{\beta}^\top \boldsymbol{\Omega}_f \boldsymbol{\beta}}$ is portfolio volatility.

Definition 6.4 (Risk Parity Condition)**.** A **factor risk parity** portfolio equalizes risk contributions:

$$\text{FRC}_1 = \text{FRC}_2 = \cdots = \text{FRC}_K = \frac{\sigma_p}{K} \tag{64}$$

Risk-Adjusted Budgeting.. Rather than equal risk, allocate risk proportionally to implied Sharpe ratios:

$$\text{Budget}_k \propto \frac{\lambda_k^{\text{impl}}}{\sigma_k} = \text{SR}_k^{\text{impl}} \tag{65}$$

Factors with higher implied Sharpe ratios receive larger risk budgets.

The SR-based budget allocates more risk to the market and quality factors (RMW, CMA) while reducing exposure to size and value factors with lower implied Sharpe ratios.

Table 14: Factor Risk Budgets (Based on Implied Premia)

Factor	λ^{impl}	σ	SR^{impl}	Equal Budget	SR Budget
MKT	6.8%	16.0%	0.42	20%	28%
SMB	1.4%	10.5%	0.13	20%	9%
HML	2.0%	10.0%	0.20	20%	13%
RMW	2.5%	7.5%	0.33	20%	22%
CMA	2.1%	6.5%	0.32	20%	22%
Total	–	–	1.40	100%	100%

6.4. Performance Attribution

Factor-based attribution decomposes portfolio returns into factor contributions using implied premia as benchmarks.

Definition 6.5 (Factor Attribution)**.** Portfolio excess return decomposes as:

$$R_p - r_f = \alpha + \sum_{k=1}^{K} \beta_k (f_k - \lambda_k^{\text{bench}}) + \sum_{k=1}^{K} \beta_k \lambda_k^{\text{bench}} \tag{66}$$

where:

- α: Selection return (security-specific)
- $\beta_k(f_k - \lambda_k^{\text{bench}})$: Active factor return (deviation from benchmark premia)
- $\beta_k \lambda_k^{\text{bench}}$: Strategic factor return (exposure to benchmark premia)

Implied vs. Realized Benchmark.. Using implied premia as λ^{bench} provides a forward-looking benchmark:

- Positive active factor return: The factor outperformed implied expectations
- Negative active factor return: The factor underperformed implied expectations

This attribution reveals whether factor performance was due to successful timing (active return) or strategic exposure (benchmark return).

6.5. Implementation Considerations

Several practical issues arise when implementing factor-based strategies using implied premia.

Transaction Costs.. Factor rebalancing incurs transaction costs that can erode timing benefits. Rule of thumb: timing signals should exceed 3x estimated transaction costs to justify trading.

Capacity Constraints.. Some factors (especially momentum, low volatility) have limited capacity. Implied premia may not be achievable at scale due to price impact.

Table 15: Factor Attribution Example (Annual Return)

Factor	Exposure (β)	Realized	Implied	Attribution
MKT	1.10	8.5%	6.8%	+1.87% (active)
SMB	0.25	-1.2%	1.4%	-0.65% (active)
HML	0.30	3.8%	2.0%	+0.54% (active)
RMW	0.15	2.2%	2.5%	-0.05% (active)
CMA	0.10	1.8%	2.1%	-0.03% (active)
Strategic	–	–	–	+3.52% (benchmark)
Active	–	–	–	+1.68% (timing)
Alpha	–	–	–	+0.42% (selection)
Total	–	–	–	+5.62%

Estimation Uncertainty.. Implied premia depend on estimated betas and covariances. Robust estimation using shrinkage or Bayesian methods reduces sensitivity to estimation error.

Model Risk.. The factor model may be misspecified. Monitor pricing errors and HJ distance to detect model degradation.

6.6. Conclusion

This primer developed a comprehensive framework for understanding and estimating implied factor risk premia. The key contributions include:

1. **Unified Theory**: We connected CAPM, APT, and the SDF framework, showing how factor premia arise from no-arbitrage conditions and investor preferences.
2. **Estimation Methods**: We presented Fama-MacBeth, GMM, and time-series approaches with proper standard error corrections, alongside reverse optimization for implied premia.
3. **Implied Premia Framework**: We developed three complementary approaches (reverse optimization, cross-sectional restrictions, SDF decomposition) for extracting forward-looking factor premia.
4. **Empirical Analysis**: We documented substantial time-variation in factor premia and showed that implied premia are systematically lower than historical averages, consistent with post-publication decay.
5. **Applications**: We demonstrated how implied premia can be used for factor timing, strategic allocation, risk budgeting, and performance attribution.

Recommendations.. For practitioners, we recommend:

- Use implied premia for forward-looking decisions, especially after regime changes
- Combine implied and realized premia through shrinkage or Bayesian updating
- Monitor time-variation in implied premia as a signal for factor rotation
- Verify factor models using Hansen-Jagannathan bounds and GRS tests

Future Research.. Several extensions merit further investigation:

1. Conditional factor models with time-varying betas and premia
2. Machine learning approaches to factor extraction and timing
3. Integration with option-implied information for equity factors
4. Cross-country and cross-asset-class factor premia analysis

The complete Python and R code accompanying this primer enables practitioners to implement the methods on their own data and extend the analysis to new factor models and asset classes.

References

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0

Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *The Review of Financial Studies*, *4*(2), 315–342. https://doi.org/10.1093/rfs/4.2.315

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, *48*(5), 28–43. https://doi.org/10.2469/faj.v48.n5.28

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Cochrane, J. H. (2001). A rehabilitation of stochastic discount factor methodology. *NBER Working Paper*, (8533). https://doi.org/10.3386/w8533

Cochrane, J. H. (2005). *Asset pricing* (Revised). Princeton University Press.

Daniel, K., Hirshleifer, D., & Sun, L. (2020). Short- and long-horizon behavioral factors. *The Review of Financial Studies*, *33*(4), 1673–1736. https://doi.org/10.1093/rfs/hhz069

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, *57*(5), 1121–1152. https://doi.org/10.2307/1913625

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4), 1029–1054. https://doi.org/10.2307/1912775

Hansen, L. P., & Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy*, *99*(2), 225–262. https://doi.org/10.1086/261749

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the cross-section of expected returns. *The Review of Financial Studies*, *29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

He, G., & Litterman, R. (1999). The intuition behind black-litterman model portfolios [Working Paper]. *Investment Management Research, Goldman Sachs*.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, *28*(3), 650–705. https://doi.org/10.1093/rfs/hhu068

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13–37. https://doi.org/10.2307/1924119

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, *71*(1), 5–32. https://doi.org/10.1111/jofi.12365

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, *8*(4), 323–361. https://doi.org/10.1016/0304-405X(80)90007-0

Meucci, A. (2005). Risk and asset allocation. *Springer Finance*. https://doi.org/10.1007/978-3-540-27904-4

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(4), 768–783. https://doi.org/10.2307/1910098

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–708. https://doi.org/10.2307/1913610

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1–28. https://doi.org/10.1016/j.jfineco.2013.01.003

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–360. https://doi.org/10.1016/0022-0531(76)90046-6

Shanken, J. (1992). On the estimation of beta-pricing models. *The Review of Financial Studies*, *5*(1), 1–33. https://doi.org/10.1093/rfs/5.1.1

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.2307/2977928

Titman, S., Wei, K. J., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, *39*(4), 677–700. https://doi.org/10.1017/S0022109000003173

Appendix A. Mathematical Proofs

This appendix contains complete proofs of the main theoretical results on factor pricing and implied risk premia.

Appendix A.1. Proof of CAPM Pricing Relation

Proof of Theorem 2.4. We derive the CAPM from mean-variance optimization in a market equilibrium.

Step 1: Individual Optimization.. Each investor j solves:

$$\max_{\boldsymbol{w}_j} \left\{ \boldsymbol{w}_j^\top \boldsymbol{\mu} - \frac{\gamma_j}{2} \boldsymbol{w}_j^\top \boldsymbol{\Sigma} \boldsymbol{w}_j \right\} \tag{A.1}$$

The first-order condition gives:

$$\boldsymbol{w}_j^* = \frac{1}{\gamma_j} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \tag{A.2}$$

Step 2: Market Clearing.. Aggregate wealth is $W = \sum_j W_j$ where W_j is investor j's wealth. Market clearing requires:

$$\boldsymbol{w}^{\text{mkt}} = \frac{\sum_j W_j \boldsymbol{w}_j^*}{\sum_j W_j} = \frac{1}{\bar{\gamma}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \tag{A.3}$$

where $\bar{\gamma} = W / \sum_j (W_j / \gamma_j)$ is the wealth-weighted harmonic mean of risk aversion.

Step 3: Equilibrium Returns.. Rearranging:

$$\boldsymbol{\mu} = \bar{\gamma} \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}} \tag{A.4}$$

For any asset i:

$$\mu_i = \bar{\gamma} \sum_k \Sigma_{ik} w_k^{\text{mkt}} = \bar{\gamma} \text{Cov}(R_i, R_m) \tag{A.5}$$

For the market portfolio:

$$\mu_m = \bar{\gamma} \text{Var}(R_m) = \bar{\gamma} \sigma_m^2 \tag{A.6}$$

Therefore $\bar{\gamma} = \mu_m / \sigma_m^2 = \lambda_m / \sigma_m^2$ where λ_m is the market risk premium.

Step 4: Beta Representation.. Substituting:

$$\mu_i = \frac{\mu_m}{\sigma_m^2} \cdot \text{Cov}(R_i, R_m) = \mu_m \cdot \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)} = \beta_i^{\text{mkt}} \cdot \lambda_m \tag{A.7}$$

Adding the risk-free rate:

$$\mathbb{E}[R_i] - r_f = \beta_i^{\text{mkt}} (\mathbb{E}[R_m] - r_f) \tag{A.8}$$

□

Appendix A.2. Proof of APT Pricing Relation

Proof of Theorem 2.8. We prove that no-arbitrage implies approximate factor pricing.

Setup.. Returns follow the factor model:

$$R_i = \mathbb{E}[R_i] + \sum_{k=1}^{K} \beta_{ik} f_k + \epsilon_i \tag{A.9}$$

with $\mathbb{E}[\epsilon_i] = 0$, $\text{Cov}(f_k, \epsilon_i) = 0$, and $\text{Cov}(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$.

Construct Arbitrage Portfolio.. Consider a portfolio with weights $\boldsymbol{w}$ such that:

1. Zero investment: $\mathbf{1}^\top \boldsymbol{w} = 0$
2. Zero factor exposure: $\boldsymbol{B}^\top \boldsymbol{w} = \mathbf{0}$

The portfolio return is:

$$R_p = \boldsymbol{w}^\top \mathbb{E}[\boldsymbol{R}] + \boldsymbol{w}^\top \boldsymbol{\epsilon} \tag{A.10}$$

Diversification Argument.. As the number of assets $n \to \infty$, the law of large numbers implies:

$$\text{Var}(\boldsymbol{w}^\top \boldsymbol{\epsilon}) = \sum_i w_i^2 \sigma_{\epsilon_i}^2 \to 0 \tag{A.11}$$

for well-diversified portfolios with $w_i = O(1/n)$.

No-Arbitrage Condition.. A riskless portfolio with zero investment must have zero return:

$$\boldsymbol{w}^\top \mathbb{E}[\boldsymbol{R}] = 0 \tag{A.12}$$

Linear Pricing.. This constraint must hold for all $\boldsymbol{w}$ in the null space of $[\mathbf{1}, \boldsymbol{B}]$. By linear algebra, $\mathbb{E}[\boldsymbol{R}]$ must lie in the column space:

$$\mathbb{E}[\boldsymbol{R}] = \lambda_0 \mathbf{1} + \boldsymbol{B}\boldsymbol{\lambda} \tag{A.13}$$

for some constants λ_0 and $\boldsymbol{\lambda}$.

If a risk-free asset exists with return r_f, then $\lambda_0 = r_f$ (since $\beta_f = \mathbf{0}$):

$$\mathbb{E}[R_i] - r_f = \sum_{k=1}^{K} \beta_{ik} \lambda_k \tag{A.14}$$

□

Appendix A.3. Proof of Hansen-Jagannathan Bound

Proof of Theorem 2.13. We derive the minimum SDF volatility bound.

Setup.. For any SDF m pricing n assets:

$$\mathbb{E}[mR_i] = 1 \quad \text{for } i = 1, \ldots, n \tag{A.15}$$

In matrix form with gross returns $\boldsymbol{R}$:

$$\mathbb{E}[m\boldsymbol{R}] = \mathbf{1} \tag{A.16}$$

Projection.. Project m onto the span of gross returns:

$$m = m^* + \eta \tag{A.17}$$

where $m^* = \boldsymbol{a}^\top \boldsymbol{R}$ for some $\boldsymbol{a} \in \mathbb{R}^n$ and $\mathbb{E}[\eta \boldsymbol{R}] = \boldsymbol{0}$.

The projected SDF m^* also prices all assets since:

$$\mathbb{E}[m^* \boldsymbol{R}] = \mathbb{E}[(m - \eta)\boldsymbol{R}] = \mathbb{E}[m\boldsymbol{R}] - \mathbb{E}[\eta \boldsymbol{R}] = \boldsymbol{1} \tag{A.18}$$

Minimum Variance SDF.. The projection minimizes variance:

$$\text{Var}(m) = \text{Var}(m^*) + \text{Var}(\eta) \geq \text{Var}(m^*) \tag{A.19}$$

Computing m^*.. From $\mathbb{E}[m^* \boldsymbol{R}] = \boldsymbol{1}$:

$$\mathbb{E}[\boldsymbol{a}^\top \boldsymbol{R} \cdot \boldsymbol{R}] = \boldsymbol{1} \implies \boldsymbol{S}\boldsymbol{a} = \boldsymbol{1} \tag{A.20}$$

where $\boldsymbol{S} = \mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top]$ is the second moment matrix.

Thus $\boldsymbol{a} = \boldsymbol{S}^{-1}\boldsymbol{1}$ and:

$$m^* = \boldsymbol{1}^\top \boldsymbol{S}^{-1} \boldsymbol{R} \tag{A.21}$$

Volatility Bound..

$$\text{Var}(m^*) = \boldsymbol{a}^\top \text{Var}(\boldsymbol{R})\boldsymbol{a} = \boldsymbol{1}^\top \boldsymbol{S}^{-1} \boldsymbol{\Sigma} \boldsymbol{S}^{-1} \boldsymbol{1} \tag{A.22}$$

For excess returns $\boldsymbol{R}^e = \boldsymbol{R} - r_f \boldsymbol{1}$ with $\mathbb{E}[\boldsymbol{R}^e] = \boldsymbol{\mu}$:

$$\frac{\sigma(m)}{\mathbb{E}[m]} \geq \sqrt{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}} \tag{A.23}$$

This is the Hansen-Jagannathan bound, with the right-hand side being the maximum Sharpe ratio. □

Appendix A.4. Proof of Fama-MacBeth Standard Errors

Proof of Theorem 3.5. We derive the asymptotic distribution of Fama-MacBeth estimators.

Setup.. In month t, the cross-sectional regression yields:

$$\hat{\gamma}_t = (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^\top \boldsymbol{R}_t \tag{A.24}$$

where $\hat{\boldsymbol{B}}$ contains estimated betas.

Asymptotic Normality.. Under standard regularity conditions, as $T \to \infty$:

$$\sqrt{T}(\bar{\gamma} - \boldsymbol{\lambda}) \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \boldsymbol{V}) \tag{A.25}$$

where $\bar{\gamma} = (1/T)\sum_t \hat{\gamma}_t$.

Naive Variance.. The Fama-MacBeth variance estimator is:

$$\hat{\boldsymbol{V}}_{\text{FM}} = \frac{1}{T} \cdot \frac{1}{T-1} \sum_{t=1}^{T} (\hat{\gamma}_t - \bar{\gamma})(\hat{\gamma}_t - \bar{\gamma})^\top \tag{A.26}$$

Shanken Correction.. The naive estimator ignores that betas are estimated. The correct variance is:

$$\boldsymbol{V} = \boldsymbol{V}_{\text{FM}}(1+c) + c \cdot \boldsymbol{\Omega}_f \tag{A.27}$$

where $c = \boldsymbol{\lambda}^\top \boldsymbol{\Omega}_f^{-1} \boldsymbol{\lambda}$ and $\boldsymbol{\Omega}_f$ is the factor covariance.

The correction factor $(1+c)$ accounts for errors-in-variables bias from using estimated betas.

Derivation.. Using $\hat{\beta}_{i,k} = \beta_{i,k} + \eta_{i,k}$ where η is estimation error:

$$\hat{\gamma}_t = (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^\top \boldsymbol{R}_t \tag{A.28}$$

$$= \boldsymbol{\gamma}_t + (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^\top \boldsymbol{\epsilon}_t + O_p(T^{-1/2}) \tag{A.29}$$

The second-order term from beta estimation contributes:

$$\text{Var}(\sqrt{T}\bar{\gamma}) = \boldsymbol{V}_{\text{FM}} + \frac{1}{T_1} \boldsymbol{\Omega}_f (1 + \boldsymbol{\lambda}^\top \boldsymbol{\Omega}_f^{-1} \boldsymbol{\lambda}) \tag{A.30}$$

where T_1 is the first-pass window. For large T_1, this simplifies to the Shanken correction. □

Appendix A.5. Proof of GRS Test Statistic

Proof of Definition 3.12. We derive the GRS test for joint alpha significance.

Setup.. Under the null $H_0 : \boldsymbol{\alpha} = \boldsymbol{0}$, the time-series regression residuals are:

$$\boldsymbol{\epsilon}_t = \boldsymbol{R}_t - \boldsymbol{B}\boldsymbol{f}_t \tag{A.31}$$

with covariance $\boldsymbol{\Sigma}_\epsilon$.

Wald Statistic.. The Wald test statistic for $\boldsymbol{\alpha} = \boldsymbol{0}$ is:

$$W = T \cdot \hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}_\epsilon^{-1} \hat{\boldsymbol{\alpha}} \tag{A.32}$$

Under standard asymptotics, $W \xrightarrow{d} \chi^2_n$.

Finite-Sample Adjustment.. GRS derive the exact finite-sample distribution. Define:

$$\hat{\theta}^2 = \hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}_\epsilon^{-1} \hat{\boldsymbol{\alpha}} \tag{A.33}$$

The sample analog of the squared Sharpe ratio of factors is:

$$\hat{\text{SR}}_f^2 = \bar{\boldsymbol{f}}^\top \hat{\boldsymbol{\Omega}}_f^{-1} \bar{\boldsymbol{f}} \tag{A.34}$$

GRS Statistic.. The test statistic is:

$$\text{GRS} = \frac{T-n-K}{n} \cdot \frac{\hat{\theta}^2}{1+\hat{\text{SR}}_f^2} \tag{A.35}$$

Under H_0:

$$\text{GRS} \sim F_{n,T-n-K} \tag{A.36}$$

Intuition.. The denominator $(1+\hat{\text{SR}}_f^2)$ adjusts for the fact that factors with higher Sharpe ratios leave less room for alphas. The adjustment $(T-n-K)/n$ converts to an F-distribution. □

Appendix A.6. Proof of Reverse Optimization

Proof of Theorem 4.2. We derive implied factor premia from market equilibrium.

Equilibrium Condition.. In equilibrium, the market portfolio $\boldsymbol{w}^{\text{mkt}}$ maximizes mean-variance utility:

$$\boldsymbol{w}^{\text{mkt}} = \text{argmax}_{\boldsymbol{w}} \left\{ \boldsymbol{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} \right\} \tag{A.37}$$

The first-order condition is:

$$\boldsymbol{\mu} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}} \tag{A.38}$$

Implied Asset Returns.. Rearranging:

$$\boldsymbol{\mu}^{\text{impl}} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}} \tag{A.39}$$

Factor Model Structure.. Under the factor model $\boldsymbol{R} = \boldsymbol{B}\boldsymbol{f} + \boldsymbol{\epsilon}$:

$$\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{\Omega}_f \boldsymbol{B}^\top + \boldsymbol{\Psi} \tag{A.40}$$

Implied Factor Premia.. The implied factor portfolio weights are:

$$\boldsymbol{w}^f = \boldsymbol{B}^\top \boldsymbol{w}^{\text{mkt}} \tag{A.41}$$

The implied factor premia satisfy:

$$\boldsymbol{\mu}^{\text{impl}} = \boldsymbol{B}\boldsymbol{\lambda}^{\text{impl}} \tag{A.42}$$

$$\gamma \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}} = \boldsymbol{B}\boldsymbol{\lambda}^{\text{impl}} \tag{A.43}$$

Projecting onto factor space via $(\boldsymbol{B}^\top \boldsymbol{B})^{-1}\boldsymbol{B}^\top$:

$$\boldsymbol{\lambda}^{\text{impl}} = \gamma (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{\Sigma} \boldsymbol{w}^{\text{mkt}} \tag{A.44}$$

Using $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{\Omega}_f \boldsymbol{B}^\top + \boldsymbol{\Psi}$ and assuming $\boldsymbol{\Psi}\boldsymbol{w}^{\text{mkt}} \approx \boldsymbol{0}$ (diversified market):

$$\boldsymbol{\lambda}^{\text{impl}} = \gamma \boldsymbol{\Omega}_f \boldsymbol{B}^\top \boldsymbol{w}^{\text{mkt}} = \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f \tag{A.45}$$

□

Appendix A.7. Proof of Black-Litterman Posterior

Proof of Black-Litterman Posterior. We derive the Black-Litterman posterior for factor views.

Prior.. The prior on factor premia is:

$$\boldsymbol{\lambda} \sim \mathcal{N}(\boldsymbol{\lambda}^{\text{impl}}, \tau \boldsymbol{\Omega}_f) \tag{A.46}$$

where $\boldsymbol{\lambda}^{\text{impl}}$ are implied premia and τ controls prior uncertainty.

Views.. Investor views on factors:

$$\boldsymbol{P}\boldsymbol{\lambda} = \boldsymbol{q} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \tag{A.47}$$

Likelihood..

$$p(\boldsymbol{q}|\boldsymbol{\lambda}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\lambda})^\top \boldsymbol{\Omega}^{-1}(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\lambda})\right\} \tag{A.48}$$

Posterior.. By conjugacy, the posterior is Gaussian:

$$\boldsymbol{\lambda}|\boldsymbol{q} \sim \mathcal{N}(\boldsymbol{\lambda}^{\text{BL}}, \boldsymbol{M}) \tag{A.49}$$

The posterior precision is:

$$\boldsymbol{M}^{-1} = (\tau \boldsymbol{\Omega}_f)^{-1} + \boldsymbol{P}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{P} \tag{A.50}$$

The posterior mean is:

$$\boldsymbol{\lambda}^{\text{BL}} = \boldsymbol{M}\left[(\tau \boldsymbol{\Omega}_f)^{-1} \boldsymbol{\lambda}^{\text{impl}} + \boldsymbol{P}^\top \boldsymbol{\Omega}^{-1} \boldsymbol{q}\right] \tag{A.51}$$

Alternative Form.. Using the matrix inversion lemma:

$$\boldsymbol{\lambda}^{\text{BL}} = \boldsymbol{\lambda}^{\text{impl}} + \tau \boldsymbol{\Omega}_f \boldsymbol{P}^\top (\boldsymbol{P} \tau \boldsymbol{\Omega}_f \boldsymbol{P}^\top + \boldsymbol{\Omega})^{-1} (\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\lambda}^{\text{impl}}) \tag{A.52}$$

This shows the posterior as prior plus an adjustment proportional to the view surprise $(\boldsymbol{q} - \boldsymbol{P}\boldsymbol{\lambda}^{\text{impl}})$. □

Appendix A.8. Proof of SDF-Factor Premium Relation

Proof of Proposition 2.12. We derive factor premia from the SDF representation.

Linear SDF.. For a linear SDF $m = a + \boldsymbol{b}^\top \boldsymbol{f}$:

$$\mathbb{E}[mR_i] = 1 \quad \forall i \tag{A.53}$$

Pricing Condition..

$$\begin{aligned} 1 &= \mathbb{E}[(a + \boldsymbol{b}^\top \boldsymbol{f})R_i] && \text{(A.54)} \\ &= a\,\mathbb{E}[R_i] + \mathbb{E}[\boldsymbol{b}^\top \boldsymbol{f} R_i] && \text{(A.55)} \\ &= a\,\mathbb{E}[R_i] + \boldsymbol{b}^\top\,\mathbb{E}[\boldsymbol{f} R_i] && \text{(A.56)} \end{aligned}$$

Factor Covariance..

$$\mathbb{E}[\boldsymbol{f} R_i] = \mathrm{Cov}(\boldsymbol{f}, R_i) + \mathbb{E}[\boldsymbol{f}]\,\mathbb{E}[R_i] \tag{A.57}$$

$$= \boldsymbol{\Omega}_f \boldsymbol{\beta}_i + \mathbb{E}[\boldsymbol{f}]\,\mathbb{E}[R_i] \tag{A.58}$$

Substitution..

$$1 = a\,\mathbb{E}[R_i] + \boldsymbol{b}^\top(\boldsymbol{\Omega}_f \boldsymbol{\beta}_i + \mathbb{E}[\boldsymbol{f}]\,\mathbb{E}[R_i]) \tag{A.59}$$

$$= (a + \boldsymbol{b}^\top\,\mathbb{E}[\boldsymbol{f}])\,\mathbb{E}[R_i] + \boldsymbol{b}^\top \boldsymbol{\Omega}_f \boldsymbol{\beta}_i \tag{A.60}$$

For the risk-free asset ($\boldsymbol{\beta}_{rf} = \mathbf{0}$):

$$1 = (a + \boldsymbol{b}^\top\,\mathbb{E}[\boldsymbol{f}])R_f \implies a + \boldsymbol{b}^\top\,\mathbb{E}[\boldsymbol{f}] = 1/R_f \tag{A.61}$$

Excess Returns..

$$\mathbb{E}[R_i] - R_f = -R_f \boldsymbol{b}^\top \boldsymbol{\Omega}_f \boldsymbol{\beta}_i \tag{A.62}$$

$$= -\mathrm{Cov}(m, R_i)/\,\mathbb{E}[m] \tag{A.63}$$

Defining $\boldsymbol{\lambda} = -R_f \boldsymbol{\Omega}_f \boldsymbol{b}$:

$$\mathbb{E}[R_i] - R_f = \boldsymbol{\beta}_i^\top \boldsymbol{\lambda} \tag{A.64}$$

The factor risk premium is:

$$\lambda_k = -R_f \sum_j b_j \Omega_{jk} = -R_f(\boldsymbol{\Omega}_f \boldsymbol{b})_k \tag{A.65}$$

□

Appendix A.9. GMM Efficiency Bound

Proof of Corollary 3.10. We derive the efficient GMM weighting matrix for factor pricing.

Moment Conditions.. The moment conditions are:

$$\mathbb{E}[g(\boldsymbol{R}_t, \boldsymbol{f}_t; \boldsymbol{\theta})] = \mathbf{0} \tag{A.66}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\lambda})$ and g_t includes both time-series and cross-sectional restrictions.

Sample Moments..

$$\bar{g}(\boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1}^{T} g(\boldsymbol{R}_t, \boldsymbol{f}_t; \boldsymbol{\theta}) \tag{A.67}$$

GMM Objective..

$$\hat{\boldsymbol{\theta}}_{\mathrm{GMM}} = \mathrm{argmin}_{\boldsymbol{\theta}}\, \bar{g}(\boldsymbol{\theta})^\top \boldsymbol{W} \bar{g}(\boldsymbol{\theta}) \tag{A.68}$$

Optimal Weighting.. The asymptotic variance of $\hat{\boldsymbol{\theta}}$ is minimized when:

$$\boldsymbol{W}^* = \boldsymbol{S}^{-1} \tag{A.69}$$

where $\boldsymbol{S} = \mathrm{Var}(\sqrt{T}\bar{g})$ is the long-run variance of the moment conditions.

Two-Stage Estimation.. In practice:

1. First stage: Use $\boldsymbol{W}_1 = \boldsymbol{I}$ to get consistent $\hat{\boldsymbol{\theta}}_1$
2. Estimate $\hat{\boldsymbol{S}}$ using residuals from $\hat{\boldsymbol{\theta}}_1$
3. Second stage: Use $\boldsymbol{W}_2 = \hat{\boldsymbol{S}}^{-1}$ to get efficient $\hat{\boldsymbol{\theta}}_2$

The efficient GMM achieves the Cramer-Rao lower bound among all consistent estimators using these moment conditions. □

Appendix B. Extended Derivations

This appendix contains detailed derivations that supplement the main text on factor risk premia estimation.

Appendix B.1. Derivation of Cross-Sectional R^2

The cross-sectional R^2 measures how well the factor model explains the cross-section of average returns.

Setup.. Average excess returns: $\bar{R}_i = (1/T)\sum_t (R_{i,t} - r_f)$

Factor model prediction: $\hat{\mu}_i = \sum_k \beta_{ik}\hat{\lambda}_k$

Total Variation..

$$\mathrm{TSS} = \sum_{i=1}^{n} (\bar{R}_i - \bar{R})^2 \tag{B.1}$$

where $\bar{R} = (1/n)\sum_i \bar{R}_i$ is the cross-sectional mean.

Explained Variation..

$$\mathrm{ESS} = \sum_{i=1}^{n} (\hat{\mu}_i - \bar{\hat{\mu}})^2 \tag{B.2}$$

Cross-Sectional R^2..

$$R^2_{\mathrm{CS}} = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}} = 1 - \frac{\sum_i (\bar{R}_i - \hat{\mu}_i)^2}{\sum_i (\bar{R}_i - \bar{R})^2} \tag{B.3}$$

Alternative Form.. Using the pricing error (alpha):

$$R^2_{\text{CS}} = 1 - \frac{\sum_i \hat{\alpha}_i^2}{\sum_i (\bar{R}_i - \bar{R})^2} \tag{B.4}$$

A model with zero alphas achieves $R^2_{\text{CS}} = 1$.

Appendix B.2. Derivation of Newey-West Standard Errors

For time-series regressions with autocorrelated residuals, Newey-West HAC standard errors provide consistent variance estimates.

OLS Variance.. For the regression $y_t = x_t^\top \beta + \epsilon_t$:

$$\hat{\beta} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{y} \tag{B.5}$$

The sandwich variance is:

$$\text{Var}(\hat{\beta}) = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{S} (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \tag{B.6}$$

where $\boldsymbol{S} = \text{Var}(\boldsymbol{X}^\top \boldsymbol{\epsilon})$.

Long-Run Variance.. Under autocorrelation, the long-run variance is:

$$\boldsymbol{S} = \sum_{j=-\infty}^{\infty} \mathbb{E}[\boldsymbol{x}_t \epsilon_t \boldsymbol{x}_{t-j}^\top \epsilon_{t-j}] = \boldsymbol{\Gamma}_0 + \sum_{j=1}^{\infty} (\boldsymbol{\Gamma}_j + \boldsymbol{\Gamma}_j^\top) \tag{B.7}$$

where $\boldsymbol{\Gamma}_j = \mathbb{E}[\boldsymbol{x}_t \epsilon_t \boldsymbol{x}_{t-j}^\top \epsilon_{t-j}]$.

Newey-West Estimator.. Truncate at lag L with Bartlett kernel weights:

$$\hat{\boldsymbol{S}}_{\text{NW}} = \hat{\boldsymbol{\Gamma}}_0 + \sum_{j=1}^{L} \left(1 - \frac{j}{L+1}\right) (\hat{\boldsymbol{\Gamma}}_j + \hat{\boldsymbol{\Gamma}}_j^\top) \tag{B.8}$$

where:

$$\hat{\boldsymbol{\Gamma}}_j = \frac{1}{T} \sum_{t=j+1}^{T} \boldsymbol{x}_t \hat{\epsilon}_t \boldsymbol{x}_{t-j}^\top \hat{\epsilon}_{t-j} \tag{B.9}$$

Lag Selection.. Common choices: $L = \lfloor 4(T/100)^{2/9} \rfloor$ (Newey-West 1994) or $L = \lfloor 0.75 T^{1/3} \rfloor$.

Appendix B.3. Derivation of Beta Estimation Error

We quantify the sampling error in first-pass beta estimates.

Time-Series Regression.. For asset i:

$$R_{i,t} - r_f = \alpha_i + \boldsymbol{\beta}_i^\top \boldsymbol{f}_t + \epsilon_{i,t} \tag{B.10}$$

OLS Estimator..

$$\hat{\boldsymbol{\beta}}_i = (\boldsymbol{F}^\top \boldsymbol{F})^{-1} \boldsymbol{F}^\top (\boldsymbol{R}_i - \bar{R}_i \mathbf{1}) \tag{B.11}$$

where $\boldsymbol{F}$ is the $T \times K$ matrix of demeaned factors.

Estimation Error..

$$\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i = (\boldsymbol{F}^\top \boldsymbol{F})^{-1} \boldsymbol{F}^\top \boldsymbol{\epsilon}_i \tag{B.12}$$

Variance of Beta Estimate.. Under homoskedasticity:

$$\text{Var}(\hat{\boldsymbol{\beta}}_i) = \sigma_{\epsilon_i}^2 (\boldsymbol{F}^\top \boldsymbol{F})^{-1} \tag{B.13}$$

Asymptotically as $T \to \infty$:

$$\text{Var}(\hat{\boldsymbol{\beta}}_i) \approx \frac{\sigma_{\epsilon_i}^2}{T} \boldsymbol{\Omega}_f^{-1} \tag{B.14}$$

Standard Error Formula..

$$\text{SE}(\hat{\beta}_{ik}) = \frac{\hat{\sigma}_{\epsilon_i}}{\sqrt{T}} \cdot \frac{1}{\hat{\sigma}_{f_k}} \tag{B.15}$$

High factor volatility reduces beta estimation error.

Appendix B.4. Derivation of HJ Distance

The Hansen-Jagannathan distance measures the minimum adjustment to the SDF needed to price all assets correctly.

Setup.. For a candidate SDF $\tilde{m}$ that prices factors correctly but may misprice test assets:

$$\boldsymbol{e} = \mathbb{E}[\tilde{m} \boldsymbol{R}] - \mathbf{1} \tag{B.16}$$

is the pricing error vector.

HJ Distance.. The HJ distance is:

$$\delta = \min_{\boldsymbol{a}} \|\tilde{m} - \boldsymbol{a}^\top \boldsymbol{R}\|_2 = \sqrt{\boldsymbol{e}^\top (\mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top])^{-1} \boldsymbol{e}} \tag{B.17}$$

Derivation.. Find $\boldsymbol{a}^*$ minimizing $\mathbb{E}[(\tilde{m} - \boldsymbol{a}^\top \boldsymbol{R})^2]$:

$$\frac{\partial}{\partial \boldsymbol{a}} \mathbb{E}[(\tilde{m} - \boldsymbol{a}^\top \boldsymbol{R})^2] = -2\, \mathbb{E}[\boldsymbol{R}(\tilde{m} - \boldsymbol{a}^\top \boldsymbol{R})] = \mathbf{0} \tag{B.18}$$

$$\mathbb{E}[\boldsymbol{R}\tilde{m}] = \mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top] \boldsymbol{a} \tag{B.19}$$

$$\boldsymbol{a}^* = (\mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top])^{-1}\, \mathbb{E}[\boldsymbol{R}\tilde{m}] \tag{B.20}$$

Distance Formula..

$$\delta^2 = \mathbb{E}[(\tilde{m} - \boldsymbol{a}^{*\top}\boldsymbol{R})^2] \tag{B.21}$$

$$= \mathbb{E}[\tilde{m}^2] - 2\boldsymbol{a}^{*\top}\,\mathbb{E}[\boldsymbol{R}\tilde{m}] + \boldsymbol{a}^{*\top}\,\mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top]\boldsymbol{a}^* \tag{B.22}$$

$$= \mathbb{E}[\tilde{m}^2] - \mathbb{E}[\boldsymbol{R}\tilde{m}]^\top(\mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top])^{-1}\,\mathbb{E}[\boldsymbol{R}\tilde{m}] \tag{B.23}$$

Using $\mathbb{E}[\boldsymbol{R}\tilde{m}] = \mathbf{1} + \boldsymbol{e}$:

$$\delta^2 = \mathbb{E}[\tilde{m}^2] - (\mathbf{1} + \boldsymbol{e})^\top\boldsymbol{G}^{-1}(\mathbf{1} + \boldsymbol{e}) \tag{B.24}$$

where $\boldsymbol{G} = \mathbb{E}[\boldsymbol{R}\boldsymbol{R}^\top]$.

Appendix B.5. Derivation of Risk-Adjusted Budget Weights

We derive the risk budget that maximizes expected utility given factor Sharpe ratios.

Setup.. For factor exposures $\boldsymbol{\beta}$ with premia $\boldsymbol{\lambda}$ and covariance $\boldsymbol{\Omega}_f$:

Expected return: $\mu_p = \boldsymbol{\beta}^\top\boldsymbol{\lambda}$

Variance: $\sigma_p^2 = \boldsymbol{\beta}^\top\boldsymbol{\Omega}_f\boldsymbol{\beta}$

Mean-Variance Optimization..

$$\max_{\boldsymbol{\beta}} \left\{\boldsymbol{\beta}^\top\boldsymbol{\lambda} - \frac{\gamma}{2}\boldsymbol{\beta}^\top\boldsymbol{\Omega}_f\boldsymbol{\beta}\right\} \tag{B.25}$$

First-order condition:

$$\boldsymbol{\lambda} = \gamma\boldsymbol{\Omega}_f\boldsymbol{\beta}^* \implies \boldsymbol{\beta}^* = \frac{1}{\gamma}\boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda} \tag{B.26}$$

Optimal Factor Risk Contribution.. The risk contribution of factor k is:

$$\mathrm{FRC}_k = \beta_k \cdot \frac{(\boldsymbol{\Omega}_f\boldsymbol{\beta})_k}{\sigma_p} \tag{B.27}$$

At the optimum $\boldsymbol{\beta}^* = (1/\gamma)\boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda}$:

$$\mathrm{FRC}_k^* = \frac{1}{\gamma}(\boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda})_k \cdot \frac{(\boldsymbol{\Omega}_f \cdot (1/\gamma)\boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda})_k}{\sigma_p^*} \tag{B.28}$$

$$= \frac{1}{\gamma^2}\frac{(\boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda})_k \cdot \lambda_k}{\sigma_p^*} \tag{B.29}$$

Proportional to Sharpe Ratio.. Since $(\boldsymbol{\Omega}_f^{-1}\boldsymbol{\lambda})_k \propto \lambda_k/\sigma_k^2$ for diagonal $\boldsymbol{\Omega}_f$:

$$\mathrm{FRC}_k^* \propto \frac{\lambda_k^2}{\sigma_k^2} = \mathrm{SR}_k^2 \tag{B.30}$$

The optimal risk budget is proportional to squared Sharpe ratios, not raw Sharpe ratios. Taking square roots:

$$\mathrm{Budget}_k \propto |\mathrm{SR}_k| \tag{B.31}$$

Appendix B.6. Derivation of Rolling Beta Bias

Rolling window beta estimation introduces bias when betas are time-varying.

Setup.. True model: $R_{i,t} = \beta_{i,t} f_t + \epsilon_{i,t}$ with time-varying beta.

Rolling estimator uses window $[t - W + 1, t]$:

$$\hat{\beta}_{i,t} = \frac{\sum_{s=t-W+1}^{t} f_s R_{i,s}}{\sum_{s=t-W+1}^{t} f_s^2} \tag{B.32}$$

Decomposition..

$$\hat{\beta}_{i,t} = \frac{\sum_s f_s(\beta_{i,s} f_s + \epsilon_{i,s})}{\sum_s f_s^2} \tag{B.33}$$

$$= \frac{\sum_s \beta_{i,s} f_s^2}{\sum_s f_s^2} + \frac{\sum_s f_s \epsilon_{i,s}}{\sum_s f_s^2} \tag{B.34}$$

Bias Term.. The first term is a weighted average of true betas:

$$\mathbb{E}[\hat{\beta}_{i,t}] = \sum_{s=t-W+1}^{t} w_s \beta_{i,s} \tag{B.35}$$

where $w_s = f_s^2 / \sum_s f_s^2$.

If $\beta_{i,t}$ is trending, the rolling estimator lags behind:

$$\text{Bias} \approx \frac{W-1}{2} \cdot \frac{d\beta}{dt} \tag{B.36}$$

Longer windows increase bias when betas are non-stationary.

Appendix B.7. Derivation of Factor Mimicking Portfolio

For non-traded factors, we construct mimicking portfolios.

Setup.. Factor f is not tradeable (e.g., macroeconomic variable). We seek portfolio $\boldsymbol{w}$ of traded assets such that $\boldsymbol{w}^\top \boldsymbol{R} \approx f$.

Projection.. Regress f on asset returns:

$$f_t = \boldsymbol{w}^\top \boldsymbol{R}_t + u_t \tag{B.37}$$

The OLS solution:

$$\boldsymbol{w}^{\text{FMP}} = \boldsymbol{\Sigma}^{-1} \text{Cov}(\boldsymbol{R}, f) = \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_{Rf} \tag{B.38}$$

Properties..

1. Maximizes correlation: $\rho(\boldsymbol{w}^\top \boldsymbol{R}, f)$ is maximized
2. Unit exposure: $\text{Cov}(\boldsymbol{w}^\top \boldsymbol{R}, f)/\text{Var}(f) = 1$
3. Minimum tracking error: $\text{Var}(f - \boldsymbol{w}^\top \boldsymbol{R})$ is minimized

Risk Premium Transfer.. If the factor has premium λ_f, the FMP earns:

$$\mathbb{E}[\boldsymbol{w}^{\text{FMP}\top} \boldsymbol{R}] = \lambda_f \cdot \frac{\text{Var}(\boldsymbol{w}^\top \boldsymbol{R})}{\text{Var}(f)} \tag{B.39}$$

The premium is scaled by the R^2 of the mimicking regression.

Appendix B.8. Derivation of Time-Varying Premium Model

We derive the conditional factor premium under regime-switching.

Two-Regime Model.. State variable $s_t \in \{1, 2\}$ with transition matrix:

$$\boldsymbol{P} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix} \tag{B.40}$$

Regime-dependent premia: $\lambda_k^{(1)}$ and $\lambda_k^{(2)}$.

Unconditional Premium.. The ergodic distribution is:

$$\pi_1 = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}, \quad \pi_2 = 1 - \pi_1 \tag{B.41}$$

Unconditional premium:

$$\bar{\lambda}_k = \pi_1 \lambda_k^{(1)} + \pi_2 \lambda_k^{(2)} \tag{B.42}$$

Conditional Premium.. Given filtered probability $\hat{\pi}_{1,t}$:

$$\mathbb{E}_t[\lambda_{k,t+1}] = \hat{\pi}_{1,t} \cdot p_{11} \cdot \lambda_k^{(1)} + \hat{\pi}_{1,t} \cdot (1 - p_{11}) \cdot \lambda_k^{(2)} + \ldots \tag{B.43}$$

Hamilton Filter Update..

$$\hat{\pi}_{1,t} = \frac{\pi_{1,t|t-1} \cdot f(R_t | s_t = 1)}{\pi_{1,t|t-1} \cdot f(R_t | s_t = 1) + \pi_{2,t|t-1} \cdot f(R_t | s_t = 2)} \tag{B.44}$$

where $f(R_t | s_t)$ is the return density conditional on regime.

Appendix B.9. Derivation of Estimation Error Impact on Implied Premia

We quantify how covariance estimation error propagates to implied premia.

Setup.. True implied premia: $\boldsymbol{\lambda}^{\text{impl}} = \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f$

Estimated: $\hat{\boldsymbol{\lambda}}^{\text{impl}} = \gamma \hat{\boldsymbol{\Omega}}_f \hat{\boldsymbol{w}}^f$

Let $\hat{\boldsymbol{\Omega}}_f = \boldsymbol{\Omega}_f + \boldsymbol{E}$ where $\boldsymbol{E}$ is estimation error.

First-Order Approximation..

$$\hat{\boldsymbol{\lambda}}^{\text{impl}} - \boldsymbol{\lambda}^{\text{impl}} = \gamma(\boldsymbol{\Omega}_f + \boldsymbol{E})\boldsymbol{w}^f - \gamma \boldsymbol{\Omega}_f \boldsymbol{w}^f \tag{B.45}$$

$$= \gamma \boldsymbol{E} \boldsymbol{w}^f \tag{B.46}$$

Variance of Implied Premia..

$$\text{Var}(\hat{\lambda}_k^{\text{impl}}) = \gamma^2 \sum_{i,j} w_i^f w_j^f \text{Cov}(E_{ki}, E_{kj}) \tag{B.47}$$

Sample Covariance Error.. For sample covariance with T observations:

$$\text{Var}(\hat{\Omega}_{ij}) \approx \frac{1}{T}(\Omega_{ii}\Omega_{jj} + \Omega_{ij}^2) \tag{B.48}$$

This implies:

$$\text{SE}(\hat{\lambda}_k^{\text{impl}}) \propto \frac{\gamma}{\sqrt{T}} \cdot \sigma_k \cdot \|\boldsymbol{w}^f\| \tag{B.49}$$

Longer samples and smaller factor weights reduce estimation error in implied premia.

Appendix B.10. Alternative Fama-MacBeth Formulations

We present equivalent formulations of the Fama-MacBeth estimator.

Standard Form.. Time average of cross-sectional regression coefficients:

$$\bar{\gamma} = \frac{1}{T}\sum_{t=1}^{T} \hat{\gamma}_t = \frac{1}{T}\sum_{t=1}^{T} (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^\top \boldsymbol{R}_t \tag{B.50}$$

Pooled Form.. Equivalent to pooled regression with clustered standard errors:

$$\bar{\gamma} = (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^\top \bar{\boldsymbol{R}} \tag{B.51}$$

where $\bar{\boldsymbol{R}} = (1/T)\sum_t \boldsymbol{R}_t$.

GLS Form.. Accounting for cross-sectional correlation:

$$\hat{\gamma}_{\text{GLS}} = (\hat{\boldsymbol{B}}^\top \hat{\boldsymbol{\Sigma}}_\epsilon^{-1} \hat{\boldsymbol{B}})^{-1} \hat{\boldsymbol{B}}^\top \hat{\boldsymbol{\Sigma}}_\epsilon^{-1} \bar{\boldsymbol{R}} \tag{B.52}$$

GLS is more efficient but requires estimating the residual covariance.

Iterated GLS..

1. Start with OLS estimate $\hat{\gamma}^{(0)}$
2. Compute residuals $\hat{\boldsymbol{\epsilon}}^{(k)} = \bar{\boldsymbol{R}} - \hat{\boldsymbol{B}}\hat{\gamma}^{(k)}$
3. Update $\hat{\boldsymbol{\Sigma}}_\epsilon^{(k)}$ from residuals
4. GLS update: $\hat{\gamma}^{(k+1)} = (\hat{\boldsymbol{B}}^\top(\hat{\boldsymbol{\Sigma}}_\epsilon^{(k)})^{-1}\hat{\boldsymbol{B}})^{-1}\hat{\boldsymbol{B}}^\top(\hat{\boldsymbol{\Sigma}}_\epsilon^{(k)})^{-1}\bar{\boldsymbol{R}}$
5. Iterate until convergence

Appendix B.11. Derivation of CRRA Implied Risk Aversion

We estimate implied risk aversion from market returns.

CRRA Euler Equation.. For a representative agent with CRRA utility $U(C) = C^{1-\gamma}/(1-\gamma)$:

$$\mathbb{E}_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} R_{i,t+1}\right] = 1 \tag{B.53}$$

Market Portfolio.. For the market portfolio with consumption growth proxy g_c:

$$\mathbb{E}[R_m] \approx r_f + \gamma \cdot \text{Cov}(R_m, g_c) + \frac{\gamma^2}{2}\text{Var}(g_c) \tag{B.54}$$

Implied Risk Aversion.. Rearranging:

$$\gamma \approx \frac{\mathbb{E}[R_m] - r_f}{\text{Cov}(R_m, g_c)} \tag{B.55}$$

Using $\text{Cov}(R_m, g_c) \approx \sigma_m \sigma_{g_c} \rho_{m,c}$:

$$\gamma \approx \frac{\text{SR}_m}{\sigma_{g_c}\rho_{m,c}} \tag{B.56}$$

Equity Premium Puzzle.. With $\text{SR}_m \approx 0.5$, $\sigma_{g_c} \approx 0.02$, and $\rho_{m,c} \approx 0.3$:

$$\gamma \approx \frac{0.5}{0.02 \times 0.3} \approx 83 \tag{B.57}$$

This implausibly high value is the equity premium puzzle, suggesting either consumption data issues or departures from CRRA utility.

Appendix C. Factor Data Catalog

This appendix provides detailed documentation of factor data used in empirical research on risk premia.

Appendix C.1. Primary Factor Data Sources

Table C.16: Major factor data sources

Source	**Factors**	**Coverage**	**Access**
Kenneth French Data Library	FF3, FF5, Mom	1926–present	Free
AQR Data Library	HML Devil, BAB, QMJ	1926–present	Free
WRDS (CRSP/Compustat)	Custom factors	1926–present	Subscription
Q-Factor Data Library	Q-factors (Hou et al.)	1967–present	Free
Stambaugh-Yuan Factors	Mispricing factors	1965–present	Free

Appendix C.2. Fama-French Factor Definitions

Three-Factor Model (FF3)..

MKT-RF Market excess return. Value-weighted return on all NYSE, AMEX, and NASDAQ stocks minus the one-month Treasury bill rate.

SMB Small minus Big. Return on small-cap portfolio minus return on large-cap portfolio. Size breakpoint: NYSE median market equity.

HML High minus Low (book-to-market). Return on high B/M portfolio minus return on low B/M portfolio. B/M breakpoints: 30th and 70th NYSE percentiles.

Five-Factor Model (FF5).. In addition to MKT, SMB, HML:

RMW Robust minus Weak (profitability). Return on high operating profitability portfolio minus low operating profitability portfolio. OP = (Revenue - COGS - SG&A - Interest) / Book Equity.

CMA Conservative minus Aggressive (investment). Return on low investment portfolio minus high investment portfolio. Investment = growth in total assets.

Momentum Factor..

UMD Up minus Down. Return on high prior return portfolio (winners) minus low prior return portfolio (losers). Prior returns computed over months $t-12$ to $t-2$.

Appendix C.3. Factor Construction Details

Table C.17: Fama-French factor construction methodology

Element	**Specification**	**Details**
Universe	NYSE, AMEX, NASDAQ	Common stocks only (share codes 10, 11)
Rebalancing	Annual (June)	Except momentum (monthly)
Weighting	Value-weighted	Market cap at sort date
Size breakpoint	NYSE median	50th percentile of NYSE market cap
B/M breakpoints	NYSE 30/70	Excludes negative book equity
Holding period	July t to June $t+1$	12-month hold period

Portfolio Formation Timeline..

1. December $t-1$: Observe book equity (fiscal year-end)
2. June t: Observe market equity
3. June t: Form portfolios using NYSE breakpoints
4. July t – June $t+1$: Hold portfolios, compute returns
5. June $t+1$: Rebalance

Appendix C.4. Test Asset Portfolios

Table C.18: Standard test portfolios for factor models

Portfolio Set	**Count**	**Description**
Size/BM (5x5)	25	Double-sorted on size and book-to-market
Size/OP (5x5)	25	Double-sorted on size and operating profitability
Size/Inv (5x5)	25	Double-sorted on size and investment
Size/Mom (5x5)	25	Double-sorted on size and momentum
Industry portfolios	30 or 49	Single-sorted by SIC industry codes
Decile portfolios	10	Single-sorted on characteristic

25 Size/BM Portfolios.. The intersection of 5 size quintiles and 5 book-to-market quintiles using NYSE breakpoints. These portfolios have substantial cross-sectional dispersion in average returns and factor loadings.

Appendix C.5. Factor Return Statistics

Table C.19: Fama-French factor statistics (July 1963 – December 2023)

	MKT-RF	SMB	HML	RMW	CMA	UMD
Monthly Statistics (%)						
Mean	0.66	0.18	0.29	0.24	0.22	0.52
Std Dev	4.48	3.06	2.81	2.23	1.98	4.12
Min	-23.24	-16.39	-11.18	-18.36	-6.86	-34.72
Max	16.10	22.00	12.87	13.38	9.56	18.38
Skewness	-0.51	0.42	0.35	-0.12	0.18	-1.24
Kurtosis	2.12	2.84	2.45	3.21	1.89	8.76
Annualized Statistics (%)						
Mean	7.92	2.16	3.48	2.88	2.64	6.24
Std Dev	15.52	10.60	9.73	7.73	6.86	14.27
Sharpe Ratio	0.51	0.20	0.36	0.37	0.39	0.44
Risk Metrics						
Max Drawdown	-54.5%	-32.1%	-51.2%	-26.8%	-18.4%	-73.8%
95% VaR	-6.9%	-4.8%	-4.3%	-3.4%	-3.0%	-5.9%

Appendix C.6. Factor Correlations

Table C.20: Factor correlation matrix (1963–2023)

	MKT	SMB	HML	RMW	CMA	UMD
MKT	1.00					
SMB	0.28	1.00				
HML	-0.27	-0.10	1.00			
RMW	-0.23	-0.36	0.06	1.00		
CMA	-0.40	-0.04	0.70	0.05	1.00	
UMD	-0.13	0.01	-0.19	0.10	0.03	1.00

Key Observations..

- HML and CMA are highly correlated ($\rho = 0.70$), reflecting overlap in value and investment strategies
- Market has negative correlation with value (HML), quality (RMW), and investment (CMA)
- SMB and RMW are negatively correlated ($\rho = -0.36$)
- Momentum (UMD) is relatively uncorrelated with other factors

Appendix C.7. Alternative Factor Specifications

Table C.21: Alternative factor definitions

Factor	**Description**	**Source**
HML Devil	Industry-adjusted HML	AQR
BAB	Betting Against Beta	AQR
QMJ	Quality minus Junk	AQR
ME	Market Equity factor	Q-factor model
I/A	Investment-to-Assets	Q-factor model
ROE	Return on Equity	Q-factor model
MGMT	Mispricing (management)	Stambaugh-Yuan
PERF	Mispricing (performance)	Stambaugh-Yuan

Appendix C.8. International Factor Data

Table C.22: International factor data availability

Region	**Start Date**	**Factors**	**Source**
Developed (ex-US)	1990	FF5 + Mom	Ken French
Europe	1990	FF5 + Mom	Ken French
Japan	1990	FF5 + Mom	Ken French
Asia Pacific (ex-Japan)	1990	FF5 + Mom	Ken French
Emerging Markets	1990	FF3 + Mom	Ken French
Global	1990	FF5 + Mom	Ken French

Appendix C.9. Risk-Free Rate Data

Table C.23: Risk-free rate specifications

Rate	**Maturity**	**Usage**
1-month T-bill	30 days	Fama-French factors, standard benchmark
3-month T-bill	90 days	Money market proxy
1-year T-bill	1 year	Longer-horizon studies
Fed Funds Rate	Overnight	Short-term financing cost
LIBOR/SOFR	Various	International studies

Appendix C.10. Data Quality Issues

Known Issues..

1. **Delisting bias**: Returns may be missing or incorrect for delisted stocks. Shumway (1997) provides delisting return adjustments.
2. **Penny stocks**: Stocks with price below $5 can have spurious returns due to bid-ask bounce.
3. **Survivorship bias**: Excluding failed firms overstates returns.
4. **Look-ahead bias**: Using information not available at portfolio formation time.
5. **Revision bias**: Using revised (not originally reported) accounting data.

Standard Filters..

- Share code 10 or 11 (common stocks only)
- Exchange code 1, 2, or 3 (NYSE, AMEX, NASDAQ)
- Price $\geq$ $1 or $5 (to exclude penny stocks)
- Positive book equity
- At least 24 months of return history (for beta estimation)

Appendix C.11. Benchmark Index Data

Table C.24: Market benchmark indices

Index	**Coverage**	**Weighting**
CRSP VW Market	All NYSE/AMEX/NASDAQ	Value-weighted
S&P 500	500 large-cap US	Value-weighted
Russell 1000	1000 largest US	Value-weighted
Russell 2000	Next 2000 US	Value-weighted
MSCI World	Developed markets	Value-weighted
MSCI ACWI	All country world	Value-weighted

Appendix C.12. Data Dictionary

Table C.25: Factor data variable definitions

Variable	**Type**	**Description**
`date`	datetime	End of month/period date
`mkt_rf`	float	Market excess return (%)
`smb`	float	Size factor return (%)
`hml`	float	Value factor return (%)
`rmw`	float	Profitability factor return (%)
`cma`	float	Investment factor return (%)
`umd`	float	Momentum factor return (%)
`rf`	float	Risk-free rate (%)
`mktcap`	float	Market capitalization ($M)
`bm`	float	Book-to-market ratio

Variable	Type	Description
`op`	float	Operating profitability
`inv`	float	Investment (asset growth)
`ret`	float	Monthly return (%)
`retx`	float	Return excl. dividends (%)
`beta_mkt`	float	Market beta
`beta_smb`	float	Size beta
`beta_hml`	float	Value beta
`alpha`	float	Jensen's alpha (%)
`r2`	float	Regression R-squared
`idiovol`	float	Idiosyncratic volatility (%)

Appendix C.13. Accessing Factor Data

Python (pandas-datareader)..

```python
import pandas_datareader as pdr

# Fama-French factors
ff_factors = pdr.get_data_famafrench('F-F_Research_Data_Factors',
                                      start='1963-07')
ff5_factors = pdr.get_data_famafrench('F-F_Research_Data_5_Factors_2x3',
                                       start='1963-07')
mom_factor = pdr.get_data_famafrench('F-F_Momentum_Factor',
                                      start='1963-07')

# 25 Size/BM portfolios
portfolios_25 = pdr.get_data_famafrench('25_Portfolios_5x5',
                                         start='1963-07')
```

R (frenchdata package)..

```r
library(frenchdata)

# Download Fama-French factors
ff_factors <- download_french_data("Fama/French 3 Factors")
ff5_factors <- download_french_data("Fama/French 5 Factors (2x3)")

# 25 portfolios
portfolios <- download_french_data("25 Portfolios Formed on Size and Book-to-Market (5 x 5)")
```

Direct Download.. Kenneth French Data Library: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

AQR Data Library: https://www.aqr.com/Insights/Datasets

Appendix C.14. Recommended Citations

When using factor data, cite the original papers:

- FF3: Fama and French (1993) "Common risk factors in the returns on stocks and bonds"
- FF5: Fama and French (2015) "A five-factor asset pricing model"
- Momentum: Carhart (1997) "On persistence in mutual fund performance"
- Q-factors: Hou et al. (2015) "Digesting anomalies"
- BAB: Frazzini and Pedersen (2014) "Betting against beta"

Appendix D. Exercises and Research Questions

This appendix provides exercises for readers and suggests directions for future research. The exercises progress from conceptual understanding to computational implementation, and finally to open research questions.

Appendix D.1. Conceptual Exercises

1. **CAPM and Factor Pricing**: Consider two assets with the following characteristics:
   - Asset A: $\beta^{\text{mkt}} = 1.2$, $\beta^{\text{SMB}} = 0.5$, $\beta^{\text{HML}} = -0.3$
   - Asset B: $\beta^{\text{mkt}} = 0.8$, $\beta^{\text{SMB}} = -0.2$, $\beta^{\text{HML}} = 0.6$

   Using factor premia $\lambda_{\text{MKT}} = 6\%$, $\lambda_{\text{SMB}} = 2\%$, $\lambda_{\text{HML}} = 3\%$:

   (a) Compute expected excess returns for both assets under the three-factor model.
   (b) What would CAPM predict for these assets if we ignore SMB and HML?
   (c) Which asset has higher expected return? Is this consistent with its risk profile?
   (d) Construct a zero-beta portfolio using A and B that has zero market exposure.

2. **APT Arbitrage**: Suppose returns follow a two-factor model but one asset is mispriced:
   - Factor premia: $\lambda_1 = 4\%$, $\lambda_2 = 2\%$
   - Asset C: $\beta_1 = 1.0$, $\beta_2 = 0.5$, $\mathbb{E}[R_C] = 5.5\%$ (correct price)
   - Asset D: $\beta_1 = 0.5$, $\beta_2 = 1.0$, $\mathbb{E}[R_D] = 4.0\%$ (mispriced)

   (a) What should Asset D's expected return be according to APT?
   (b) Is Asset D overpriced or underpriced?
   (c) Construct an arbitrage portfolio using C, D, and the risk-free asset that exploits the mispricing.
   (d) What are the weights in your arbitrage portfolio?

3. **Hansen-Jagannathan Bounds**: Given the following test assets:
   - Portfolio 1: $\mathbb{E}[R] = 8\%$, $\sigma = 16\%$
   - Portfolio 2: $\mathbb{E}[R] = 5\%$, $\sigma = 10\%$
   - Correlation: $\rho_{12} = 0.4$
   - Risk-free rate: $r_f = 2\%$

(a) Compute the maximum Sharpe ratio attainable from these assets.
(b) What is the minimum SDF volatility implied by these assets?
(c) If a candidate SDF has $\sigma(m)/\mathbb{E}[m] = 0.3$, does it satisfy the HJ bound?
(d) How does adding a third uncorrelated asset with $\mathbb{E}[R] = 6\%$, $\sigma = 12\%$ change the bound?

4. **Fama-MacBeth Procedure**: You observe the following data over $T = 120$ months for $n = 25$ portfolios:
   - Average cross-sectional regression coefficient: $\bar{\gamma} = 0.5\%$ monthly
   - Time-series standard deviation of γ_t: $s_\gamma = 1.8\%$
   - Factor variance: $\sigma_f^2 = 0.04$ (monthly)

   (a) Calculate the Fama-MacBeth standard error of $\bar{\gamma}$.
   (b) Is the risk premium statistically significant at the 5% level?
   (c) Compute the Shanken correction factor, assuming $c = \lambda^2/\sigma_f^2 = 0.25$.
   (d) How does the Shanken-corrected t-statistic compare to the naive FM t-statistic?

5. **Implied vs. Realized Premia**: A factor has the following characteristics:
   - Historical average return: $\bar{f} = 4\%$ annually
   - Factor volatility: $\sigma_f = 8\%$
   - Market's implied factor portfolio weight: $w^f = 0.15$
   - Risk aversion: $\gamma = 3$

   (a) Calculate the implied factor premium using $\lambda^{\text{impl}} = \gamma \cdot \sigma_f^2 \cdot w^f/\sigma_f^2$.
   (b) Is the implied premium higher or lower than the realized premium?
   (c) List three possible explanations for the difference.
   (d) How would you combine implied and realized premia in practice?

Appendix D.2. Computational Exercises

These exercises use Python and the Fama-French data library.

1. **Data Download and Summary Statistics**:
   (a) Download Fama-French five-factor data from 1963 to present.
   (b) Compute monthly and annualized mean returns for each factor.
   (c) Calculate Sharpe ratios with and without risk-free rate adjustment.
   (d) Create a correlation matrix and identify highly correlated factor pairs.
   (e) Plot cumulative returns for all factors on a log scale.

2. **Fama-MacBeth Implementation**: Using the 25 Size/BM portfolios:
   (a) Estimate 60-month rolling betas for all portfolios against FF3 factors.
   (b) Run monthly cross-sectional regressions of returns on betas.
   (c) Compute Fama-MacBeth average premia and standard errors.
   (d) Implement the Shanken correction.
   (e) Compare your estimates to the published factor returns.

3. **GRS Test**: Evaluate the Fama-French five-factor model:

(a) Run time-series regressions for 25 portfolios on FF5 factors.
(b) Collect the alphas and residual covariance matrix.
(c) Compute the GRS test statistic.
(d) Test at the 5% significance level—is the model rejected?
(e) Repeat for FF3 and compare results.

4. **Implied Factor Premia**: Implement reverse optimization:
   (a) Download market cap weights for the 25 Size/BM portfolios.
   (b) Estimate factor betas for each portfolio.
   (c) Compute implied factor portfolio weights: $\boldsymbol{w}^f = \boldsymbol{B}^\top \boldsymbol{w}^{\text{mkt}}$.
   (d) Apply reverse optimization with $\gamma \in \{2, 3, 4, 5\}$.
   (e) Compare implied premia to historical averages.

5. **Time-Varying Premia**: Analyze the evolution of factor premia:
   (a) Compute 5-year rolling Fama-MacBeth premia for each factor.
   (b) Plot the time series of rolling premia.
   (c) Identify periods of negative value (HML) or size (SMB) premia.
   (d) Test for structural breaks using Chow or Andrews-Ploberger tests.
   (e) Correlate rolling premia with business cycle indicators.

6. **Factor Timing Strategy**: Implement a simple timing strategy:
   (a) Compute the timing signal $z_{k,t} = (\lambda_k^{\text{impl}} - \bar{\lambda}_k)/\sigma(\lambda_k)$.
   (b) Create long/short factor positions based on signal strength.
   (c) Backtest the strategy from 1980 to 2023.
   (d) Calculate Sharpe ratio, maximum drawdown, and turnover.
   (e) Compare to a static factor allocation.

Appendix D.3. Advanced Exercises

1. **GMM Estimation**: Implement GMM for factor pricing:
   (a) Define moment conditions: $\mathbb{E}[\boldsymbol{f}_t \otimes (\boldsymbol{R}_t - \boldsymbol{B}\boldsymbol{f}_t)] = \mathbf{0}$ and $\mathbb{E}[\boldsymbol{R}_t - \boldsymbol{B}\boldsymbol{\lambda}] = \mathbf{0}$.
   (b) Estimate using the identity weighting matrix (first-stage GMM).
   (c) Compute the optimal weighting matrix from first-stage residuals.
   (d) Re-estimate using efficient GMM.
   (e) Compute Hansen's J-test and interpret the result.

2. **SDF-Based Implied Premia**: Extract implied premia from the SDF:
   (a) Estimate SDF loadings $\boldsymbol{b}$ using $\min_{\boldsymbol{b}} \mathbb{E}[(1 - \boldsymbol{b}^\top \boldsymbol{f})^2]$ subject to $\mathbb{E}[(1 - \boldsymbol{b}^\top \boldsymbol{f})\boldsymbol{R}] = \mathbf{0}$.
   (b) Convert to factor premia: $\boldsymbol{\lambda} = -R_f \boldsymbol{\Omega}_f \boldsymbol{b}$.
   (c) Compare SDF-implied premia to Fama-MacBeth estimates.
   (d) Test whether the SDF satisfies Hansen-Jagannathan bounds.
   (e) Analyze how SDF loadings change over time.

3. **Black-Litterman for Factors**: Implement factor-level Black-Litterman:
   (a) Use implied factor premia as the prior.
   (b) Express views: "Value will outperform by 3%" (confidence 70%).

   (c) Compute posterior factor premia.
   (d) Map to optimal asset weights.
   (e) Sensitivity analysis: How does posterior change with confidence?

4. **Cross-Country Factor Premia**: Compare factor premia internationally:
   (a) Download Fama-French factors for US, Europe, Japan, and Emerging Markets.
   (b) Estimate premia for each region using Fama-MacBeth.
   (c) Test whether premia are equal across regions.
   (d) Analyze correlation of factor returns across regions.
   (e) Construct a global factor portfolio and assess diversification.

5. **Factor Model Horse Race**: Compare alternative factor models:
   (a) Estimate FF3, FF5, Q-factor (Hou-Xue-Zhang), and Stambaugh-Yuan models.
   (b) Compute GRS statistics for each model on common test assets.
   (c) Calculate Hansen-Jagannathan distances.
   (d) Rank models by explanatory power and parsimony.
   (e) Discuss economic interpretation of each model's factors.

Appendix D.4. Research Questions

The following open questions merit further investigation:

1. **Post-Publication Decay**: Factor premia appear to decline after academic publication. Is this due to arbitrage eliminating anomalies, data mining in the original study, or time-varying true premia? Can we predict which anomalies will decay fastest?
2. **Implied vs. Realized Premia Gap**: Implied premia from reverse optimization are systematically lower than historical averages. What explains this gap? Are investors irrational, or do they have information not captured by historical returns?
3. **Factor Timing Predictability**: Do implied premia predict future realized premia? What is the optimal forecast horizon? Can machine learning improve factor timing relative to simple linear models?
4. **Risk Aversion Estimation**: The implied risk aversion parameter γ ranges from 2 to 10 depending on methodology. Can we pin down γ using options data, surveys, or revealed preference? How sensitive are implied premia to γ?
5. **Conditional Factor Models**: How should we model time-variation in factor betas and premia? Are regime-switching models or smooth transition models better? Can we identify economic state variables that drive premium dynamics?
6. **Factor Model Spanning**: Do new factors provide incremental pricing information, or are they spanned by existing factors? How should we test for true factor versus repackaged exposure?
7. **Transaction Costs and Capacity**: What are the capacity constraints of factor strategies? How do implied premia change when accounting for realistic transaction costs and price impact?
8. **Firm-Level vs. Portfolio-Level Estimation**: Is it better to estimate factor premia using firm-level returns or sorted portfolios? What are the statistical trade-offs? How does the choice affect inference?

9. **Option-Implied Factor Premia**: Can we extract factor premia from equity index options? What information do option-implied moments provide about factor risk that is not in historical returns?
10. **ESG and Factor Premia**: How do environmental, social, and governance factors interact with traditional risk factors? Do ESG constraints affect implied factor premia? Is there an ESG risk premium?

Appendix D.5. Solutions and Hints

Selected solutions and hints are available:

Exercise 1a.. Expected excess return for Asset A:

$$\begin{aligned}\mathbb{E}[R_A] - r_f &= \beta_A^{\text{mkt}}\lambda_{\text{MKT}} + \beta_A^{\text{SMB}}\lambda_{\text{SMB}} + \beta_A^{\text{HML}}\lambda_{\text{HML}} \\ &= 1.2 \times 6\% + 0.5 \times 2\% + (-0.3) \times 3\% \\ &= 7.2\% + 1.0\% - 0.9\% = 7.3\%\end{aligned}$$

For Asset B:

$$\mathbb{E}[R_B] - r_f = 0.8 \times 6\% + (-0.2) \times 2\% + 0.6 \times 3\% = 6.2\%$$

Exercise 2a.. APT-implied expected return for Asset D:

$$\mathbb{E}[R_D] = r_f + \beta_1\lambda_1 + \beta_2\lambda_2 = r_f + 0.5 \times 4\% + 1.0 \times 2\% = r_f + 4\%$$

If the actual expected return is 4%, then Asset D is overpriced (expected return is 0% above risk-free, but should be 4%).

Exercise 3a.. Maximum Sharpe ratio from two assets:

$$\text{SR}_{\max} = \sqrt{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}}$$

With $\boldsymbol{\mu} = (0.06, 0.03)^\top$ (excess returns) and $\boldsymbol{\Sigma}$ computed from given volatilities and correlation:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 0.0256 & 0.0064 \\ 0.0064 & 0.0100 \end{pmatrix}$$

Computing $\boldsymbol{\Sigma}^{-1}$ and $\text{SR}_{\max} \approx 0.51$.

Exercise 4a-b.. Fama-MacBeth standard error:

$$\text{SE}_{\text{FM}} = \frac{s_\gamma}{\sqrt{T}} = \frac{0.018}{\sqrt{120}} = 0.00164 = 0.164\%$$

t-statistic: $t = 0.5\%/0.164\% = 3.05$, significant at 5% level.

Exercise 4c-d.. Shanken correction factor: $(1 + c) = 1.25$

Corrected SE: $\text{SE}_{\text{Shanken}} = 0.164\% \times \sqrt{1.25} = 0.183\%$

Corrected t-statistic: $t = 0.5\%/0.183\% = 2.73$, still significant but lower.

Full Solutions.. Complete solutions for all exercises, including Python code and detailed explanations, are available in the companion Jupyter notebook at:

`notebooks/factor_premia_notebook.ipynb`

The notebook includes:

- Executable code for all computational exercises
- Step-by-step solutions for conceptual exercises
- Extended commentary on research questions
- Additional data visualizations and sensitivity analyses