Cross-Sectional Asset Pricing with Unsystematic Risk*

Massimo Dello-Preite Raman Uppal Paolo Zaffaroni Irina Zviadadze

February 5, 2024

Abstract

Our objective is to price the cross-section of asset returns. In contrast to existing models that allow expected excess returns to reflect compensation only for systematic risk, we derive a stochastic discount factor (SDF) implied by the Arbitrage Pricing Theory and consistent with the equilibrium model of Merton (1987), in which there is compensation *also* for unsystematic risk. Empirically, we find that more than seventy percent of this SDF's variation is explained by *unsystematic risk*. Our SDF dominates traditional factor models and the state-of-the-art models of latent systematic risk in pricing the cross-section of asset returns in and out of sample.

JEL classification: C52, C58, G11, G12.

Keywords: Unsystematic risk, weak factors, factor models, model misspecification.

*Dello-Preite is affiliated with Imperial College London; email: `m.dello-preite18@imperial.ac.uk`. Uppal is affiliated with EDHEC Business School and CEPR; email: `Raman.Uppal@edhec.edu`. Zaffaroni is affiliated with Imperial College London; email: `P.Zaffaroni@imperial.ac.uk`. Irina Zviadadze is affiliated with HEC Paris and CEPR; email: `zviadadze@hec.fr`. We gratefully acknowledge comments from Torben Andersen, Harjoat Bhamra, Svetlana Bryzgalova, John Campbell, Ines Chaieb, Pierre Collin-Dufresne, Magnus Dahlquist, Victor DeMiguel, René Garcia, Lars Hansen, Valentin Haddad, Ravi Jagannathan, Christian Julliard, Robert Korajczyk, Stefan Nagel, Markus Pelger, Valentina Raponi, Cesare Robotti, Simon Rottke, Olivier Scaillet, Alberto Teguia, Viktor Todorov, Fabio Trojani, Dimitri Vayanos, Dacheng Xiu, and seminar and conference participants at Bauer College of Business, Bocconi University, Carey Business School, EDHEC, EPFL, Erasmus University Rotterdam, ESSEC, Federal Reserve Bank of New York, Fulcrum Asset Management, HEC Paris, HEC Lausanne, HEC Montreal, Inquire Europe, Imperial College, INSEAD, Judge Business School, Kellogg School of Management, London Business School, London School of Economics, Lund University, Saïd Business School, SKEMA, Stockholm Business School, Stockholm School of Economics, Tinbergen Institute, Toulouse Business School, University of Chicago, University of Geneva, University of Lausanne, University of Sussex, University of York, Conference in Financial Economics Research at Arison School of Business, AFA, Asset Pricing Conference at York University, Bristol Financial Markets Conference, CEPR Advanced Forum for Financial Economics, Corporate Policies and Asset Prices (COAP) Conference, Durham Asset Pricing Workshop, EEA-ESEM Barcelona, EFA, Financial Econometrics Conference at Lancaster University, Financial Econometrics Conference at Toulouse School of Economics, Financial Markets Conference at Bristol University, Inquire Europe, International Conference on Computational and Financial Econometrics, Junior European Finance Seminar, SAFE Asset Pricing Workshop, SoFiE Conference, SFS Cavalcade North America, UBC Finance Summer Conference, and Women in Economics Seminar Series. An earlier version of this manuscript was circulated under the title "What is Missing in Asset Pricing Factor Models."

1 Introduction

A major challenge in asset pricing is to explain the cross-section of asset returns. The earliest model proposed to explain the cross-section of stock returns is the Capital Asset Pricing Model (CAPM) of Sharpe (1964). In the world of frictionless markets assumed by the CAPM, investors hold a perfectly diversified portfolio of risky assets, which, in equilibrium, is the market portfolio. In the data, however, risk compensation for asset exposures to the market factor, which is the only source of systematic risk in the CAPM, performs poorly in explaining the cross-section of expected stock returns. Consequently, researchers have continued to assume that in equilibrium investors hold perfectly diversified portfolios and have empirically examined a large number of alternative proxies for systematic risk, leading to a factor zoo (Cochrane, 2011). However, virtually all models featuring factors from this zoo have sizable pricing errors (Bryzgalova, Huang, and Julliard, 2023).

Merton (1987), in contrast to Sharpe (1964), relaxes the assumption of frictionless markets and derives an equilibrium in which investors hold portfolios that are not fully diversified, consistent with a large body of empirical evidence.[1] We show that the SDF of an agent who can trade all assets in the economy of Merton (1987) loads on both systematic risk (i.e., common risk factors) and unsystematic risk (i.e., asset-return shocks unexplained by systematic risk factors). Motivated by the risk-return tradeoff implied by Merton (1987), in this paper, we study how allowing compensation for unsystematic risk provides an avenue for explaining the cross-section of expected asset returns and resolving the factor zoo.

As the foundation for our analysis of the role of unsystematic risk in pricing assets, we use the Arbitrage Pricing Theory (APT) of Ross (1976, 1977), Chamberlain (1983), and Chamberlain and Rothschild (1983). The APT provides an ideal framework because, in contrast to what is commonly believed, it allows for *both* systematic and *asset-specific* components in expected excess returns. The asset-specific components are unrelated to compensation for asset exposures to systematic risk factors and satisfy an asymptotic no-arbitrage restriction. Thus, the APT permits us to entertain, in a no-arbitrage setting, the possibility that these asset-specific components in expected excess returns represent compensation for unsystematic risk. We allow unsystematic risk to include both purely asset-specific (i.e., idiosyncratic) risk and weak factors (Lettau and Pelger, 2020).

[1]For evidence on portfolio underdiversification, see Campbell (2006) and the literature that builds on it.

Our first contribution is to derive the APT-implied SDF. We decompose this SDF into two parts: a traditional *systematic SDF component* that reflects systematic risk factors and a new *unsystematic SDF component* that reflects asset-return shocks unexplained by systematic risk factors.[2] We show that for an asymptotically large number of assets, the structure of the SDF implied by the APT coincides with that of the SDF in the equilibrium model of Merton (1987). We then demonstrate that the asset-specific components in expected excess returns that are typically interpreted as pricing errors from the perspective of the traditional systematic SDF component, instead, reflect compensation for unsystematic risk under the APT-implied SDF.

Our second contribution is to provide empirical support for the insight that unsystematic risk is priced and to quantify its importance. To do this, we first overcome the challenge that the APT-implied SDF, being dependent on latent systematic factors and unsystematic shocks, is empirically infeasible. We develop a projection-based SDF, which is the APT-implied SDF projected on a set of basis assets and a risk-free asset. Next, we show that the projection-based SDF converges in probability to the APT-implied SDF as the number of basis assets increases.

Then, using data for monthly returns on 202 portfolios of stocks analyzed in Giglio and Xiu (2017), we estimate the APT model of asset returns and construct the corresponding SDF. To guard against in-sample overfitting, we use cross-validation to estimate the two key parameters of the APT: the number of latent factors and the no-arbitrage bound on the compensation for unsystematic risk. The number of latent factors determines the size and variability of the systematic SDF component, whereas the no-arbitrage bound determines whether unsystematic risk is compensated. We use the Hansen and Jagannathan (1997) (HJ) distance as a selection metric for these two parameters.

Our key finding is that compensation for unsystematic risk is economically significant, with the unsystematic SDF component explaining 72.60 percent of the APT-implied SDF's variation. Thus, unsystematic risk plays a major role in pricing the cross-section of asset returns, despite the risk premia associated with unsystematic shocks being small on average. Furthermore, we find that the two constituents of unsystematic risk, purely asset-specific

[2] In the APT, latent factors capture systematic risk, and thus, ex-ante, we are agnostic about which observable variables drive common variation in asset returns. Our definition of systematic factors allows for both fundamental risk factors and behavioral factors to capture common variation in asset returns, as in Kozak, Nagel, and Santosh (2018).

risk and weak factors, contribute about equally to the unsystematic SDF component's variation.

Our third contribution is identifying the reasons for the poor performance of popular candidate factor models used to price a cross-section of stock returns. We study three models traditionally used in the literature: (i) a model with the market return as the only systematic risk factor, as suggested by the CAPM of Sharpe (1964), (ii) a model with the consumption-mimicking portfolio as the only systematic risk factor, as implied by the Consumption Capital Asset Pricing Model (C-CAPM) of Breeden (1979), and (iii) the three-factor model (FF3) of Fama and French (1993).[3] Both theoretically and empirically, we identify and characterize the wedge between the APT-implied SDF and the SDFs implied by these candidate factor models. We document, in sample and out of sample, that although these factor models omit some sources of systematic risk, the major source of model misspecification is *unsystematic* risk, which, in these candidate models, similarly to virtually all other factor models, is assumed to have zero compensation.

To investigate the robustness of our findings, we undertake three exercises. First, we confirm our findings by estimating the APT-implied SDFs on different sets of basis assets. Specifically, we consider the dataset for fifty anomaly portfolios used in Kozak, Nagel, and Santosh (2020) and the dataset for seventy-four characteristic-based portfolios used in Lettau and Pelger (2020). We find that the relative importance of unsystematic risk is even higher than that documented in our original dataset of 202 portfolios of stock returns. Second, we show that, compared to models based on either observable candidate factors or latent factors, our APT-implied SDF has smaller *out-of-sample* pricing errors in both time series and cross-section.

Third, we build on the recent work of Kozak et al. (2020) and Lettau and Pelger (2020) that has generalized the conventional definition of systematic risk to include not just standard factors explaining the covariance matrix of asset returns but also those factors that have a high price of risk. We find that both Kozak et al. (2020) and Lettau and Pelger (2020), consistent with their definitions of systematic risk, capture, at least partially, weak-factor risk. However, when we correct the SDFs implied by their models to obtain the APT-implied SDF, we document that (i) the correction component is still sizable and

[3]We do not analyze other candidate factor models because the conclusions we draw from our empirical analysis of these three candidate factor models apply to virtually any candidate factor model that is based on the premise that only systematic risk is compensated in financial markets.

represents omitted sources of priced unsystematic risk and (ii) our APT-implied SDF has substantially smaller pricing errors than their SDFs, both in sample and out of sample.

To understand the implications of the sizable importance of unsystematic risk for investment decisions, we examine 457 trading strategies. We find that, of the strategies with large compensation for unsystematic risk, some capture behavioral biases—for example, the performance factor (Stambaugh and Yuan, 2017), the long-horizon financial factor (Daniel, Hirshleifer, and Sun, 2020a), the factor reflecting expectations about future earnings (La Porta, 1996), and the momentum factor (Jegadeesh and Titman, 1993)—while others reflect market frictions—for example, the betting-against-beta factor (Frazzini and Pedersen, 2014) and distress risk (Campbell, Hilscher, and Szilagyi, 2008). The compensation offered by these strategies for bearing unsystematic risk is large; for instance, the premium for bearing the unsystematic risk for the 12-month momentum strategy of Jegadeesh and Titman (1993) is 8.27 percent per annum.

Our finding that unsystematic risk is priced also sheds light on the development of cross-sectional asset pricing models. When the CAPM failed to explain a cross-section of stock returns, the response was to search for additional proxies for systematic factors. For instance, momentum (Jegadeesh and Titman, 1993), value (Fama and French, 2015), and investment (Hou, Xue, and Zhang, 2015) have attracted attention as successful proxies. We find, however, that these variables correlate more highly with the unsystematic SDF component rather than the systematic SDF component. That is, these observable variables represent weak factors that matter primarily for explaining the cross-section of expected returns, as opposed to strong factors, which also explain the covariance of asset returns. Empirically, we show that a factor model that includes even a large number of these observable weak factors fails to capture priced unsystematic risk.

Recently, Daniel, Mota, Rottke, and Santos (2020b) have proposed removing unpriced risk from returns on characteristic-sorted portfolios that are used as risk factors in asset-pricing models. We find that these newly constructed factors do not capture all priced risks. The adjusted five factors of Daniel et al. (2020b) explain 41.07 percent of the variation in the systematic SDF component but only 13.41 percent of the variation in the unsystematic SDF component and 31.65 percent of the variation of the overall APT-implied SDF.

Our work relates to several strands of the literature. First, we contribute to the literature that uses a large cross-section of asset returns to examine the risk-return tradeoff implied by factor models. Unlike existing studies, our focus lies in examining a distinct aspect of the risk-return tradeoff. Specifically, we allow compensation for asset-specific risk, in addition to compensation for systematic risk factors (Kozak et al., 2018; Pelger, 2020; Giglio and Xiu, 2021), weak factors (Kozak et al., 2020; Lettau and Pelger, 2020), and semi-strong factors (Giglio, Xiu, and Zhang, 2021b). Our empirical results reinforce the message of Daniel and Titman (1997) that common factors do not explain the cross-section of asset returns and strengthen it by illustrating that this is true even in the no-arbitrage setting. A major benefit of our no-arbitrage framework is that it allows us to construct an SDF, which can then be used to quantify the importance of *unsystematic risk* for pricing assets that, as described above, we find is substantial.

Second, because our methodology allows us to correct the misspecified SDF implied by a given candidate factor model, we contribute to the literature that studies misspecification of the SDF and develops methods to characterize the wedge between the misspecified and *admissible* SDFs, that is, SDFs pricing assets without error. Our approach complements existing methods, such as those by Hansen and Jagannathan (1997), Almeida and Garcia (2012), Ghosh, Julliard, and Taylor (2017), Sandulescu and Schneider (2021), and Korsaye, Quaini, and Trojani (2021), by providing a method for quantifying the contribution of priced unsystematic risk in correcting misspecified SDFs.

Third, we contribute to the factor-zoo literature (see, e.g., Cochrane, 2011; Harvey, Liu, and Zhu, 2015; Kogan and Tian, 2015), which has proposed hundreds of variables that can potentially proxy for systematic risk priced in the cross-section of asset returns. Our contribution is to identify that what asset-pricing factor models are missing is compensation for unsystematic risk rather than a yet-undiscovered proxy for systematic risk.

Furthermore, our paper complements methodological advances aimed at taming the factor zoo. Feng, Giglio, and Xiu (2020), Freyberger, Neuhierl, and Weber (2020), Giglio, Liao, and Xiu (2021a), and Bryzgalova et al. (2023) propose model-selection methods to discipline the proliferation of factors and account for data snooping when performing multiple-hypothesis testing in linear asset-pricing models. Our focus is different; however, as a by-product of our analysis, we provide a method that establishes whether a set of arbitrary observed variables span the systematic and unsystematic SDF components and quantifies the risk prices of these variables.

Our work is also related to the work on the idiosyncratic-volatility puzzle, which studies the relation between the compensation for asset-specific risk and the *volatility* of asset-specific shocks; see, e.g., Fama and MacBeth (1973) and Ang, Hodrick, Xing, and Zhang (2006), and the comprehensive review by Bali, Engle, and Murray (2016). Complementary to this empirical literature, we construct an SDF and show that compensation for unsystematic risk represents the negative covariance between this SDF and unsystematic shocks (rather than their *volatility*). Furthermore, we find that the returns of the idiosyncratic-volatility factors of Ali, Hwang, and Trombley (2003) and Ang et al. (2006) explain less than 10 percent of the variation of our unsystematic SDF component. Thus, what the traditional risk-return tradeoff is missing is not just the idiosyncratic-volatility factor.

The rest of the paper is organized as follows. Section 2 presents our theoretical results for constructing an APT-implied SDF that allows for nonzero compensation for both systematic and unsystematic risk. Section 3 explains how to estimate this SDF. Section 4 describes the data we use to undertake our empirical analysis. Section 5 presents our empirical findings. Section 6 shows that these findings are robust out-of-sample, for different sets of basis assets, and for more general definitions of systematic risk. Section 7 provides an example of an equilibrium model in which unsystematic risk is priced. We conclude in Section 8. The Internet Appendix reports proofs, the estimation algorithm, and additional results.

2 Constructing the APT-implied SDF

In this section, and throughout the paper, we consider two cases. In the first case, we derive the SDF under the classical APT, in which the systematic risk factors are latent. In the second case, we construct an APT-implied SDF by correcting for misspecification an arbitrary candidate linear factor model that specifies the sources of systematic risk. We correct for misspecification arising from omitted: (i) sources of systematic risk, (ii) compensation for unsystematic risk, and (iii) time variation in risk premia. In the empirical application, the first case allows us to shed light on the quantitative importance of unsystematic risk, and the second case allows us to investigate the main sources of misspecification in popular factor models. For both cases, we explain how to ensure the positivity and empirical feasibility of the APT-implied SDF.

In our analysis, we use the following notation. Let an N-dimensional vector $R_{t+1} = (R_{1,t+1}, R_{2,t+1}, \dots, R_{N,t+1})'$ denote the vector of gross returns of N risky assets between t and $t+1$. Let R_f be the gross return on a risk-free asset over the same period.[4] Let $\mathbb{E}(\cdot)$ denote the expectation operator and 1_N indicate an $N \times 1$ vector of ones, so that $\mathbb{E}(R_{t+1} - R_f 1_N)$ represents the vector of expected excess returns on the N assets. Let f_{t+1} be a $K \times 1$ vector of systematic risk factors with a finite number $K < N$ and a $K \times K$ positive definite covariance matrix V_f. Let $\beta = (\beta_1, \beta_2, \dots, \beta_N)'$ be an $N \times K$ full-rank matrix of loadings of asset returns on the systematic factors f_{t+1}. The notation 0_N indicates an $N \times 1$ vector of zeros. For deterministic sequences $\{a_N\}$ and $\{b_N\}$ the notation $a_N = O(b_N)$ means that, as $N \to \infty$, $|a_N|/b_N < \delta$, where $\delta > 0$ is some finite number, and $a_N = o(b_N)$ means that $|a_N|/b_N \to 0$.

2.1 The SDF under the Arbitrage Pricing Theory (APT)

The APT of Ross (1976, 1977), Chamberlain (1983), and Chamberlain and Rothschild (1983) is our working assumption about the true data-generating process for asset returns. There are several advantages to choosing the APT as the null model. First, the APT is a flexible model that does not take a stand on systematic risk factors. Second, it is a no-arbitrage model; the absence of arbitrage opportunities implies the existence of an SDF. Third, and more importantly for our purpose, the APT allows for asset-specific components in expected excess returns that are unrelated to compensation for systematic risk.

The classical APT builds on the following two assumptions.

Assumption 2.1 (Linear Factor Model)**.** *The vector R_{t+1} of gross asset returns satisfies*

$$R_{t+1} = \mathbb{E}(R_{t+1}) + \beta\big(f_{t+1} - \mathbb{E}(f_{t+1})\big) + e_{t+1},$$

where the vector e_{t+1} has $\mathbb{E}(e_{t+1}) = 0_N$ and a positive-definite $N \times N$ covariance matrix V_e, whose eigenvalues are uniformly bounded and bounded away from zero. The shocks e_{t+1} constitute unsystematic risk and are uncorrelated with the K systematic factors f_{t+1}, implying the covariance matrix of returns is

$$V_R = \beta V_f \beta' + V_e. \tag{1}$$

[4] If a risk-free asset does not exist, one can use the return on the minimum-variance portfolio or the return on a zero-beta portfolio instead.

The systematic factors f_{t+1} are often referred to as strong factors, that is, factors that explain substantial comovement in asset returns and, for which, as $N \to \infty$, $N^{-1}\beta' V_e{}^{-1}\beta \longrightarrow E$, where E is some arbitrary symmetric positive definite $K \times K$ matrix.

If V_e is diagonal, the unsystematic risk arises from purely asset-specific shocks. However, in the APT of Chamberlain (1983) and Chamberlain and Rothschild (1983), V_e is not restricted to be diagonal. The case of a non-diagonal matrix V_e can accommodate, in addition to purely asset-specific shocks, the presence of weak factors f_{t+1}^{weak}. We define weak factors, as in Lettau and Pelger (2020), as factors that affect only a subset of the underlying assets or all assets but marginally, such that the sum of squared risk exposures remains bounded.[5]

Assumption 2.2 (Asymptotic No Arbitrage)**.** *There is no sequence of portfolios containing N risky assets with weights $w = (w_1, w_2, \ldots, w_N)'$, for which, as $N \to \infty$:*

$$\mathrm{var}(R_{t+1}' w) \to 0 \qquad \textit{and} \qquad (\mathbb{E}(R_{t+1}) - R_f 1_N)' w \geq \delta > 0,$$

where δ denotes an arbitrary positive scalar.

Assumptions 2.1 and 2.2 imply that, under the APT, asset excess returns satisfy

$$R_{t+1} - R_f 1_N = a + \beta\lambda + \beta(f_{t+1} - \mathbb{E}(f_{t+1})) + e_{t+1}, \tag{2}$$

with expected excess returns,

$$\mathbb{E}(R_{t+1} - R_f 1_N) = a + \beta\lambda, \tag{3}$$

containing two components: a and $\beta\lambda$. The $K \times 1$ vector of risk premia λ represents the compensations for a unit of assets' exposures to the factors f_{t+1}.[6] The $N \times 1$ vector $a = (\mathbb{E}(R_{t+1}) - R_f 1_N) - \beta\lambda$, typically referred to as the vector of pricing errors, satisfies the following *no-arbitrage restriction*, where δ_{apt} is some arbitrary positive scalar:

$$a' V_e^{-1} a \leq \delta_{\text{apt}} < \infty. \tag{4}$$

We now provide the APT-implied SDF and show that for this SDF, the vector a represents compensation for bearing unsystematic risk e_{t+1}.

[5]Mathematically, if β^{weak} is a matrix of loadings of returns R_{t+1} on K^{weak} weak factors f_{t+1}^{weak}, then, as $N \to \infty$, this matrix satisfies $\beta^{\text{weak}\prime}\beta^{\text{weak}} \to E$, where E is some symmetric positive-definite $K^{\text{weak}} \times K^{\text{weak}}$ matrix (see Onatski, 2012).

[6]Ingersoll (1984) derives the precise condition for λ to exist and shows the result that $\lambda = \lim_{N\to\infty} \left(\beta' V_e^{-1}\beta\right)^{-1} \beta' V_e^{-1}(\mathbb{E}(R_{t+1}) - R_f 1_N)$.

Proposition 1 (The APT-implied SDF). *The APT-implied SDF M_{t+1} is*

$$M_{t+1} = M^{\beta}_{t+1} + M^{a}_{t+1}, \quad \textit{where} \tag{5}$$

$$M^{\beta}_{t+1} = \frac{1}{R_f} - \frac{\lambda' V_f^{-1}}{R_f}(f_{t+1} - \mathbb{E}(f_{t+1})) \quad \textit{and}$$

$$M^{a}_{t+1} = -\frac{a' V_e^{-1}}{R_f} e_{t+1},$$

with $\operatorname{cov}(M^{\beta}_{t+1}, M^{a}_{t+1}) = 0$ *and*

$$a = -\operatorname{cov}(M_{t+1}, e_{t+1}) \times R_f = -\operatorname{cov}(M^{a}_{t+1}, R_{t+1}) \times R_f. \tag{6}$$

The component M^{β}_{t+1} of the APT-implied SDF is a linear function of the systematic risk factors f_{t+1}, and therefore we refer to M^{β}_{t+1} as the *systematic* SDF component. The component M^{a}_{t+1} of the APT-implied SDF is a linear function of the unsystematic shocks e_{t+1}, and therefore we refer to M^{a}_{t+1} as the *unsystematic* SDF component.

The presence of the unsystematic component M^{a}_{t+1} in the APT-implied SDF M_{t+1} is a deviation from the traditional approach relying only on systematic risk factors. This deviation leads to a key insight underlying our approach, which is the *interpretation* of the vector a. The traditional approach assumes that the SDF consists only of the systematic SDF component M^{β}_{t+1}. Because the expectation of the product of the systematic SDF component with excess returns is equal to a instead of a vector of zeros,

$$a = \mathbb{E}(M^{\beta}_{t+1}(R_{t+1} - R_f 1_N)) \times R_f, \tag{7}$$

in the traditional approach, the vector a has the interpretation of pricing errors. In contrast, expression (6) indicates that the vector a is the negative covariance between the APT-implied SDF and unsystematic shocks. Thus, by definition of the risk premium, the vector a represents the vector of risk compensations associated with unsystematic shocks e_{t+1}. This interpretation of the vector a paves the way for a quantitative assessment of *priced* unsystematic risk in financial markets that we undertake in our empirical analysis.

2.2 Constructing the APT-implied SDF Empirically

Empirically, there are two challenges in constructing the APT-implied SDF. First, the linear SDF (5) may not always be strictly positive, which could result in negative asset prices leading to arbitrage opportunities. To guarantee the SDF's positivity, following Gourieroux and Monfort (2007) and Ghosh et al. (2017), we model the SDF as an exponential function of payoffs

$$
\begin{aligned}
M_{\exp,t+1} &= M^{\beta}_{\exp,t+1} \times M^{a}_{\exp,t+1}, \quad \text{where} \qquad (8)\\
M^{\beta}_{\exp,t+1} &= \frac{1}{R_f} \times \exp\Big(- \lambda' V_f^{-1} (f_{t+1} - \mathbb{E}(f_{t+1})) - \frac{1}{2} \lambda' V_f^{-1} \lambda \Big) \quad \text{and}\\
M^{a}_{\exp,t+1} &= \exp\Big(- a' V_e^{-1} e_{t+1} - \frac{1}{2} a' V_e^{-1} a \Big).
\end{aligned}
$$

The second challenge is that the SDF (8) is not feasible empirically because it depends on the unobserved factors f_{t+1} and shocks e_{t+1}. To overcome this challenge, we replace $M^{\beta}_{\exp,t+1}$ and $M^{a}_{\exp,t+1}$ with an exponential function of the linear projections of M^{a}_{t+1} and M^{β}_{t+1} on the set of the risk-free and risky assets:

$$
\hat{M}^{\beta}_{\exp,t+1} = \frac{1}{R_f} \times \exp\Big(- (\beta\lambda)' V_R^{-1} (R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2} (\beta\lambda)' V_R^{-1} \beta\lambda \Big) \quad \text{and} \qquad (9)
$$

$$
\hat{M}^{a}_{\exp,t+1} = \exp\Big(- a' V_R^{-1} (R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2} a' V_R^{-1} a \Big), \qquad (10)
$$

where the symbol $\hat{\cdot}$ denotes a function of a linear projection, the covariance matrix of asset returns V_R satisfies equation (1), and the expected excess returns $\mathbb{E}(R_{t+1}) - R_f 1_N$ satisfy equation (3). This leads to a feasible and positive SDF.

$$
\hat{M}_{\exp,t+1} = \hat{M}^{\beta}_{\exp,t+1} \times \hat{M}^{a}_{\exp,t+1}. \qquad (11)
$$

The next proposition shows that, as $N \to \infty$, our feasible SDF (11) recovers the admissible SDF (8):

Proposition 2 (Asymptotic Properties of the Feasible SDF)**.** *Under Assumptions 2.1 and 2.2 of the APT and the assumption that the systematic factors f_{t+1} and unsystematic shocks e_{t+1} are jointly Gaussian, the SDF in equation (8) is admissible. Furthermore, if $\beta' V_e^{-1} a = o(N^{\frac{1}{2}})$, then, as $N \to \infty$, the following results hold*[7]

$$
\hat{M}^{a}_{\exp,t+1} - M^{a}_{\exp,t+1} \xrightarrow{p} 0, \quad \hat{M}^{\beta}_{\exp,t+1} - M^{\beta}_{\exp,t+1} \xrightarrow{p} 0, \quad \mathrm{cov}(\hat{M}^{\beta}_{\exp,t+1}, \hat{M}^{a}_{\exp,t+1}) \to 0.
$$

Proposition 2 does not rely on V_e being diagonal and, therefore, allows for the presence of weak factors in asset-return shocks e_{t+1}. Thus, our methodology characterizes the pricing implications of weak factors *without* needing to estimate them separately from purely

[7] The the no-arbitrage constraint (4) and the Cauchy–Schwarz inequality imply $\beta' V_e a = O(N^{\frac{1}{2}})$, which is a mild form of asymptotic orthogonality. We need a slightly stronger condition to derive the exact limit of the projection-based SDF.

asset-specific shocks. This is a major advantage of our approach given that weak latent factors cannot be estimated consistently (Lettau and Pelger, 2020).[8] Internet Appendix IA.6 discusses explicitly the case of weak factors in e_{t+1} and shows that we can construct the unsystematic SDF component even in the presence of weak factors.

2.3 What is Missing in Popular SDF Models

So far, we have considered the case of the APT model in which the systematic risk factors are latent. In this section, we show how our APT model can also be used to shed light on the poor performance of an arbitrary candidate factor model based on a given set of observable factors as sources of systematic risk. We use the superscript "can" for variables related to a candidate factor model.

A standard candidate factor model has K^{can} observable risk factors f_{t+1}^{can} implying that the expected excess asset returns reflect compensation for exposures to these risk factors. A classic example of a candidate factor model is the CAPM with $K^{\text{can}} = 1$ systematic risk factor represented by the market excess return and compensation for unsystematic risk $a^{\text{can}} = 0_N$. Viewed through the lenses of the APT, candidate factor models suffer from possibly three sources of misspecification. First, these models may omit systematic risk factors. Second, these models may omit compensation for unsystematic risk. Candidate models may also be misspecified because they do not account for time variation in prices of risk or risk exposures. We address the first two sources of misspecification in Section 2.3.1 and the third source of misspecification in Section 2.3.2.

2.3.1 Accounting for omitted systematic risk factors or omitted compensation for unsystematic risk

Let β^{can} denote an $N \times K^{\text{can}}$ matrix of loadings of asset returns on the candidate factors f_{t+1}^{can} with the factor covariance matrix $V_{f^{\text{can}}}$, and λ^{can} denote a $K^{\text{can}} \times 1$ vector of risk premia for unit exposures to these factors.[9] Given this candidate factor model, asset returns satisfy

$$R_{t+1} - R_f 1_N = \alpha + \beta^{\text{can}}\lambda^{\text{can}} + \beta^{\text{can}}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) + \varepsilon_{t+1}, \tag{12}$$

[8] Lettau and Pelger (2020) develop an approach to pin down noisy estimates of weak factors, and then use these to construct the SDF. Giglio et al. (2021b) provide a consistent estimator of the SDF spanned by latent semi-strong factors (as defined by Chudik, Pesaran, and Tosetti, 2011).

[9] We assume that the vector of candidate factors does not include spurious factors.

where the vector $\alpha = (\mathbb{E}(R_{t+1}) - R_f 1_N) - \beta^{\text{can}}\lambda^{\text{can}}$ captures the cross-sectional variation in expected excess returns left unexplained by compensation for asset exposures to systematic risk factors f_{t+1}^{can}, and the vector ε_{t+1}, with positive-definite covariance matrix V_ε, captures the return variation not explained by the candidate factors f_{t+1}^{can}.

The candidate factor model implies a linear SDF

$$M_{t+1}^{\beta,\text{can}} = \frac{1}{R_f} - \frac{(\lambda^{\text{can}})' V_{f^{\text{can}}}^{-1}}{R_f}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})),$$

which, if $\alpha \neq 0_N$, values asset returns with pricing errors $\alpha = \mathbb{E}(M_{t+1}^{\beta,\text{can}}(R_{t+1} - R_f 1_N)) \times R_f$.

Traditionally, the presence of pricing errors in a factor model has been associated with omitted systematic risk factors. If the candidate factor model omits K^{mis} systematic factors f_{t+1}^{mis} with the positive-definite covariance matrix $V_{f^{mis}}$ (that is, the first K^{mis} eigenvalues of covariance matrix V_ε are unbounded with the remaining eigenvalues uniformly bounded, and the smallest eigenvalue strictly positive), then ε_{t+1} exhibits the following factor structure

$$\varepsilon_{t+1} = \beta^{\text{mis}}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) + e_{t+1},$$

where an $N \times K^{\text{mis}}$ matrix β^{mis} denotes the matrix of asset-return exposures to these omitted risk factors. Consequently, the covariance matrix of ε_{t+1}, V_ε, satisfies[10]

$$V_\varepsilon = \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e. \tag{13}$$

Let the $K^{\text{mis}} \times 1$ vector λ^{mis} denote the compensation for unit exposures of asset returns to the omitted factors f_{t+1}^{mis}. The vector α includes the term $\beta^{\text{mis}}\lambda^{\text{mis}}$ reflecting risk premia for omitted sources of systematic risk. Our approach, in contrast to conventional thinking, allows α to *also* include compensation for unsystematic risk, a, that is

$$\alpha = \beta^{\text{mis}}\lambda^{\text{mis}} + a, \tag{14}$$

where, by no arbitrage, for some constant δ^*_{apt}, we have:[11]

$$\alpha' V_\varepsilon^{-1} \alpha \leq \delta^*_{\text{apt}}. \tag{15}$$

Next, we show how to correct the linear SDF $M_{t+1}^{\beta,\text{can}}$ of the misspecified candidate factor model of asset returns to obtain the APT-implied SDF M_{t+1}.

---

[10] Without loss of generality, we assume that these omitted factors f_{t+1}^{mis} are orthogonal to the candidate factors f_{t+1}^{can}. Internet Appendix IA.8 shows that the admissible SDF is invariant to our assumption about the correlation structure between the candidate and missing factors.

[11] Note that $\delta^*_{\text{apt}} = \delta_{\text{apt}} + \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} \lambda^{\text{mis}} + o(1)$, with δ_{apt} defined in equation (4). The no-arbitrage restriction in equation (15) is asymptotically equivalent to that in equation (4), in the sense that δ^*_{apt} is finite if and only if δ_{apt} is finite. A formal proof of this result is available upon request.

Proposition 3 (Correcting a Misspecified Linear SDF). *Under Assumptions 2.1 and 2.2 of the APT, given the candidate SDF $M_{t+1}^{\beta,\text{can}}$, there exists an SDF M_{t+1} that prices assets without any pricing errors, such that*

$$M_{t+1} = M_{t+1}^{\beta,\text{can}} + M_{t+1}^{\alpha} = M_{t+1}^{\beta,\text{can}} + \underbrace{(M_{t+1}^{\beta,\text{mis}} + M_{t+1}^{a})}_{=M_{t+1}^{\alpha}}, \quad \textit{where}$$

$$M_{t+1}^{\beta,\text{mis}} = -\frac{(\lambda^{\text{mis}})' V_{f^{\text{mis}}}^{-1}}{R_f}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) \qquad \textit{and} \qquad M_{t+1}^{a} = -\frac{a' V_e^{-1}}{R_f} e_{t+1},$$

with $\operatorname{cov}(M_{t+1}^{\beta,\text{can}}, M_{t+1}^{a}) = 0$*,* $\operatorname{cov}(M_{t+1}^{a}, M_{t+1}^{\beta,\text{mis}}) = 0$*, and* $\operatorname{cov}(M_{t+1}^{\beta,\text{can}}, M_{t+1}^{\beta,\text{mis}}) = 0$*.*

Given that asset returns satisfy the assumptions of the APT, the SDF M_{t+1} that prices assets without any errors is the APT-implied SDF. The wedge between the APT-implied SDF M_{t+1} and the candidate SDF $M_{t+1}^{\beta,\text{can}}$ is a correction term labeled M_{t+1}^{α} that includes *two* components: $M_{t+1}^{\beta,\text{mis}}$ and M_{t+1}^{a}. The first component $M_{t+1}^{\beta,\text{mis}}$ captures pricing of the *systematic* risk factors f_{t+1}^{mis} omitted in the candidate factor model. The second component M_{t+1}^{a} captures pricing of the *unsystematic* sources of risk e_{t+1}.

The presence of the unsystematic SDF component M_{t+1}^{a} in the correction term M_{t+1}^{α} changes the direction of the quest for an asset-pricing model that explains the cross-section of expected excess returns. Candidate factor models may be misspecified because of missing systematic risk, the search for which has been the focus of the existing literature. However, Proposition 3 shows that candidate models may also be misspecified because they omit compensation for unsystematic risk. What matters most—omitted systematic risk or nonzero compensation for unsystematic risk—is an empirical question that we answer in this paper.

In practice, for the same reasons as explained in Section 2.2, recovering the positive and empirically feasible APT-implied SDF after correcting a misspecified candidate SDF requires using the exponential function of the linear projection of M_{t+1}^{a} and $M_{t+1}^{\beta,\text{mis}}$ on the set of the risk-free and risky assets, that is,

$$\hat{M}_{\exp,t+1}^{\beta,\text{mis}} = \exp\Big(-(\beta^{\text{mis}}\lambda^{\text{mis}})' V_R^{-1}(R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2}(\beta^{\text{mis}}\lambda^{\text{mis}})' V_R^{-1} \beta^{\text{mis}}\lambda^{\text{mis}} \Big), \tag{16}$$

$$\hat{M}_{\exp,t+1}^{a} = \exp\Big(-a' V_R^{-1}(R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2} a' V_R^{-1} a \Big), \quad \text{where} \tag{17}$$

$\mathbb{E}(R_{t+1} - 1_N R_f) = a + \beta^{\text{mis}}\lambda^{\text{mis}} + \beta^{\text{can}}\lambda^{\text{can}}$ and $V_R = \beta^{\text{can}} V_{f^{\text{can}}} \beta^{\text{can}\prime} + \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e$.

Specifying in exponential form also the component of the SDF based on observed systematic risk factors, we get

$$M^{\beta,\text{can}}_{\exp,t+1} = \frac{1}{R_f} \times \exp\Big(- \lambda^{\text{can}\prime} V^{-1}_{f^{\text{can}}} (f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) - \frac{1}{2}\lambda^{\text{can}\prime} V^{-1}_{f^{\text{can}}} \lambda^{\text{can}} \Big) \tag{18}$$

which then leads to

$$\hat{M}_{\exp,t+1} = M^{\beta,\text{can}}_{\exp,t+1} \times \hat{M}^{\beta,\text{mis}}_{\exp,t+1} \times \hat{M}^{a}_{\exp,t+1}. \tag{19}$$

We show in Internet Appendix IA.4 that as $N \to \infty$, the projection version of the corrected SDF specified in equation (19) recovers the APT-implied SDF specified in exponential form.

2.3.2 Accounting for Time-Variation in Risk Premia

So far, we have considered factor models with constant prices of risk and risk exposures. However, in practice, one may wonder whether the vector a arises as a consequence of time variation in prices of risk or risk exposures. Below, we demonstrate that an arbitrary model of asset returns that has time-varying prices of risk or risk exposures is nested in the classical APT and that the interpretation of a as compensation for unsystematic risk is preserved. To distinguish models with constant parameters from those with time-varying parameters, we use a tilde $\tilde{\ }$ to denote all the elements of the models with time-variation. To facilitate our discussion, we consider a model with only time-varying risk exposures; the analysis of a model with time-varying prices of risk is similar and is omitted for brevity.

Without loss of generality, assume that the true model for asset returns is a conditional model with a single factor $\tilde{f}_{t+1}$ and zero compensation for unsystematic risk $\tilde{e}_{t+1}$

$$R_{t+1} - \mathbb{E}_t(R_{t+1}) = \tilde{\beta}_t \tilde{f}_{t+1} + \tilde{e}_{t+1}, \tag{20}$$

where $\tilde{f}_{t+1}$ is a factor with unconditional risk premium $\tilde{\lambda}$, $\mathbb{E}_t(\tilde{f}_{t+1}) = 0$, $\tilde{\beta}_t$ is an $N \times 1$ vector of risk exposures of asset returns R_{t+1} to the factor $\tilde{f}_{t+1}$, and $\tilde{e}_{t+1}$ is an $N \times 1$ vector of unsystematic shocks with a positive-definite covariance matrix $V_{\tilde{e}}$. We consider two cases.

Case 1: Common source of variation in risk exposures. Assume that

$$\tilde{\beta}_t = \tilde{\beta}_0 + \tilde{\beta}_1 \tilde{g}_t,$$

where $\tilde{g}_t$ is a *common* source of time-variation in assets' exposures $\tilde{\beta}_t$ to the risk factor $\tilde{f}_{t+1}$. Without loss of generality, assume that $\mathbb{E}(\tilde{g}_t) = 0$. Given these assumptions, the true data-generating process for asset returns is

$$R_{t+1} - R_f 1_N = \tilde{\beta}_0 \tilde{\lambda} + \tilde{\beta}_1 \tilde{g}_t \tilde{\lambda} + \tilde{\beta}_0 \tilde{f}_{t+1} + \tilde{\beta}_1 \tilde{g}_t \tilde{f}_{t+1} + \tilde{e}_{t+1},$$

or equivalently

$$R_{t+1} - \mathbb{E}(R_{t+1}) = \tilde{\beta}_0 \tilde{f}_{t+1} + \tilde{\beta}_1 \tilde{g}_t \tilde{\lambda} + \tilde{\beta}_1 (\tilde{g}_t \tilde{f}_{t+1} - \mathbb{E}(\tilde{g}_t \tilde{f}_{t+1})) + \tilde{e}_{t+1}.$$

Thus, the true factor model (20) with the single factor $\tilde{f}_{t+1}$, time variation in risk premia driven by one common variable $\tilde{g}_t$, and zero compensation for unsystematic risk $\tilde{e}_{t+1}$, is observationally equivalent to the APT model of asset returns with $a = 0_N$ and three systematic factors, $f_{t+1} = (\tilde{f}_{t+1},\ \tilde{g}_t,\ \tilde{g}_t \tilde{f}_{t+1})'$.[12] Therefore, if researchers assume a misspecified candidate model with the single factor $f^{\text{can}}_{t+1} = \tilde{f}_{t+1}$ and constant risk exposures $\beta^{\text{can}} = \tilde{\beta}_0$, we can use our approach developed in Section 2.3.1 to correct this candidate model and obtain the APT-implied SDF that prices assets without errors. In this APT-implied SDF, the component $M^{\beta,\text{mis}}_{t+1}$ is a function of the omitted factors $f^{\text{mis}}_{t+1} = (\tilde{g}_t,\ \tilde{g}_t \tilde{f}_{t+1})'$. Furthermore, $M^{\beta,\text{mis}}_{t+1}$ captures completely the wedge between the admissible SDF and the SDF implied by the candidate factor model, so that $M^{a}_{t+1} = 0$.

Case 2: Asset-specific source of time-variation in risk exposures. Now, assume that

$$\tilde{\beta}_t = \tilde{\beta}_0 + \tilde{\beta}_1 \odot \tilde{G}_t,$$

where $\tilde{G}_t = (\tilde{g}_{1t},\ \tilde{g}_{2t},\ \cdots,\ \tilde{g}_{Nt})'$ is a vector of *asset-specific* sources of time-variation in risk exposures $\tilde{\beta}_t$ to the risk factor $\tilde{f}_{t+1}$, and the symbol $\odot$ denotes the Hadamard element-wise product. Without loss of generality, assume that $\mathbb{E}(\tilde{g}_{it}) = 0$ for each $1 \le i \le N$. Given these assumptions, the true data-generating process for asset returns is

$$R_{t+1} - R_f 1_N = (\tilde{\beta}_0 + \tilde{\beta}_1 \odot \tilde{G}_t)\tilde{\lambda} + (\tilde{\beta}_0 + \tilde{\beta}_1 \odot \tilde{G}_t)\tilde{f}_{t+1} + \tilde{e}_{t+1}$$

or equivalently

$$R_{t+1} - \mathbb{E}(R_{t+1}) = \tilde{\beta}_0 \tilde{f}_{t+1} + \underbrace{(\tilde{\beta}_1 \odot \tilde{G}_t)\tilde{\lambda} + (\tilde{\beta}_1 \odot \tilde{G}_t)\tilde{f}_{t+1} - \mathbb{E}((\tilde{\beta}_1 \odot \tilde{G}_t)\tilde{f}_{t+1}) + \tilde{e}_{t+1}}_{\tilde{\eta}_{t+1}}.$$

Thus, the true factor model (20) with the single factor $\tilde{f}_{t+1}$, time variation in risk premia driven by asset-specific variables $\tilde{G}_t$, and zero compensation for unsystematic risk $\tilde{e}_{t+1}$, is observationally equivalent to the APT model of asset returns with the single systematic factor $f_{t+1} = \tilde{f}_{t+1}$ and unsystematic shocks $\tilde{\eta}_{t+1}$. In the APT, unsystematic shocks $\tilde{e}_{t+1}$ have zero compensation, exactly as in the true factor model, while unsystematic shocks $e_{t+1} = \tilde{\eta}_{t+1} - \tilde{e}_{t+1}$ are compensated if $\mathbb{E}(\tilde{G}_t \tilde{f}_{t+1}) \neq 0_N$.[13] Therefore, if researchers assume a

[12] Note that $\tilde{g}_t \tilde{f}_{t+1}$ can be interpreted as a scale factor of Gagliardini, Ossola, and Scaillet (2016).

[13] This example illustrates that priced unsystematic shocks may represent the product of a systematic risk factor with asset-specific drivers of risk premia.

misspecified candidate model with the single factor $f^{\text{can}}_{t+1} = \tilde{f}_{t+1}$ and constant risk exposures $\beta^{\text{can}} = \tilde{\beta}_0$, we can use our approach developed in Section 2.3.1 to correct this candidate model and obtain the APT-implied SDF. In this APT-implied SDF, the component M^{a}_{t+1} is a function of unsystematic shocks e_{t+1} (not $\tilde{e}_{t+1}$). Furthermore, M^{a}_{t+1} captures completely the wedge between the admissible SDF implied by the APT and the SDF implied by the candidate factor model, so that $M^{\beta,\text{mis}}_{t+1} = 0$.

3 Estimation Details

In this section, we describe our approach for estimating the APT-implied SDF for two cases. In the first case, we estimate the APT model of asset returns, in which the systematic factors are latent. In the second case, we explain how to estimate the APT-implied SDF by correcting a candidate factor model of asset returns with K^{can} observable factors. For both cases, we explain how to identify the number of latent factors and the corresponding no-arbitrage bound δ_{apt} in equation (4). We also highlight the role played by the no-arbitrage restriction and describe how to estimate the covariance matrix V_e for unsystematic shocks.

3.1 Our Estimation Approach

For both cases described above, we recover the APT-implied SDF in two steps. For the first case, we start by specifying values for the parameters K and δ_{apt}. For the parameter K, we consider values ranging from 1 to 10. For the parameter δ_{apt}, we consider a grid ranging from 0 to 0.25 that corresponds to Sharpe ratios ranging from 0 to $\sqrt{0.25} = 0.5$ per month for the portfolio associated with unsystematic risk.[14] To choose the optimal values for K and δ_{apt}, we use ten-fold cross-validation with the HJ distance as a selection metric. For each particular combination of K and δ_{apt}, we use a constrained maximum-likelihood estimator to estimate the APT model of asset returns given in (2), subject to the no-arbitrage restriction (4), on all but one fold, which we use for validation.[15] Then, we use

[14] The APT theory is silent about the value of δ_{apt}; Ross (1977) suggests using a bound that is a multiple of the Sharpe ratio for the market portfolio, which is about 0.12 per month.

[15] We adopt the MLE approach thanks to its properties and computational convenience. First, the MLE approach is efficient. Second, it delivers closed-form solutions for many parameters of the model, leading to a reduction in computation time (see Appendix IA.7). Third, the MLE approach applied to the model without compensation for unsystematic risk is asymptotically equivalent to the standard PCA that is used to estimate the risk factors in leading empirical SDF-models (see, e.g., Kozak et al., 2018, 2020; Lettau and Pelger, 2020; Giglio and Xiu, 2021).

the parameter estimates and the formulas (9), (10), and (11), where the covariance matrix of unsystematic risk satisfies equation (1) and expected excess returns satisfy equation (3), to construct the positive feasible APT-implied SDF on the validation fold. We repeat this procedure ten times. We compute the HJ distance for all the validation folds. Having obtained the HJ distances for all combinations of K and δ_{apt}, we choose the optimal K and δ_{apt} that deliver the smallest HJ distance. We then use the optimal K and δ_{apt} to estimate the APT-implied SDF using the entire sample. Internet Appendix IA.7 provides a more detailed description of the estimation approach.

For the second case, in which we correct an arbitrary candidate factor model with K^{can} observable factors, we start by specifying values for the parameters K^{mis} and δ_{apt}.[16] For K^{mis}, we consider values ranging from 0 to 5, and for δ_{apt}, just as before, we consider a grid ranging from 0 to 0.25. To choose the optimal values for K^{mis} and δ_{apt}, we use ten-fold cross-validation with the HJ distance as a selection metric. For each particular combination of K^{mis} and δ_{apt}, we use a constrained maximum-likelihood estimator to estimate the model of asset returns given in expression (12), where V_ε and α are defined in (13) and (14), subject to the no-arbitrage restriction (4).[17] Then, we use the extended version of Proposition 2, which is formally presented and proved in Internet Appendix IA.4, and formulas (16), (17), (18), and (19) to recover the positive feasible APT-implied SDF on the validation folds. Just as in the first case, having obtained the HJ distances for all the combinations of K^{mis} and δ_{apt}, we choose the optimal K^{mis} and δ_{apt} that deliver the smallest HJ distance. We then use the optimal K^{mis} and δ_{apt} to estimate the corrected SDF using the entire sample.

We use the HJ distance to select K (or K^{mis}) and δ_{apt} because it is a widely recognized economically meaningful metric of pricing performance. Moreover, the HJ distance depends on the SDF that provides, under the null hypothesis of the APT, the correct interpretation of the vector a as compensation for unsystematic risk. Furthermore, under the assumption that asset returns are Gaussian, the HJ distance summarizes how competing asset-pricing models fit both the first *and* second moments of the return distribution.[18] This is in contrast

[16]Without loss of generality, we consider candidate models with tradable factors that represent either factor returns (for example, the market factor and returns on long-minus-short strategies) or excess returns on factor-mimicking portfolios. It is straightforward to extend the estimation algorithm to the case of candidate factor models with nontradable systematic risk factors.

[17]We use asymptotic equivalence of the no-arbitrage restrictions in equations (4) and (15).

[18]Because we assume for tractability that asset returns are Gaussian, the SDF and HJ distance depend only on the mean and variance of excess returns. The APT-implied SDF in formulas (9)–(10) and (16)–(17) depends on the model-implied quantities for $\mathbb{E}(R_{t+1} - R_f 1_N)$ and V_R.

to other metrics, for example, the cross-sectional R^2, that assess how competing models fit only average excess returns. Thus, as a model-diagnostic measure, the HJ distance sets a higher hurdle for the APT model of asset returns, in which the choice of the number of systematic factors and the value of the no-arbitrage bound have implications for both expected excess returns and the return covariance matrix.

The prior literature has used other methods for selecting the number of common risk factors in SDF models. For example, Giglio and Xiu (2021) use a statistical information criterion similar to Bai and Ng (2002). Lettau and Pelger (2020) and Kozak et al. (2020) use economic restrictions relating expected returns to the covariance of returns with factors, in addition to time-series information on the variation in asset returns.[19] Because none of these approaches applies directly to a model with nonzero compensation for unsystematic risk, i.e., $a \neq 0_N$, we face a choice: either to use a two-stage estimation approach that pins down K (or K^{mis}) in the first step and δ_{apt} in the second step or to design our own method. We choose the latter and optimize an objective function that explicitly incorporates the no-arbitrage restriction while simultaneously selecting the number of systematic risk factors and the no-arbitrage bound to minimize the HJ distance.

In Sections 5 and 6, we show that our main empirical insight about unsystematic risk being priced is robust to picking a different number of common factors, estimating the APT-implied SDF on a different set of basis assets, or using different methods for estimating the systematic risk factors.

3.2 The No-Arbitrage Restriction

The no-arbitrage restriction in (4) on the vector a serves several purposes. First, economically, it rules out asymptotic arbitrage. Equivalently, the no-arbitrage restriction constrains the Sharpe ratio of the so-called alpha portfolio of Raponi, Uppal, and Zaffaroni (2022). In our setting, $a'V_e^{-1}a$ is approximately equal to the square of the Sharpe ratio associated with investing in a portfolio that represents the unsystematic SDF component $M^a_{\text{exp},t+1}$,

[19] Even though our objective function is similar to that of Lettau and Pelger (2020), our approach has two crucial differences. First, as explained above, from the perspective of the APT model, α is not a pricing error; rather, it represents a crucial component of the SDF. Thus, α does not need to be the null vector, and therefore, our aim is not to compress it as much as possible but only to ensure that the no-arbitrage restriction holds. Second, our objective function does not explicitly include a pricing metric measuring goodness of fit, such as the HJ distance. If we augmented our log-likelihood function with an additional penalty term represented by the HJ distance, it would have made our problem computationally intensive.

that is, $\delta_{\text{apt}} \approx \text{var}(\log(M^a_{\exp,t+1}))$. In the same spirit, Kozak et al. (2020) rule out near-arbitrage opportunities by restricting the maximum squared Sharpe ratio implied by the overall SDF.[20]

Second, statistically, when estimating the APT model of asset returns, the no-arbitrage restriction (when binding) leads to the identification of the vectors $\beta\lambda$ and a, and therefore the systematic and unsystematic SDF components $\hat{M}^{\beta}_{\exp,t+1}$ and $\hat{M}^{a}_{\exp,t+1}$.[21] Specifically, at the estimation stage, the no-arbitrage restriction provides the N conditions that allow us to identify separately the estimates of $\beta\lambda$ and a (see Proposition IA.7.5 in Internet Appendix IA.7). Similarly, when correcting a candidate factor model, the no-arbitrage condition leads to the identification of $\beta^{\text{mis}}\lambda^{\text{mis}}$ and a that is necessary for constructing the missing systematic and unsystematic components of the SDF in (19), $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$ and $\hat{M}^{a}_{\exp,t+1}$, respectively. This use of the asymptotic no-arbitrage restriction is specific to our paper, which, to the best of our knowledge, is the first to estimate the SDF with both systematic and unsystematic risk.

Finally, under the no-arbitrage restriction, the estimator of a has the form of a ridge estimator, as shown in Proposition IA.7.5 of Internet Appendix IA.7. The ridge estimator has the appealing property of mitigating estimation noise. In our case, this property is especially valuable because the vector a is a component of expected returns, which, as is well known (Merton, 1980), are difficult to estimate. It also means that our estimator of a is not simply a plug for the sample average of expected excess returns net of factor risk premia. As we will show below, without the APT restriction, it would not be possible to obtain a precise estimate of the vector a, and hence pin down the compensation for bearing unsystematic risk.

3.3 Estimating the Covariance Matrix of Unsystematic Shocks

As pointed out earlier, the covariance matrix V_e of unsystematic shocks e_{t+1} in equation (2) or (13) does not have to be diagonal but must have bounded eigenvalues. That is, the shocks e_{t+1} do not have to be uncorrelated across basis assets but may include weak latent

[20]There is an important difference in the economic implications of the no-arbitrage constraint in our paper and that in Kozak et al. (2020). In our paper, the no-arbitrage restriction constrains only the prices of unsystematic risk—those of weak factors and purely asset-specific shocks—but not of common factors. In Kozak et al. (2020), the no-arbitrage restriction constrains the prices of all factors.

[21]Even in population, the no-arbitrage restriction can be influenced by the presence of financial frictions (Korsaye et al., 2021, sec. 2).

factors. Because it is not possible to obtain consistent estimates of weak factors (Lettau and Pelger, 2020, prop. 2), estimating V_e in the presence of weak factors is challenging.

Motivated by the shrinkage approach of Ledoit and Wolf (2004a,b), we develop the following two-step estimator for the matrix V_e. In the first step, we assume initially that V_e is a diagonal matrix. Given this assumption, for each K (if estimating the APT model of asset returns) or K^{mis} (if correcting a candidate factor model with K^{can} factors) and each value of δ_{apt}, we optimize the log-likelihood of asset returns subject to the no-arbitrage restriction. As a result, we obtain the first-step estimate $V_e^{(1)}$ of V_e. Next, we check whether the covariance matrix $V_{e^{\text{fit}}}$ of the fitted residuals e^{fit} of the asset-return model (2) or the candidate factor model corrected for missing systematic risk factors, is diagonal. If it is not, we proceed to the second step, in which we estimate $V_e = V_e^{(2)}$ as a linear combination of $V_e^{(1)}$ and $V_{e^{\text{fit}}}$,

$$V_e^{(2)} = \theta V_e^{(1)} + (1-\theta) V_{e^{\text{fit}}},$$

where we choose θ so that $\delta_{\text{apt}} \approx \text{var}(\log(\hat{M}^a_{\text{exp},t+1}))$. Effectively, we shrink the empirical covariance matrix of shocks e_{t+1} towards a diagonal matrix $V_e^{(1)}$, and we choose the degree of shrinkage, $1-\theta$, to preserve the economic interpretation of δ_{apt} as the squared Sharpe ratio of the portfolio that loads only on unsystematic risk.

4 Data

This section describes the data we use for our empirical analysis. First, we describe the basis assets we use to estimate the APT-implied SDF. Then, we examine variables that could potentially span the estimated SDF and its components.

4.1 Basis Assets

We construct a projection of the SDF on a large set of standard characteristics-based portfolios of U.S. stocks. As in Giglio and Xiu (2017), we use monthly returns data for 202 portfolios from Kenneth French's website, which we label GX monthly data. The data includes returns on 25 portfolios sorted by size and book-to-market ratio (ME & BM), 17 industry portfolios (Ind), 25 portfolios sorted by operating profitability and investment (OP & INV), 25 portfolios sorted by size and variance (ME & VAR), 35 portfolios sorted by size and net issuance (ME & NetISS), 25 portfolios sorted by size and accruals (ME & ACCR), 25 portfolios sorted by size and beta (ME & BETA), and 25 portfolios sorted by size and momentum (ME & MOM). The sample runs from July 1963 to August 2019.

While our theoretical results apply to any type of asset—individual assets or portfolios—as basis assets we use portfolios rather than individual assets for two related reasons. First, portfolios exhibit a more stable factor structure (Lettau and Pelger, 2020; Giglio and Xiu, 2021). Second, given that portfolios exhibit a more stable factor structure, the state-of-the-art models based on latent systematic risk (Kozak et al., 2020; Lettau and Pelger, 2020), which we use for comparison in our robustness analysis, are also estimated and tested on portfolio data. Because, we are using portfolios as basis assets, in our empirical investigation, purely asset-specific risk refers to idiosyncratic risk specific to these portfolios.

In Section 6, we also use as basis assets other datasets that have been employed in related work. Specifically, we estimate the APT-implied SDF on fifty anomaly portfolios (for both daily- and monthly-return data) used in Kozak et al. (2020), and then monthly-return data for the seventy-four characteristic-based portfolios used in Lettau and Pelger (2020). For ease of reference, we refer to these datasets as KNS daily data, KNS monthly data, and LP monthly data, respectively. The KNS daily and monthly data include returns from November 1973 to December 2017 and the LP monthly data include returns from November 1963 to December 2017.

4.2 Variables Potentially Spanning the SDF

To interpret our results and understand which economic variables may explain the APT-implied SDF's variation, we collect a comprehensive set of variables available at a monthly frequency. Our dataset includes both macroeconomic and financial indicators and returns on trading strategies. In the factor-zoo literature, the returns on these trading strategies are also known as factors or anomalies. We consider returns on 457 trading strategies studied in Novy-Marx (2013), Kozak et al. (2020), Chen and Zimmermann (2022), Jensen, Kelly, and Pedersen (2022), Hou, Mo, Xue, and Zhang (2021), and Bryzgalova et al. (2023). Furthermore, we consider 103 macroeconomic and financial indicators, which include aggregate consumption growth, inflation, various measures of economy-wide sentiment, disagreement, and volatility. We provide the details regarding the data sources and construction of these variables in Internet Appendix IA.9.

5 Empirical Analysis

In this section, first, we analyze the estimated APT-implied SDF and characterize its components, thereby establishing the relative importance of systematic versus unsystematic risk. Then, we examine the reasons for the poor performance of traditional candidate factor models: the market model, the model with consumption growth as the sole factor, and the FF3 model. For each candidate model, we characterize the *missing* systematic and unsystematic components of the corresponding SDFs. We conclude this section by explaining that our finding regarding the importance of the missing unsystematic SDF component applies to virtually any other asset-pricing model with only systematic risk factors.

5.1 The SDF under the APT Model of Asset Returns

To analyze the APT-implied SDF, we first estimate the APT model of asset returns specified in equations (2) and (4). As explained in Section 3.1, we use a cross-validation procedure to determine the number of latent systematic factors K and the no-arbitrage bound, δ_{apt}.

5.1.1 Number of Latent Systematic Risk Factors and the No-Arbitrage Bound

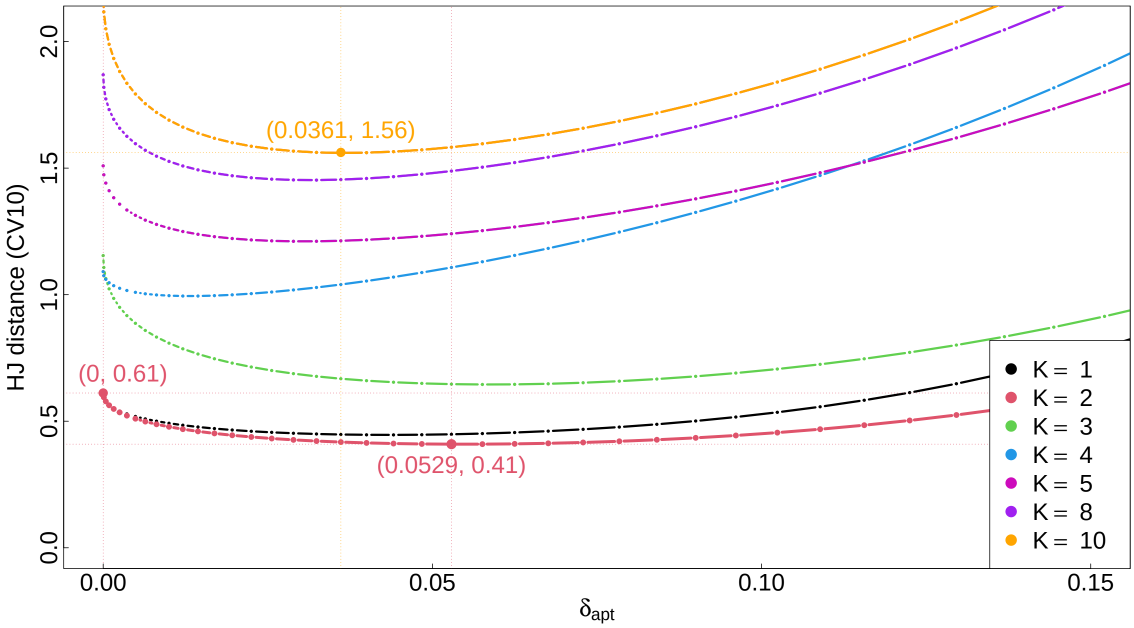
Figure 1 shows how the HJ distance of the APT model, which is evaluated on the *validation folds* of the cross-validation procedure, changes as we use different K and δ_{apt} in the estimation. We see that the combination of $K = 2$ latent factors and $\delta_{\text{apt}} = 0.0529$ achieves the smallest HJ distance of 0.41.[22] The low value of K is consistent with the evidence on low-dimensional latent factor models in Kozak, Nagel, and Santosh (2018, 2020) and Lettau and Pelger (2020). The nonzero value of the optimal δ_{apt}, which bounds the vector a ($\delta_{\text{apt}} \neq 0$ if and only if $a \neq 0_N$), indicates that, contrary to conventional wisdom, unsystematic risk is priced in the stock market.

To understand the economic importance of accounting for compensation for unsystematic risk, we analyze two extreme situations. First, we explore how the HJ distance changes quantitatively in a model with $K = 2$ latent systematic factors if we set $\delta_{\text{apt}} = 0$, which implies that $a = 0_N$. Figure 1 shows that the HJ distance increases by an economically

[22] We see that the HJ distance obtained for a model with $K = 1$ is very similar to that obtained for a model with $K = 2$, if the value of δ_{apt} is not too large. We provide an economic explanation of this finding in Section 5.1.4.

Figure 1: APT model selection using the HJ distance

This figure illustrates how the HJ distance changes with the key parameters of the APT model, K and δ_{apt}. The figure shows, for different combinations of K and δ_{apt}, the results of the cross-validation exercise in which we split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The numbers reported in the figure are $(\delta_{\text{apt}}, \text{HJ distance})$. The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors and W is the inverse of the second-order moment matrix of asset returns.

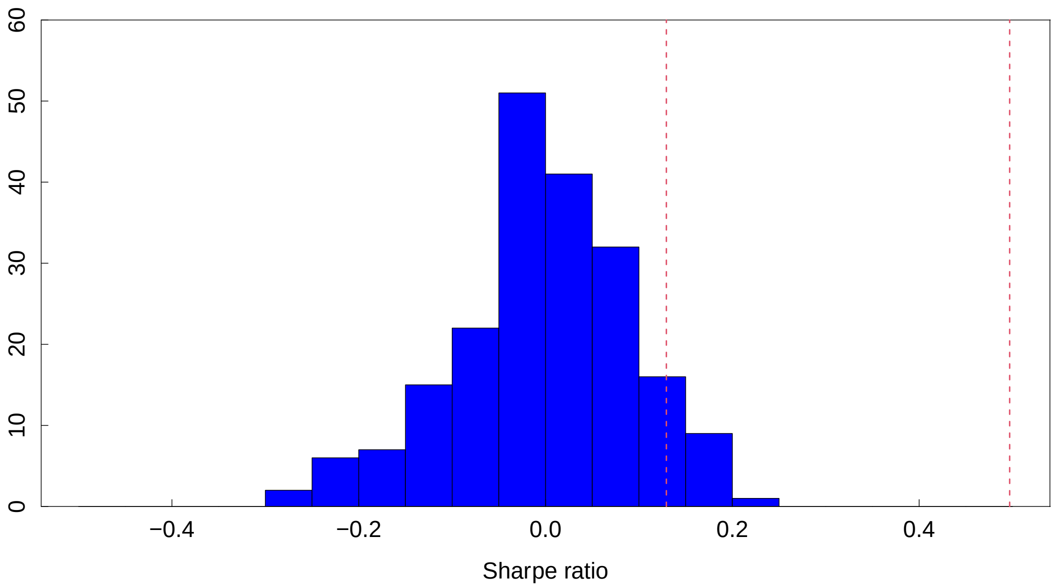


significant amount—49% $= (0.61/0.41 - 1)$ if $\delta_{\text{apt}} = 0$.[23] Moreover, as δ_{apt} gets close to zero, the HJ distance increases dramatically for any number of latent factors K.

Second, we analyze whether increasing the number of systematic factors reduces the optimal δ_{apt} to zero. Figure 1 shows that even if we were to assume that the APT model had a much larger number of factors than the optimal $K = 2$, the compensation for unsystematic risk would remain sizable. For example, if we set the number of systematic factors to be $K = 10$ and then choose only δ_{apt} in our cross-validation exercise, the HJ distance is minimized at $\delta_{\text{apt}} = 0.0361$ rather than $\delta_{\text{apt}} = 0$. This finding confirms that our main insight about unsystematic risk being priced is robust with respect to the choice of the number of systematic factors. In Section 6, we also show that our empirical analysis leads to the same conclusion if we use a different definition of systematic risk (e.g., that of Kozak

[23] Figure IA.1 in Internet Appendix IA.11 shows that the large increase in the HJ distance when we set $\delta_{\text{apt}} = 0$ is true not just for the cross-validation exercise but also when the SDF is estimated using the full sample.

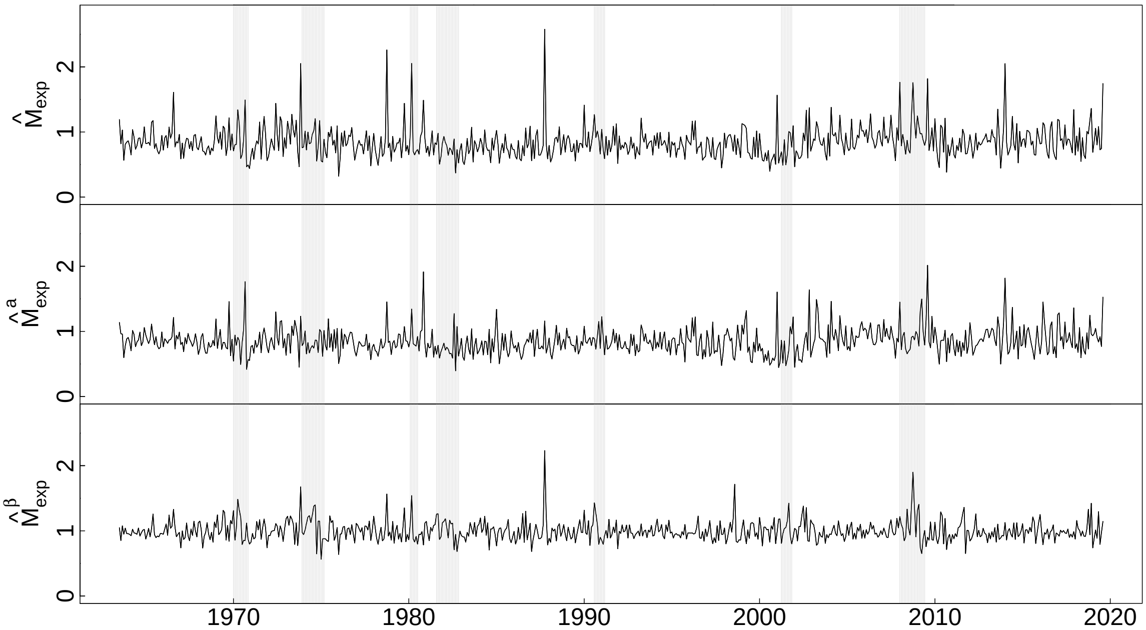
Figure 2: Estimated Sharpe ratios for individual sources of risks
The histogram in this figure shows the annualized Sharpe ratios, a_i/σ_i, associated with shocks e_i and the dotted vertical lines indicate the Sharpe ratios associated with the $K = 2$ systematic factors.



et al. (2020) or Lettau and Pelger (2020)) or estimate the model on different sets of basis assets. The analysis of these two extreme situations explains why prior work that only investigated pricing of *systematic risk* could not explain the cross-section of expected stock returns.

Even though our finding about nonzero compensation for unsystematic risk challenges the conventional view that expected excess asset returns compensate only for asset exposures to systematic risk factors, our empirical results are consistent with the idea that the factors that explain most of the comovements of asset returns earn the highest compensation (Kozak et al., 2018). Figure 2 displays the histogram of the Sharpe ratios associated with individual unsystematic shocks e_i, along with two dotted lines showing the Sharpe ratios of the $K = 2$ systematic factors that explain most of the variation in the cross-section of asset returns. We see from this figure that the source of risk that exhibits the highest Sharpe ratio is indeed one of the systematic factors.

Figure 3: Time-series behavior of the APT-implied SDF and its components
This figure has three panels. The top, middle, and bottom panels show the dynamics of the estimated annualized APT-implied SDF $\hat{M}_{\text{exp},t+1}$, its unsystematic component $\hat{M}^{a}_{\text{exp},t+1}$, and its systematic component $\hat{M}^{\beta}_{\text{exp},t+1}$, respectively. Gray bars indicate the NBER recession periods.



5.1.2 Time-Series and Business-Cycle Properties of SDF and its Components

Having estimated the APT model of asset returns, we study the time-series properties of the implied SDF, $\hat{M}_{\text{exp},t+1}$ specified in (11), and its components, $\hat{M}^{\beta}_{\text{exp},t+1}$ and $\hat{M}^{a}_{\text{exp},t+1}$ specified in (9) and (10), respectively. Figure 3 shows that both $\hat{M}^{\beta}_{\text{exp},t+1}$ and $\hat{M}^{a}_{\text{exp},t+1}$, exhibit sizable volatility during recessions and also during normal times. Furthermore, we see that different components of the SDF dominate its variation in different periods. For example, the increase in $\hat{M}^{\beta}_{\text{exp},t+1}$ in October 1987 shows that systematic risk factors were responsible for the dramatic increase in the level and volatility of the SDF. On the other hand, in the early 2000s (following the dot-com bubble), the increase in the unsystematic component $\hat{M}^{a}_{\text{exp},t+1}$ generated the spike in the volatility of the SDF. Thus, both systematic and unsystematic risk contribute to explaining asset valuations.

Next, we analyze the relative importance of the two SDF components for the SDF's variance. Table 1 reports the standard deviation of the APT-implied SDF, $\log(\hat{M}_{\text{exp},t+1})$, and its two components, $\log(\hat{M}^{a}_{\text{exp},t+1})$ and $\log(\hat{M}^{\beta}_{\text{exp},t+1})$. These standard deviations correspond to annual Sharpe ratios associated with exposure to the overall SDF and its unsystematic

Table 1: The APT-implied SDF and its components

This table reports two sets of quantities for the APT model: (1) The annualized Sharpe ratio of the SDF and its components, where the Sharpe ratios are approximated by the standard deviation (sd) of the log SDF and its components; and (2) the variance decomposition of the log SDF.

	Sharpe ratio (p.a.)			Variance decomp. (%)	
Model	$\mathrm{sd}(\log(\hat{M}_{\exp,t+1}))$	$\mathrm{sd}(\log(\hat{M}^{a}_{\exp,t+1}))$	$\mathrm{sd}(\log(\hat{M}^{\beta}_{\exp,t+1}))$	$\log(\hat{M}^{a}_{\exp,t+1})$	$\log(\hat{M}^{\beta}_{\exp,t+1})$
APT	0.89	0.79	0.51	72.60	27.40

and systematic components and are 0.89, 0.79, and 0.51, respectively.[24] Thus, strikingly, a unit exposure to unsystematic risk is compensated more prominently in financial markets than a unit exposure to systematic risk. Similarly, we find that of the total variation of the SDF, the unsystematic component contributes 72.60 percent, while the systematic component contributes only 27.40 percent. Thus, any model based on only systematic risk factors implies an SDF that is too smooth. The dominant role of unsystematic risk in the SDF's variation that we document is consistent also with the puzzling evidence in Daniel and Titman (1997), Herskovic, Moreira, and Muir (2019), Chaieb, Langlois, and Scaillet (2021), and Lopez-Lira and Roussanov (2022) that a substantial portion of expected excess returns is left unexplained by factor risk premia. Our work shows quantitatively that expected excess returns are explained largely by compensation for *unsystematic* risk.

We conclude by exploring the business-cycle properties of the estimated SDF and its components. To this end, we run a regression analysis of the log SDF and its components on macroeconomic and financial indicators; we do the log transformation because our SDF is in exponential form. We find that $\log(\hat{M}^{a}_{\exp,t+1})$ is largely acyclical: it does not significantly correlate with the NBER recession indicator.[25] The macroeconomic and financial indicators it correlates most with are intermediary constraints (He, Kelly, and Manela, 2017), the sentiment indices (Baker and Wurgler, 2006; Huang, Jiang, Tu, and Zhou, 2015), shocks in VIX, and shocks in credit spread (Gilchrist and Zakrajšek, 2012). Individually, each of these variables explain less than 3.5 percent of the variation in $\log(\hat{M}^{a}_{\exp,t+1})$. Panel A of Table IA.1 in Internet Appendix IA.10 provides further details regarding these results.

[24]In population, the sum of the squares of the standard deviations of the components of the log SDF must add up to the square of the standard deviation of the log SDF itself. But, in a finite sample, the components of the log SDF are not perfectly orthogonal to one another. Therefore, the sum of the squares of the standard deviations of the components deviates slightly from the square of the standard deviation of the SDF.

[25]Throughout the manuscript, we use the 5 percent cutoff to judge statistical significance.

In contrast to $\log(\hat{M}^a_{\exp,t+1})$, the systematic component $\log(\hat{M}^\beta_{\exp,t+1})$ significantly correlates with the NBER recession indicator. In addition, the systematic SDF component correlates with the Chicago Fed National Financial Condition index, intermediary constraints (He et al., 2017), and shocks to aggregate liquidity (Pástor and Stambaugh, 2003), credit spread (Gilchrist and Zakrajšek, 2012), dividend yield, financial uncertainty (Jurado, Ludvigson, and Ng, 2015), VIX, and the TED spread. Among these variables, shocks to intermediary constraints have the largest individual explanatory power for $\log(\hat{M}^\beta_{\exp,t+1})$: $R^2 = 55\%$; for further details, see Panel B of Table IA.1 in Internet Appendix IA.10.

5.1.3 The Unsystematic SDF Component

The unsystematic SDF component $M^a_{\exp,t+1}$ is, *by construction*, a weak factor for any cross-section of basis assets. That is, even if a cross-section of basis assets has exposures to only systematic factors and purely asset-specific shocks (i.e., the covariance matrix V_e is diagonal), the unsystematic SDF component satisfies the properties of a weak factor.[26] This observation motivates us to examine whether it is weak factors in e_{t+1} or shocks specific to individual basis assets, which in our exercise are characteristic-sorted portfolios, that play a major role in the unsystematic SDF component. Answering this question is not straightforward because weak factors cannot be estimated consistently (Lettau and Pelger, 2020). To circumvent this problem, we assume that the returns on the 325 trading strategies available for the entire sample out of the 457 strategies described in Section 4.2 represent an exhaustive set of possible weak factors in the cross-section of our basis assets. Armed with this assumption, we split the fitted residuals e^{fit}_{t+1} from the estimated APT model specified in equations (2) and (4) into two parts: one representing weak factors and the other characteristic-sorted portfolio-specific shocks (CSP-specific shocks). To identify the CSP-specific shocks, we use their key property: by definition, these shocks, being specific to each individual basis asset, have a diagonal covariance matrix.

[26] Observe that the exposures of asset returns to the unsystematic SDF component M^a_{t+1} are equal to

$$\beta^a = \frac{\text{cov}(M^a_{t+1}, R_{t+1} - R_f 1_N)}{\text{var}(M^a_{t+1})} = \frac{\text{cov}\left(-\frac{a'V_e^{-1}}{R_f}e_{t+1}, R_{t+1} - R_f 1_N\right)}{\text{var}(M^a_{t+1})} = -\frac{a'R_f}{a'V_e^{-1}a}.$$

Thus, $\beta^{a\prime}\beta^a = R_f^2(a'a)/(a'V_e^{-1}a)^2$, which, together with the no-arbitrage restriction (4), the boundness of δ_{apt} away from zero, and the boundedness of the eigenvalues of the covariance matrix V_e, implies that $\beta^{a\prime}\beta^a = O(1)$, that is, $\beta^{a\prime}\beta^a$ is bounded. As a result, M^a_{t+1} satisfies the definition of a weak factor.

Empirically, we regress the fitted residuals e^{fit}_{t+1} of the APT model on the returns of all trading strategies (orthogonalized with respect to the two latent factors of the APT) that substantially reduce the cross-sectional dependence in these residuals, collected in the vector f^{weak}_{t+1}:

$$e^{\text{fit}}_{t+1} = \gamma_0 + \gamma' f^{\text{weak}}_{t+1} + \xi_{t+1}. \tag{21}$$

We find that out of the 325 trading strategies available for the entire sample, 35 reduce the number of the significant off-diagonal terms in the covariance matrix of e^{fit}_{t+1} by 68 percent, leaving only 21 percent of the off-diagonal elements in the 202×202 covariance matrix of the fitted residuals ξ^{fit}_{t+1} statistically significantly different from zero. We use this regression to split the fitted residuals e^{fit}_{t+1} in equation (21) into two parts: one explained by the 35 trading strategies, or weak factors e^{weak}_{t+1}, and the other representing CSP-specific shocks e^{csp}_{t+1}, where the vectors e^{weak}_{t+1} and e^{csp}_{t+1} are the estimated values of $\gamma_0 + \gamma' f^{\text{weak}}_{t+1}$ and ξ_{t+1}, respectively.

Next, we express $M^a_{\exp,t+1}$ as

$$M^a_{\exp,t+1} = \exp\left(-a'V_e^{-1}e^{\text{weak}}_{t+1} - a'V_e^{-1}e^{\text{csp}}_{t+1} - \frac{1}{2}a'V_e^{-1}a\right)$$

and compute the standard deviations of $-a'V_e^{-1}e^{\text{weak}}_{t+1}$ and $-a'V_e^{-1}e^{\text{csp}}_{t+1}$, which are approximately equal to the Sharpe ratios of the strategies that invest in weak factors and CSP-specific shocks, respectively. We find that these values are 0.55 and 0.56 per annum, respectively. This result has two implications. First, an investor earns sizable compensation for exposure to both types of unsystematic risk. Second, CSP-specific risk and weak factors contribute almost equally to the overall variation of the unsystematic SDF component.

Having established the quantitative importance of the unsystematic SDF component, we now examine if, in practice, there are trading strategies, whose expected excess returns reflect compensation for exposure to unsystematic risk. To this end, first, we measure exposures of trading strategies to priced unsystematic risk by running individual regressions of $\log(\hat{M}^a_{\exp,t+1})$ on the excess returns of the 457 strategies described in Section 4.2. We find that a large number of strategies considered in the literature—335 out of 457—correlate statistically significantly with the unsystematic SDF component. The returns on the five trading strategies that have the highest explanatory power for $\log(\hat{M}^a_{\exp,t+1})$ and are available for the entire sample are: one-year share issuance (Pontiff and Woodgate, 2008) with

$R^2 = 17.82\%$, one-year momentum (Jegadeesh and Titman, 1993) with $R^2 = 14.08\%$, residual momentum (Blitz, Huij, and Martens, 2011) with $R^2 = 13.22\%$, betting-against-beta (Frazzini and Pedersen, 2014) with $R^2 = 13.19\%$, and net payout yield (Richardson, Sloan, Soliman, and Tuna, 2005) with $R^2 = 13.03\%$.

We also check the explanatory power of returns of the idiosyncratic-volatility factors of Ali et al. (2003) and Ang et al. (2006) for the unsystematic SDF component and find it to be low: 6.44 percent and 9.57 percent, respectively. Thus, the low explanatory power of the idiosyncratic-volatility factor shows that our unsystematic SDF component is not just the idiosyncratic-volatility factor.

Recently, Daniel et al. (2020b) have proposed removing unpriced risk from returns on characteristic-sorted portfolios that are used as risk factors in asset-pricing factor models. We examine how the five adjusted factors of Daniel et al. (2020b) (adjusted Market, Size, Value, Profitability, and Investment) relate to the unsystematic SDF component. We find that these risk factors cannot capture priced unsystematic risk either: they explain only 13.41 percent of the unsystematic SDF component.

Second, instead of looking at individual trading strategies, we ask whether returns on the universe of available trading strategies can span the variation in $\log(\hat{M}^a_{\exp,t+1})$.[27] To answer this question, we run 325 regressions, corresponding to the trading strategies with returns available over the entire sample. In each regression, the dependent variable is $\log(\hat{M}^a_{\exp,t+1})$, whereas the number of independent variables grows from 1 to 325. The first regression includes the return on a trading strategy that explains most of the unsystematic SDF component. Each subsequent regression includes an extra trading strategy whose return adds the largest amount in explaining the dependent variable. Then, we select a linear model with the smallest Bayesian information criterion (BIC).

We find that 39 trading strategies must be included to explain 66.45 percent of variation in the unsystematic SDF component. Any further increase in R^2 leads to overfitting because then BIC deteriorates.[28] We find that these 39 trading strategies explain 80.22 percent of the variation in the component of $\log(\hat{M}^a_{\exp,t+1})$ driven by the weak factors, $a'V_e^{-1}e^{\text{weak}}_{t+1}$, but only 28.28 percent of the variation in the component of $\log(\hat{M}^a_{\exp,t+1})$ driven by the

[27]See Internet Appendix IA.5 for a formal result justifying the spanning exercise for the unsystematic SDF component.

[28]Figure IA.2 in Internet Appendix IA.11 shows how the R^2 and BIC change as we increase the number of trading strategies in the regression.

Table 2: Strategies with high unsystematic risk premium RP^a

This table reports 25 selected trading strategies whose returns reflect large premia for unsystematic risk. The first column, using the classification scheme in Jensen et al. (2022), gives the name of the cluster to which the strategy belongs. If a strategy is not in the list of Jensen et al. (2022), we assign it to the cluster Unclassified. The second column gives the source. The third column shows the name of the variable, as in Chen and Zimmermann (2022), Jensen et al. (2022), or Bryzgalova et al. (2023). The last column reports the risk premium per annum in percent. The clusters, and within each cluster, the sources, are listed in alphabetical order.

Cluster name	Source	Variable name	RP^a (%)
Low risk	Ang, Hodrick, Xing, and Zhang (2006)	IdioVol3F	6.07
	Bali, Cakici, and Whitelaw (2010)	MaxRet	6.45
	Bradshaw, Richardson, and Sloan (2006)	XFIN	5.95
	Bradshaw, Richardson, and Sloan (2006)	NetEquityFinance	4.98
	Frazzini and Pedersen (2014)	BAB	3.34
Momentum	Avramov, Chordia, Jostova, and Philipov (2007)	Mom6mJunk	5.43
	Jegadeesh and Titman (1993)	Mom12m	8.27
	Jegadeesh and Titman (1993)	Mom6m	7.21
	Moskowitz and Grinblatt (1999)	indmom	4.07
Profit Growth	Heston and Sadka (2008)	Mom12mOffSeason	7.24
	Novy-Marx (2013)	valmom	4.87
	Novy-Marx (2013)	valmomprof	4.62
Profitability	Chordia, Subrahmanyam, and Anshuman (2001)	std_turn	4.65
Quality	Haugen and Baker (1996)	VolMkt	4.32
	Stambaugh and Yuan (2017)	PERF	4.10
Value	Boudoukh, Michaely, Richardson, and Roberts (2007)	NetPayoutYield	4.24
Unclassified	Campbell, Hilscher, and Szilagyi (2008)	DISSTR	6.27
	Cooper, Gulen, and Schill (2008)	betaarb	4.82
	Daniel, Hirshleifer, Sun (2019)	BEH_FIN	4.38
	Dichev and Piotroski (2001)	CredRatDG	4.33
	Diether, Malloy, and Scherbina (2002)	ForecastDispersion	4.85
	Easley, Hvidkjaer, and O'Hara (2002)	ProbInformedTrading	5.91
	La Porta (1996)	fgr5yrLag	5.36
	Prakash and Sinha (2012)	DelDRC	4.95
	Ritter (1991)	AgeIPO	4.56

CSP-specific shocks, $a'V_e^{-1}e_{t+1}^{\text{csp}}$. Because the CSP-specific shocks drive about half of the variation in the unsystematic SDF component, a large proportion of its variation is left unexplained by these trading strategies, implying that $\log(\hat{M}^a_{\text{exp},t+1})$ cannot be spanned even by a large number of observable variables, many of which have been used as proxies for factor risk in the previous literature.

Finally, we compute the risk premia associated with compensation for the exposures of the trading strategies to the unsystematic SDF component as the negative covariance of the return on the strategy and $\hat{M}^{a}_{\exp,t+1}$:[29]

$$\mathrm{RP}^{a}_{\text{strategy}} = -\operatorname{cov}(R_{\text{strategy},t+1}, \hat{M}^{a}_{\exp,t+1}) \times \mathbb{E}(M^{\beta}_{\exp,t+1})/\mathbb{E}(\hat{M}_{\exp,t+1}).$$

Table 2 lists 25 selected strategies with high compensation for unsystematic risk. In the literature, some of these 25 strategies have been interpreted as being behavioral—for example, the performance factor (Stambaugh and Yuan, 2017), the long-horizon financial factor (Daniel et al., 2020a), the factor reflecting expectations about future earnings in growth (La Porta, 1996), and the momentum factor (Jegadeesh and Titman, 1993)—while others as reflecting market frictions—for example, the betting-against-beta factor (Frazzini and Pedersen, 2014) and distress risk (Campbell et al., 2008).

Summarizing our analysis of the unsystematic SDF component, we emphasize three novel findings about unsystematic risk, which we have demonstrated is priced in the stock market. First, many trading strategies featured in the existing literature correlate with the unsystematic SDF component. Second, the strategies correlated with the unsystematic SDF component that earn high-risk premia are related to market frictions and behavioral biases. Third, weak factors and purely CSP-specific shocks contribute about equally to the unsystematic SDF component.

5.1.4 The Systematic SDF Component

We now turn our attention to the systematic SDF component. We find that the strategy exhibiting the highest explanatory power for $\log(\hat{M}^{\beta}_{\exp,t+1})$ is the return on the market portfolio, with an $R^2 = 95.22\%$. It is remarkable that, despite all the criticism of the CAPM, when we consider only the systematic component of the SDF, the market return explains a large proportion of its time-series variation. Such a prominent role of the market factor explains why in Figure 1, the APT model with $K = 1$ factor has a similar HJ distance to that of the model with $K = 2$ factors. We find that four other trading strategies—sales-to-market (Barbee Jr, Mukherji, and Raines, 1996), dollar trading volume (Brennan, Chordia,

[29] We use the definition of risk premia and the result in Brillinger (2001, thm. 2.3.2) to obtain the risk premium decomposition on an asset i as compensation for systematic and unsystematic risk, as follows

$$\begin{aligned}\mathbb{E}(R_{it+1} - R_f) &= -\operatorname{cov}(M_{\exp,t+1}, (R_{it+1} - R_f)) \\ &= -\operatorname{cov}(M^{\beta}_{\exp,t+1}, R_{it+1} - R_f) \times \mathbb{E}(M^{a}_{\exp,t+1})/\mathbb{E}(M_{\exp,t+1}) \\ &\quad - \operatorname{cov}(M^{a}_{\exp,t+1}, R_{it+1} - R_f) \times \mathbb{E}(M^{\beta}_{\exp,t+1})/\mathbb{E}(M_{\exp,t+1}).\end{aligned}$$

Table 3: Correlations of tradable factors with SDF components

This table reports correlations of 16 selected tradable factors with the unsystematic and systematic SDF components. The tradable factors are listed in chronological order based on the publication date of the source.

Variable name	Correlation with $\log(\hat{M}^{a}_{\text{exp},t+1})$	Correlation with $\log(\hat{M}^{\beta}_{\text{exp},t+1})$	Source
Market	0.15	−0.98	Sharpe (1964), Lintner (1965)
Size	0.00	−0.36	Fama and French (1992)
Value	−0.23	0.14	Fama and French (1992)
Momentum	−0.36	0.18	Jegadeesh and Titman (1993)
Illiquidity	0.00	−0.27	Amihud (2002)
Operating profitability (RMW)	−0.25	0.17	Fama and French (2015)
Investment (CMA)	−0.32	0.31	Fama and French (2015)
Management (MGMT)	−0.37	0.48	Stambaugh and Yuan (2016)
Performance (PERF)	−0.37	0.27	Stambaugh and Yuan (2016)
Short-horizon underreaction (PEAD)	−0.23	0.14	Daniel, Hirshleifer, and Sun (2019)
Financing (FIN)	−0.41	0.42	Daniel, Hirshleifer, and Sun (2019)
Market*	−0.10	−0.56	Daniel, Mota, Rottke, and Santos (2020)
Size*	−0.13	−0.12	Daniel, Mota, Rottke, and Santos (2020)
Value*	−0.14	0.09	Daniel, Mota, Rottke, and Santos (2020)
RMW*	−0.13	0.02	Daniel, Mota, Rottke, and Santos (2020)
CMA*	−0.10	0.21	Daniel, Mota, Rottke, and Santos (2020)

and Subrahmanyam, 1998), bid-ask spread (Amihud and Mendelson, 1986), and days with zero trades (Liu, 2006)—explain an additional 4 percent of the variation in $\log(\hat{M}^{\beta}_{\text{exp},t+1})$, bringing the overall R^2 to 99.05 percent.

Next, we examine how sixteen variables most often used as systematic tradable factors in popular asset-pricing models correlate with our SDF components. Table 3 reports these correlations. Four factors stand out: the market factor is almost perfectly negatively correlated with the systematic SDF component, the adjusted profitability factor (RMW*) of Daniel et al. (2020b) has nearly zero correlation with the systematic SDF component, while the size factor (Fama and French, 1992) and illiquidity factor (Amihud, 2002) have zero correlation with the unsystematic SDF component. The other tradable factors correlate sizably with both SDF components. These findings are consistent with the results of Holcblat, Lioui, and Weber (2022) that the market and size factors seem to represent risk in a frictionless economy, whereas most of the other tradable factors reflect frictions.

5.2 Candidate Factor Models

In the section above, we have established the importance of compensation for unsystematic risk for explaining the cross-section of asset returns. We now examine the reasons for the poor performance of popular candidate factor models in pricing the cross-section of asset returns. We consider three traditional candidate models—those implied by the CAPM of Sharpe (1964), the Consumption-CAPM (C-CAPM) of Breeden (1979), and the three-factor model of Fama and French (1993). For the SDFs $M^{\beta,\text{can}}_{\exp,t+1}$ implied by each of these candidate factor models, we estimate the correction terms $\hat{M}^{a}_{\exp,t+1}$ and $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$ that are required to obtain the APT-implied SDFs. We limit our analysis to only three candidate factor models because we find that the primary source of misspecification is the omitted compensation for unsystematic risk, and, therefore, other candidate factor models with only systematic risk factors would be subject to the same misspecification.

5.2.1 The CAPM

We consider a candidate model with the market return as its sole factor ($K^{\text{can}} = 1$) and the vector $a^{\text{can}} = 0_N$, which we refer to as the CAPM. Conditional on this candidate model, our estimation procedure selects $K^{\text{mis}} = 1$ and $\delta_{\text{apt}} = 0.0529$ (see Figure IA.3 in Internet Appendix IA.11). The obtained number of missing factors to correct the market model is consistent with our earlier finding that two latent factors summarize the common variation in asset returns, with one factor being a proxy for the market factor. The nonzero value of δ_{apt} indicates that the CAPM is misspecified not only because of missing systematic risk factors but also because it omits compensation for unsystematic risk. The value of $\delta_{\text{apt}} = 0.0529$, which is the same as for the APT model, implies an annual Sharpe ratio associated with the exposure to the unsystematic SDF component equal to 0.80.

The importance of allowing nonzero compensation for unsystematic risk and accounting for an additional source of systematic risk when correcting the CAPM is evident from Table 4. This table's second and third columns show that *after* we correct the CAPM for misspecification, the relative HJ distance drops by 78.87 percentage points (= 82.96% − 4.09%). The last three columns of this table show that the lion's share of the reduction in the HJ distance is attributable to nonzero compensation for unsystematic risk. Specifically, of the variation in $\log(\hat{M}_{\exp,t+1})$, 74.14 percent is due to the *unsystematic* component, while only 18.48 percent is due to market risk and 7.38 percent due to missing systematic risk in the CAPM.[30]

[30]In Internet Appendix IA.11, Figure IA.4 shows the estimated time-series of the SDF and its components obtained after correcting the candidate CAPM model; Figure IA.5 shows the pricing errors before and after correcting the CAPM.

Table 4: Analysis of models before and after correction for misspecification
The first column of the table lists the three candidate factor models considered: CAPM, C-CAPM, and FF3. Then, the table reports three sets of quantities: (1) The HJ distances of alternative models, relative to the HJ distance of the APT model, $(\text{HJ}^{\text{model}}/\text{HJ}^{\text{APT}} - 1) \times 100\%$, before and after the model is corrected for misspecification; (2) the annualized Sharpe ratio of the corrected SDF for each of the models along with its components, where the Sharpe ratios are approximated by the standard deviation (sd) of the log SDF and its components; and (3) the variance decomposition of the log SDF.

	Relative HJ (%)		Sharpe ratio (p.a.)				Variance decomposition (%)		
			standard deviation of the log of				log of		
Model	Before correction	After correction	$\hat{M}_{\exp,t+1}$	$\hat{M}^{a}_{\exp,t+1}$	$\hat{M}^{\beta,\text{can}}_{\exp,t+1}$	$\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$	$\hat{M}^{a}_{\exp,t+1}$	$\hat{M}^{\beta,\text{can}}_{\exp,t+1}$	$\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$
CAPM	82.96	4.09	0.89	0.80	0.42	0.27	74.14	18.48	7.38
C-CAPM	83.12	0.90	0.92	0.79	0.36	0.42	66.05	15.92	18.03
FF3	84.11	0.16	0.99	0.80	0.67	0.27	55.49	38.30	6.21

Next, we analyze which variables can explain the variation in the missing *systematic* SDF component of the CAPM. We find that the size factor (Fama and French, 1993) and illiquidity factor (Amihud, 2002) explain most of the variation in $\log(\hat{M}^{\beta,\text{mis}}_{\exp,t+1})$: the R^2 of a linear regression of $\log(\hat{M}^{\beta,\text{mis}}_{\exp,t+1})$ on the size factor or the illiquidity factor is 88 percent. Such a prominent role of the size factor in $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$ explains the success of the models developed in Fama and French (1993, 2015) relative to the CAPM of Sharpe (1964).[31] Among business-cycle indicators, shocks in the credit spread (Gilchrist and Zakrajšek, 2012) have the largest, yet very small ($R^2 = 4.41\%$), explanatory power for the missing systematic SDF component, while the NBER recession indicator does not significantly correlate with it (because the candidate SDF component already includes the market factor).

We conclude our analysis of the CAPM by highlighting that our approach successfully corrects this model's SDF to obtain the APT-implied SDF. We see from the left panel of Table 5 that the corrected SDF is almost perfectly correlated with the APT-implied SDF. Similarly, we see from the right panel of Table 5 that the unsystematic SDF components of the SDFs implied by the APT and the corrected CAPM are almost perfectly correlated.

[31]Note that the illiquidity factor of Amihud (2002) and size factor of Fama and French (1993) are highly correlated, at about 92.63 percent.

Table 5: Correlation matrix of the corrected SDFs
This table reports the correlation matrix of log SDFs and their unsystematic components either obtained under the APT or after correcting different candidate models: CAPM, C-CAPM, and FF3.

		$\log(\hat{M}_{\exp,t+1})$				$\log(\hat{M}^a_{\exp,t+1})$			
			Corrected				Corrected		
		APT	CAPM	C-CAPM	FF3	APT	CAPM	C-CAPM	FF3
	APT	1.00	0.99	0.97	0.98	1.00	0.97	1.00	0.94
Corrected	CAPM	0.99	1.00	0.96	0.97	0.97	1.00	0.97	0.93
Corrected	C-CAPM	0.97	0.96	1.00	0.94	1.00	0.97	1.00	0.93
Corrected	FF3	0.98	0.97	0.94	1.00	0.94	0.93	0.93	1.00

5.2.2 The C-CAPM

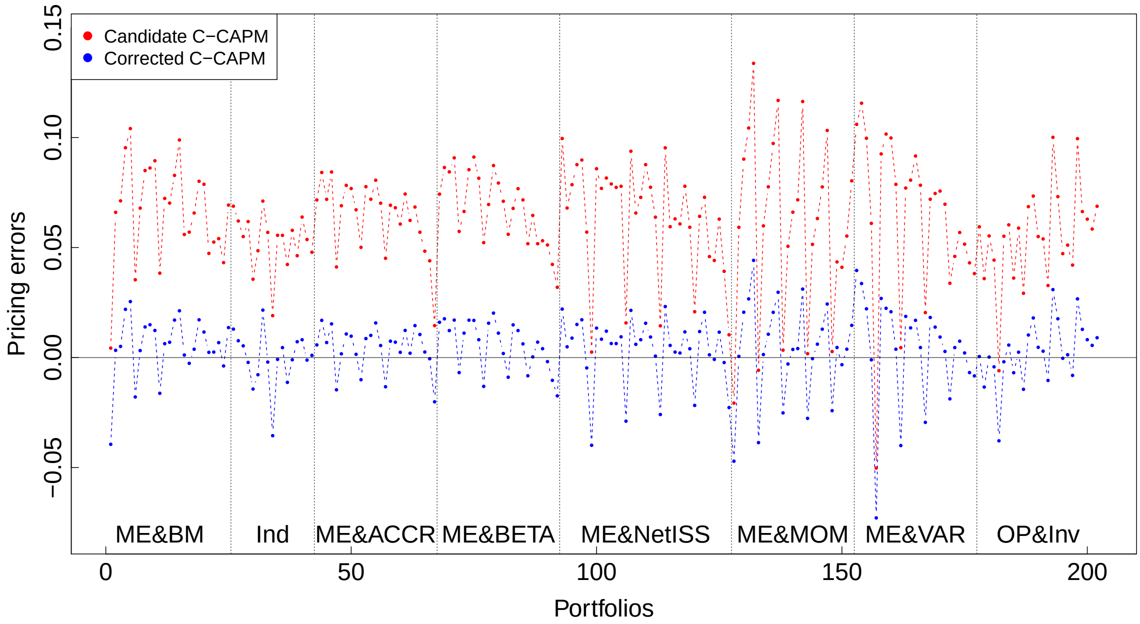
We now consider a candidate model with the return on a consumption-mimicking portfolio as its sole factor and the vector $a^{\text{can}} = 0_N$, which we refer to as the C-CAPM. We follow the standard approach of Breeden, Gibbons, and Litzenberger (1989) for constructing the consumption-mimicking portfolio.[32]

Conditional on the C-CAPM being the candidate factor model, the estimation procedure selects $K^{\text{mis}} = 2$ latent factors and $\delta_{\text{apt}} = 0.0529$ (see Figure IA.6 in Internet Appendix IA.11). The consumption-mimicking portfolio does not correlate highly with either of the two latent factors of the APT model of asset returns (the correlations are 0.30 and -0.02). Thus, two additional latent factors are still necessary to capture the common variation in asset returns. The value of δ_{apt} is 0.0529, which is the same as for the APT and (corrected) CAPM.

The second and third columns of Table 4 show that augmenting the consumption-mimicking-portfolio factor with two latent factors and allowing compensation for unsystematic risk lead to a large drop in the relative HJ distance of 82.22 percentage points ($= 83.12\% - 0.90\%$). The last three columns of the table show that, just like for the corrected CAPM, most of this drop is accounted for by the SDF's *unsystematic* component: missing systematic risk explains a much smaller proportion of the variation in the corrected

[32]As outlined in Giglio and Xiu (2021), construction of factor-mimicking portfolios can be sensitive to the choice of basis assets. They propose a three-stage procedure, insensitive to the choice of basis assets. However, their procedure does not allow compensation for unsystematic risk, which we document plays a major role in the risk-return tradeoff.

Figure 4: Pricing errors in the candidate and corrected C-CAPM
The red dots in this figure indicate the annualized pricing errors for the 202 basis assets using the C-CAPM as the *candidate* model. The blue dots indicate the annualized pricing errors using the *corrected* C-CAPM.



SDF, compared to its unsystematic component—18.03 percent versus 66.05 percent.[33] However, compared to the CAPM, the missing systematic SDF component in the C-CAPM is larger.

When we analyze the pricing errors, we observe from Figure 4 that the C-CAPM is missing a level factor: the pricing errors are centered around 6 percent in the candidate C-CAPM, whereas they are centered around zero in the corrected model. Next, we explore which observable variable explains most of the variation in $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1})$ and find, not surprisingly, that it is the market factor, with $R^2 = 92.36\%$. We find that among financial and macroeconomic indicators, shocks to intermediary constraints (He et al., 2017) and VIX innovations individually explain most of the variation in the missing systematic SDF component, with $R^2 = 55.08\%$ and 55.24%, respectively. The missing systematic SDF component has only a modest correlation with the NBER recession indicator, which is unsurprising, given that the candidate factor model already includes consumption growth.

[33] Figure IA.7 in Internet Appendix IA.11 shows the estimated time-series of the APT-implied SDF and its components obtained after correcting the candidate C-CAPM.

The left-hand-side panel of Table 5 shows that our approach for correcting misspecification in the C-CAPM model leads to an SDF highly correlated with that implied by the APT and the corrected-CAPM models. The right-hand-side panel shows that the unsystematic SDF components obtained when correcting the C-CAPM for misspecification and when estimating the APT model of asset returns are perfectly correlated.

5.2.3 The Three-Factor Model of Fama and French (1993)

We consider a candidate model with the three factors of Fama and French (1993), market, size, and value, and the vector $a^{\text{can}} = 0_N$, and we refer to this model as FF3. Conditional on FF3 being the candidate model for asset returns, our estimation method selects $K^{\text{mis}} = 1$ systematic missing latent factor and an optimal $\delta_{\text{apt}} = 0.0529$ (see Figure IA.8 in Internet Appendix IA.11). This value of δ_{apt} is the same as for the previously discussed models.

The third row of Table 4 shows that augmenting the FF3 model with one latent factor and nonzero compensation for unsystematic risk leads to a substantial improvement in pricing performance: the relative HJ distance drops by 83.95 percentage points ($= 84.11\% - 0.16\%$). The main improvement is attributable to the inclusion of nonzero compensation for unsystematic risk, as suggested by the variance decomposition of the (log of the) corrected SDF in the last three columns of Table 4. Thus, similar to Stambaugh and Yuan (2017), Clarke (2022), and Bryzgalova et al. (2023), among others, we document sizable misspecification in the FF3 model. In contrast to these papers, however, we attribute the misspecification mainly to omitted compensation for unsystematic risk rather than missing systematic risk.[34]

We analyze which observable variables can explain the variation in the missing systematic SDF component. We find that the operating profitability factor (Fama and French, 2006), return on equity (Haugen and Baker, 1996), and total accruals (Richardson et al., 2005) explain most of the variation in $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1})$: the R^2 of a univariate linear regression of $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1})$ on each of these variables is about 30 percent. We find no relation of $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1})$ with the NBER recession indicator.

[34]In Internet Appendix IA.11, Figure IA.9 displays the time-series behavior of the SDF obtained from correcting the original FF3 model; Figure IA.10 displays the pricing errors before and after correcting the FF3 model.

The left-hand-side panel of Table 5 shows that our approach for correcting misspecification in the original FF3 model leads to an SDF highly correlated with that implied by the APT model and also those obtained after correcting the CAPM and C-CAPM candidate factor models. The right-hand-side panel of the table confirms that the unsystematic SDF components obtained when correcting the FF3 for misspecification and when estimating the APT model of asset returns are almost perfectly correlated.

5.2.4 Discussion: Compensation for Unsystematic Risk

So far, we have shown the importance of including compensation for unsystematic risk in three popular asset-pricing factor models. Of course, we could repeat our empirical analysis for other candidate factor models. However, our main conclusion will not change because when correcting a given candidate SDF, it is the *unsystematic* component of the SDF that accounts for the lion's share of pricing of the cross-section of asset returns. Our conclusion remains robust for the two reasons. First, we have already illustrated in Section 5.1.3 that the unsystematic SDF component remains unspanned by virtually all known proxies for risk factors proposed in the literature so far. Second, we have also shown that adding extra *systematic* risk factors to a candidate factor model, without including compensation for unsystematic risk, cannot proxy for the unsystematic SDF component.

Our main conclusion offers a resolution for the empirical findings of Bryzgalova et al. (2023) and Bretscher, Lewis, and Santosh (2023). Bryzgalova et al. (2023) undertake a large-scale search for a *factor* model that prices a cross-section of asset returns but find none. Bretscher et al. (2023) use institutional holdings data to show that expected excess returns on stocks and bonds are determined by the covariance of asset returns not with common risk factors but with the returns on institutional investors' portfolios that is consistent with partially segmented markets, as in, for example, the model of Merton (1987), in which purely asset-specific risk is priced.

6 Robustness: Compensation for Unsystematic Risk

To illustrate the robustness of our conclusion about the importance of unsystematic risk and its compensation, we undertake three exercises. In the first exercise, we show that the importance of compensation for unsystematic risk is also present in other datasets that have been studied in related work; in fact, the importance of unsystematic risk is even greater in these alternative datasets. In the second exercise, we show that also out of sample, the APT-implied SDF has smaller pricing errors compared to the SDFs based either on observable candidate factors or latent factors. In the third exercise, we demonstrate that, even when more general methods are used to identify the systematic risk factors, compensation for unsystematic risk continues to play an important role in explaining the cross-section of expected asset returns.

6.1 Different basis assets

We use three datasets—KNS daily data, KNS monthly data, and LP monthly data—to estimate the SDFs implied by the APT, three traditional factor models, and models with only latent systematic factors (i.e., with $a = 0_N$).

For each of these three datasets, when we estimate the APT model of asset returns using cross-validation with ten folds, we confirm that unsystematic risk has sizable compensation, that is, the optimal $\delta_{\text{apt}} \neq 0$ (see Figures IA.11, IA.12, and IA.13 in Internet Appendix IA.11). The bound on the no-arbtrage constraint is $\delta_{\text{apt}} = 0.0324$ for the KNS daily data, $\delta_{\text{apt}} = 0.1369$ for the KNS monthly data, and $\delta_{\text{apt}} = 0.2401$ for the LP monthly data.[35] Furthermore, we find that for all three datasets, the optimal number of systematic factors is $K = 4$.

To highlight the importance of non-zero compensation for unsystematic risk, in Table 6, we report the HJ distance of the SDFs implied by three traditional factor models and also models based exclusively on latent systematic factors (that is, the APT model, in which we use a different number of systematic factors K but set $a = 0_N$) relative to the APT-implied SDF with $a \neq 0_N$. We include models based on latent factors because these models are agnostic about the systematic risk factors and, therefore, nest linear candidate factor models that feature different observable proxies for systematic risk.

We see from Table 6 that all the models with $a = 0_N$ have a substantially larger HJ distance than the APT model of asset returns, in which $a \neq 0_N$. The importance of accounting for compensation for unsystematic risk is even greater for the KNS daily and monthly data and LP monthly data relative to the GX monthly data. We also notice that

[35]The estimates of the bounds on the no-arbitrage constraint obtained on the monthly and daily data are not directly comparable because of different data frequency.

Table 6: Robustness with respect to different sets of basis assets

This table reports the HJ distances of alternative models, relative to the HJ distance of the benchmark APT model with priced unsystematic risk, $(\text{HJ}^{\text{model}}/\text{HJ}^{\text{APT}} - 1) \times 100\%$, for four datasets: GX monthly data, the daily and monthly datasets of Kozak et al. (2020), and the monthly dataset of Lettau and Pelger (2020). The number of factors K and the no-arbitrage bound δ_{apt} for the benchmark APT model are estimated in cross-validation on each of the four datasets. We examine the performance of the three candidate traditional factor models and also models based on only latent systematic risk (i.e., APT models with $a = 0_N$) using the entire sample. The HJ distance of the C-CAPM is not available for the daily dataset of Kozak et al. (2020) because consumption data is not available at this frequency. A positive number indicates that the corresponding model performs worse than our benchmark APT model with priced unsystematic risk (i.e., $a \neq 0_N$).

	Relative HJ distance (%)			
	GX monthly data	KNS daily data	KNS monthly data	LP monthly data
Panel A: Candidate models				
CAPM	82.96	1007.84	252.43	433.16
C-CAPM	83.12	NA	243.91	431.51
FF3	84.12	1007.41	257.46	433.96
Panel B: Latent-factor models with $a = 0_N$				
$K = 1$,	83.30	1009.92	244.20	432.41
$K = 2$	83.44	1003.56	244.30	432.28
$K = 3$	89.85	999.40	243.75	440.52
$K = 4$	139.38	977.38	241.73	465.23
$K = 5$	135.61	853.18	235.02	466.99

adding more latent factors does not always improve the pricing implications of the restricted APT model with $a = 0_N$, because adding another common risk factor does not substitute for nonzero compensation for unsystematic risk.

6.2 Out-of-Sample Analysis

We undertake two out-of-sample exercises in which we compare the performance of the APT-implied SDF to the SDFs implied by the three candidate factor models (CAPM, C-CAPM, and FF3) and latent-factor models without compensation for unsystematic risk.

The first exercise is a *time-series* out-of-sample analysis. We run this analysis on four datasets: GX monthly data, KNS daily data, KNS monthly data, and LP monthly data. For each dataset, first, we split the sample into two equal parts: one part consisting of odd-numbered observations, and the other of even-numbered observations. Next, for each dataset, we estimate each model by ten-fold cross-validation on one part of the sample

Table 7: Time-Series Out-Of-Sample Pricing Performance

This table reports the HJ distances of alternative models, relative to the HJ distance of the APT model with priced unsystematic risk, $(\mathrm{HJ}^{\mathrm{model}}/\mathrm{HJ}^{\mathrm{APT}} - 1) \times 100\%$ on four datasets: GX monthly data, KNS daily data, KNS monthly data, and LP monthly data. For each dataset, we estimate each model by cross-validation with ten folds on half of the available observations (odd or even) and then evaluate it on the other half. Then, we swap the estimation and evaluation subsamples of the data and repeat our exercise. Finally, we compute the average performance, which is the quantity reported. A positive number indicates that the corresponding model performs worse than our benchmark APT model with priced unsystematic risk (i.e., $a \neq 0_N$).

	Relative HJ distance (%)			
Model	GX monthly data	KNS daily data	KNS monthly data	LP monthly data
Panel A: Candidate models				
CAPM	10.68	180.60	49.48	61.04
C-CAPM	10.72	NA	64.36	64.93
FF3	13.52	186.34	52.47	64.61
Panel B: Latent-factor models with $a = 0_N$				
$K = 1$	11.72	182.27	52.14	61.01
$K = 2$	15.94	182.75	56.83	67.76
$K = 3$	19.34	184.31	52.79	73.50
$K = 4$	186.60	179.82	57.65	132.17
$K = 5$	272.13	154.26	50.70	149.15

and use the corresponding parameter estimates to form the SDF. We then evaluate the performance of the SDF for the part of the sample not used in the estimation. Then, we swap the subsamples that we use for estimation and evaluation. Finally, we average the results from these two out-of-sample evaluations and report them in Table 7.

The second exercise we undertake is to run a *cross-sectional* out-of-sample analysis. In this exercise, we estimate the APT model, the three candidate factor models, and latent-factor models without compensation for unsystematic risk (i.e., $a = 0_N$) on GX monthly data and then evaluate the performance of the SDFs implied by each of these models on the KNS monthly data and LP monthly data that are not used in the estimation.[36] We report the corresponding results in Table 8.

Tables 7 and 8 show how much larger the HJ distance is for various candidate models of asset returns with $a = 0_N$ relative to that for the APT model with priced unsystematic risk. First, we see that the HJ distance implied by the APT model is the smallest in both

[36] We do not use the KNS *daily* dataset for this cross-sectional out-of-sample evaluation, because the frequency of the estimated APT-implied SDF is monthly.

Table 8: Cross-Sectional Out-Of-Sample Pricing Performance

This table reports the HJ distances of alternative models, relative to the HJ distance of the APT model with priced unsystematic risk, $(\text{HJ}^{\text{model}}/\text{HJ}^{\text{APT}} - 1) \times 100\%$. We estimate each model on the GX monthly data. The pricing performance of each model is then evaluated on two datasets not used in the estimation (KNS monthly data and LP monthly data). A positive number indicates that the corresponding model performs worse than our benchmark APT model with priced unsystematic risk (i.e., $a \neq 0_N$).

	Relative HJ distance (%)	
Model	KNS monthly data	LP monthly data
Panel A: Candidate models		
CAPM	60.09	55.41
C-CAPM	58.77	53.12
FF3	60.02	53.61
Panel B: Latent-factor models with $a = 0_N$		
$K = 1$	60.00	55.59
$K = 2$	60.13	55.89
$K = 3$	70.37	65.09
$K = 4$	148.65	128.71
$K = 5$	141.83	123.68

time-series and cross-sectional out-of-sample exercises. A second noteworthy observation from these tables, as previously pointed out, is that including extra latent factors in the APT model with $a = 0_N$ can lead to a deterioration in the out-of-sample performance because of overfitting.

6.3 Different Methods for Identifying Systematic Risk Factors

Kozak et al. (2020) and Lettau and Pelger (2020) propose novel approaches for identifying systematic risk in asset returns with the goal of building SDFs spanned by this risk. Kozak et al. (2020) define systematic risk as factors that explain most of the variation in asset returns *and* earn high expected returns. Lettau and Pelger (2020) define systematic risk as factors that explain most of the comovement in asset returns *and* also the cross-sectional variation in expected asset returns. Both approaches attribute to systematic risk more variables than just those representing *strong* risk factors: Lettau and Pelger (2020) explicitly include *weak* factors, whereas Kozak et al. (2020) capture weak factors indirectly. In contrast, our definition of systematic risk is based only on strong risk factors. In this section, we show that even if one were to adopt the more general definitions of systematic risk used in Kozak et al. (2020) and Lettau and Pelger (2020), the compensation for unsystematic risk would remain substantial.

In our analysis, we use the factor models of Kozak et al. (2020) and Lettau and Pelger (2020) as candidate factor models and correct their implied SDFs to obtain the APT-implied SDF. To preserve the design of the experiments of Kozak et al. (2020) and Lettau and Pelger (2020), we estimate the SDFs on their original data. We use only the daily data for Kozak et al. (2020) because their method requires a large number of observations to obtain an accurate estimate of the covariance matrix of asset returns and subsequently interpret it as known. To be consistent with no arbitrage, we construct the *exponential* versions of the SDFs of Kozak et al. (2020) and Lettau and Pelger (2020).

We find that: (i) the methods of Kozak et al. (2020) and Lettau and Pelger (2020) accurately capture strong risk factors, (ii) both methods capture some weak factors, and (iii) the APT-implied SDF has smaller HJ distance relative to those of Kozak et al. (2020) and Lettau and Pelger (2020), which is exclusively because of the presence of the *unsystematic* SDF component in the APT-implied SDF.

6.3.1 The SDF of Kozak et al. (2020)

The approach of Kozak et al. (2020) selects four principal components (PC1, PC2, PC4, and PC9) as systematic factors to explain the asset returns on fifty anomaly portfolios netted from market exposure. Thus, the SDF of Kozak et al. (2020) that explains raw asset returns on fifty anomaly portfolios (without orthogonalizing them for the market exposure) and market return depends on five factors: the four PCs mentioned above and the market factor. We use this five-factor model as a candidate factor model and correct it using our approach. We find that this candidate model omits compensation for unsystematic risk (the optimal $\delta_{\text{apt}} = 0.0784$) and $K^{\text{mis}} = 3$ missing latent factors. These two sources of misspecification lead to the HJ distance of the SDF of Kozak et al. (2020) being four times larger than that of the APT-implied SDF (0.04 versus 0.01). Figure IA.14 in Internet Appendix IA.11 illustrates these results.

When we analyze what drives the improvement in the HJ distance after correcting the SDF of Kozak et al. (2020), we find, as expected, that it is compensation for unsystematic risk. The three latent factors missing in the model of Kozak et al. (2020) account for a tiny

Table 9: The APT-implied SDF versus the SDFs of Kozak et al. (2020) and Lettau and Pelger (2020)

This table reports the correlations between the log of the APT-implied SDF and its systematic and unsystematic components with the log of the SDFs of Kozak et al. (2020) and Lettau and Pelger (2020). We consider the sparse and non-sparse SDFs of Kozak et al. (2020). We also analyze how the SDFs of Kozak et al. (2020) and Lettau and Pelger (2020) relate to the APT-implied SDF and its components once the weak factors are excluded from the systematic risk identified by Kozak et al. (2020) and Lettau and Pelger (2020). All SDFs are exponential. The SDFs of Kozak et al. (2020) and Lettau and Pelger (2020) are estimated on the data from the original papers. The APT-implied SDF is estimated on each of these datasets.

	Correlation		
	$\log(\hat{M}^{\text{APT}}_{\text{exp},t+1})$	$\log(\hat{M}^{a}_{\text{exp},t+1})$	$\log(\hat{M}^{\beta}_{\text{exp},t+1})$
$\log(\hat{M}^{\text{KNS}}_{\text{exp},t+1})$: sparse case	0.50	0.40	0.42
$\log(\hat{M}^{\text{KNS}}_{\text{exp},t+1})$: non-sparse case	0.59	0.48	0.48
$\log(\hat{M}^{\text{LP}}_{\text{exp},t+1})$	0.77	0.60	0.45
$\log(\hat{M}^{\text{KNS}}_{\text{exp},t+1})$: sparse case without weak factors	0.28	0.09	0.75
$\log(\hat{M}^{\text{LP}}_{\text{exp},t+1})$: without weak factors	0.54	0.08	0.94

1.17 percent of the variation in the APT-implied SDF, while unsystematic risk accounts for 88.82 percent of its variation.

Given that Kozak et al. (2020) adopt a more general definition of systematic risk than ours, it is informative to measure how their SDF correlates with both the systematic and unsystematic components of the APT-implied SDF. The first row of Table 9 shows that the SDF of Kozak et al. (2020) has a correlation of about 0.50 with the APT-implied SDF and a correlation of 0.40 with the unsystematic SDF component of the APT-implied SDF. The former correlation is below one because, as we have shown above, the SDF of Kozak et al. (2020) does not capture a big chunk of priced unsystematic risk. The latter correlation is sizable because the definition of systematic risk adopted in Kozak et al. (2020) implicitly incorporates weak factors into systematic risk. When we exclude factors that contribute relatively little to explaining the comovement in returns on asset anomalies (e.g., PC4 and PC9) from the systematic risk identified by Kozak et al. (2020), we find that the resulting SDF has a much higher correlation of 0.75 with the *systematic* component of the APT-implied SDF and a low correlation of 0.09 with its *unsystematic* component. This finding confirms that our approach identifies similar sources of strong-factor risk to those identified

Table 10: Out-of-sample pricing Performance of the SDFs of Kozak et al. (2020) and Lettau and Pelger (2020)

This table reports the HJ distances of the SDFs of Kozak et al. (2020) and Lettau and Pelger (2020), relative to the HJ distance of the APT model, $(\text{HJ}^{\text{model}}/\text{HJ}^{\text{APT}} - 1) \times 100\%$. We estimate the competing models by cross-validation with ten folds on half of the available observations (odd or even) and then evaluate them on the other half. Then, we swap the estimation and evaluation subsamples of the data and repeat our exercise. Finally, we compute the average performance, which is the quantity reported. The SDFs of Kozak et al. (2020) and Lettau and Pelger (2020) are estimated on the data from the original papers. The APT-implied SDF is estimated for each of these datasets.

Model	Relative HJ distance (%)
Kozak et al. (2020) (sparse case)	149.85
Kozak et al. (2020) (non-sparse case)	137.83
Lettau and Pelger (2020)	129.11

in Kozak et al. (2020) and complements the latter approach by identifying priced purely asset-specific risk and the remaining sources of priced weak-factor risk.

We conclude our analysis by assessing the time-series out-of-sample performance of the APT-implied SDF and the SDF that uses the more general definition of systematic risk in Kozak et al. (2020) on their daily data. We see from Table 10 that the APT-implied SDF, which allows compensation for unsystematic risk, has substantially smaller pricing errors than the SDF of Kozak et al. (2020), regardless of whether the sparsity constraint is imposed or not.[37]

6.3.2 The SDF of Lettau and Pelger (2020)

We now use the model of Lettau and Pelger (2020), which includes the first five Risk Premia Principal Components (RP-PC) factors, as a candidate factor model that explains expected monthly returns of seventy-four characteristic-based portfolios. We then correct this model using our approach described in Section 2.3. We find that this candidate model does not omit any sources of systematic risk, but it is missing compensation for unsystematic risk (the optimal $\delta_{\text{apt}} = 0.1024$). The nonzero value for δ_{apt} indicates that there is either priced weak-factor risk not identified in Lettau and Pelger (2020) and/or priced purely asset-specific risk. The contribution of the missing component to explaining the SDF variation is only 20.6 percent, yet the Sharpe ratio associated with missing unsystematic risk is a sizable

[37]For details of the non-sparse SDF, see Kozak et al. (2020).

1.11. Because the model is missing, at least partially, compensation for unsystematic risk, the HJ distance of the candidate SDF is three times larger than that of the APT-implied SDF (0.39 versus 0.12). Figure IA.15 in Internet Appendix IA.11 illustrates these results.

To shed further light on how the SDF of Lettau and Pelger (2020) compares with the APT-implied SDF and its components, we calculate the correlations between them. Table 9 shows that, despite the model of Lettau and Pelger (2020) missing a bulk of priced unsystematic risk, the correlation between the two SDFs is a sizable 0.77. The SDF of Lettau and Pelger (2020) is correlated with both components of the APT-implied SDF, and more so with the unsystematic SDF component. If we remove weak factors (RP-PC5) from the candidate factor model, the resulting candidate SDF has a very high correlation of 0.94 with the *systematic* SDF component of the APT-implied SDF and almost zero correlation with its *unsystematic* SDF component. This analysis is reassuring because it highlights that alternative approaches agree on important sources of strong and weak-factor risk.

We conclude our analysis by assessing the time-series out-of-sample performance of the APT-implied SDF and the SDF that uses the more general definition of systematic risk in Lettau and Pelger (2020) on their monthly data. From Table 10, we see that the SDF of Lettau and Pelger (2020) implies a substantially larger pricing error than that of the APT-implied SDF. This result highlights the importance of accounting for priced *purely asset-specific* risk, that is, the component of unsystematic risk not spanned by weak factors.

7 Microfoundations for Priced Unsystematic Risk

In Sections 5 and 6, we have presented strong empirical evidence that factor models need to include compensation for unsystematic risk and that it is the unsystematic component of the SDF that accounts for most of its variation. Below, we present an example of an equilibrium model that provides microfoundations for the notion that unsystematic risk is priced. Our example relies on the well-known static model of Merton (1987), which has a finite number of assets N. We show that, if N is asymptotically large, then the equilibrium asset returns and SDF in this model have the same functional forms as those we have in our APT model.

In Merton (1987), investors are aware of only a subset of the available securities in which they invest. This type of "incomplete information" implies that not only systematic risk factors but also shocks specific to each security are priced. The modeling framework of Merton (1987) can be motivated by the presence of market segmentation, institutional restrictions, transaction costs, illiquidity, or imperfect divisibility of securities, that lead investors to invest in only a subset of available securities. While the incomplete information of Merton (1987) may not be the only reason why unsystematic risk is compensated, it is an appealing argument given the substantial empirical evidence documenting that both retail (Polkovnichenko, 2005; Campbell, 2006; Goetzmann and Kumar, 2008) and institutional investors (Koijen and Yogo, 2019, table 2) invest in only a small number of available stocks.

In Merton (1987), and as shown in Internet Appendix IA.4, equilibrium asset returns satisfy

$$R_i - R_f = a_i - \beta_i a_m + \beta_i(\mathbb{E}(R_m) - R_f) + \frac{b_i}{b}(R_m - \mathbb{E}(R_m)) + e_i,$$

where R_m representing the market return is the only systematic risk factor, β_i denotes the beta of asset i with respect to the market return, $e_i = \sigma_i \epsilon_i$ are asset-specific shocks with the diagonal covariance matrix containing the elements σ_i^2, b_i and σ_i are functions of the parameters of the firm's i production technology, $b = \sum_{i=1}^N x_i b_i$ with x_i being the fraction of the market portfolio invested in asset i, $a_i = (1-q_i)(\mathbb{E}(R_i) - R_f - b_i(\mathbb{E}(R_{N+1} - R_f))$ with q_i denoting the fraction of investors who know about the security i, R_{N+1} being the return on the $(N+1)$th security which is in zero net supply and which combines the risk-free security and a forward contract with cash settlements on the only systematic risk factor R_m, and $a_m = \sum_{i=1}^N x_i a_i$.

We now derive the SDF in this economy when $N \to \infty$. We assume that x_i, the fraction of the market portfolio invested in any asset i, is infinitesimally small.

Proposition 4. *When the number of assets $N \to \infty$, equilibrium asset returns are*

$$R_i - R_f = a_i + \beta_i(\mathbb{E}(R_m) - R_f) + \beta_i(R_m - \mathbb{E}(R_m)) + e_i, \tag{22}$$

$$= a_i + \beta_i(R_m - R_f) + e_i, \tag{23}$$

with the market return asymptotically orthogonal to asset-specific shocks e_i, and the equilibrium SDF is

$$M = \underbrace{\frac{-a' V_e^{-1}}{R_f} e}_{M^a} + \underbrace{\frac{1}{R_f} - \frac{\mathbb{E}(R_m) - R_f}{R_f \times \text{var}(R_m)}(R_m - \mathbb{E}(R_m))}_{M^\beta}. \tag{24}$$

Note that the model of asset returns (23) coincides with the APT model of asset returns in equation (2) with $K = 1$ systematic factor that is $f = R_m$. Similarly, the SDF in (24) coincides with the SDF in (5), given that the price of market risk is $\lambda = \mathbb{E}(R_m) - R_f$. Thus, in the model of Merton (1987) with an infinite number of assets, the SDF consists of two components: one representing unsystematic risk, M^a, and the other systematic risk, M^β, just as under the APT.

Note that, consistent with the definition of the risk premium, a_i in (22) represents the compensation for unsystematic risk, because

$$a_i = -\operatorname{cov}\left(R_i - R_f, -\frac{a'V_e^{-1}}{R_f}e_i)\right) \times R_f,$$

which coincides with the elements of the vector a in the APT. Naturally, the other part of the risk premium in (22), $\beta_i(\mathbb{E}(R_m) - R_f)$, is compensation for exposure to systematic risk, represented by market risk because of the assumption of a single systematic factor:

$$\beta_i(\mathbb{E}(R_m) - R_f) = -\operatorname{cov}\left(R_i - R_f, -\frac{\mathbb{E}(R_m) - R_f}{R_f \times \operatorname{var}(R_m)}(R_m - \mathbb{E}(R_m))\right) \times R_f.$$

If all investors are fully informed about all N assets, that is, $q_i = 1$, then $a_i = 0$, and the results in (22) and (24) simplify to the expressions for security returns and the SDF under the CAPM, respectively. Moreover, the no-arbitrage APT restriction in expression (4) is equivalent to stating that in the Merton (1987) model when $N \to \infty$ there are only a small number of assets that do not belong to the common information set of investors, that is, $q_i < 1$ for some of the assets but not all, or that there are only a small number of investors who are unaware of each asset, that is, for each i, q_i is approximately 1.

8 Conclusion

A fundamental challenge in finance is to price the cross-section of assets. The main difficulty when pricing assets is to determine the relevant sources of risk and quantify how to adjust assets' returns for these risks. The literature has proposed many proxies for *systematic risk* factors and developed factor models based on these proxies to explain the cross-sectional risk-return tradeoff. However, despite the proliferation of systematic risk factors, referred to as the factor zoo (Cochrane, 2011), there is still a sizable pricing error called alpha.

We challenge the conventional wisdom that only systematic risk receives compensation in financial markets by showing that also *unsystematic risk* is compensated. That is, the pricing error alpha implied by popular factor models includes compensation not only for omitted systematic risk factors but also for unsystematic risk. Theoretically, we demonstrate this key insight through the lens of the SDF under the assumptions of the APT and support it by demonstrating that an equilibrium model such as Merton (1987) is consistent with our insight. Empirically, we show that the component of the APT-implied SDF reflecting unsystematic risk, which is a linear combination of *unsystematic shocks*, accounts for more than 70 percent of the variation in the APT-implied SDF. Furthermore, the Sharpe ratio associated with the investment strategy exposed to only unsystematic risk is 0.8 per annum.

Our results indicate that what is missing in cross-sectional asset-pricing factor models is compensation for unsystematic risk. This insight is crucial both for empiricists wanting to resolve the factor zoo and theorists wishing to develop microfounded asset-pricing models.

References

Ali, Ashiq, Lee-Seok Hwang, and Mark A. Trombley, 2003, Arbitrage risk and the book-to-market anomaly, *Journal of Financial Economics* 69, 355–373.

Almeida, Caio, and René Garcia, 2012, Assessing misspecified asset pricing models with empirical likelihood estimators, *Journal of Econometrics* 170, 519–537.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Bai, Jushan, and Kunpeng Li, 2012, Statistical analysis of factor models of high dimension, *Annals of Statistics* 40, 436–465.

Bai, Jushan, and Serena Ng, 2002, Determining the number of factors in approximate factor models, *Econometrica* 70, 191–221.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *Quarterly Journal of Economics* 131, 1593–1636.

Baker, Scott R., Nicholas Bloom, Steven J. Davis, and Kyle J. Kost, 2019, Policy news and stock market volatility, NBER Working Paper.

Bali, Turan G., Robert F. Engle, and Scott Murray, 2016, *Empirical Asset Pricing: Cross Section of Stock Returns* (John Wiley & Sons, Hoboken, New Jersey).

Barbee Jr, William C., Sandip Mukherji, and Gary A. Raines, 1996, Do sales-price and debt-equity explain stock returns better than book-market and firm size?, *Financial Analysts Journal* 52, 56–60.

Blitz, David, Joop Huij, and Martin Martens, 2011, Residual momentum, *Journal of Empirical Finance* 18, 506–521.

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

Breeden, Douglas T., Michael R. Gibbons, and Robert H. Litzenberger, 1989, Empirical test of the consumption-oriented CAPM, *Journal of Finance* 44, 231–262.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345–373.

Bretscher, Lorenzo, Ryan Lewis, and Shrihari Santosh, 2023, Investor betas, Available at SSRN 3739424.

Brillinger, David R., 2001, *Time Series: Data Analysis and Theory* (Cambridge University Press, Cambridge).

Bryzgalova, Svetlana, Jiantao Huang, and Christian Julliard, 2023, Bayesian solutions for the factor zoo: We just ran two quadrillion models, *Journal of Finance* 78, 487–557.

Campbell, John Y, 1996, Understanding risk and return, *Journal of Political Economy* 104, 298–345.

Campbell, John Y., 2006, Household finance, *Journal of Finance* 61, 1553–1604.

Campbell, John Y, Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Chaieb, Ines, Hugues Langlois, and Olivier Scaillet, 2021, Factors and risk premia in individual international stock returns, *Journal of Financial Economics* 141, 669–692.

Chamberlain, Gary, 1983, Funds, factors and diversification in arbitrage pricing models, *Econometrica* 51, 1305–1324.

Chamberlain, Gary, and Michael Rothschild, 1983, Arbitrage, factor structure and mean-variance analysis on large asset markets, *Econometrica* 51, 1281–1304.

Chen, Andrew Y., and Tom Zimmermann, 2022, Open source cross-sectional asset pricing, *Critical Finance Review* 11, 207–264.

Chudik, Alexander, M. Hashem Pesaran, and Elisa Tosetti, 2011, Weak and strong cross-section dependence and estimation of large panels, *The Econometrics Journal* 14, C45–C90.

Clarke, Charles, 2022, The level, slope and curve factor model for stocks, *Journal of Financial Economics* 143, 159–187.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of Finance* 66, 1047–1108.

Daniel, Kent, David Hirshleifer, and Lin Sun, 2020a, Short-and long-horizon behavioral factors, *Review of Financial Studies* 33, 1673–1736.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020b, The cross-section of risk and returns, *Review of Financial Studies* 33, 1927–1979.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1–33.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2020, Taming the factor zoo: A test of new factors, *Journal of Finance* 75, 1327–1370.

Frazzini, Andrea, and Lasse H. Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2020, Dissecting characteristics nonparametrically, *Review of Financial Studies* 33, 2326–2377.

Gagliardini, Patrick, Elisa Ossola, and Olivier Scaillet, 2016, Time-varying risk premium in large cross-sectional equity data sets, *Econometrica* 84, 985–1046.

Ghosh, Anisha, Christian Julliard, and Alex P. Taylor, 2017, What is the consumption-CAPM missing? An information-theoretic framework for the analysis of asset pricing models, *Review of Financial Studies* 30, 442–504.

Giglio, Stefano, Yuan Liao, and Dacheng Xiu, 2021a, Thousands of alpha tests, *Review of Financial Studies* 34, 3456–3496.

Giglio, Stefano, and Dacheng Xiu, 2017, Inference on risk premia in the presence of omitted factors, Working Paper, University of Chicago.

Giglio, Stefano, and Dacheng Xiu, 2021, Asset pricing with omitted factors, *Journal of Political Economy* 129, 1947–1990.

Giglio, Stefano, Dacheng Xiu, and Dake Zhang, 2021b, Test assets and weak factors, Chicago Booth Research Paper.

Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit spreads and business cycle fluctuations, *American Economic Review* 102, 1692–1720.

Goetzmann, William N., and Alok Kumar, 2008, Equity portfolio diversification, *Review of Finance* 12, 433–463.

Gourieroux, Christian, and Alain Monfort, 2007, Econometric specification of stochastic discount factor models, *Journal of Econometrics* 136, 509–530.

Hansen, Lars P., and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557–590.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2015, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

Herskovic, Bernard, Alan Moreira, and Tyler Muir, 2019, Hedging risk factors, Available at SSRN 3148693.

Holcblat, Benjamin, Abraham Lioui, and Michael Weber, 2022, Anomaly or possible risk factor? Simple-to-use tests, Chicago Booth Research Paper No. 22-11.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2021, An augmented q-factor model with expected growth, *Review of Finance* 25, 1–41.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Huang, Dashan, Fuwei Jiang, Jun Tu, and Guofu Zhou, 2015, Investor sentiment aligned: A powerful predictor of stock returns, *Review of Financial Studies* 28, 791–837.

Huang, Dashan, Jiangyuan Li, and Liyao Wang, 2021, Are disagreements agreeable? Evidence from information aggregation, *Journal of Financial Economics* 141, 83–101.

Ingersoll, Jonathan E., Jr., 1984, Some results in the theory of arbitrage pricing, *Journal of Finance* 39, 1021–1039.

Jegadeesh, N., and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jensen, Theis I., Bryan T. Kelly, and Lasse H. Pedersen, 2022, Is there a replication crisis in finance?, *Journal of Finance* forthcoming.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng, 2015, Measuring uncertainty, *American Economic Review* 105, 1177–1216.

Kogan, Leonid, and Mary H. Tian, 2015, Firm characteristics and empirical factor models: A data-mining experiment, MIT Working Paper.

Koijen, Ralph S. J., and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Korsaye, Sofonias, Alberto Quaini, and Fabio Trojani, 2021, Smart stochastic discount factors, University of Geneva, Working Paper.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kozak, Serhyi, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

La Porta, Rafael, 1996, Expectations and the cross-section of stock returns, *Journal of Finance* 51, 1715–1742.

Ledoit, Olivier, and Michael Wolf, 2004a, Honey, I shrunk the sample covariance matrix, *Journal of Portfolio Management* 30, 110–119.

Ledoit, Olivier, and Michael Wolf, 2004b, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88, 365–411.

Lettau, Martin, and Markus Pelger, 2020, Factors that fit the time-series and cross-section of stock returns, *Review of Financial Studies* 33, 2274–2325.

Liu, Weimin, 2006, A liquidity-augmented capital asset pricing model, *Journal of Financial Economics* 82, 631–671.

Lopez-Lira, Alejandro, and Nikolai Roussanov, 2022, Do common factors really explain the cross-section of stock returns?, Jacobs Levy Equity Management Center for Quantitative Financial Research Paper.

McCracken, Michael W., and Serena Ng, 2016, FRED-MD: A monthly database for macroeconomic research, *Journal of Business & Economic Statistics* 34, 574–589.

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.

Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483–510.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Onatski, Alexei, 2012, Asymptotics of the principal components estimator of large factor models with weakly influential factors, *Journal of Econometrics* 168, 244–258.

Pasquariello, Paolo, 2014, Financial market dislocations, *Review of Financial Studies* 27, 1868–1914.

Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political economy* 111, 642–685.

Pelger, Markus, 2020, Understanding systematic risk: A high-frequency approach, *Journal of Finance* 75, 2179–2220.

Polkovnichenko, Valery, 2005, Household portfolio diversification: A case for rank-dependent preferences, *Review of Financial Studies* 18, 1467–1502.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921–945.

Raponi, Valentina, Raman Uppal, and Paolo Zaffaroni, 2022, Robust portfolio choice, Working Paper, Imperial College London.

Richardson, Scott A., Richard G. Sloan, Mark T. Soliman, and Irem Tuna, 2005, Accrual reliability, earnings persistence and stock prices, *Journal of Accounting and Economics* 39, 437–485.

Ross, S.A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Ross, Stephen A., 1977, Return, risk, and arbitrage, in Irwin Friend and J.L. Bicksler, (eds.) *Risk and Return in Finance* (Ballinger, Cambridge, MA).

Sandulescu, Mirela, and Paul Schneider, 2021, International mispricing and arbitrage premia, Swiss Finance Institute Research Paper.

Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 1919, 425–442.

Stambaugh, Robert F., and Yu Yuan, 2017, Mispricing factors, *Review of Financial Studies* 30, 1270–1315.

Internet Appendix

In Section IA.1, we define the notation we will use in the Internet Appendix. Section IA.2 lists the assumptions used to prove the lemmas in Section IA.3 and propositions in Section IA.4. Section IA.5 contains proofs for the propositions related to the spanning of different SDF components using observable variables. Section IA.6 presents the results for weak factors. Section IA.7 gives the details of how we estimate the APT model of asset returns. Section IA.8 discusses the case where the candidate factors are not assumed to be orthogonal to the missing sources of systematic risk. Section IA.9 provides the details of the data we use in our analysis. Section IA.10 collects additional tables and Section IA.11 additional figures that are related to the results reported in the main text of the manuscript.

IA.1 Notation

We adopt the following notation in the manuscript and appendix. $\mathbb{E}(\cdot)$ denotes the expectation operator. Capital letters denote matrices, while lowercase letters denote scalars or vectors. The notation 0_N and 1_N indicates an $N \times 1$ vector of zeros and ones, respectively. The notation I_K and O_K denotes the $K \times K$ identity matrix and matrix of zeros, respectively. For an arbitrary matrix A, the expression $A > 0$ means that A is a positive-definite matrix, $\|A\|$ denotes the Frobenius norm $\|A\| = (\text{tr}(A'A))^{\frac{1}{2}}$, where $\text{tr}(\cdot)$ is the trace operator, and $|A|$ is the determinant when A is a square matrix. For deterministic sequences $\{a_N\}$ and $\{b_N\}$, the notation $a_N = O(b_N)$ means that $|a_N|/b_N < \delta$, where $\delta > 0$ is a finite constant, and $a_N = o(b_N)$ means that $|a_N|/b_N \to 0$, as $N \to \infty$. The notation $a_N = O(b_N)$ and $a_N = o(b_N)$ is adopted for scalars and finite-dimensional vectors and matrices (whose number of rows and columns are not a function of N). Finally, the notation $a_N = O_p(b_N)$ and $a_N = o_p(b_N)$ means that the previous statements hold in probability. The notation $vec(A)$ for an arbitrary matrix A stands for an operator that transforms the matrix A into a column vector by vertically stacking the columns of the matrix. The notation $vech(A)$ for an arbitrary symmetric matrix A indicates an operator that transforms a symmetric matrix into a column vector by stacking the elements in the lower triangular part of A. We use $\otimes$ to denote the Kronecker product.

IA.2 Assumptions

This section provides a set of assumptions we use in the lemmas and propositions of Sections IA.3 and IA.4, respectively.

Assumption IA.2.1 (Systematic candidate factors)**.** *We assume that a candidate model contains* K^{can} *systematic factors* f_t^{can}*, that is,* $\beta^{\text{can}\prime} V_e^{-1} \beta^{\text{can}}/N \to D$ *, where* $D > 0$ *is a* $K^{\text{can}} \times K^{\text{can}}$ *matrix.*

Assumption IA.2.2 (Asymptotic orthogonality of β^{can} and a)**.** *We assume that* $\beta^{\text{can}\prime} V_e^{-1} a = o(N^{\frac{1}{2}})$*.*

Assumption IA.2.3 (Systematic missing factors)**.** *We assume that a candidate factor model is missing* K^{mis} *systematic factors* f_t^{mis}*, that is,* $\beta^{\text{mis}\prime} V_e^{-1} \beta^{\text{mis}}/N \to E$*, where* $E > 0$ *is some* $K^{\text{mis}} \times K^{\text{mis}}$ *matrix.*

Assumption IA.2.4 (Asymptotic orthogonality of β^{mis} and a)**.** *We assume that* $\beta^{\text{mis}\prime} V_e^{-1} a = o(N^{\frac{1}{2}})$*.*

Remark. Assumptions IA.2.2 and IA.2.4 represent asymptotic orthogonality conditions because they imply that as $N \to \infty$, $\beta^{\text{can}\prime} V_e^{-1} a/N \to 0$ and $\beta^{\text{mis}\prime} V_e^{-1} a/N \to 0$.[38]

IA.3 Lemmas

We now provide a set of lemmas that will be useful for proving our propositions.

Lemma IA.3.1. *For a normally-distributed vector* $z \sim N(\mu_z, \Sigma_z)$*, and a constant vector* d*:*

$$\mathbb{E}(z e^{d' z}) = \mu^* e^{\frac{1}{2}(\mu^{*\prime} \Sigma_z^{-1} \mu^* - \mu_z' \Sigma_z^{-1} \mu_z)}, \quad \textit{where} \quad \mu^* = (\mu_z + \Sigma_z d).$$

Proof: Denote by n_z the dimension of the vector z. Use the definition of the mathematical expectation to obtain

$$\mathbb{E}(z e^{d' z}) = \frac{1}{(\sqrt{2\pi})^{n_z} |\Sigma_z|^{\frac{1}{2}}} \int_{-\infty}^{\infty} z e^{d' z} e^{-\frac{1}{2}(z-\mu_z)' \Sigma_z^{-1} (z-\mu_z)} dz.$$

Note that

$$\begin{aligned}
e^{d' z} e^{-\frac{1}{2}(z-\mu_z)' \Sigma_z^{-1} (z-\mu_z)} &= e^{d' z - \frac{1}{2} z' \Sigma_z^{-1} z - \frac{1}{2} \mu_z' \Sigma_z^{-1} \mu_z + \mu_z' \Sigma_z^{-1} z} \\
&= e^{-\frac{1}{2} z' \Sigma_z^{-1} z - \frac{1}{2} \mu_z' \Sigma_z^{-1} \mu_z + (\Sigma_z d + \mu_z)' \Sigma_z^{-1} z} \\
&= e^{-\frac{1}{2} z' \Sigma_z^{-1} z - \frac{1}{2} \mu_z' \Sigma_z^{-1} \mu_z + \mu^{*\prime} \Sigma_z^{-1} z} \\
&= e^{-\frac{1}{2} \mu_z' \Sigma_z^{-1} \mu_z + \frac{1}{2} \mu^{*\prime} \Sigma_z^{-1} \mu^*} e^{-\frac{1}{2} z' \Sigma_z^{-1} z + \mu^{*\prime} \Sigma_z^{-1} z - \frac{1}{2} \mu^{*\prime} \Sigma_z^{-1} \mu^*} \\
&= e^{-\frac{1}{2} \mu_z' \Sigma_z^{-1} \mu_z + \frac{1}{2} \mu^{*\prime} \Sigma_z^{-1} \mu^*} e^{-\frac{1}{2}(z-\mu^*)' \Sigma_z^{-1} (z-\mu^*)},
\end{aligned}$$

[38] Assumptions IA.2.1 and IA.2.3, together with asymptotic no arbitrage, by the Cauchy-Schwarz inequality, imply that $\beta^{\text{can}\prime} V_e^{-1} a = O(N^{1/2})$ and $\beta^{\text{mis}\prime} V_e^{-1} a = O(N^{1/2})$ but we need a slightly slower rate.

implying that

$$\mathbb{E}(ze^{d'z}) = e^{-\frac{1}{2}\mu_z'\Sigma_z^{-1}\mu_z + \frac{1}{2}\mu^{*\prime}\Sigma_z^{-1}\mu^*} \times \left(\frac{1}{(\sqrt{2\pi})^{n_z}|\Sigma_z|^{\frac{1}{2}}} \int_{-\infty}^{\infty} ze^{-\frac{1}{2}(z-\mu^*)'\Sigma_z^{-1}(z-\mu^*)} dz \right)$$
$$= e^{-\frac{1}{2}\mu_z'\Sigma_z^{-1}\mu_z + \frac{1}{2}\mu^{*\prime}\Sigma_z^{-1}\mu^*}\mu^*. \qquad \square$$

Lemma IA.3.2. *Under Assumptions IA.2.1 and IA.2.3:*

$$\beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{can}} = O(N).$$

Proof: We apply the Cauchy-Schwarz inequality for matrices and obtain

$$0 \leq ||\beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{can}}|| \leq ||\beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}}||^{\frac{1}{2}} \times ||\beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}}||^{\frac{1}{2}} = O(N). \qquad \square$$

Lemma IA.3.3. *Under Assumptions IA.2.1 and IA.2.3:*

$$\beta^{\text{can}\prime}V_\varepsilon^{-1}\beta^{\text{can}} = O(N).$$

Proof: Recall that $V_\varepsilon = \beta^{\text{mis}}V_{f^{\text{mis}}}\beta^{\text{mis}\prime} + V_e$ and apply the Sherman-Morrison-Woodbury formula to V_ε^{-1} to obtain

$$\beta^{\text{can}\prime}V_\varepsilon^{-1}\beta^{\text{can}} = \beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}} - \beta^{\text{can}\prime}V_e^{-1}\beta^{\text{mis}}(V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}})^{-1}\beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{can}}$$
$$= O(N) + O(N) \times [O(1) + O(N)]^{-1} \times O(N) = O(N). \qquad \square$$

Lemma IA.3.4. *Under Assumption IA.2.3:*

$$\beta^{\text{mis}\prime}V_\varepsilon^{-1}\beta^{\text{mis}} \to V_{f^{\text{mis}}}^{-1} \quad as \quad N \to \infty,$$

implying that $\beta^{\text{mis}\prime}V_\varepsilon^{-1}\beta^{\text{mis}} = O(1)$.

Proof: Recall that $V_\varepsilon = \beta^{\text{mis}}V_{f^{\text{mis}}}\beta^{\text{mis}\prime} + V_e$ and apply the Sherman-Morrison-Woodbury formula to V_ε^{-1} to obtain

$$\beta^{\text{mis}\prime}V_\varepsilon^{-1}\beta^{\text{mis}} = \beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}} - \beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}}(V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}})^{-1}\beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}}$$
$$= V_{f^{\text{mis}}}^{-1}(V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}})^{-1}\beta^{\text{mis}\prime}V_e^{-1}\beta^{\text{mis}}$$
$$= V_{f^{\text{mis}}}^{-1} \times [O(1) + O(N)]^{-1} \times O(N) \to V_{f^{\text{mis}}}^{-1}. \qquad \square$$

Lemma IA.3.5. *Under Assumptions IA.2.1 and IA.2.3:*

$$\beta^{\text{mis}\prime}V_\varepsilon^{-1}\beta^{\text{can}} = O(1).$$

Proof: Recall that $V_\varepsilon = \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e$ and apply the Sherman-Morrison-Woodbury formula to $V_\varepsilon{}^{-1}$ to obtain

$$\begin{aligned}
\beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}} &= \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{can}} - \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{can}} \\
&= V_{f^{\text{mis}}}^{-1} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{can}} \\
&= O(1) \times [O(1) + O(N)]^{-1} \times O(N) = O(1). \qquad \square
\end{aligned}$$

Lemma IA.3.6. *Under Assumptions IA.2.3 and IA.2.4:*

$$a' V_\varepsilon{}^{-1} a - a' V_e{}^{-1} a \to 0 \quad \textit{as} \quad N \to \infty.$$

Proof: Recall that $V_\varepsilon = \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e$ and apply the Sherman-Morrison-Woodbury formula to $V_\varepsilon{}^{-1}$ to obtain

$$\begin{aligned}
a' V_\varepsilon{}^{-1} a &= a' V_e{}^{-1} a - a' V_e{}^{-1} \beta^{\text{mis}} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e{}^{-1} a \\
&= a' V_e{}^{-1} a + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times o(N^{\frac{1}{2}}) \\
&= a' V_e{}^{-1} a + o(1),
\end{aligned}$$

where $a' V_e{}^{-1} a = O(1)$. $\square$

Lemma IA.3.7. *Under Assumptions IA.2.1, IA.2.2, IA.2.3 and IA.2.4:*

$$\beta^{\text{can}\prime} V_\varepsilon{}^{-1} a = o(N^{\frac{1}{2}}).$$

Proof: Recall that $V_\varepsilon = \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e$ and apply the Sherman-Morrison-Woodbury formula to $V_\varepsilon{}^{-1}$ to obtain

$$\begin{aligned}
\beta^{\text{can}\prime} V_\varepsilon{}^{-1} a &= \beta^{\text{can}\prime} V_e{}^{-1} a - \beta^{\text{can}\prime} V_e{}^{-1} \beta^{\text{mis}} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e{}^{-1} a \\
&= o(N^{\frac{1}{2}}) + O(N) \times [O(1) + O(N)]^{-1} \times o(N^{\frac{1}{2}}) \\
&= o(N^{\frac{1}{2}}). \qquad \square
\end{aligned}$$

Lemma IA.3.8. *Under Assumptions IA.2.3 and IA.2.4:*

$$\beta^{\text{mis}\prime} V_\varepsilon{}^{-1} a = o(N^{-\frac{1}{2}}).$$

Proof: Recall that $V_\varepsilon = \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e$ and apply the Sherman-Morrison-Woodbury formula to $V_\varepsilon{}^{-1}$ to obtain

$$\begin{aligned}
\beta^{\text{mis}\prime} V_\varepsilon{}^{-1} a &= \beta^{\text{mis}\prime} V_e{}^{-1} a - \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e{}^{-1} a \\
&= V_{f^{\text{mis}}}^{-1} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e{}^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e{}^{-1} a \\
&= O(1) \times [O(1) + O(N)]^{-1} \times o(N^{\frac{1}{2}}) = o(N^{-\frac{1}{2}}). \qquad \square
\end{aligned}$$

Lemma IA.3.9. *Let e be an $N \times 1$ random vector with zero mean and covariance matrix V_e. Under Assumptions IA.2.1 and IA.2.3:*

$$\beta^{\mathrm{can}\prime} V_\varepsilon^{-1} e = O_p(N^{\frac{1}{2}}).$$

Proof: For any random variable X with a finite second moment, we have that $X = O_p((\mathbb{E}(X^2))^{\frac{1}{2}})$. If $X = \beta^{\mathrm{can}\prime} V_e^{-1} e$, then

$$\mathbb{E}(\beta^{\mathrm{can}\prime} V_e^{-1} e e' V_e^{-1} \beta^{\mathrm{can}}) = \beta^{\mathrm{can}\prime} V_e^{-1} \beta^{\mathrm{can}} = O(N),$$

and therefore, $\beta^{\mathrm{can}\prime} V_e^{-1} e = O_p(N^{\frac{1}{2}})$. Similarly, we can show that $\beta^{\mathrm{mis}\prime} V_e^{-1} e = O_p(N^{\frac{1}{2}})$. Apply the Sherman-Morrison-Woodbury formula to V_ε^{-1} and use Lemma IA.3.2 to obtain

$$\begin{aligned}
\beta^{\mathrm{can}\prime} V_\varepsilon^{-1} e &= \beta^{\mathrm{can}\prime} V_e^{-1} e - \beta^{\mathrm{can}\prime} V_e^{-1} \beta^{\mathrm{mis}} (V_{f^{\mathrm{mis}}}^{-1} + \beta^{\mathrm{mis}\prime} V_e^{-1} \beta^{\mathrm{mis}})^{-1} \beta^{\mathrm{mis}\prime} V_e^{-1} e \\
&= O_p(N^{\frac{1}{2}}) + O(N) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) = O_p(N^{\frac{1}{2}}). \qquad \square
\end{aligned}$$

Lemma IA.3.10. *Under Assumption IA.2.3:*

$$\beta^{\mathrm{mis}\prime} V_\varepsilon^{-1} e = O_p(N^{-\frac{1}{2}}).$$

Proof: From the proof of Lemma IA.3.9, $\beta^{\mathrm{mis}\prime} V_e^{-1} e = O_p(N^{\frac{1}{2}})$. Apply the Sherman-Morrison-Woodbury formula to V_ε^{-1} and use Lemma IA.3.2 to obtain

$$\begin{aligned}
\beta^{\mathrm{mis}\prime} V_\varepsilon^{-1} e &= \beta^{\mathrm{mis}\prime} V_e^{-1} e - \beta^{\mathrm{mis}\prime} V_e^{-1} \beta^{\mathrm{mis}} (V_{f^{\mathrm{mis}}}^{-1} + \beta^{\mathrm{mis}\prime} V_e^{-1} \beta^{\mathrm{mis}})^{-1} \beta^{\mathrm{mis}\prime} V_e^{-1} e \\
&= V_{f^{\mathrm{mis}}}^{-1} (V_{f^{\mathrm{mis}}}^{-1} + \beta^{\mathrm{mis}\prime} V_e^{-1} \beta^{\mathrm{mis}})^{-1} \beta^{\mathrm{mis}\prime} V_e^{-1} e \\
&= O(1) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) = O_p(N^{-\frac{1}{2}}). \qquad \square
\end{aligned}$$

IA.4 Proofs of Propositions

In this section, we provide the proofs for the propositions in the manuscript.

Proof of Proposition 1

We use a guess-and-verify method to derive the SDF. Specifically, we guess that the SDF has the following functional form

$$M_{t+1} = \mathbb{E}(M_{t+1}) + b'(f_{t+1} - \mathbb{E}(f_{t+1})) + c' e_{t+1},$$

where b is a $K \times 1$ vector and c is an $N \times 1$ vector. We identify the unknown vector b and c by using the Law of One Price. Specifically, because we assume the existence of the risk-free asset, to determine the mean of the SDF we use the following condition:

$$\mathbb{E}(M_{t+1}) = \frac{1}{R_f}.$$

Next, because λ represents the vector of prices of risk of f_{t+1}, we have that

$$-\operatorname{cov}(M_{t+1}, f_{t+1}) \times R_f = \lambda.$$

These K conditions identify b:

$$b = -\frac{V_f^{-1}\lambda}{R_f}.$$

Finally, it must be the case that the SDF M_{t+1} prices the N assets:

$$\mathbb{E}\big(M_{t+1}(R_{t+1} - R_f 1_N)\big) = 0_N.$$

These N equations identify c:

$$c = -\frac{V_e^{-1}a}{R_f}.$$

Taken together

$$M_{t+1} = M_{t+1}^{\beta} + M_{t+1}^{a},$$

where

$$M_{t+1}^{\beta} = \frac{1}{R_f} - \frac{\lambda' V_f^{-1}}{R_f}(f_{t+1} - \mathbb{E}(f_{t+1})) \quad \text{and}$$

$$M_{t+1}^{a} = -\frac{a' V_e^{-1}}{R_f} e_{t+1}.$$

Pairwise uncorrelatedness of f_t and e_t implies that the covariance between M_{t+1}^{β} and M_{t+1}^{a} is zero. □

Proof of Proposition 2

First, we prove that the exponential SDF specified in equation (8) is the APT-implied SDF. We use a guess-and-verify method. We guess that the SDF has the following functional form:

$$M_{\exp,t+1} = \exp\left[\mu_+ + b_+'(f_{t+1} - \mathbb{E}(f_{t+1})) + c_+' e_{t+1}\right],$$

with unknown vectors b_+ and c_+, as well as an unknown scalar μ_+.

To identify the unknowns and verify our guess we use the following $K + N + 1$ equations, which are implications of the Law of One Price:

$$\begin{aligned} -\operatorname{cov}(M_{\exp,t+1}, f_{t+1}) \times R_f &= \lambda, \\ \mathbb{E}(M_{\exp,t+1}(R_{t+1} - R_f 1_N)) &= 0_N, \\ \mathbb{E}(M_{\exp,t+1}) &= \frac{1}{R_f}. \end{aligned}$$

The first K equations imply that

$$-\mathbb{E}(M_{\exp,t+1}(f_{t+1}-\mathbb{E}(f_{t+1})))=\mathbb{E}(M_{\exp,t+1})\times\lambda,$$

which, along with Lemma IA.3.1, gives

$$b_+=-V_f^{-1}\lambda.$$

The next N equations and Lemma IA.3.1 imply that

$$\begin{aligned}0_N&=\mathbb{E}(M_{\exp,t+1}(R_{t+1}-R_f1_N))=\mathbb{E}(M_{\exp,t+1}(a+\beta\lambda+\beta(f_{t+1}-\mathbb{E}(f_{t+1}))+e_{t+1}))\\&=(a+\beta\lambda)\mathbb{E}(M_{\exp,t+1})+\mathbb{E}(M_{\exp,t+1}e_{t+1})+\mathbb{E}(M_{\exp,t+1}\beta(f_{t+1}-\mathbb{E}(f_{t+1})))\\&=(a+\beta\lambda)\mathbb{E}(M_{\exp,t+1})+V_ec_+\mathbb{E}(M_{\exp,t+1})-\beta\lambda\mathbb{E}(M_{\exp,t+1})=(a+V_ec_+)\mathbb{E}(M_{\exp,t+1}).\end{aligned}$$

As a result,

$$c_+=-V_e^{-1}a.$$

Finally, the last identifying condition implies

$$\begin{aligned}R_f^{-1}&=\mathbb{E}(M_{\exp,t+1})\\&=\mathbb{E}(\exp[\mu_++b_+'(f_{t+1}-\mathbb{E}(f_{t+1}))+c_+'e_{t+1}])=\exp\left[\mu_++b_+'V_fb_+/2+c_+'V_ec_+/2\right].\end{aligned}$$

Thus,

$$\exp(\mu_+)=R_f^{-1}\times\exp\left[-\lambda'V_f^{-1}\lambda/2-a'V_e^{-1}a/2\right].$$

Collecting all these results, we obtain

$$M_{\exp,t+1}=M^{\beta}_{\exp,t+1}\times M^{a}_{\exp,t+1},$$

where

$$\begin{aligned}M^{\beta}_{\exp,t+1}&=R_f^{-1}\times\exp\left[-\lambda'V_f^{-1}(f_{t+1}-\mathbb{E}(f_{t+1}))-\lambda'V_f^{-1}\lambda/2)\right],\\M^{a}_{\exp,t+1}&=\exp\left[-a'V_e^{-1}e_{t+1}-a'V_e^{-1}a/2\right].\end{aligned}$$

Next, we prove that, as $N\to\infty$, the feasible SDF given in equation (11) recovers the exponential SDF (8). We start by analyzing the exponent of $\hat{M}^{a}_{\exp,t+1}$:

$$-a'V_R^{-1}(R_{t+1}-\mathbb{E}[R_{t+1}])-\frac{1}{2}a'V_R^{-1}a=-a'V_R^{-1}\beta(f_{t+1}-\mathbb{E}(f_{t+1}))-a'V_R^{-1}e_{t+1}-\frac{1}{2}a'V_R^{-1}a.$$

We apply the Sherman-Morrison-Woodbury formula to V_R^{-1}, use Assumptions IA.2.1 and IA.2.2, and Lemma IA.3.9 to obtain[39]

$$a'V_R^{-1}\beta=a'V_e^{-1}\beta-a'V_e^{-1}\beta(V_f^{-1}+\beta'V_e^{-1}\beta)^{-1}\beta'V_e^{-1}\beta$$

[39] The APT model contains K systematic factors f_{t+1}, so it satisfies Assumptions IA.2.1 and IA.2.2, where one replaces K^{can} by K, f^{can}_{t+1} by f_{t+1}, and β^{can} by β.

$$
\begin{aligned}
&= a'V_e^{-1}\beta(V_f^{-1} + \beta'V_e^{-1}\beta)^{-1}V_f^{-1} \\
&= o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O(1) \\
&= o(N^{-1/2}), \\
a'V_R^{-1}e_{t+1} &= a'V_e^{-1}e_{t+1} - a'V_e^{-1}\beta(V_f^{-1} + \beta'V_e^{-1}\beta)^{-1}\beta'V_e^{-1}e_{t+1} \\
&= a'V_e^{-1}e_{t+1} + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) \\
&= a'V_e^{-1}e_{t+1} + o_p(1), \\
a'V_R^{-1}a &= a'V_e^{-1}a - a'V_e^{-1}\beta(V_f^{-1} + \beta'V_e^{-1}\beta)^{-1}\beta'V_e^{-1}a \\
&= a'V_e{}^{-1}a + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times o(N^{\frac{1}{2}}) \\
&= a'V_e^{-1}a + o(1).
\end{aligned}
$$

These three results imply that

$$
-a'V_R^{-1}(R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2}a'V_R^{-1}a = -a'V_e^{-1}e_{t+1} - \frac{1}{2}a'V_e^{-1}a + o_p(1),
$$

and therefore, by Slutzky's theorem,

$$
\hat{M}^a_{\exp,t+1} - M^a_{\exp,t+1} \xrightarrow{p} 0 \qquad \text{as} \quad N \to \infty.
$$

Next, we analyze the exponent of $\hat{M}^\beta_{\exp,t+1}$:

$$
\begin{aligned}
&-(\beta\lambda)'V_R^{-1}(R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2}(\beta\lambda)'V_R^{-1}\beta\lambda \\
&\quad = -(\beta\lambda)'V_R^{-1}\beta(f_{t+1} - \mathbb{E}(f_{t+1})) - (\beta\lambda)'V_R^{-1}e_{t+1} - \frac{1}{2}(\beta\lambda)'V_R^{-1}\beta\lambda.
\end{aligned}
$$

We apply the Sherman-Morrison-Woodbury formula to V_R^{-1}, use Assumptions IA.2.1 and Lemma IA.3.9 to obtain

$$
\begin{aligned}
\beta'V_R^{-1}\beta &= \beta'V_e^{-1}\beta - \beta'V_e^{-1}\beta(V_f^{-1} + \beta'V_e^{-1}\beta)^{-1}\beta'V_e^{-1}\beta \\
&= V_f^{-1} + o(1),
\end{aligned}
$$

$$
\begin{aligned}
\beta'V_R^{-1}e_{t+1} &= \beta'V_e^{-1}e_{t+1} - \beta'V_e^{-1}\beta(V_f^{-1} + \beta'V_e^{-1}\beta)^{-1}\beta'V_e^{-1}e_{t+1} \\
&= O_p(N^{-\frac{1}{2}}) + O(1) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) \\
&= O_p(N^{-\frac{1}{2}}).
\end{aligned}
$$

These two results imply that

$$
-(\beta\lambda)'V_R^{-1}(R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2}(\beta\lambda)'V_R^{-1}\beta\lambda
$$

$$= -\lambda' V_f^{-1}(f_{t+1} - \mathbb{E}(f_{t+1})) - \frac{1}{2}\lambda' V_f^{-1}\lambda + o_p(1)$$

and therefore

$$\hat{M}^{\beta}_{\exp,t+1} - M^{\beta}_{\exp,t+1} \xrightarrow{p} 0 \quad \text{as} \quad N \to \infty.$$

Independence of f_t and e_t implies that the covariance between $M^{\beta}_{\exp,t+1}$ and $M^{a}_{\exp,t+1}$ is zero. The same remains true asymptotically for the projected versions, $\hat{M}^{a}_{\exp,t+1}$ and $\hat{M}^{\beta}_{\exp,t+1}$, thanks to the asymptotic-in-N equivalences proven above. $\square$

Proof of an extended version of Proposition 2

The following proposition extends Proposition 2 to the case in which we correct a misspecified candidate factor model to obtain the APT-implied SDF. The correction includes the systematic risk factors missing in the candidate factor model and compensation for unsystematic risk. This case is described in Section 2.3 of the manuscript.

Proposition IA.4.1 (Feasible Admissible SDF Constructed by Correcting a Candidate Factor Model)**.** *Consider a candidate factor model with K^{can} factors f^{can}_{t+1}. Suppose the first K^{mis} eigenvalues of the covariance matrix V_ε are unbounded when $N \to \infty$, the remaining eigenvalues are uniformly bounded, and the smallest eigenvalue is strictly positive. Under Assumptions 2.1 and 2.2 of the APT and the assumption that the factors f^{can}_{t+1} and f^{mis}_{t+1} and unsystematic shocks e_{t+1} are jointly Gaussian, the SDF*

$$M_{\exp,t+1} = M^{a}_{\exp,t+1} \times M^{\beta,\text{can}}_{\exp,t+1} \times M^{\beta,\text{mis}}_{\exp,t+1} \quad \textit{with} \tag{IA1}$$

$$M^{a}_{\exp,t+1} = \exp\left(-a' V_e^{-1} e_{t+1} - \frac{1}{2} a' V_e^{-1} a\right),$$

$$M^{\beta,\text{can}}_{\exp,t+1} = \frac{1}{R_f} \times \exp\left(-\lambda^{\text{can}\prime} V^{-1}_{f^{\text{can}}}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1}) - \frac{1}{2}\lambda^{\text{can}\prime} V^{-1}_{f^{\text{can}}}\lambda^{\text{can}}\right),$$

$$M^{\beta,\text{mis}}_{\exp,t+1} = \exp\left(-\lambda^{\text{mis}\prime} V^{-1}_{f^{\text{mis}}}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1}) - \frac{1}{2}\lambda^{\text{mis}\prime} V^{-1}_{f^{\text{mis}}}\lambda^{\text{mis}})\right),$$

prices assets correctly. Furthermore, under Assumptions IA.2.1–IA.2.4, as $N \to \infty$, the following results hold

$$\hat{M}^{a}_{\exp,t+1} - M^{a}_{\exp,t+1} \xrightarrow{p} 0, \quad \hat{M}^{\beta,\text{mis}}_{\exp,t+1} - M^{\beta,\text{mis}}_{\exp,t+1} \xrightarrow{p} 0, \quad \text{cov}(\hat{M}^{\beta,\text{mis}}_{\exp,t+1}, \hat{M}^{a}_{\exp,t+1}) \to 0,$$

where

$$\hat{M}^{a}_{\exp,t+1} = \exp\Big(-a' V_R^{-1}(R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2} a' V_R^{-1} a\Big), \qquad \textit{and}$$

$$\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1} = \exp\Big(- (\beta^{\text{mis}}\lambda^{\text{mis}})' V_R^{-1}(R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2}(\beta^{\text{mis}}\lambda^{\text{mis}})' V_R^{-1}\beta^{\text{mis}}\lambda^{\text{mis}}\Big),$$

implying that

$$\hat{M}_{\text{exp},t+1} = M^{\beta,\text{can}}_{\text{exp},t+1} \times \hat{M}^{\beta,\text{mis}}_{\text{exp},t+1} \times \hat{M}^{a}_{\text{exp},t+1} \tag{IA2}$$

is a feasible positive SDF that prices assets correctly, when $N \to \infty$.

Proof: First, we prove that the exponential SDF specified in equation (IA1) prices assets correctly. To this end, we use a guess-and-verify method. We guess that the exponential SDF is

$$M_{\text{exp},t+1} = \exp\left[\mu_+ + b^{\text{can}\prime}_+(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) + b^{\text{mis}\prime}_+(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})) + c'_+ e_{t+1}\right] \tag{IA3}$$

with unknown vectors b^{can}_+, b^{mis}_+, and c_+, as well as an unknown scalar μ_+.

To identify the unknowns and verify our guess we use the following $K^{\text{can}} + K^{\text{mis}} + N + 1$ equations, which are the implications of the Law of One Price:

$$\begin{aligned}
-\text{cov}(M_{\text{exp},t+1}, f^{\text{can}}_{t+1}) \times R_f &= \lambda^{\text{can}},\\
-\text{cov}(M_{\text{exp},t+1}, f^{\text{mis}}_{t+1}) \times R_f &= \lambda^{\text{mis}},\\
\mathbb{E}(M_{\text{exp},t+1}(R_{t+1} - R_f 1_N)) &= 0_N,\\
\mathbb{E}(M_{\text{exp},t+1}) &= \frac{1}{R_f}.
\end{aligned}$$

The first K^{can} equations imply that

$$-\mathbb{E}(M_{\text{exp},t+1}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1}))) = \mathbb{E}(M_{\text{exp},t+1}) \times \lambda^{\text{can}},$$

which, along with Lemma IA.3.1, gives:

$$b^{\text{can}}_+ = -V^{-1}_{f^{\text{can}}}\lambda^{\text{can}}.$$

Similarly, the next K^{mis} equations imply that

$$-\mathbb{E}(M_{\text{exp},t+1}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1}))) = \mathbb{E}(M_{\text{exp},t+1}) \times \lambda^{\text{mis}},$$

which, along with Lemma IA.3.1, leads to:

$$b^{\text{mis}}_+ = -V^{-1}_{f^{\text{mis}}}\lambda^{\text{mis}}.$$

Then, we use the next N equations and Lemma IA.3.1 to obtain

$$\begin{aligned}
0_N &= \mathbb{E}(M_{\text{exp},t+1}(R_{t+1} - R_f 1_N))\\
&= \mathbb{E}(M_{\text{exp},t+1}(a + \beta^{\text{mis}}\lambda^{\text{mis}} + \beta^{\text{can}}\lambda^{\text{can}} + \beta^{\text{can}}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1}))
\end{aligned}$$

$$
\begin{aligned}
&\quad + \beta^{\text{mis}}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})) + e_{t+1}))\\
&= (a + \beta^{\text{mis}}\lambda^{\text{mis}} + \beta^{\text{can}}\lambda^{\text{can}})\mathbb{E}(M_{\exp,t+1}) + \mathbb{E}(M_{\exp,t+1}e_{t+1})\\
&\quad + \mathbb{E}(M_{\exp,t+1}\beta^{\text{can}}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1}))) + \mathbb{E}(M_{\exp,t+1}\beta^{\text{mis}}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})))\\
&= (a + \beta^{\text{mis}}\lambda^{\text{mis}} + \beta^{\text{can}}\lambda^{\text{can}})\mathbb{E}(M_{\exp,t+1}) + V_e c_+ \mathbb{E}(M_{\exp,t+1})\\
&\quad - \beta^{\text{can}}\lambda^{\text{can}}\mathbb{E}(M_{\exp,t+1}) - \beta^{\text{mis}}\lambda^{\text{mis}}\mathbb{E}(M_{\exp,t+1})\\
&= (a + V_e c_+)\mathbb{E}(M_{\exp,t+1}).
\end{aligned}
$$

As a result,

$$
c_+ = -V_e^{-1}a.
$$

Finally, the last identifying condition implies

$$
\begin{aligned}
R_f^{-1} &= \mathbb{E}(M_{\exp,t+1})\\
&= \mathbb{E}(\exp[\mu_+ + b^{\text{can}\prime}_+(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) + b^{\text{mis}\prime}_+(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})) + c'_+ e_{t+1}])\\
&= \exp\left[\mu_+ + b^{\text{can}\prime}_+ V_{f^{\text{can}}} b^{\text{can}}_+/2 + b^{\text{mis}\prime}_+ V_{f^{\text{mis}}} b^{\text{mis}}_+/2 + c'_+ V_e c_+/2\right].
\end{aligned}
$$

Thus,

$$
\exp(\mu_+) = R_f^{-1} \times \exp\big[-\lambda^{\text{can}\prime}V^{-1}_{f^{\text{can}}}\lambda^{\text{can}}/2 - \lambda^{\text{mis}\prime}V^{-1}_{f^{\text{mis}}}\lambda^{\text{mis}}/2 - a'V_e^{-1}a/2\big].
$$

We substitute b^{can}_+, b^{mis}_+, c_+, and μ_+ in equation (IA3) and obtain the exponential SDF given in equation (IA1). Thus, we have successfully verified our guess.

Next, we prove that, as $N \to \infty$, the feasible SDF given in equation (IA2) recovers the exponential SDF specified in equation (IA1). We start by analyzing the exponent of $M^a_{\exp,t+1}$. First, we note that

$$
\begin{aligned}
-a'V_R^{-1}(R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2}a'V_R^{-1}a = &-a'V_R^{-1}\beta^{\text{can}}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1}))\\
&- a'V_R^{-1}\beta^{\text{mis}}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1}))\\
&- a'V_R^{-1}e_{t+1}\\
&- \frac{1}{2}a'V_R^{-1}a.
\end{aligned}
$$

We analyze the four right-hand-side terms one-by-one. We apply the Sherman-Morrison-Woodbury formula to V_R^{-1} and V_ε^{-1} and use Lemmas IA.3.3, IA.3.5–IA.3.9, and the proof of Lemma IA.3.10 to obtain

$$
\begin{aligned}
a'V_R^{-1}\beta^{\text{can}} &= a'V_\varepsilon{}^{-1}\beta^{\text{can}} - a'V_\varepsilon{}^{-1}\beta^{\text{can}}(V^{-1}_{f^{\text{can}}} + \beta^{\text{can}\prime}V_\varepsilon{}^{-1}\beta^{\text{can}})^{-1}\beta^{\text{can}\prime}V_\varepsilon{}^{-1}\beta^{\text{can}}\\
&= a'V_\varepsilon^{-1}\beta^{\text{can}}(V^{-1}_{f^{\text{can}}} + \beta^{\text{can}\prime}V_\varepsilon^{-1}\beta^{\text{can}})^{-1}V^{-1}_{f^{\text{can}}}
\end{aligned}
$$

$$
\begin{aligned}
&= o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O(1) \\
&= o(N^{-1/2}), \\
a' V_R^{-1} \beta^{\text{mis}} &= a' V_\varepsilon{}^{-1} \beta^{\text{mis}} - a' V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{mis}} \\
&= o(N^{-\frac{1}{2}}) + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O(1) = \\
&= o(N^{-1/2}), \\
a' V_R^{-1} e_{t+1} &= a' V_\varepsilon{}^{-1} e_{t+1} - a' V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} \beta^{\text{can}\prime} V_\varepsilon{}^{-1} e_{t+1} \\
&= a' V_\varepsilon{}^{-1} e_{t+1} + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) \\
&= a' V_\varepsilon{}^{-1} e_{t+1} + o_p(1) \\
&= a' V_e^{-1} e_{t+1} - a' V_e^{-1} \beta^{\text{mis}} (V_{f^{\text{mis}}}^{-1} + \beta^{\text{mis}\prime} V_e^{-1} \beta^{\text{mis}})^{-1} \beta^{\text{mis}\prime} V_e^{-1} e_{t+1} + o_p(1) \\
&= a' V_e^{-1} e_{t+1} + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) + o_p(1) \\
&= a' V_e^{-1} e_{t+1} + o_p(1), \\
a' V_R^{-1} a &= a' V_\varepsilon{}^{-1} a - a' V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} \beta^{\text{can}\prime} V_\varepsilon{}^{-1} a \\
&= a' V_e{}^{-1} a + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times o(N^{\frac{1}{2}}) + o(1) \\
&= a' V_e^{-1} a + o(1).
\end{aligned}
$$

These four results imply that

$$
- a' V_R^{-1} (R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2} a' V_R^{-1} a = -a' V_e^{-1} e_{t+1} - \frac{1}{2} a' V_e^{-1} a + o_p(1),
$$

and therefore we obtain

$$
\hat{M}^a_{\exp,t+1} - M^a_{\exp,t+1} \xrightarrow{p} 0 \qquad \text{as} \quad N \to \infty.
$$

Next, we analyze the exponent of $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$:

$$
\begin{aligned}
&- (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} (R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2} (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} \beta^{\text{mis}} \lambda^{\text{mis}} \\
&\quad = -(\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} \beta^{\text{can}} (f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) \\
&\qquad - (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} \beta^{\text{mis}} (f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})) \\
&\qquad - (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} e_{t+1} \\
&\qquad - \frac{1}{2} (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} \beta^{\text{mis}} \lambda^{\text{mis}}.
\end{aligned}
$$

We apply the Sherman-Morrison-Woodbury formula and Lemmas IA.3.3, IA.3.4, IA.3.5, IA.3.9, and IA.3.10 to the first three terms above:

$$
\beta^{\text{mis}\prime} V_R^{-1} \beta^{\text{can}} = \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}} - \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}}
$$

$$\begin{aligned}
&= \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} V_{f^{\text{can}}}^{-1} \\
&= O(1) \times [O(1) + O(N)]^{-1} \times O(1) \\
&= O(N^{-1}),
\end{aligned}$$

$$\begin{aligned}
\beta^{\text{mis}\prime} V_R^{-1} \beta^{\text{mis}} &= \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{mis}} - \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{mis}} \\
&= (V_{f^{\text{mis}}}^{-1} + o(1)) + O(1) \times [O(1) + O(N)]^{-1} \times O(1) \\
&= V_{f^{\text{mis}}}^{-1} + o(1), \quad \text{and}
\end{aligned}$$

$$\begin{aligned}
\beta^{\text{mis}\prime} V_R^{-1} e_{t+1} &= \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} e_{t+1} - \beta^{\text{mis}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}} (V_{f^{\text{can}}}^{-1} + \beta^{\text{can}\prime} V_\varepsilon{}^{-1} \beta^{\text{can}})^{-1} \beta^{\text{can}\prime} V_\varepsilon{}^{-1} e_{t+1} \\
&= O_p(N^{-\frac{1}{2}}) + O(1) \times [O(1) + O(N)]^{-1} \times O_p(N^{\frac{1}{2}}) \\
&= O_p(N^{-\frac{1}{2}}).
\end{aligned}$$

These three results imply that

$$\begin{aligned}
&- (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} (R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2} (\beta^{\text{mis}} \lambda^{\text{mis}})' V_R^{-1} \beta^{\text{mis}} \lambda^{\text{mis}} \\
&\quad = -\lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} (f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) - \frac{1}{2} \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} \lambda^{\text{mis}} + o_p(1),
\end{aligned}$$

and therefore we obtain

$$\hat{M}_{\exp,t+1}^{\beta,\text{mis}} - M_{\exp,t+1}^{\beta,\text{mis}} \xrightarrow{p} 0 \quad \text{as} \quad N \to \infty.$$

Pairwise uncorrelatedness (and independence by Gaussianity) of $f_{t+1}^{\text{can}}, f_{t+1}^{\text{mis}}$, and e_{t+1} implies that the pairwise covariances between $M_{\exp,t+1}^{\beta,\text{can}}, M_{\exp,t+1}^{\beta,\text{mis}}$, and $M_{\exp,t+1}^{a}$ are all zero. The same remains true asymptotically for the projected versions, $\hat{M}_{\exp,t+1}^{a}$ and $\hat{M}_{\exp,t+1}^{\beta,\text{mis}}$, thanks to the asymptotic-in-N equivalences proven above. □

Remark: This proposition assumes the presence of at least one omitted systematic risk factor, that is, $K^{\text{mis}} > 0$. If instead $K^{\text{mis}} = 0$, that is, all eigenvalues of V_ε are bounded, then the data-generating process of asset returns with K^{can} factors given in expression (12) satisfies the assumptions of the classical APT provided in Section 2.1.

Proof of Proposition 3

We use a guess-and-verify method to derive the SDF. We guess that the SDF has the following functional form

$$M_{t+1} = \mathbb{E}(M_{t+1}) + b^{\text{can}\prime} (f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) + b^{\text{mis}\prime} (f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) + c' e_{t+1},$$

where b^{can} is a $K^{\text{can}} \times 1$ vector , b^{mis} is a $K^{\text{mis}} \times 1$ vector, and c is an $N \times 1$ vector. We identify the unknown vectors b^{can}, b^{mis}, and c by using the Law of One Price. Specifically, because we assume the existence of the risk-free asset, to determine the mean of the SDF we use the following condition:

$$\mathbb{E}(M_{t+1}) = \frac{1}{R_f}.$$

Next, because λ^{can} represents a vector of prices of risk of f^{can}_{t+1}, we have that

$$-\operatorname{cov}(M_{t+1}, f^{\text{can}}_{t+1}) \times R_f = \lambda^{\text{can}}.$$

These K^{can} conditions identify b^{can}:

$$b^{\text{can}} = -\frac{V^{-1}_{f^{\text{can}}} \lambda^{\text{can}}}{R_f}.$$

Similarly, λ^{mis} is the price of risk associated with factors f^{mis}_{t+1}, or equivalently,

$$-\operatorname{cov}(M_{t+1}, f^{\text{mis}}_{t+1}) \times R_f = \lambda^{\text{mis}}.$$

These K^{mis} conditions identify b^{mis}:

$$b^{\text{mis}} = -\frac{V^{-1}_{f^{\text{mis}}} \lambda^{\text{mis}}}{R_f}.$$

Finally, it must be the case that the SDF prices the N assets

$$\mathbb{E}\big(M_{t+1}(R_{t+1} - R_f 1_N)\big) = 0_N.$$

These N equations identify c:

$$c = -\frac{V_e^{-1} a}{R_f}.$$

Taken together

$$M_{t+1} = M^{\beta,\text{can}}_{t+1} + M^{\beta,\text{mis}}_{t+1} + M^{a}_{t+1},$$

where

$$\begin{aligned}
M^{\beta,\text{can}}_{t+1} &= \frac{1}{R_f} - \frac{\lambda^{\text{can}\prime} V^{-1}_{f^{\text{can}}}}{R_f}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})), \\
M^{\beta,\text{mis}}_{t+1} &= -\frac{\lambda^{\text{mis}\prime} V^{-1}_{f^{\text{mis}}}}{R_f}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})), \\
M^{a}_{t+1} &= -\frac{a' V_e^{-1}}{R_f} e_{t+1}.
\end{aligned}$$

Pairwise uncorrelatedness of f^{can}_{t+1}, f^{mis}_{t+1}, and e_{t+1} implies that the pairwise covariances between $M^{\beta,\text{can}}_{t+1}$, $M^{\beta,\text{mis}}_{t+1}$, and M^{a}_{t+1} are all zero. $\square$

Proof of Proposition 4

Below we summarize the main assumptions of the model in Merton (1987) and then analyze its equilibrium implications for the SDF and expected excess returns. For details of the model, we refer the reader to Merton (1987).

Assume that there are N firms in the economy whose end-of-period cash flows are:

$$C_i = I_i \left[\mu_i + \eta_i Y + s_i \epsilon_i\right],$$

where, for simplicity, it is assumed that there is a single random systematic factor Y with $E(Y) = 0$ and $E(Y^2) = 1$, with $E(\varepsilon_i) = E\left(\varepsilon_i \mid \varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_{i+1}, \ldots, \varepsilon_N, Y\right) = 0$, for $i = \{1, \ldots, N\}$, where ε_i are asset-specific shocks.[40] Here, I_i is the amount of physical investment in firm i and μ_i, η_i, and s_i represent parameters of firm i 's production technology.

Let V_i denote the equilibrium value of firm i at the beginning of the period. If R_i is the equilibrium return per dollar from investing in firm i over the period, then $R_i = C_i/V_i$, and

$$R_i = \mathbb{E}(R_i) + b_i Y + \sigma_i \epsilon_i, \tag{IA4}$$

where b_i and σ_i are functions of the parameters of firm i's production technology.

There are two additional securities in the economy, both in zero net supply: a security that is risk-free with return R_f and the $(N+1)$-th risky security, which combines the risk-free security and a forward contract with cash settlements on the factor Y. Without loss of generality, the forward price of the contract is assumed to be such that the standard deviation of the equilibrium returns on the security is unity. As a result, its return is

$$R_{N+1} = \mathbb{E}(R_{N+1}) + Y. \tag{IA5}$$

There is a sufficiently large number of investors with a sufficiently dispersed distribution of wealth so that each investor acts as a price taker. Each investor is risk averse and has mean-variance preferences over the end-of-period wealth:

$$U^j = E(R^j W^j) - \frac{\gamma^j}{2W^j}\,\mathrm{var}(R^j W^j),$$

where W^j denotes the value of the initial endowment of investor j evaluated at equilibrium prices, R^j denotes the return per dollar on investor j's optimal portfolio, and $\gamma^j > 0$ is the risk-aversion of investor j.

[40] We have made the following changes to the notation used in Merton (1987) so that it is consistent with the notation in our paper. We denote an investor's risk aversion by γ instead of δ; we denote the total number of assets by N instead of n; we index individual assets by i instead of k; and we denote the unsystematic risk premium by a_i instead of λ_k.

Investors differ in their information sets. The common part of investors' information sets includes: (i) the return on the risk-free security, (ii) the structure of securities' return given in expression (IA4), and (iii) the expected return and variance of the forward-contract security given in (IA5). However, different investors have knowledge about the parameters b_i and σ_i for different *subsets* of securities. The investors who know about security i agree on its characteristics. To simplify the analysis, investors are assumed to have identical risk aversion $\gamma^j = \gamma$ and identical initial wealth $W^j = W$.

The optimal solution to each investor's portfolio problem allows us to obtain the aggregate demand for every security. Equating this to the aggregate supply for every security leads to the equilibrium expected return for asset i (Merton, 1987, eq. (16)):

$$\mathbb{E}(R_i) = R_f + \gamma b_i b + \gamma x_i \sigma_i^2 / q_i, \quad \text{for} \quad i = \{1, \ldots, N\}, \tag{IA6}$$

where x_i is the fraction of the market portfolio invested in asset i,

$$b = \sum_{i=1}^{N} x_i b_i,$$

and q_i is the fraction of investors who know about security i.

Denoting the return on the market as $R_m = \sum_{i=1}^{N} x_i R_i$, Merton (1987, eq. (24)) obtains the equilibrium expected excess return on the market:

$$\mathbb{E}(R_m) - R_f = \gamma \operatorname{var}(R_m) + a_m, \tag{IA7}$$

where $a_m = \sum_{i=1}^{N} x_i a_i$,

$$\begin{aligned} a_i &= (1 - q_i)\Delta_i, \\ \Delta_i &= \mathbb{E}(R_i) - R_f - b_i(\mathbb{E}(R_{N+1}) - R_f). \end{aligned}$$

Equations (IA4) and (IA7) then imply

$$R_i - R_f = \beta_i(\mathbb{E}(R_m) - R_f) + a_i - \beta_i a_m + b_i Y + \sigma_i \epsilon_i, \tag{IA8}$$

where β_i denotes the covariance of the return on security i with the return on the market portfolio, divided by the variance of the market return. Equation (IA8) contains Y on the right-hand side. We substitute out Y by using the definition of the market portfolio return along with equations (IA4) and (IA6), to obtain

$$R_i - R_f = a_i - \beta_i a_m + \beta_i(\mathbb{E}(R_m) - R_f) + \frac{b_i}{b}(R_m - \mathbb{E}(R_m)) + \sigma_i \epsilon_i.$$

The equilibrium process for asset returns, given by equations (2) and (25) in Merton (1987), is

$$R_i - R_f = \beta_i(\mathbb{E}(R_m) - R_f) + a_i - \beta_i a_m + b_i Y + \sigma_i \varepsilon_i. \tag{IA9}$$

We posit that the SDF M has the following form,

$$M = \xi + \chi Y + \sum_{i=1}^{N} \zeta_i \varepsilon_i,$$

where ξ, χ, and ζ_i, $i = \{1, \dots, N\}$, are determined using the $N+2$ equations for the Law of One Price:

$$\mathbb{E}[M] = \frac{1}{R_f}, \tag{IA10}$$

$$\mathbb{E}[M(R_{N+1} - R_f)] = 0 \tag{IA11}$$

$$\mathbb{E}[M(R_i - R_f)] = 0, \quad \text{for} \quad i = \{1, \dots, N\}, \tag{IA12}$$

where, from (3) and (11) in Merton (1987),

$$R_{N+1} = R_f + \gamma b + Y.$$

From expression (IA10), we get

$$\xi = \frac{1}{R_f}.$$

From expression (IA11), we get

$$\chi = -\frac{\gamma b}{R_f}.$$

From expression (IA12), for each $i = \{1, \dots, N\}$, we have

$$\xi \beta_i (\mathbb{E}(R_m) - R_f) + \xi(a_i - \beta_i a_m) + \chi \beta_i + \zeta \sigma_i = 0.$$

As a result,

$$\zeta = -\frac{1}{R_f} \frac{\beta_i (\mathbb{E}(R_m) - R_f) + a_i - \beta_i a_m - b_i \gamma b}{\sigma_i}.$$

Recalling that

$$R_m = \sum_{i=1}^{N} x_i R_i$$

and using (2) and (16) from Merton (1987), we have

$$\begin{aligned} R_m - R_f &= \sum_{i=1}^{N} x_i(\gamma b_i b + \gamma x_i \sigma_i^2 / q_i) + \sum_{i=1}^{N} x_i b_i Y + \sum_{i=1}^{N} x_i \sigma_i \varepsilon_i \\ &= \gamma b^2 + \gamma \sum_{i=1}^{N} x_i^2 \sigma_i^2 / q_i + bY + \sum_{i=1}^{N} x_i \sigma_i \varepsilon_i. \end{aligned}$$

From the last expression, we obtain

$$bY = (R_m - R_f) - \gamma b^2 - \gamma \sum_{i=1}^{N} x_i^2 \sigma_i^2 / q_i - \sum_{i=1}^{N} x_i \sigma_i \varepsilon_i.$$

As a result, the SDF is

$$\begin{aligned} M = &\frac{1}{R_f} - \frac{\gamma}{R_f} \left((R_m - R_f) - b^2\gamma - \gamma \sum_{i=1}^{N} x_i^2 \sigma_i^2 / q_i - \sum_{i=1}^{N} x_i \sigma_i \varepsilon_i \right) \\ &- \frac{1}{R_f} \sum_{t=1}^{N} \frac{\beta_i(\mathbb{E}(R_m) - R_f) + a_i - \beta_i a_m - b_i \gamma b}{\sigma_i} \varepsilon_i. \end{aligned}$$

Grouping together similar terms, we obtain

$$\begin{aligned} M = &\frac{1}{R_f} + \frac{\gamma^2 b^2}{R_f} + \frac{\gamma^2 \sum_{i=1}^{N} x_i^2 \sigma_i^2 / q_i}{R_f} - \frac{\gamma}{R_f}(R_m - R_f) \\ &- \frac{1}{R_f} \sum_{i=1}^{N} \left(\frac{\beta_i(\mathbb{E}(R_m) - R_f) + a_i - \beta_i a_m - b_i \gamma b - \gamma x_i \sigma_i^2}{\sigma_i} \varepsilon_i \right). \end{aligned}$$

Finally, we use expressions (22) and (24) in Merton (1987) to simplify the loading of M on ε_i and obtain

$$-\frac{1}{R_f} \sum_{i=1}^{N} \left(\frac{\beta_i(\mathbb{E}(R_m) - R_f) + a_i - \beta_i a_m - b_i \gamma b - \gamma x_i \sigma_i^2}{\sigma_i} \right) = -\frac{1}{R_f} \sum_{i=1}^{N} \frac{a_i}{\sigma_i}.$$

Using the demeaned return on the market portfolio as a factor in the SDF, along with expressions (15), (19), and (24) in Merton (1987), we obtain

$$\begin{aligned} M &= -\frac{1}{R_f} \sum_{i=1}^{N} \left(\frac{a_i}{\sigma_i} \varepsilon_i \right) + \frac{1}{R_f} - \frac{(\mathbb{E}(R_m) - R_f)}{R_f \operatorname{var}(R_m)} (R_m - \mathbb{E}(R_m)) \\ &= \underbrace{-\frac{a' V_e^{-1}}{R_f} e_{t+1}}_{M^a} + \underbrace{\frac{1}{R_f} - \frac{(\mathbb{E}(R_m) - R_f)}{R_f \operatorname{var}(R_m)} (R_m - \mathbb{E}(R_m))}_{M^\beta}, \end{aligned} \tag{IA13}$$

where $e_i = \sigma_i \epsilon_i$ and V_e is the covariance matrix of e with σ_i^2 on its diagonal.

To characterize the limiting behavior of this economy, as $N \to \infty$, assume that $x_i \to 0$, that is the fraction of market portfolio invested in each asset i is infinitesimally small. Then, as $N \to \infty$, we have

$$\beta_i = \frac{b_i b + x_i \sigma_i^2}{b^2 + \sum_{i=1}^{N} x_i^2 \sigma_i^2} \to \frac{b_i}{b^*}, \quad \text{where} \quad b \to b^*,$$

$$a_m = \sum_{i=1}^{N} x_i a_i = \sum_{i=1}^{N} x_i(1-q_i)\Delta_i = \sum_{i=1}^{N} \gamma x_i^2 \sigma_i^2 \frac{(1-q_i)}{q_i} \to 0,$$

$$\operatorname{cov}\left(\sum_{i=1}^{N} x_i\sigma_i\varepsilon_i, \varepsilon_i\right) = \sum_{i=1}^{N} x_i\sigma_i \to 0.$$

Thus, given $N \to \infty$, we have: (i) $\beta_i \to b_i/b^*$, (ii) $a_m \to 0$, and (iii) the market return is asymptotically orthogonal to all unsystematic shocks, e_i. Making these substitutions in equations (IA9) and (IA13) gives the results in (22) and (24). □

IA.5 Spanning of the SDF Components

Proposition IA.4.1 implies that, as $N \to \infty$, the log of the estimated SDF component $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1})$ converges to a linear function of the missing systematic factors. Proposition IA.5.2 below shows how to determine whether a vector of observable variables g_t represents missing sources of systematic risk in the candidate factor model, and if so, how to estimate the prices of risk associated with these missing risk factors.

Let f_t^{mis} be the vector of true systematic risk factors that are missing in the candidate factor model. Consider the regression of $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t})$ on an intercept and the vector g_t,

$$\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t}) = \gamma_0 + \gamma_1' g_t + u_t.$$

Denote by γ_1^{ols} the OLS-estimator of γ_1 and by R_g^2 the coefficient of determination in the corresponding regression.

Proposition IA.5.2 (Detecting Missing Systematic Factors)**.** *Under the assumptions of the extended Proposition IA.4.1 and if $g_t = Qf_t^{\text{mis}}$, for some nonsingular Q, as $N \to \infty$ we have*

$$\gamma_1^{\text{ols}} \xrightarrow{p} -(Q')^{-1} V_{f^{\text{mis}}}^{-1} \lambda^{\text{mis}} \quad \textit{and} \quad R_g^2 \xrightarrow{p} 1.$$

On the other hand, if g_t is orthogonal to f_t^{mis} then

$$\gamma_1^{\text{ols}} \xrightarrow{p} 0_{K^{\text{mis}}} \quad \textit{and} \quad R_g^2 \xrightarrow{p} 0.$$

Proof: Collect the values of the vector g_t for each t in a matrix $G = (g_1 \cdots g_T)'$. Likewise, collect the values of the vector f_t^{mis} for each t in a matrix $F^{\text{mis}} = (f_1^{\text{mis}} \cdots f_T^{\text{mis}})'$. For each t, collect the values of the systematic component $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t+1})$ of the SDF in a vector $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp}}) = (\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},1}) \cdots \log(\hat{M}^{\beta,\text{mis}}_{\text{exp},T}))'$. Then, the R^2 of the regression of $\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t})$ on an intercept and the vector g_t,

$$\log(\hat{M}^{\beta,\text{mis}}_{\text{exp},t}) = \gamma_0 + \gamma_1' g_t + u_t,$$

is

$$R_g^2 = \frac{\gamma_1^{\text{ols}\prime} G'(I_T - 1_T 1_T'/T) G \gamma_1^{\text{ols}}}{\log(\hat{M}_{\text{exp}}^{\beta,\text{mis}})'(I_T - 1_T 1_T'/T)\log(\hat{M}_{\text{exp}}^{\beta,\text{mis}})},$$

where $\gamma_1^{\text{ols}} = (G'(I_T - 1_T 1_T'/T)G)^{-1} G'(I_T - 1_T 1_T'/T)\log(\hat{M}_{\text{exp}}^{\beta,\text{mis}})$.

In Proposition IA.4.1, we have showen that

$$\log(\hat{M}_{\text{exp},t+1}^{\beta,\text{mis}}) \xrightarrow{p} -\lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) - \frac{1}{2}\lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}}.$$

For simplicity, we set $M_{1_T} = I_T - 1_T 1_T'/T$ and, given that $M_{1_T} 1_T = 0_T$, we obtain

$$\begin{aligned}
\gamma_1^{\text{ols}} &\xrightarrow{p} -(G' M_{1_T} G)^{-1} G' M_{1_T} (F^{\text{mis}} - 1_T \mathbb{E}(f_{t+1}^{\text{mis}\prime})) V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}} \\
&= -(Q F^{\text{mis}\prime} M_{1_T} F^{\text{mis}} Q')^{-1} Q F^{\text{mis}\prime} M_{1_T} F^{\text{mis}} V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}} \\
&= -(Q')^{-1} V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}}.
\end{aligned}$$

The limiting behavior of the numerator of R_g^2 is as follows

$$\begin{aligned}
\gamma_1^{\text{ols}\prime}(G' M_{1_T} G)\gamma_1^{\text{ols}} &\xrightarrow{p} \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} Q^{-1} Q (F^{\text{mis}\prime} M_{1_T} F^{\text{mis}}) Q' (Q')^{-1} V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}} \\
&= \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} (F^{\text{mis}\prime} M_{1_T} F^{\text{mis}}) V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}}.
\end{aligned}$$

The limiting behavior of the denominator of R_g^2 is as follows

$$\begin{aligned}
&\log(\hat{M}_{\text{exp}}^{\beta,\text{mis}})'(I_T - 1_T 1_T'/T)\log(\hat{M}_{\text{exp}}^{\beta,\text{mis}}) \\
&\quad \xrightarrow{p} \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} (F^{\text{mis}} - 1_T \mathbb{E}(f_{t+1}^{\text{mis}\prime}))' M_{1_T} (F^{\text{mis}} - 1_T \mathbb{E}(f_{t+1}^{\text{mis}\prime})) V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}} \\
&\quad = \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} (F^{\text{mis}\prime} M_{1_T} F^{\text{mis}}) V_{f^{\text{mis}}}^{-1}\lambda^{\text{mis}}.
\end{aligned}$$

Given that the limit of the numerator equals the limit of the denominator, $R_g^2 \xrightarrow{p} 1$.

The proof of the case of G being orthogonal to F^{mis}, that is, when $G'(I_T - 1_T 1_T'/T)F^{\text{mis}} = O_{K^{\text{mis}}}$, is straightforward, and therefore, omitted. □

Along the same lines, if weak factors span the unsystematic component of the SDF $M_{\text{exp},t+1}^a$, Proposition IA.5.3 below shows how to determine whether a vector of observable variables h_t is a linear combination of these weak factors and if so how to estimate their prices of risk. Assume that $\log(M_{\text{exp},t+1}^a) = -a' V_e^{-1} e_t - a' V_e^{-1} a/2 = \gamma_{\text{weak}}' f_t^{\text{weak}}$, where f_t^{weak} is a vector of true latent K^{weak} weak factors with the identity covariance matrix $V_{f^{\text{weak}}} = I_{K^{\text{weak}}}$.

Proposition IA.5.3 (Detecting Missing Weak Factors)**.** *Consider the regression of* $\log(\hat{M}_{\text{exp},t}^a)$ *on an intercept and the vector* h_t*,*

$$\log(\hat{M}_{\text{exp},t}^a) = \gamma_0 + \gamma_1' h_t + u_t.$$

Denote by γ_1^{ols} an OLS estimator of γ_1 and R_h^2 the coefficient of determination in the corresponding regression.

Under the assumptions of Proposition IA.4.1 and if $h_t = Qf_t^{\text{weak}}$ for some nonsingular Q, as $N \to \infty$ we have

$$\hat{\gamma}_1 \xrightarrow{p} -(Q')^{-1}\gamma_{\text{weak}} \quad \textit{and} \quad R_h^2 \xrightarrow{p} 1.$$

On the other hand, if h_t is orthogonal to f_t^{weak} then

$$\hat{\gamma}_1 \xrightarrow{p} 0_{K^{\text{weak}}} \quad \textit{and} \quad R_h^2 \xrightarrow{p} 0.$$

Proof: Collect the values of the vector h_t for each t in a matrix $H = (h_1 \cdots h_T)'$. Likewise, collect the values of the vector f_t^{weak} for each t in a matrix $F^{\text{weak}} = (f_1^{\text{weak}} \cdots f_T^{\text{weak}})'$. For each t, collect the values of $\log(\hat{M}^a_{\exp,t})$ in a vector $\log(\hat{M}^a_{\exp}) = (\log(\hat{M}^a_{\exp,1}) \cdots \log(\hat{M}^a_{\exp,T}))'$. Then, the R^2 of the regression of $\log(\hat{M}^a_{\exp,t})$ on an intercept and the vector h_t,

$$\log(\hat{M}^a_{\exp,t}) = \gamma_0 + \gamma_1' h_t + u_t,$$

is given by

$$R_h^2 = \frac{\gamma_1^{\text{ols}\prime} H'(I_T - 1_T 1_T'/T) H \gamma_1^{\text{ols}}}{\log(\hat{M}^a_{\exp})'(I_T - 1_T 1_T'/T)\log(\hat{M}^a_{\exp})},$$

where $\gamma_1^{\text{ols}} = (H'(I_T - 1_T 1_T'/T)H)^{-1} H'(I_T - 1_T 1_T'/T)\log(\hat{M}^a_{\exp})$.

In Proposition IA.4.1, we showed that, as $N \to \infty$,

$$\log(\hat{M}^a_{\exp,t+1}) - aV_e^{-1}e_{t+1} - \frac{1}{2}a'V_e^{-1}a \xrightarrow{p} 0.$$

For simplicity, we set $M_{1_T} = I_T - 1_T 1_T'/T$. Given that $M_{1_T} 1_T = 0_T$, we obtain

$$\begin{aligned}
\gamma_1^{\text{ols}} &\xrightarrow{p} -(H'M_{1_T}H)^{-1}H'M_{1_T}(F^{\text{weak}} - 1_T\mathbb{E}(f_{t+1}^{\text{weak}\prime}))\gamma_{\text{weak}} \\
&= -(QF^{\text{weak}\prime}M_{1_T}F^{\text{weak}}Q')^{-1}QF^{\text{weak}\prime}M_{1_T}F^{\text{weak}}\gamma_{\text{weak}} \\
&= -(Q')^{-1}\gamma_{\text{weak}}, \text{ when } N \to \infty.
\end{aligned}$$

The limiting behavior of the numerator of R_h^2 is as follows

$$\begin{aligned}
\gamma_1^{\text{ols}\prime}(H'M_{1_T}H)\gamma_1^{\text{ols}} &\xrightarrow{p} \gamma_{\text{weak}}'Q^{-1}Q(F^{\text{weak}\prime}M_{1_T}F^{\text{weak}})Q'(Q')^{-1}\gamma_{\text{weak}} \\
&= \gamma_{\text{weak}}'(F^{\text{weak}\prime}M_{1_T}F^{\text{weak}})\gamma_{\text{weak}}, \quad \text{when } N \to \infty.
\end{aligned}$$

The limiting behavior of the denominator of R_h^2 is as follows

$$\log(\hat{M}^a_{\exp})'(I_T - 1_T 1_T'/T)\log(\hat{M}^a_{\exp})$$

$$\xrightarrow{p} \gamma_{\text{weak}}'(F^{\text{weak}} - 1_T\mathbb{E}(f_{t+1}^{\text{weak}\prime}))' M_{1_T}(F^{\text{weak}} - 1_T\mathbb{E}(f_{t+1}^{\text{weak}\prime}))\gamma_{\text{weak}}$$
$$= \gamma_{\text{weak}}'(F^{\text{weak}\prime} M_{1_T} F^{\text{mis}})\gamma_{\text{weak}}, \quad \text{when } N \to \infty.$$

Given that the limit of the numerator equals the limit of the denominator, $R_h^2 \xrightarrow{p} 1$.

The proof when H is orthogonal to F^{weak}, that is, when $H'(I_T - 1_T 1_T'/T)F^{\text{weak}} = O_{K^{\text{weak}}}$, is straightforward, and therefore, omitted. □

A major strength of our approach is that we do not need to estimate the exposures of asset returns to an observable variable in order to define whether this variable represents a systematic or weak factor in the given cross-section of asset returns and quantify its price of risk. Thus, Propositions IA.5.2 and IA.5.3 complement the three-pass method of Giglio and Xiu (2021) and the supervised PCA of Giglio et al. (2021b) for describing the role of systematic and weak factors in pricing a cross-section of asset returns, and estimating the corresponding risk premia.[41]

IA.6 Weak Factors

Thanks to the assumption of the approximate factor structure of asset returns, our methodology accommodates weak factors in the unsystematic shocks e_{t+1}. The approximate factor structure implies that V_e can be non-diagonal, with the only constraint being that the maximum eigenvalue of V_e is uniformly bounded. Even though we have already mentioned that our theoretical results hold regardless of the presence of weak factors in e_{t+1}, we now prove Proposition IA.4.1 where we *explicitly* allow for weak factors in e_{t+1}. This proof strengthens the relevance of our methodology for identifying the importance of compensation for unsystematic risk, in particular, including that arising from weak factors.

Specifically, we assume now that

$$e_{t+1} = \beta^{\text{weak}} f_{t+1}^{\text{weak}} + e_{t+1}^{\text{as}}, \tag{IA14}$$

where e_{t+1}^{as} is a vector of asset-specific ("as") shocks with a diagonal covariance matrix $V_{e^{\text{as}}}$ that has a bounded maximum eigenvalue, f_{t+1}^{weak} is a vector of K^{weak} latent weak factors with covariance matrix V_f^{weak}, and β^{weak} is the matrix of assets' exposures to the weak factors f_{t+1}^{weak}. We define weak factors in accordance with the definition of Lettau and Pelger (2020): f_{t+1}^{weak} is a weak factor, if the following condition holds $\beta^{\text{weak}\prime}\beta^{\text{weak}} \to A > 0$, if $N \to \infty$, where A is some constant matrix. This condition can be written as $\beta^{\text{weak}\prime}\beta^{\text{weak}} = O(1)$,

[41] The large-N results of Propositions IA.5.2 and IA.5.3 abstract from estimation uncertainty, unlike Giglio and Xiu (2021) and Giglio et al. (2021b), who allow for sampling variability by developing their analysis under both large N and large T.

in contrast to the condition for systematic factors $\beta^{\text{can}\prime} V_e^{-1} \beta^{\text{can}}/N = O(1)$, as indicated in Assumption IA.2.1. In practice, the condition $\beta^{\text{weak}\prime}\beta^{\text{weak}} = O(1)$ holds when either the factors f_{t+1}^{weak} affect only a subset of the asset returns or the factors f_{t+1}^{weak} affect all asset returns but only marginally.

Without loss of generality, we assume that the weak factors f_{t+1}^{weak} are orthogonal to the asset-specific shocks e_{t+1}^{as} and $V_{f^{\text{weak}}} = I_{K^{\text{weak}}}$, thus $V_e = \beta^{\text{weak}}\beta^{\text{weak}\prime} + V_{e^{\text{as}}}$.

Denote the vector of prices of unit assets' exposures to weak factors by λ^{weak}, thus compensation for the unsystematic risk includes compensation for exposure to weak factors and compensation for asset-specific shocks, $a = a^{\text{as}} + \beta^{\text{weak}}\lambda^{\text{weak}}$, such that the no-arbitrage constraint holds, that is, $a^{\text{as}\prime} V_{e^{\text{as}}}^{-1} a^{\text{as}} < \delta_{\text{apt}}^{\text{as}} < \infty$, for some constant $\delta_{\text{apt}}^{\text{as}} > 0$.

For simplicity, assume that the candidate factor model includes all systematic risk factors, that is, $K^{\text{mis}} = 0$, implying that $V_\varepsilon = V_e$.

Assumption IA.6.1. *The following assumptions explicitly incorporate weak factors in the unsystematic shocks, imposing more structure on the covariance matrix V_e:*[42]

$$\begin{aligned}
&N^{-1}\beta^{\text{can}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{can}} \longrightarrow D > 0, \;\; \textit{as} \quad N \to \infty,\\
&\beta^{\text{weak}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{weak}} \longrightarrow E > 0, \;\; \textit{as} \quad N \to \infty,\\
&\beta^{\text{can}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{weak}} = o(N^{\frac{1}{2}}),\\
&\beta^{\text{can}\prime} V_{e^{\text{as}}}^{-1} a^{\text{as}} = o(N^{\frac{1}{2}}).
\end{aligned}$$

Lemma IA.6.1.

$$\beta^{\text{can}\prime} V_e{}^{-1}\beta^{\text{can}} = O(N).$$

Proof: The Sherman-Morrison-Woodbury formula applied to $V_e{}^{-1}$ leads to

$$\begin{aligned}
\beta^{\text{can}\prime} V_e{}^{-1}\beta^{\text{can}} &= \beta^{\text{can}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{can}} - \beta^{\text{can}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{weak}} (V_{f^{\text{weak}}}^{-1} + \beta^{\text{weak}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{weak}})^{-1}\beta^{\text{weak}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{can}}\\
&= O(N) + o(N^{\frac{1}{2}}) \times [O(1) + O(1)]^{-1} \times o(N^{\frac{1}{2}})\\
&= O(N) + o(N) = O(N). \qquad \square
\end{aligned}$$

Lemma IA.6.2.

$$\beta^{\text{weak}\prime} V_e{}^{-1}\beta^{\text{can}} = o(N^{\frac{1}{2}}).$$

[42] Because the matrix $V_{e^{\text{as}}}^{-1}$ has uniformly bounded eigenvalues, the definition of a weak factor of Lettau and Pelger (2020) can be equivalently written as $\beta^{\text{weak}\prime} V_{e^{\text{as}}}^{-1}\beta^{\text{weak}} \longrightarrow E > 0$, as $N \to \infty$.

Proof: The Sherman-Morrison-Woodbury formula applied to ${V_e}^{-1}$, leads to

$$\begin{aligned}\beta^{\text{weak}\,\prime}{V_e}^{-1}\beta^{\text{can}} &= \beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{can}} - \beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}}(V_{f^{\text{weak}}}^{-1} + \beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}})^{-1}\beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{can}} \\ &= V_{f^{\text{weak}}}^{-1}(V_{f^{\text{weak}}}^{-1} + \beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}})^{-1}\beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{can}} \\ &= O(1) \times [O(1) + O(1)]^{-1} \times o(N^{\frac{1}{2}}) = o(N^{\frac{1}{2}}).\end{aligned}$$

□

Lemma IA.6.3.

$$\beta^{\text{can}\,\prime}{V_e}^{-1}a = o(N^{\frac{1}{2}}).$$

Proof: The Sherman-Morrison-Woodbury formula applied to ${V_e}^{-1}$, given $a = a^{\text{as}} + \beta^{\text{weak}}\lambda^{\text{weak}}$, leads to

$$\begin{aligned}\beta^{\text{can}\,\prime}{V_e}^{-1}a^{\text{as}} &= \beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}a^{\text{as}} - \beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}}(V_{f^{\text{mis}}}^{-1} + \beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}})^{-1}\beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}a^{\text{as}} \\ &= o(N^{\frac{1}{2}}) + o(N^{\frac{1}{2}}) \times [O(1) + O(1)]^{-1} \times O(1) \\ &= o(N^{\frac{1}{2}}).\end{aligned}$$

Thus, $\beta^{\text{can}\,\prime}{V_e}^{-1}\beta^{\text{weak}}\lambda^{\text{weak}} = o(N^{\frac{1}{2}})$ by Lemma IA.6.2. □

Lemma IA.6.4. *Let e^{as} be an $N \times 1$ random vector with zero mean and covariance matrix $V_{e^{\text{as}}}$.*

$$\beta^{\text{can}\,\prime}{V_e}^{-1}e^{\text{as}} = O_p(N^{\frac{1}{2}}).$$

Proof: For any random variable X with a finite second moment, we have that $X = O_p((\mathbb{E}(X^2))^{\frac{1}{2}})$. If $X = \beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}e^{\text{as}}$, then

$$\mathbb{E}(\beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}e^{\text{as}}e^{\text{as}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{can}}) = \beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{can}} = O(N),$$

and therefore, $\beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}e^{\text{as}} = O_p(N^{\frac{1}{2}})$. Similarly, we can show that $\beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}e^{\text{as}} = O_p(1)$. Finally, we apply the Sherman-Morrison-Woodbury formula to ${V_e}^{-1}$ and obtain

$$\begin{aligned}\beta^{\text{can}\,\prime}{V_e}^{-1}e^{\text{as}} &= \beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}e^{\text{as}} - \beta^{\text{can}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}}(V_{f^{\text{weak}}}^{-1} + \beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}\beta^{\text{weak}})^{-1}\beta^{\text{weak}\,\prime}V_{e^{\text{as}}}^{-1}e^{\text{as}} \\ &= O_p(N^{\frac{1}{2}}) + o(N^{\frac{1}{2}}) \times [O(1) + O(1)]^{-1} \times O_p(1) = O_p(N^{\frac{1}{2}}).\end{aligned}$$

□

We now generalize Proposition IA.4.1 for the case in which the unsystematic shocks explicitly include weak factors, that is, when (IA14) holds. To do this, the only result we need to prove is that the presence of weak factors does not have implications for the limiting behavior of $\hat{M}^a_{\exp,t+1}$ and $M^a_{\exp,t+1}$.

Proposition IA.6.4 (Properties of $\hat{M}^a_{\exp,t+1}$, when Shocks e_{t+1} Include Weak Factors)**.** *Under Assumptions 2.1 and 2.2, the assumption that returns R_{t+1} are Gaussian, Assumptions IA.6.1, and the assumption that $K^{\text{mis}} = 0$, we have*

$$\hat{M}^a_{\exp,t+1} - M^a_{\exp,t+1} \xrightarrow{p} 0 \quad \textit{as} \quad N \to \infty.$$

Proof: Recall that

$$\hat{M}^a_{\exp,t+1} = \exp\Big[- aV_R^{-1}(R_{t+1} - \mathbb{E}(R_{t+1})) - \frac{1}{2}a'V_R^{-1}a\Big] \text{ and}$$
$$M^a_{\exp,t+1} = \exp\Big[- a'V_e^{-1}e_{t+1} - \frac{1}{2}a'V_e^{-1}a\Big].$$

The exponent of $\hat{M}^a_{\exp,t+1}$, given that $K^{\text{mis}} = 0$, is

$$\begin{aligned}
&- a'V_R^{-1}(R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2}a'V_R^{-1}a = \\
&- a'V_R^{-1}\beta^{\text{can}}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) - a'V_R^{-1}e_{t+1} - \frac{1}{2}a'V_R^{-1}a.
\end{aligned}$$

We analyze the three terms of the exponent of $\hat{M}^a_{\exp,t+1}$ one-by-one. We use Lemmas IA.6.1 and IA.6.3 and apply the Sherman-Morrison-Woodbury formula to V_R^{-1}:

$$\begin{aligned}
a'V_R^{-1}\beta^{\text{can}} &= a'V_e^{-1}\beta^{\text{can}} - a'V_e^{-1}\beta^{\text{can}}(V^{-1}_{f^{\text{can}}} + \beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}})^{-1}\beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}} \\
&= a'V_e^{-1}\beta^{\text{can}}(V^{-1}_{f^{\text{can}}} + \beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}})^{-1}V^{-1}_{f^{\text{can}}} \\
&= o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times O(1) \\
&= o(N^{-\frac{1}{2}}).
\end{aligned}$$

Next, by Lemmas IA.6.1, IA.6.3, and IA.6.4, and by taking into account that $e_{t+1} = \beta^{\text{weak}}f^{\text{weak}}_{t+1} + e^{\text{as}}_t$ and $a = a^{\text{as}} + \beta^{\text{weak}}\lambda^{\text{weak}}$, we obtain

$$\begin{aligned}
a'V_R^{-1}e_{t+1} &= a'V_e^{-1}e_{t+1} - a'V_e^{-1}\beta^{\text{can}}(V^{-1}_{f^{\text{can}}} + \beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}})^{-1}\beta^{\text{can}\prime}V_e^{-1}e_{t+1} \\
&= a'V_e^{-1}e_{t+1} + o_p(N^{\frac{1}{2}})[O(1) + O(N)]^{-1}O_p(N^{\frac{1}{2}}) = a'V_e^{-1}e_{t+1} + o_p(1).
\end{aligned}$$

Finally, by Lemmas IA.6.1 and IA.6.3,

$$\begin{aligned}
a'V_R^{-1}a &= a'V_e^{-1}a - a'V_e^{-1}\beta^{\text{can}}(V^{-1}_{f^{\text{can}}} + \beta^{\text{can}\prime}V_e^{-1}\beta^{\text{can}})^{-1}\beta^{\text{can}\prime}V_e^{-1}a \\
&= a'V_e^{-1}a + o(N^{\frac{1}{2}}) \times [O(1) + O(N)]^{-1} \times o(N^{\frac{1}{2}}) \\
&= a'V_e^{-1}a + o(1).
\end{aligned}$$

Putting these results together, we obtain

$$-a'V_R^{-1}(R_{t+1} - \mathbb{E}[R_{t+1}]) - \frac{1}{2}a'V_R^{-1}a = -a'V_e^{-1}e_{t+1} - \frac{1}{2}a'V_e^{-1}a + o_p(1),$$

implying that

$$\hat{M}^a_{\exp,t+1} - M^a_{\exp,t+1} \xrightarrow{p} 0 \quad \text{as} \quad N \to \infty. \qquad \square$$

Remark: Proposition IA.6.4 shows that our methodology is still valid if expected excess returns include compensation for exposure to weak factors that are present in unsystematic shocks e_{t+1}. This is an important result because of the challenges associated with identifying weak factors. For example, it is well known that weak factors cannot be estimated consistently (Lettau and Pelger, 2020). Our methodology does not require estimating weak factors but can still accurately characterize the importance of unsystematic risk that includes weak factors, which makes our approach compelling.

IA.7 Estimation

We start this section by discussing the identification conditions that fix the rotation of risk factors. These identification conditions do not have any implications for the SDF but allow us to estimate the model of asset returns. Next, we show how to estimate the model of asset returns. In the empirical analysis, we use an observable time-varying risk-free rate R_{ft} in place of R_f.

IA.7.1 Identification conditions

In a candidate model, the loadings of asset returns on the missing factors, and the missing factors themselves, are unique up to a rotation. Similarly, identifying the loadings of asset returns on the latent factors in the APT model is unique up to a rotation. Thus, at the estimation stage, we need to impose identification conditions. These identification conditions affect the interpretation of latent factors but not the estimated SDF.

Below, we detail the identification strategy, which we use to correct a candidate model of asset returns. The difference between identifying missing factors in the candidate factor model and identifying latent factors of the APT model is only because of the presence of observable factors in the candidate model. Thus, the identification strategy for the APT model is equivalent to that described below but in which $K^{\text{can}} = 0$ and $K^{\text{mis}} = K$.

We follow the identification strategy of Bai and Li (2012) and adapt it to the case in which a model has K^{can} observable and K^{mis} latent risk factors. Denote F^{can} a matrix $T \times K^{\text{can}}$ that collects candidate factors column by column. Denote by F^{mis} a matrix $T \times K^{\text{mis}}$ that collects missing factors column by column. Combine these matrices in a

$T \times (K^{\text{can}} + K^{\text{mis}})$ matrix $F = [F^{\text{can}},\ F^{\text{mis}}]$. Note that the rotation of this matrix is defined by a squared invertible matrix of a dimension $(K^{\text{can}} + K^{\text{mis}}) \times (K^{\text{can}} + K^{\text{mis}})$, and therefore, the rotation is pinned down by $(K^{\text{can}} + K^{\text{mis}})^2$ parameters.

At the estimation stage, we impose the following $(K^{\text{can}} + K^{\text{mis}})^2$ identification conditions to fix the rotation:

- The first K^{can} columns of the rotation matrix are fixed because F^{can} includes only observable factors. This is equivalent to $K^{\text{can}} \times (K^{\text{can}} + K^{\text{mis}})$ restrictions being already imposed.
- $V_{f^{\text{mis}}} = I_{K^{\text{mis}}}$ introduces $K^{\text{mis}} \times (K^{\text{mis}} + 1)/2$ restrictions.
- $\beta^{\text{mis}\prime} V_e^{-1} \beta^{\text{mis}}$ is a diagonal matrix that is equivalent to imposing $(K^{\text{mis}} - 1) \times K^{\text{mis}}/2$ restrictions. We also introduce an order restriction that requires that the diagonal elements of the matrix $\beta^{\text{mis}\prime} V_e^{-1} \beta^{\text{mis}}$ follow in descending order. In addition, we require the eigenvectors of $\beta^{\text{mis}\prime} V_e^{-1} \beta^{\text{mis}}$ to have positive means to identify the latent factors uniquely, rather than up to a sign.
- Candidate factors f_{t+1}^{can} are uncorrelated with missing factors f_{t+1}^{mis}. This requirement is equivalent to imposing $K^{\text{can}} \times K^{\text{mis}}$ additional restrictions.

IA.7.2 Constrained Maximum-Likelihood (ML) Estimator

This section describes how to estimate the candidate factor model and its required correction to obtain the APT-implied SDF. The underlying problem is described in Section 2.3.

To simplify exposition, introduce the following notation

$$\bar{R} = T^{-1} \sum_{t=1}^{T} R_t, \qquad \bar{R}_f = T^{-1} \sum_{t=1}^{T} R_{ft}, \qquad \bar{f}^{\text{can}} = T^{-1} \sum_{t=1}^{T} f_t^{\text{can}},$$

$$\hat{Q}_{f^{\text{can}}} = T^{-1} \sum_{t=1}^{T} f_t^{\text{can}} f_t^{\text{can}\prime}, \quad \text{and} \quad \hat{Q}_{Rf^{\text{can}}} = \frac{1}{T} \sum_{t=1}^{T} (R_t - R_{ft-1} 1_N) f_t^{\text{can}\prime}.$$

For this section only, we use the notation $\hat{\cdot}$ to denote an estimator.

Without loss of generality, assume that the candidate factors f_{t+1}^{can} are tradable factors in the form of excess returns on investment strategies (if any candidate factor is not tradable, we use its factor-mimicking portfolio, as in Breeden et al. (1989)).

For a generic vector Θ that collects all the unknown parameters of the corrected model of asset returns, $\Theta = (a',\ vech(V_e)',\ vech(V_{f^{\text{can}}})',\ vec(\beta^{\text{can}})',\ \lambda^{\text{can}\prime},\ vec(\beta^{\text{mis}})',\ \lambda^{\text{mis}\prime}),$

denote $L(\Theta)$ the (up to a constant) Gaussian joint likelihood of the vector of asset returns in excess of the risk-free rate, $R_{t+1} - R_{ft}1_N$, and observable factors f_{t+1}^{can} scaled by the number of time-series observations T

$$\log(L(\Theta)) = -\frac{1}{2}\log(|V_\varepsilon|) - \frac{1}{2}\log(|V_f^{\text{can}}|) - \frac{1}{2T}\sum_{t=0}^{T-1}\varepsilon_{t+1}' V_\varepsilon^{-1}\varepsilon_{t+1}$$
$$- \frac{1}{2T}\sum_{t=0}^{T-1}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}}))' V_{f^{\text{can}}}^{-1}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})), \tag{IA15}$$

where $\varepsilon_{t+1} = R_{t+1} - R_{ft}1_N - a - \beta^{\text{mis}}\lambda^{\text{mis}} - \beta^{\text{can}}\lambda^{\text{can}} - \beta^{\text{can}}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}}))$ and $V_\varepsilon = \beta^{\text{mis}} V_{f^{\text{mis}}} \beta^{\text{mis}\prime} + V_e$.

We maximize this log-likelihood function (IA15) subject to the no-arbitrage restriction (4). Without loss of generality, we replace the no-arbitrage restriction (4) with the expression

$$a' V_\varepsilon^{-1} a \leq \delta_{\text{apt}}$$

that is, replacing V_e with V_ε, is more convenient when deriving the first-order conditions.

We use the Karush-Kuhn-Tucker (KKT) multiplier method to solve the resulting constrained optimization problem,

$$\hat{\Theta} = \text{argmax}\ \log(L(\Theta)) \quad \text{subject to} \quad a' V_\varepsilon^{-1} a \leq \delta_{\text{apt}}, \tag{IA16}$$

and denote the KKT multiplier by $\kappa/2$.

The optimization problem for estimating the parameters of the APT-implied SDF is identical to that formulated in expression (IA16), in which there are no candidate factors, $K^{\text{can}} = 0$, and the missing factors f_{t+1}^{mis} are replaced with latent factors f_{t+1}. Correspondingly, the parameters characterizing the missing factors f_{t+1}^{mis}, such as β^{mis}, λ^{mis}, and K^{mis}, are replaced with the parameters characterizing the latent factors f_{t+1}, which are β, λ, and K, respectively.

Proposition IA.7.5 (Constrained ML Estimator)**.** *Suppose that the assumptions of Proposition IA.4.1 hold. Assume that the number K^{mis} of missing factors in the candidate model and the no-arbitrage bound δ_{apt} are known and that the sample covariance matrix $\hat{V}_{f^{\text{can}}}$ of candidate factors is nonsingular. Then the estimators of λ^{can} and $V_{f^{\text{can}}}$ coincide with the sample mean and sample covariance of the candidate factors f_t^{can}:*

$$\hat{\lambda}^{\text{can}} = \bar{f}^{\text{can}},$$
$$\hat{V}_f^{\text{can}} = \hat{Q}_{f^{\text{can}}} - \bar{f}^{\text{can}}\bar{f}^{\text{can}\prime}.$$

The estimators $\hat{\beta}^{\text{mis}}$ and $\hat{V}_e$ of β^{mis} and V_e do not admit a closed-form solution.

(i) If the optimal value of the Karush-Kuhn-Tucker multiplier $\hat{\kappa}$ is greater than zero, the estimators of β^{can}, λ^{mis}, and a, are

$$\begin{aligned}
\text{vec}(\hat{\beta}^{\text{can}}) &= \left(\hat{Q}_{f^{\text{can}}} \otimes I_N - \bar{f}^{\text{can}}\bar{f}^{\text{can}\prime} \otimes \hat{G}\right)^{-1} \\
&\qquad \times \text{vec}\left(\hat{Q}_{Rf^{\text{can}}} - \hat{G}(\bar{R} - \bar{R}_f 1_N)\bar{f}^{\text{can}\prime}\right) \qquad \text{(IA17)} \\
\hat{\lambda}^{\text{mis}} &= (\hat{\beta}^{\text{mis}\prime}\hat{V}_\varepsilon^{-1}\hat{\beta}^{\text{mis}})^{-1}\hat{\beta}^{\text{mis}\prime}\hat{V}_\varepsilon^{-1}\left(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}}\hat{\lambda}^{\text{can}}\right), \quad \textit{and} \\
\hat{a} &= \frac{1}{\hat{\kappa}+1}\left(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}}\hat{\lambda}^{\text{can}} - \hat{\beta}^{\text{mis}}\hat{\lambda}^{\text{mis}}\right),
\end{aligned}$$

where

$$\hat{\kappa} = \left(\frac{(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}}\hat{\lambda}^{\text{can}} - \hat{\beta}^{\text{mis}}\hat{\lambda}^{\text{mis}})' V_\varepsilon^{-1}(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}}\hat{\lambda}^{\text{can}} - \hat{\beta}^{\text{mis}}\hat{\lambda}^{\text{mis}})}{\delta_{\text{apt}}} - 1\right)^{1/2},$$

$$\begin{aligned}
\hat{G} &= \frac{1}{(\hat{\kappa}+1)} I_N + \frac{\hat{\kappa}}{(\hat{\kappa}+1)}\hat{\beta}^{\text{mis}}(\hat{\beta}^{\text{mis}\prime}\,\hat{V}_\varepsilon^{-1}\,\hat{\beta}^{\text{mis}})^{-1}\hat{\beta}^{\text{mis}\prime}\,\hat{V}_\varepsilon^{-1}, \quad \textit{and} \\
\hat{V}_\varepsilon &= \hat{\beta}^{\text{mis}}\hat{\beta}^{\text{mis}\prime} + \hat{V}_e.
\end{aligned}$$

(ii) If the optimal value of the Karush-Kuhn-Tucker multiplier satisfies $\hat{\kappa} = 0$, it is possible to estimate the vector α

$$\hat{\alpha} = \bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}}\hat{\lambda}^{\text{can}}$$

but not its components, a and $\beta^{\text{mis}}\lambda^{\text{mis}}$. The estimator of $\text{vec}(\beta^{\text{can}})$ *is given by expression* (IA17) *with $\hat{\kappa} = 0$.*

Proof: The Lagrangian for our optimization problem is

$$\begin{aligned}
L_p(\Theta) = &-\frac{\kappa}{2}(a'V_\varepsilon^{-1}a - \delta_{\text{apt}}) - \frac{1}{2}\log(|V_\varepsilon|) - \frac{1}{2}\log(|V_{f^{\text{can}}}|) - \frac{1}{2T}\sum_{t=0}^{T-1}\varepsilon_{t+1}'V_\varepsilon^{-1}\varepsilon_{t+1} \\
&- \frac{1}{2T}\sum_{t=0}^{T-1}(f_{t+1}^{\text{can}} - \mathbb{E}(f_t^{\text{can}}))'V_{f^{\text{can}}}^{-1}(f_{t+1}^{\text{can}} - \mathbb{E}(f_t^{\text{can}})). \qquad \text{(IA18)}
\end{aligned}$$

Recall that the candidate factors f_t^{can} represent excess returns on tradable investment strategies, that is, $\mathbb{E}(f_t^{\text{can}}) = \lambda^{\text{can}}$. The first-order condition for λ^{can} results in

$$\hat{\lambda}^{\text{can}} = \frac{1}{T}\sum_{t=1}^{T} f_t^{\text{can}}.$$

Similarly, the first-order condition for V_f^{can} gives

$$\hat{V}_{f^{\text{can}}} = \frac{1}{T}\sum_{t=1}^{T}(f_t^{\text{can}} - \hat{\lambda}^{\text{can}})(f_t^{\text{can}} - \hat{\lambda}^{\text{can}})'.$$

Next, we consider two cases, $\hat{\kappa} > 0$ and $\hat{\kappa} = 0$.

First, suppose that $\hat{\kappa} > 0$, and therefore $a' V_\varepsilon^{-1} a = \delta_{\text{apt}}$. We differentiate the Lagrangian in equation (IA18) with respect to λ^{mis} and a and obtain the following $K^{\text{mis}} + N$ first-order conditions:

$$\begin{pmatrix} \hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1} \\ I_N \end{pmatrix} \left(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}} \hat{\lambda}^{\text{can}} \right) = \begin{pmatrix} \hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1} \hat{\beta}^{\text{mis}} & \hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1} \\ \hat{\beta}^{\text{mis}\prime} & (1+\hat{\kappa}) I_N \end{pmatrix} \begin{pmatrix} \hat{\lambda}^{\text{mis}} \\ \hat{a} \end{pmatrix}.$$

The matrix premultiplying the vector $(\hat{\lambda}^{\text{mis}\prime}, \hat{a}')'$ is nonsingular when the no-arbitrage restriction binds, implying that $\hat{\lambda}^{\text{mis}}$ and $\hat{a}$ are identified separately :

$$\hat{\lambda}^{\text{mis}} = (\hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1} \left(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}} \hat{\lambda}^{\text{can}} \right), \tag{IA19}$$

$$\hat{a} = \frac{1}{\hat{\kappa} + 1} \left(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}} \hat{\lambda}^{\text{can}} - \hat{\beta}^{\text{mis}} \hat{\lambda}^{\text{mis}} \right). \tag{IA20}$$

Next, we use equation (IA20) and the binding no-arbitrage restriction $a' V_\varepsilon^{-1} a = \delta_{\text{apt}}$ to obtain

$$\hat{\kappa} = \left(\frac{(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}} \hat{\lambda}^{\text{can}} - \hat{\beta}^{\text{mis}} \hat{\lambda}^{\text{mis}})' \hat{V}_\varepsilon^{-1} (\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}} \hat{\lambda}^{\text{can}} - \hat{\beta}^{\text{mis}} \hat{\lambda}^{\text{mis}})}{\delta_{\text{apt}}} - 1 \right)^{1/2}. \tag{IA21}$$

Finally, we consider the first-order condition with respect to the generic (i, j)th element of β^{can}, denoted by β_{ij}^{can} with $1 \leq i \leq N$, $1 \leq j \leq K^{\text{can}}$, and obtain

$$-\frac{1}{T} \sum_{t=1}^{T} \left(R_t - R_{ft} 1_N - \hat{\beta}^{\text{mis}} \hat{\lambda}^{\text{mis}} - \hat{a} - \hat{\beta}^{\text{can}} f_t^{\text{can}} \right)' \hat{V}_\varepsilon^{-1} \left(- \frac{\partial \beta^{\text{can}}}{\partial \beta_{ij}^{\text{can}}} |_{\beta^{\text{can}} = \hat{\beta}^{\text{can}}} f_t^{\text{can}} \right) = 0,$$

which can be rearranged by stacking together the first-order conditions as

$$\hat{Q}_{R f^{\text{can}}} - (\hat{a} + \hat{\beta}^{\text{mis}} \hat{\lambda}^{\text{mis}}) \bar{f}^{\text{can}\prime} - \hat{\beta}^{\text{can}} \hat{Q}_{f^{\text{can}}} = 0_{N \times K^{\text{can}}}. \tag{IA22}$$

Next, we define

$$\hat{G} = \frac{1}{(\hat{\kappa}+1)} I_N + \frac{\hat{\kappa}}{(\hat{\kappa}+1)} \hat{\beta}^{\text{mis}} (\hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} \hat{V}_\varepsilon^{-1},$$

and use the formulas (IA19) and (IA20) to rewrite equation (IA22) as follows

$$\hat{\beta}^{\text{can}} \hat{Q}_{f^{\text{can}}} - \hat{G} \hat{\beta}^{\text{can}} \bar{f}^{\text{can}} \bar{f}^{\text{can}\prime} = \hat{Q}_{R f^{\text{can}}} - G(\bar{R} - \bar{R}_f 1_N) \bar{f}^{\text{can}\prime}.$$

Then, we take the vec operator and solve for $\hat{\beta}^{\text{can}}$ to obtain

$$\text{vec}(\hat{\beta}^{\text{can}}) = \left(\hat{Q}_{f^{\text{can}}} \otimes I_N - \bar{f}^{\text{can}} \bar{f}^{\text{can}\prime} \otimes \hat{G} \right)^{-1} \times \text{vec} \left(\hat{Q}_{R f^{\text{can}}} - G(\bar{R} - \bar{R}_f 1_N) \bar{f}^{\text{can}\prime} \right) \tag{IA23}$$

The solution for $\hat{\beta}^{\text{can}}$ exists because the matrix $\left(\hat{Q}_{f^{\text{can}}} \otimes I_N - \bar{f}^{\text{can}}\bar{f}^{\text{can}\prime} \otimes \hat{G}\right)$ is nonsingular. We note that

$$\hat{Q}_{f^{\text{can}}} \otimes I_N - \bar{f}^{\text{can}}\,\bar{f}^{\text{can}\prime} \otimes \hat{G} = \hat{V}_{f^{\text{can}}} \otimes I_N + \bar{f}^{\text{can}}\,\bar{f}^{\text{can}\prime} \otimes (I_N - \hat{G}),$$

where $\hat{V}_{f^{\text{can}}}$, being a covariance matrix, is positive-definite, and $\bar{f}^{\text{can}}\,\bar{f}^{\text{can}\prime} \otimes (I_N - \hat{G})$ is positive semi-definite, because

$$\begin{aligned}
I_N - \hat{G} &= I_N - \frac{1}{(\hat{\kappa}+1)} I_N - \left(\frac{\hat{\kappa}}{1+\hat{\kappa}}\right) \hat{\beta}^{\text{mis}} (\hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \\
&= \left(\frac{\hat{\kappa}}{1+\hat{\kappa}}\right) (I_N - \hat{\beta}^{\text{mis}} (\hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1}) \\
&= \left(\frac{\hat{\kappa}}{1+\hat{\kappa}}\right) \hat{V}_{\varepsilon} (\hat{V}_{\varepsilon}^{-1} - \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}} (\hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1}) \\
&= \left(\frac{\hat{\kappa}}{1+\hat{\kappa}}\right) \hat{V}_{\varepsilon} (\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}} (I_N - (\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}} \hat{\beta}^{\text{mis}} (\hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} (\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}}) (\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}}
\end{aligned}$$

is the product of the positive-definite matrices $I_N - (\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}} \hat{\beta}^{\text{mis}} (\hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}})^{-1} \hat{\beta}^{\text{mis}\prime} (\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}}$ (projection matrix), $\hat{V}_{\varepsilon}$, and $(\hat{V}_{\varepsilon}^{-1})^{\frac{1}{2}}$. Note that $\hat{\lambda}^{\text{mis}}$, $\hat{a}$, $\hat{\kappa}$ are functions of $\hat{\beta}^{\text{mis}}$, $\hat{V}_e$, and $\hat{\beta}^{\text{can}}$:

$$\hat{\lambda}^{\text{mis}} = \hat{\lambda}^{\text{mis}}(\hat{\beta}^{\text{mis}}, \hat{\beta}^{\text{can}}, \hat{V}_e,), \quad \hat{a} = \hat{a}(\hat{\beta}^{\text{mis}}, \hat{\beta}^{\text{can}}, \hat{V}_e), \quad \hat{\kappa} = \hat{\kappa}(\hat{\beta}^{\text{mis}}, \hat{\beta}^{\text{can}}, \hat{V}_e). \quad \text{(IA24)}$$

However, we cannot obtain the explicit representation of $\hat{\lambda}^{\text{mis}}$, $\hat{a}$, and $\hat{\kappa}$ in terms of fewer parameters, for example, only $\hat{\beta}^{\text{mis}}$ and $\hat{V}_e$. This is because substituting $\hat{\kappa}$ into expression (IA23) for $\hat{\beta}^{\text{can}}$ creates a fixed-point problem for $\hat{\beta}^{\text{can}}$.

Because a fixed-point problem slows down substantially the optimization routine, we do not use the closed-form solution (IA23) for $\hat{\beta}^{\text{can}}$ and instead substitute expressions (IA24) into $L_p(\Theta)$ to obtain the concentrated log-likelihood function, which is a function of only β^{mis}, β^{can}, and V_e. We maximize the concentrated log-likelihood numerically, thereby obtaining the estimates of β^{mis}, β^{can}, and V_e, which also imply the optimal values of the other parameters. Finally, we verify if equation (IA23) holds, thereby checking convergence of our optimization algorithm.

If equation (IA21) implies that $\hat{\kappa} < 0$ then we ignore all the obtained above results and move to the next case of $\hat{\kappa} = 0$.

Consider the second case, in which the Karush-Kuhn-Tucker multiplier is zero: $\hat{\kappa} = 0$. In this case, a feasible solution to the optimization problem satisfies $a' V_{\varepsilon}^{-1} a < \delta_{\text{apt}}$.

The first-order conditions with respect to λ^{mis} and a imply the following singular system of $K^{\text{mis}} + N$ equations

$$\begin{pmatrix} \hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \\ I_N \end{pmatrix} \left(\bar{R} - \bar{R}_f 1_N - \hat{\beta}^{\text{can}} \hat{\lambda}^{\text{can}}\right) = \begin{pmatrix} \hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \hat{\beta}^{\text{mis}} & \hat{\beta}^{\text{mis}\prime} \hat{V}_{\varepsilon}^{-1} \\ \hat{\beta}^{\text{mis}} & I_N \end{pmatrix} \begin{pmatrix} \hat{\lambda}^{\text{mis}} \\ \hat{a} \end{pmatrix}.$$

The matrix

$$\begin{pmatrix} \hat{\beta}^{\text{mis}\prime}\hat{V}_{\varepsilon}^{-1}\hat{\beta}^{\text{mis}} & \hat{\beta}^{\text{mis}\prime}\hat{V}_{\varepsilon}^{-1} \\ \hat{\beta}^{\text{mis}} & I_N \end{pmatrix}$$

is of dimension $(N+K^{\text{mis}})\times(N+K^{\text{mis}})$ but of rank N, and therefore it is noninvertible. As a result, we cannot identify separately a and λ^{mis}, implying that if $\hat{\kappa}=0$, we can only identify the sum $a+\beta^{\text{mis}}\lambda^{\text{mis}}$ but not its two components separately. All the other parameters of the vector Θ are identified separately, and their expressions follow from differentiating the Lagrangian (IA18) and solving the resulting first-order conditions. For instance, the formula for $\hat{\beta}^{\text{can}}$ follows by setting $\hat{G}=I_N$ into (IA17).

When both cases, $\hat{\kappa}>0$ and $\hat{\kappa}=0$, are feasible, we choose the one under which the log-likelihood $L_p(\Theta)$ is larger. □

IA.8 The SDF with Nonorthogonal Components

In the main text of the manuscript, we assumed that the candidate risk factors f_{t+1}^{can} are orthogonal to the missing sources of systematic risk f_{t+1}^{mis} and unsystematic shocks e_{t+1}. This assumption is without loss of generality because if the (observable) systematic risk factors f_{t+1}^{mis} that the candidate model omits are correlated with f_{t+1}^{can}, there exists an *observationally equivalent* representation of the SDF M_{t+1}, such that the factors f_{t+1}^{can} are orthogonal to some latent systematic risk factors (residuals from an orthogonal projection of omitted observable risk factors onto the candidate factors). Thus, the assumption of orthogonality affects the interpretation of the missing factors but not the admissibility of the pricing kernel.

In particular,

$$\begin{aligned} M_{t+1} &= \frac{1}{R_f} + b^{\text{can}\prime}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) + b^{\text{mis}\prime}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) + c' e_{t+1} \\ &= \frac{1}{R_f} + \tilde{b}^{\text{can}\prime}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) + b^{\text{mis}\prime}(\tilde{f}_{t+1}^{\text{mis}} - \mathbb{E}(\tilde{f}_{t+1}^{\text{mis}})) + c' e_{t+1}, \end{aligned}$$

where $Q=\text{cov}(f_{t+1}^{\text{can}},\ f_{t+1}^{\text{mis}\prime})$ is a $K^{\text{can}}\times K^{\text{mis}}$ matrix of covariances and

$$\begin{aligned} \tilde{b}^{\text{can}} &= b^{\text{can}} + V_{f^{\text{can}}}^{-1} Q b^{\text{mis}}, \\ \tilde{f}_{t+1}^{\text{mis}} - \mathbb{E}(\tilde{f}_{t+1}^{\text{mis}}) &= (f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) - Q' V_{f^{\text{can}}}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})). \end{aligned}$$

Notice that by construction $\text{cov}(f_{t+1}^{\text{can}}, \tilde{f}_{t+1}^{\text{mis}\prime})$ is a $K^{\text{can}}\times K^{\text{mis}}$ matrix of zeros, because $\tilde{f}_{t+1}^{\text{mis}}$ represent the linear-projection residual from projecting $f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})$ on $f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})$.

We now show how, starting from a candidate factor model with factors f_{t+1}^{can} that are correlated with the missing systematic factors f_{t+1}^{mis}, we can construct the APT-implied SDF.

Proposition IA.8.6 (SDF: Correlated case)**.** *Under Assumptions 2.1 and 2.2 of the APT, there exists an SDF of the form*

$$M_{t+1} = \frac{1}{R_f} + b^{\text{can}\,\prime}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) + b^{\text{mis}\,\prime}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})) + c' e_{t+1}, \quad \textit{where}$$

$$b^{\text{can}\,\prime} = \left(-\frac{\lambda^{\text{can}\,\prime} V^{-1}_{f^{\text{can}}}}{R_f} + \frac{\lambda^{\text{mis}\,\prime} V^{-1}_{f^{\text{mis}}}}{R_f} Q' V^{-1}_{f^{\text{can}}}\right) \times \left(I_{K^{\text{can}}} - Q V^{-1}_{f^{\text{mis}}} Q' V^{-1}_{f^{\text{can}}}\right)^{-1},$$

$$b^{\text{mis}\,\prime} = \left(-\frac{\lambda^{\text{mis}} V^{-1}_{f^{\text{mis}}}}{R_f} + \frac{\lambda^{\text{can}\,\prime} V^{-1}_{f^{\text{can}}}}{R_f} Q V^{-1}_{f^{\text{mis}}}\right) \times \left(I_{K^{\text{mis}}} - Q' V^{-1}_{f^{\text{can}}} Q V^{-1}_{f^{\text{mis}}}\right)^{-1},$$

$$c' = -\frac{a V_e^{-1}}{R_f}.$$

Proof: We guess that the SDF has the following functional form

$$M_{t+1} = \mathbb{E}(M_{t+1}) + b^{\text{can}\,\prime}(f^{\text{can}}_{t+1} - \mathbb{E}(f^{\text{can}}_{t+1})) + b^{\text{mis}\,\prime}(f^{\text{mis}}_{t+1} - \mathbb{E}(f^{\text{mis}}_{t+1})) + c' e_{t+1},$$

where b^{can} is a $K^{\text{can}} \times 1$ vector, b^{mis} is a $K^{\text{mis}} \times 1$ vector, and c is an $N \times 1$ vector. We identify the unknown vectors b^{can}, b^{mis}, and c by using the Law of One Price. Specifically, because we assume the existence of the risk-free asset, to determine the mean of the SDF, we use the condition

$$\mathbb{E}(M_{t+1}) = \frac{1}{R_f}.$$

Next, because λ^{can} represents a vector of prices of risk of f^{can}_{t+1} we have that

$$-\operatorname{cov}(M_{t+1}, f^{\text{can}}_{t+1}) \times R_f = \lambda^{\text{can}}.$$

These K^{can} conditions identify b^{can}:

$$b^{\text{can}\,\prime} = -\frac{1}{R_f} \lambda^{\text{can}\,\prime} V^{-1}_{f^{\text{can}}} - b^{\text{mis}\,\prime} Q' V^{-1}_{f^{\text{can}}}. \tag{IA25}$$

Similarly, λ^{mis} is the price of risk associated with factors f^{mis}_{t+1}, or equivalently,

$$-\operatorname{cov}(M_{t+1}, f^{\text{mis}}_{t+1}) \times R_f = \lambda^{\text{mis}}.$$

These K^{mis} conditions identify b^{mis}:

$$b^{\text{mis}\,\prime} = -\frac{\lambda^{\text{mis}\,\prime} V^{-1}_{f^{\text{mis}}}}{R_f} - b^{\text{can}\,\prime} Q V^{-1}_{f^{\text{mis}}}. \tag{IA26}$$

Putting together expressions (IA25) and (IA26), we obtain

$$b^{\text{can}\,\prime} = \left(-\frac{\lambda^{\text{can}\,\prime} V^{-1}_{f^{\text{can}}}}{R_f} + \frac{\lambda^{\text{mis}\,\prime} V^{-1}_{f^{\text{mis}}}}{R_f} Q' V^{-1}_{f^{\text{can}}}\right) \times \left(I_{K^{\text{can}}} - Q V^{-1}_{f^{\text{mis}}} Q' V^{-1}_{f^{\text{can}}}\right)^{-1},$$

$$b^{\text{mis}\prime} = \left(-\frac{\lambda^{\text{mis}}V_{f^{\text{mis}}}^{-1}}{R_f} + \frac{\lambda^{\text{can}\prime}V_{f^{\text{can}}}^{-1}}{R_f}QV_{f^{\text{mis}}}^{-1}\right) \times \left(I_{K^{\text{mis}}} - Q'V_{f^{\text{can}}}^{-1}QV_{f^{\text{mis}}}^{-1}\right)^{-1}.$$

Finally, it must be the case that the SDF prices the N basis assets:

$$\mathbb{E}(M_{t+1}(R_{t+1} - R_f 1_N)) = 0_N.$$

These N equations identify c. Given expressions (IA25) and (IA26), we obtain

$$c' = -\frac{aV_e^{-1}}{R_f}. \qquad \square$$

Next, we provide a non-negative SDF.

Proposition IA.8.7 (Nonnegative SDF: Correlated case)**.** *Under Assumptions 2.1 and 2.2 of the APT and the assumption that returns R_{t+1} are Gaussian, there exists an SDF $M_{\exp,t+1}$*

$$M_{\exp,t+1} = M_{\exp,t+1}^{\beta,\text{can}} \times M_{\exp,t+1}^{a} \times M_{\exp,t+1}^{\beta,\text{mis}} \quad \textit{where}$$

$$M_{\exp,t+1}^{\beta,\text{can}} = \frac{1}{R_f}\exp\left(b_+^{\text{can}\prime}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) - \frac{1}{2}b_+^{\text{can}\prime}V_{f^{\text{can}}}b_+^{\text{can}} - \frac{1}{2}b_+^{\text{can}\prime}Qb_+^{\text{mis}}\right)$$

$$M_{\exp,t+1}^{\beta,\text{mis}} = \exp\left(b_+^{\text{mis}\prime}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) - \frac{1}{2}b_+^{\text{mis}\prime}V_{f^{\text{mis}}}b_+^{\text{mis}} - \frac{1}{2}b_+^{\text{can}\prime}Qb_+^{\text{mis}}\right)$$

$$M_{\exp,t+1}^{a} = \exp\left(-a'V_e^{-1}e_{t+1} - \frac{1}{2}a'V_e^{-1}a\right), \quad \textit{where}$$

$$b_+^{\text{can}\prime} = (-\lambda^{\text{can}\prime}V_{f^{\text{can}}} + \lambda^{\text{mis}\prime}V_{f^{\text{mis}}}^{-1}Q'V_{f^{\text{can}}}^{-1}) \times (I_{K^{\text{can}}} - QV_{f^{\text{mis}}}^{-1}Q'V_{f^{\text{can}}}^{-1})^{-1},$$

$$b_+^{\text{mis}\prime} = (-\lambda^{\text{mis}\prime}V_{f^{\text{mis}}} + \lambda^{\text{can}\prime}V_{f^{\text{can}}}^{-1}QV_{f^{\text{mis}}}^{-1}) \times (I_{K^{\text{mis}}} - Q'V_{f^{\text{can}}}^{-1}QV_{f^{\text{mis}}}^{-1})^{-1}.$$

Proof: We use a guess-and-verify method to derive a nonnegative SDF. We guess that the SDF has the following functional form

$$M_{\exp,t+1} = \exp\left[\mu_+ + b_+^{\text{can}\prime}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) + b_+^{\text{mis}\prime}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) + c_+' e_{t+1}\right]$$

with unknown vectors b_+^{can}, b_+^{mis}, and c_+, as well as an unknown scalar μ_+. To identify the unknowns and verify our guess, we use the following $K^{\text{can}} + K^{\text{mis}} + N + 1$ equations, which are implications of the Law of One Price:

$$-\operatorname{cov}(M_{\exp,t+1}, f_{t+1}^{\text{can}}) \times R_f = \lambda^{\text{can}},$$

$$-\operatorname{cov}(M_{\exp,t+1}, f_{t+1}^{\text{mis}}) \times R_f = \lambda^{\text{mis}},$$

$$\mathbb{E}(M_{\exp,t+1}(R_{t+1} - R_f 1_N)) = 0_N$$

$$\mathbb{E}(M_{\exp,t+1}) = R_f^{-1}.$$

The first K^{can} equations imply that

$$-\mathbb{E}(M_{\exp,t+1}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}}))) = \mathbb{E}(M_{\exp,t+1}) \times \lambda^{\text{can}},$$

which. along with Lemma IA.3.1, give

$$V_{f^{\text{can}}}(b_{+}^{\text{can}} + V_{f^{\text{can}}}^{-1} Q b_{+}^{\text{mis}}) = -\lambda^{\text{can}}. \tag{IA27}$$

Similarly, the next K^{mis} equations imply that

$$-\mathbb{E}(M_{\exp,t+1}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}}))) = \mathbb{E}(M_{\exp,t+1}) \times \lambda^{\text{mis}},$$

which, along with Lemma IA.3.1, lead to:

$$V_{f^{\text{mis}}}(b_{+}^{\text{mis}} + V_{f^{\text{mis}}}^{-1} Q' b_{+}^{\text{mis}}) = -\lambda^{\text{mis}}. \tag{IA28}$$

From expressions (IA27) and (IA28), we obtain

$$\begin{aligned} b_{+}^{\text{can}\prime} &= (-\lambda^{\text{can}\prime} V_{f^{\text{can}}} + \lambda^{\text{mis}\prime} V_{f^{\text{mis}}}^{-1} Q' V_{f^{\text{can}}}^{-1}) \times (I_{K^{\text{can}}} - Q V_{f^{\text{mis}}}^{-1} Q' V_{f^{\text{can}}}^{-1})^{-1}, \\ b_{+}^{\text{mis}\prime} &= (-\lambda^{\text{mis}\prime} V_{f^{\text{mis}}} + \lambda^{\text{can}\prime} V_{f^{\text{can}}}^{-1} Q V_{f^{\text{mis}}}^{-1}) \times (I_{K^{\text{mis}}} - Q' V_{f^{\text{can}}}^{-1} Q V_{f^{\text{mis}}}^{-1})^{-1}. \end{aligned}$$

Next, we use the condition that the SDF prices the N basis assets and Lemma IA.3.1 to derive:

$$c_{+}' = -a' V_e^{-1}.$$

Finally, the last identification condition implies

$$\begin{aligned} &\frac{1}{R_f} = \mathbb{E}(M_{\exp,t+1}) \\ &= \mathbb{E}(\exp(\mu_{+} + b_{+}^{\text{can}\prime}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})) + b_{+}^{\text{mis}\prime}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})) + c_{+}' e_{t+1})) \\ &= \exp\left(\mu_{+} + (b_{+}^{\text{can}} + V_{f^{\text{can}}}^{-1} Q b_{+}^{\text{mis}})' V_{f^{\text{can}}} (b_{+}^{\text{can}} + V_{f^{\text{can}}}^{-1} Q b_{+}^{\text{mis}})/2 + b_{+}^{\text{mis}\prime} V_{f^{\text{mis}}} b_{+}^{\text{mis}}/2 + c_{+}' V_e c_{+}/2\right) \\ &= \exp(\mu_{+} + b_{+}^{\text{can}\prime} V_{f^{\text{can}}} b_{+}^{\text{can}}/2 + b_{+}^{\text{mis}\prime} V_{f^{\text{mis}}} b_{+}^{\text{mis}}/2 + a' V_e^{-1} a/2 + b_{+}^{\text{can}\prime} Q b_{+}^{\text{mis}}). \end{aligned}$$

In the last equation, we use $V_{f^{\text{mis}}} = Q' V_{f^{\text{can}}}^{-1} Q + V_{\tilde{f}^{\text{mis}}}$, where f_{t+1}^{can} and $\tilde{f}_{t+1}^{\text{mis}}$ are orthogonal, and $c' = -a' V_e^{-1}$. As a result,

$$\exp(\mu_{+}) = R_f^{-1} \times \exp(-b_{+}^{\text{can}\prime} V_{f^{\text{can}}} b_{+}^{\text{can}}/2 - b_{+}^{\text{mis}\prime} V_{f^{\text{mis}}} b_{+}^{\text{mis}}/2 - a' V_e^{-1} a/2 - b_{+}^{\text{can}\prime} Q b_{+}^{\text{mis}}). \quad \square$$

Next, let us introduce the projection version of the SDF $M_{\exp,t+1}$. First, note that it is convenient to express $M_{\exp,t+1}$ as

$$M_{\exp,t+1} = \frac{1}{R_f} \times \exp\left(m_{t+1} - \frac{1}{2} m\right) \tag{IA29}$$

$$= \frac{1}{R_f} \times \exp\big(m_{t+1}^{\beta,\mathrm{can}} + m_{t+1}^{\beta,\mathrm{mis}} + m_{t+1}^{a} - \frac{1}{2} m^{\beta,\mathrm{can}} - \frac{1}{2} m^{\beta,\mathrm{mis}} - \frac{1}{2} m^{a}\big),$$

where

$$\begin{aligned}
m_{t+1} &= m_{t+1}^{\beta,\mathrm{can}} + m_{t+1}^{\beta,\mathrm{mis}} + m_{t+1}^{a},\\
m &= m^{\beta,\mathrm{can}} + m^{\beta,\mathrm{mis}} + m^{a},\\
m_{t+1}^{\beta,\mathrm{can}} &= b_{+}^{\mathrm{can}\,\prime}(f_{t+1}^{\mathrm{can}} - \mathbb{E}(f_{t+1}^{\mathrm{can}})),\\
m_{t+1}^{\beta,\mathrm{mis}} &= b_{+}^{\mathrm{mis}\,\prime}(f_{t+1}^{\mathrm{mis}} - \mathbb{E}(f_{t+1}^{\mathrm{mis}})),\\
m_{t+1}^{a} &= -a' V_e^{-1} e_{t+1},\\
m^{\beta,\mathrm{can}} &= b_{+}^{\mathrm{can}\,\prime} V_{f^{\mathrm{can}}} b_{+}^{\mathrm{can}} + b_{+}^{\mathrm{can}\,\prime} Q b_{+}^{\mathrm{mis}},\\
m^{\beta,\mathrm{mis}} &= b_{+}^{\mathrm{mis}\,\prime} V_{f^{\mathrm{mis}}} b_{+}^{\mathrm{mis}} + b_{+}^{\mathrm{can}\,\prime} Q b_{+}^{\mathrm{mis}},\\
m^{a} &= a' V_e^{-1} a.
\end{aligned}$$

Second, set $X_{t+1} = R_{t+1} - R_f 1_N - \mu$ with $\mu = \mathbb{E}(R_{t+1} - R_f 1_N)$, $f_t = (f_t^{\mathrm{can}\,\prime}, f_t^{\mathrm{mis}\,\prime})'$, and $\beta = (\beta^{\mathrm{can}}, \beta^{\mathrm{mis}})$ and notice that $V_R = \beta V \beta' + V_e$ with $V = \begin{pmatrix} V_{f^{\mathrm{can}}} & Q \\ Q' & V_{f^{\mathrm{mis}}} \end{pmatrix}$.

Finally, define the projected non-negative SDF as

$$\hat{M}_{\exp,t+1} = \frac{1}{R_f} \times \exp\left(\hat{m}_{t+1} - \frac{1}{2}\hat{m}\right), \quad \text{where} \tag{IA30}$$

$$\begin{aligned}
\hat{m}_{t+1} &= \mathbb{E}(m_{t+1} X_{t+1}') \mathbb{E}(X_{t+1} X_{t+1}')^{-1} X_{t+1},\\
\hat{m} &= \frac{1}{2} Var(\hat{m}_{t+1}).
\end{aligned}$$

Thus,

$$\begin{aligned}
\hat{m}_{t+1} &= \hat{m}_{t+1}^{\beta,\mathrm{can}} + \hat{m}_{t+1}^{\beta,\mathrm{mis}} + \hat{m}_{t+1}^{a} \quad \text{and} \quad \hat{m} = \hat{m}^{\beta,\mathrm{can}} + \hat{m}^{\beta,\mathrm{mis}} + \hat{m}^{a}, \quad \text{where}\\
\hat{m}_{t+1}^{\beta,\mathrm{can}} &= b_{+}^{\mathrm{can}\,\prime}(V_{f^{\mathrm{can}}} \beta^{\mathrm{can}\,\prime} + Q\beta^{\mathrm{mis}\,\prime}) V_R^{-1} X_{t+1} = b_{+}^{\mathrm{can}\,\prime}(V_{f^{\mathrm{can}}},\ Q)\beta' V_R^{-1} X_{t+1},\\
\hat{m}_{t+1}^{\beta,\mathrm{mis}} &= b_{+}^{\mathrm{mis}\,\prime}(Q'\beta^{\mathrm{can}\,\prime} + V_{f^{\mathrm{mis}}} \beta^{\mathrm{mis}\,\prime}) V_R^{-1} X_{t+1} = b_{+}^{\mathrm{mis}\,\prime}(Q',\ V_f^{\mathrm{mis}})\beta' V_R^{-1} X_{t+1},\\
\hat{m}_{t+1}^{a} &= c_{+}' V_e V_R^{-1} X_{t+1},\\
\hat{m}^{\beta,\mathrm{can}} &= b_{+}^{\mathrm{can}\,\prime}(V_{f^{\mathrm{can}}}, Q)\beta' V_R^{-1} \beta (V_{f^{\mathrm{can}}}, Q)' b_{+}^{\mathrm{can}} + b_{+}^{\mathrm{can}\,\prime}(V_{f^{\mathrm{can}}}, Q)\beta' V_R^{-1} \beta (Q', V_{f^{\mathrm{mis}}})' b_{+}^{\mathrm{mis}},\\
\hat{m}^{\beta,\mathrm{mis}} &= b_{+}^{\mathrm{mis}\,\prime}(Q', V_{f^{\mathrm{mis}}})\beta' V_R^{-1} \beta (Q', V_{f^{\mathrm{mis}}})' b_{+}^{\mathrm{mis}} + b_{+}^{\mathrm{can}\,\prime}(V_{f^{\mathrm{can}}}, Q)\beta' V_R^{-1} \beta (Q', V_{f^{\mathrm{mis}}})' b_{+}^{\mathrm{mis}},\\
\hat{m}^{a} &= c_{+}' V_e V_R^{-1} V_e c_{+}.
\end{aligned}$$

Proposition IA.8.8 (Asymptotic Properties of the SDF Projections: Correlated Case)**.** *Under the assumptions of Proposition IA.4.1, as $N \to \infty$, $M_{\exp,t+1}$ and $\hat{M}_{\exp,t+1}$ of* (IA29) *and* (IA30) *satisfy*

$$\hat{M}_{\exp,t+1} - M_{\exp,t+1} \xrightarrow{p} 0.$$

Proof: We have, as $N \to \infty$,

$$c_+' V_e V_R^{-1} e_{t+1} - c_+' e_{t+1} \xrightarrow{p} 0,$$
$$\beta' V_e V_R^{-1} c_+ \longrightarrow 0_{K^{\text{can}}+K^{\text{mis}}},$$
$$\beta' V_R^{-1} e_{t+1} \xrightarrow{p} 0_{K^{\text{can}}+K^{\text{mis}}},$$
$$\beta' V_R^{-1} \beta \longrightarrow V_f^{-1},$$

and therefore,

$$\begin{aligned} \hat{m}_{t+1}^{\beta,\text{can}} &= b_+^{\text{can}\,\prime}(V_{f^{\text{can}}}, \quad Q)\beta' V_R^{-1} X_{t+1} \xrightarrow{p} b_+^{\text{can}\,\prime}(V_{f^{\text{can}}}, \quad Q)V^{-1}(f_{t+1} - \mathbb{E}(f_{t+1})) \\ &= b_+^{\text{can}\,\prime}(I_{K^{\text{can}}}, \quad O_{K^{\text{can}}})(f_{t+1} - \mathbb{E}(f_{t+1})) = b_+^{\text{can}\,\prime}(f_{t+1}^{\text{can}} - \mathbb{E}(f_{t+1}^{\text{can}})), \end{aligned} \tag{IA31}$$

and

$$\begin{aligned} \hat{m}_{t+1}^{\beta,\text{mis}} &= b_+^{\text{mis}\,\prime}(Q, \quad V_{f^{\text{mis}}})\beta' V_R^{-1} X_{t+1} \xrightarrow{p} b_+^{\text{mis}\,\prime}(Q', \quad V_{f^{\text{can}}})V^{-1}(f_{t+1} - \mathbb{E}(f_{t+1})) \\ &= b_+^{\text{mis}\,\prime}(O_{K^{\text{mis}}}, \quad I_{K^{\text{mis}}})(f_{t+1} - \mathbb{E}(f_{t+1})) = b_+^{\text{mis}\,\prime}(f_{t+1}^{\text{mis}} - \mathbb{E}(f_{t+1}^{\text{mis}})). \end{aligned} \tag{IA32}$$

Given that $\beta' V_e V_R^{-1} c_+ \longrightarrow 0_{K^{\text{can}}+K^{\text{mis}}}$ and $c_+' V_e V_R^{-1} e_{t+1} - c_+' e_{t+1} \xrightarrow{p} 0$, as $N \to \infty$, then

$$c_+' V_e V_R^{-1} X_{t+1} - c_+' e_{t+1} \xrightarrow{p} 0, \quad \text{as} \quad N \to \infty. \tag{IA33}$$

As a result, expressions (IA31), (IA32), and (IA33) imply that

$$\hat{m}_{t+1} - m_{t+1} \xrightarrow{p} 0, \quad \text{as} \quad N \to \infty, \tag{IA34}$$

Similarly, the result that, as $N \to \infty$, $\beta' V_R^{-1} \beta \longrightarrow V_f^{-1}$, implies that

$$\hat{m}^{\beta\,\text{can}} \longrightarrow b_+^{\text{can}\,\prime} V_{f^{\text{can}}} b_+^{\text{can}} + b_+^{\text{can}\,\prime} Q b_+^{\text{mis}} = m^{\beta,\text{can}}, \quad \text{as} \quad N \to \infty, \quad \text{and} \tag{IA35}$$

$$\hat{m}^{\beta\,\text{mis}} \longrightarrow b_+^{\text{mis}\,\prime} V_{f^{\text{mis}}} b_+^{\text{mis}} + b_+^{\text{can}\,\prime} Q b_+^{\text{mis}} = m^{\beta,\text{mis}}, \quad \text{as} \quad N \to \infty. \tag{IA36}$$

Notice that

$$\hat{m}^a = c_+' V_e V_R^{-1} V_e c_+ = a' V_e^{-1} V_e V_R^{-1} V_e V_e^{-1} a = a' V_R^{-1} a$$

and recall that the proof of Proposition IA.4.1 shows that $a' V_R^{-1} a - a' V_e^{-1} a \longrightarrow 0$, as $N \to \infty$, which implies that

$$\hat{m}^a - m^a \longrightarrow 0, \quad \text{as} \quad N \to \infty, \tag{IA37}$$

Expressions (IA35), (IA36), and (IA37) imply that

$$\hat{m} - m \longrightarrow 0, \quad \text{as} \quad N \to \infty. \tag{IA38}$$

From the results in expressions (IA34) and (IA38), we obtain

$$\hat{M}_{\exp,t+1} - M_{\exp,t+1} \xrightarrow{p} 0, \quad \text{as} \quad N \to \infty. \qquad \square$$

Remark: Note that because f_{t+1}^{can} are observable factors, in empirical work, we may use the exact component $m_{t+1}^{\beta,\text{can}} + m^{\beta,\text{can}}$ rather than its projected counterpart $\hat{m}_{t+1}^{\beta,\text{can}} + \hat{m}^{\beta,\text{can}}$.

IA.9 Data description

To examine which economic variables may explain variation in the SDF, we collect the returns on a set of 457 trading strategies and 103 macroeconomic and financial indicators.

The set of trading strategies includes:

- 205 strategies from Chen and Zimmermann (2022).
- 153 strategies in the Global Factor Dataset from Jensen et al. (2022).
- 55 strategies from Kozak et al. (2020).
- 35 strategies from Bryzgalova et al. (2023). The sources of these strategies are specified in their Internet Appendix. Their dataset includes 34 trading strategies, but we consider two versions of the size strategy, one from Fama and French (1993) and the other from Fama and French (2015).
- We add the following nine strategies:
  - Industry-adjusted value, momentum, and profitability factors; intra-industry value, momentum, and profitability factors; profitable-minus-unprofitable factor from Novy-Marx (2013), available from `http://rnm.simon.rochester.edu`.
  - Expected-growth factor of Hou et al. (2021), available from `https://global-q.org/index.html`.
  - Up-minus-down (UMD) factor from the AQR data library, available from `https://www.aqr.com/Insights/Datasets`.

The set of macroeconomic and financial indicators includes:

- 53 variables constructed from 17 variables from Bryzgalova et al. (2023). Below we explain how we get to 53 variables.
  - For indices of financial uncertainty, real uncertainty, and macroeconomic uncertainty, we consider time horizons of 1, 3, and 12 months. We use these variables in levels and consider their AR(1) innovations, for a total of 18 variables.
  - For the investor-sentiment measures of Baker and Wurgler (2006) and Huang et al. (2015), labeled as BW_INV_SENT and HJTZ_INV_SENT, respectively, we consider both the orthogonalized and non-orthogonalized versions. We use these variables in levels and consider AR(1) innovations of these variables for a total of 8 variables.

  - For other persistent variables, such as the term spread (TERM), change in the difference between a 10-year Treasury bond yield and a 3-month Treasury bill yield (DELTA_SLOPE), credit spread (CREDIT), dividend yield (DIV), price-earnings ratio (PE), unemployment rate (UNRATE), the growth rate of industrial production (IND_PROD), the monthly growth rate of the Producer Price Index for Crude Petroleum (OIL), we look at both levels and first-order differences, for a total of 16 variables.
  - Real per capita consumption growth on nondurable goods and services separately and jointly. We also include the 3-year consumption growth (nondurable goods and services) and its AR(1) innovations, for a total of 5 variables.
  - Inflation, computed as the log-difference in the price index for both nondurable goods and services and its AR(1) innovations, for a total of 2 variables.
  - The level of the intermediary-capital ratio and its innovations, for a total of 2 variables.
  - The level of the aggregate liquidity factor and its innovations.

- The first 3 principal components and their VAR(1) innovations for the 279 macroeconomic variables from Jurado et al. (2015), for a total of 6 variables.

- The first eight principal components and their VAR(1) innovations for the 128 macroeconomic variables from the FRED-MD dataset of McCracken and Ng (2016), gives a total of 16 variables. We obtain these macro variables from `https://research.stlouisfed.org/econ/mccracken/fred-databases` and use the data vintage for February 2021. We exclude four variables, ACOGNO, ANDENOx, TWEXAFEGSMTHx, and UMCSENT, which have missing observations at the start of the sample.

- Consumer sentiment and its first-order differences.

- The market-dislocation index of Pasquariello (2014), its first-order differences, and AR(1) innovations.

- The disagreement index of Huang, Li, and Wang (2021) and its first-order differences.

- The Chicago Board Options Exchange (CBOE) volatility index (VIX) available on the website of the CBOE, its first-order differences, and AR(1) innovations.

- The U.S. economic policy uncertainty index (EPU) of Baker, Bloom, and Davis (2016) and the equity market volatility (EMV) tracker of Baker, Bloom, Davis, and Kost (2019), which are available from `www.policyuncertainty.com`. For both indices, we also consider their first-order differences and AR(1) innovations.

- The U.S. business-confidence index, the U.S. consumer-confidence index, and the U.S. composite leading indicator from the OECD library.
- The coincident economic-activity index and its first-order differences from https://fred.stlouisfed.org/series/USPHCI.
- The NBER recession index from https://fred.stlouisfed.org/series/USREC.
- The TED spread from https://fred.stlouisfed.org/series/TEDRATE.
- The effective federal funds rate and the real federal funds rate from https://fred.stlouisfed.org/series/FEDFUNDS.
- The credit-spread index (Gilchrist and Zakrajšek, 2012) and its first order differences.
- The Chicago Fed National Financial Condition Index from https://fred.stlouisfed.org/series/NFCI.

IA.10 Additional Tables

This section contains additional tables supporting the interpretation of our empirical results described in the manuscript.

Table IA.1: The APT-implied SDF and observable variables

This table reports the explanatory power of the selected variables for the unsystematic (Panel A) and systematic (Panel B) components of the APT-implied SDF.

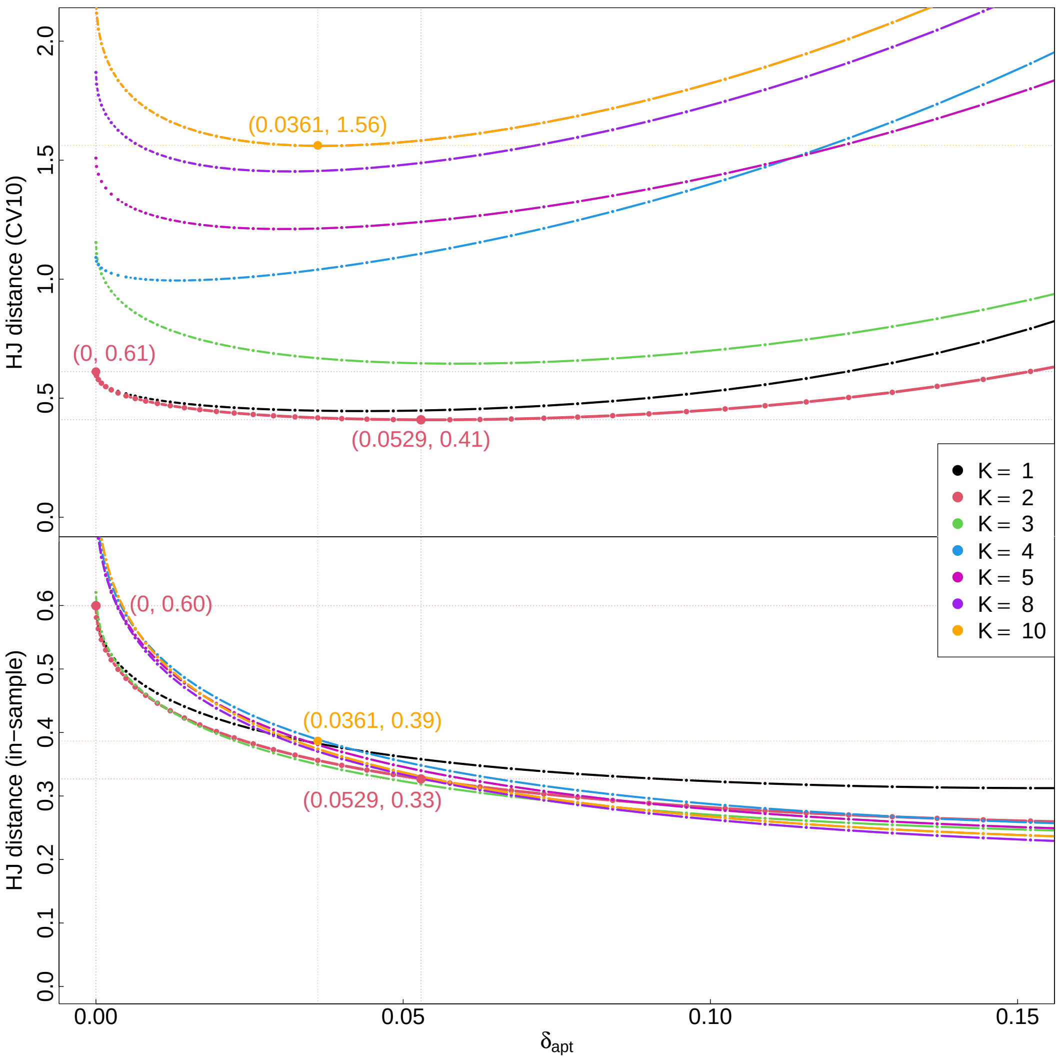
	R^2(%)	p-value
Panel A: $\log(M^{a}_{\exp,t+1})$		
NBER recession indicator	0.18	0.27
Intermediary constraints (He et al., 2017)	2.71	0.00
Sentiment index (Baker and Wurgler, 2006)	2.60	0.00
Sentiment index (Huang et al., 2015)	3.41	0.00
Shocks in credit spread (Gilchrist and Zakrajšek, 2012)	1.92	0.00
Shocks in VIX	2.24	0.00
Panel B: $\log(M^{\beta}_{\exp,t+1})$		
NBER recession indicator	0.76	0.02
Chicago Fed National Financial Condition Index	2.64	0.00
Intermediary constraints (He et al., 2017)	55.15	0.00
Shocks in aggregate liquidity (Pástor and Stambaugh, 2003)	11.12	0.00
Shocks in credit spread (Gilchrist and Zakrajšek, 2012)	13.79	0.00
Shocks in dividend yield (Campbell, 1996)	40.39	0.00
Shocks in financial uncertainty (Jurado et al., 2015)	11.09	0.00
Shocks in VIX	54.32	0.00
TED spread	4.39	0.00

IA.11 Additional Figures

This section contains additional figures supporting the interpretation of our empirical results described in the manuscript.

Figure IA.1: APT model selection: Cross-validation vs. in-sample

This figure illustrates how the HJ distance changes with the key parameters of the APT model, K and δ_{apt}. The top panel, which is the same as Figure 1 in the main text, shows the results of the cross-validation exercise for different combinations of K and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results when the model is estimated on the full sample. The numbers reported in the two plots are $(\delta_{\text{apt}}, \text{HJ distance})$. The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors and W is the inverse of the second-order moment matrix of asset returns.



The top panel of Figure IA.1, which is the same as Figure 1 in the main text, shows how the HJ distance of the APT model, which is evaluated on the *validation folds* of the cross-validation procedure, changes as we use different K and δ_{apt} in the estimation. The bottom panel of Figure IA.1 shows how the HJ distance of the APT model changes as we use different K and δ_{apt} in the estimation on the *full sample*.

The bottom panel shows that if we were to naively choose K and δ_{apt} based on an *in-sample* analysis instead of cross-validation, we would have selected much larger values for these parameters. This is because the larger number of factors K fits better the in-sample covariance matrix of returns, while the larger δ_{apt} fits better the in-sample cross-sectional variation in expected excess returns. However, choosing K and δ_{apt} based on in-sample fit leads to overfitting and, consequently, an inferior fit of asset returns out-of-sample.

Figure IA.2: Spanning the unsystematic SDF component

The blue curve shows the R^2 (left axis) of 325 regressions of $\log(\hat{M}^a_{\exp,t+1})$ on the returns of trading strategies that are available for the entire sample. The first regression includes the return on the trading strategy that explains the most variation in $\log(\hat{M}^a_{\exp,t+1})$; each subsequent regression includes the return on an extra trading strategy that adds the most to explaining the variation in $\log(\hat{M}^a_{\exp,t+1})$. The red curve shows the value of the Bayesian Information Criteria (BIC) (right axis) associated with these regressions. The minimal BIC$= -552.86$ is for the regression with 39 explanatory variables; the associated $R^2 = 66.45\%$.

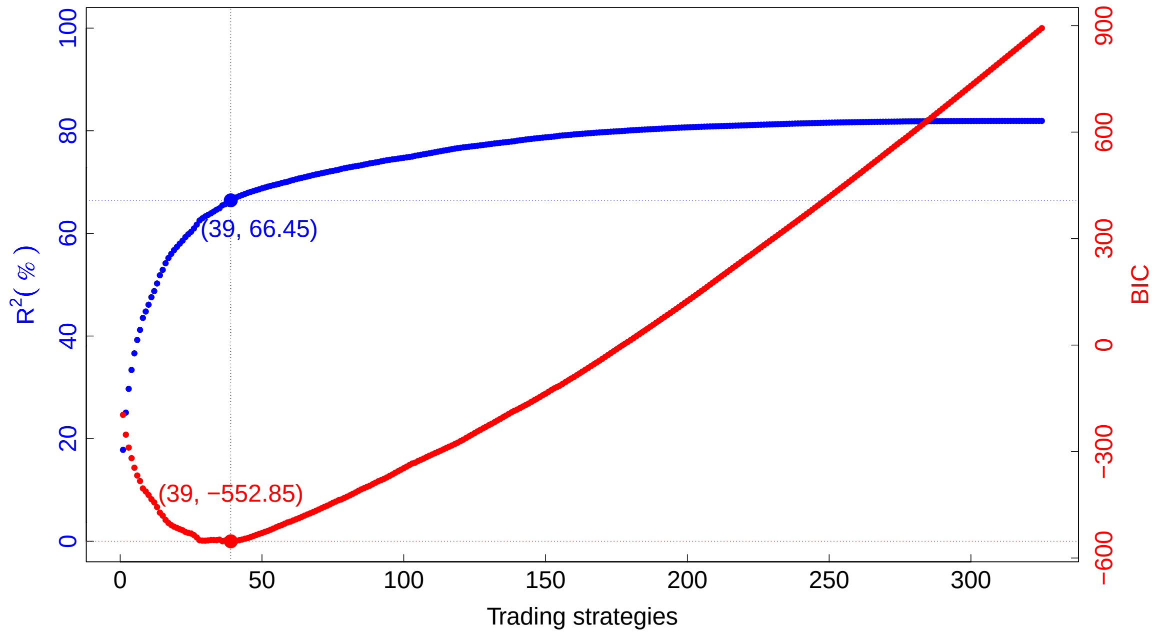


Figure IA.3: Correcting the CAPM

This figure illustrates how the HJ distance changes with K^{mis} and δ_{apt} when we correct for misspecification the candidate model, which includes only the market return as a systematic factor. The top panel shows the results of the cross-validation exercise for different combinations of K^{mis} and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are (δ_{apt}, HJ distance). The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\text{exp},t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors, W is the inverse of the second-order moment matrix of asset returns, and $\hat{M}_{\text{exp},t+1}$ is the SDF implied by the CAPM (brown triangle in the bottom panel) or the corrected SDFs for different values of K^{mis} and δ_{apt}.

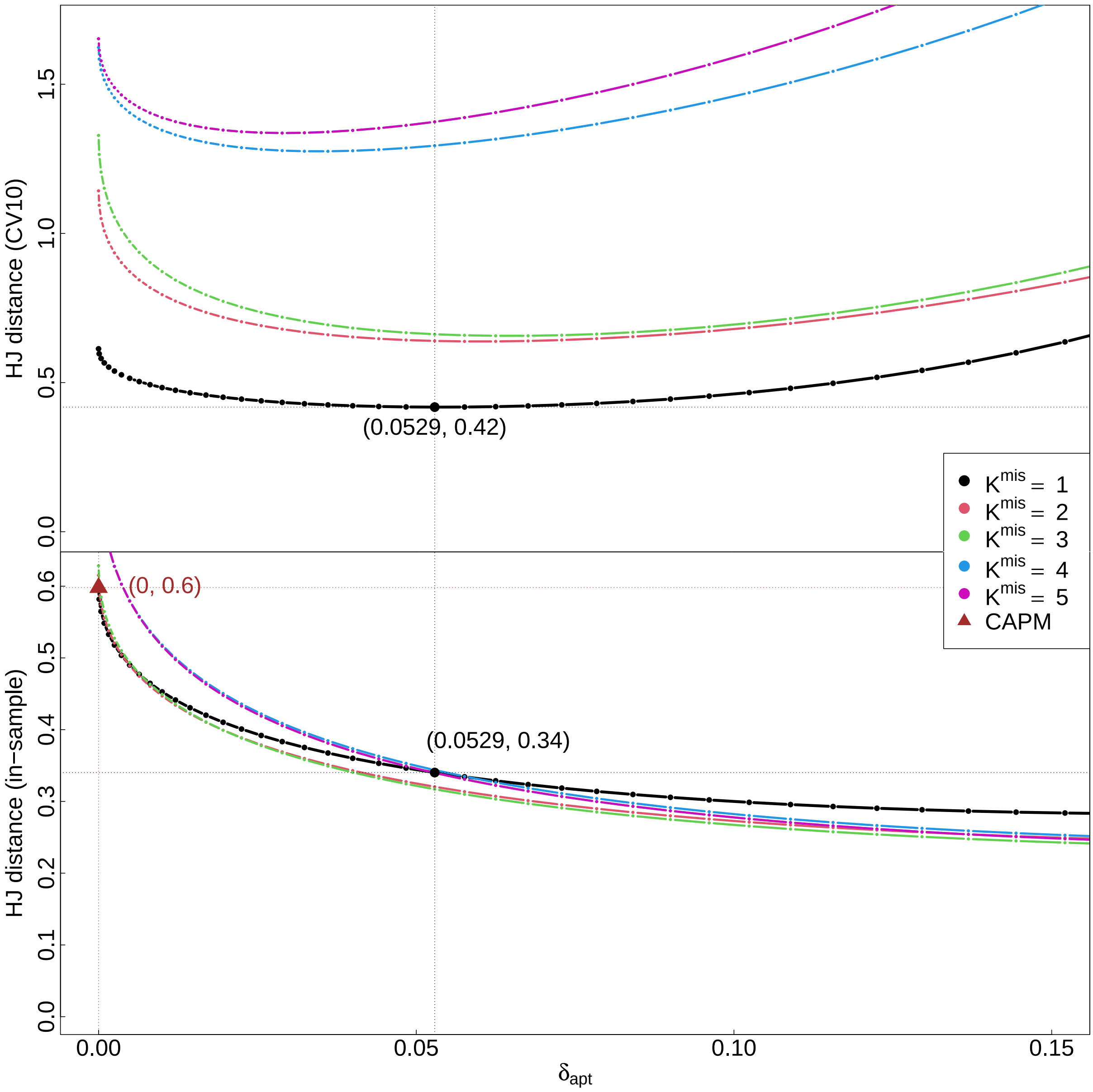


Figure IA.4: Time-series behavior of the corrected SDF and its components when the candidate model is the CAPM

This figure has four panels, which show the dynamics of the corrected SDF, $\hat{M}_{\exp,t+1}$ and its three components: the unsystematic component $\hat{M}^{a}_{\exp,t+1}$, the component $\hat{M}^{\beta,\text{can}}_{\exp,t+1}$ corresponding to the candidate model with the market factor, and the missing systematic component $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$. Gray bars indicate NBER recession periods.

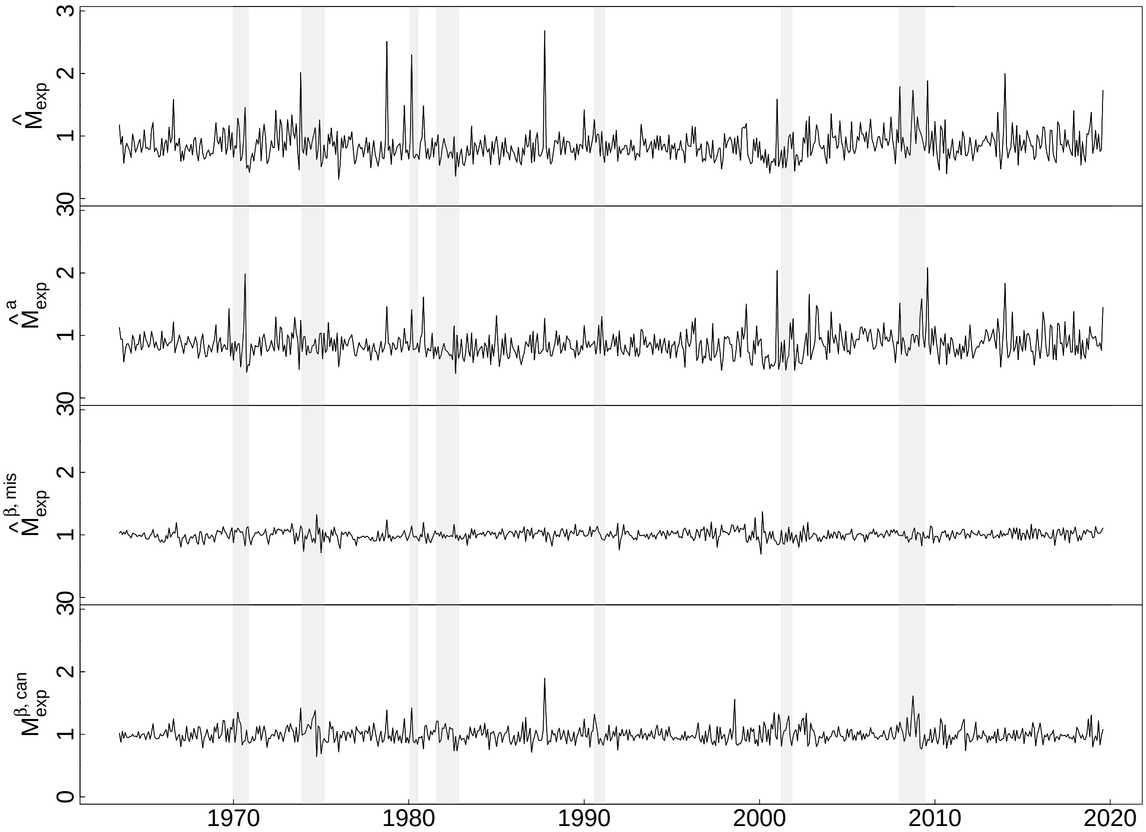


Figure IA.5: Pricing errors in the candidate and corrected CAPM

The red dots in this figure indicate the annualized pricing errors for the 202 basis assets using the CAPM as the candidate model. The blue dots indicate the annualizes pricing errors using the *corrected* CAPM.

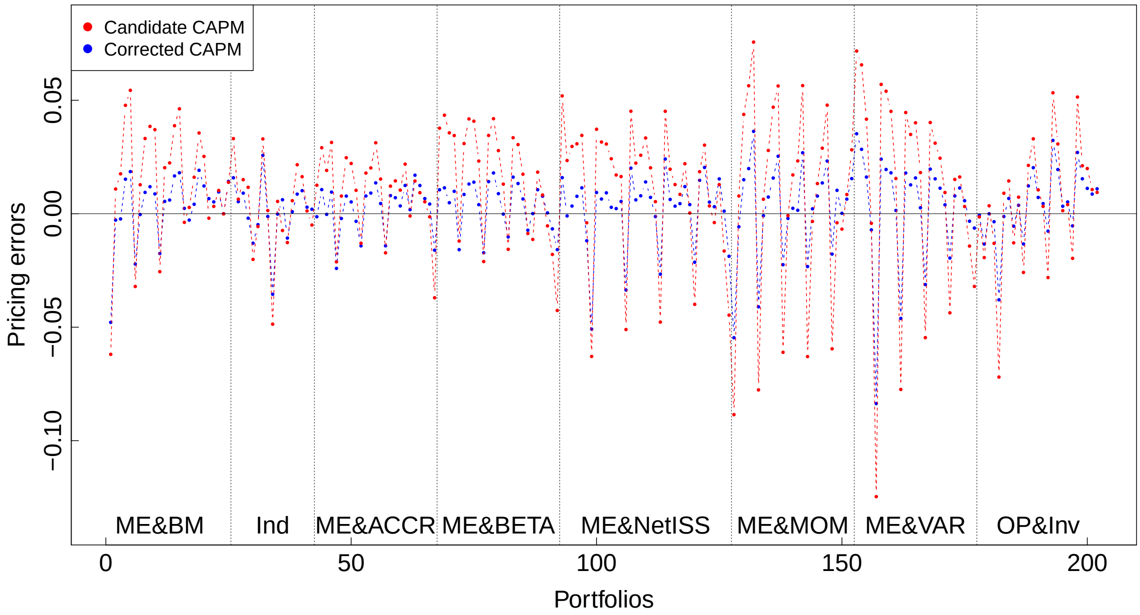


Figure IA.6: Correcting the C-CAPM

This figure illustrates how the HJ distance changes with K^{mis} and δ_{apt} when we correct for misspecification the candidate model, which includes only the return on the consumption-mimicking portfolio of Breeden et al. (1989) as a systematic factor. The top panel shows the results of the cross-validation exercise for different combinations of K^{mis} and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are (δ_{apt}, HJ distance). The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\text{exp},t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors, W is the inverse of the second-order moment matrix of asset returns, and $\hat{M}_{\text{exp},t+1}$ is the SDF implied by the C-CAPM (brown triangle in the bottom panel) or the corrected SDFs with different values of K^{mis} and δ_{apt}.

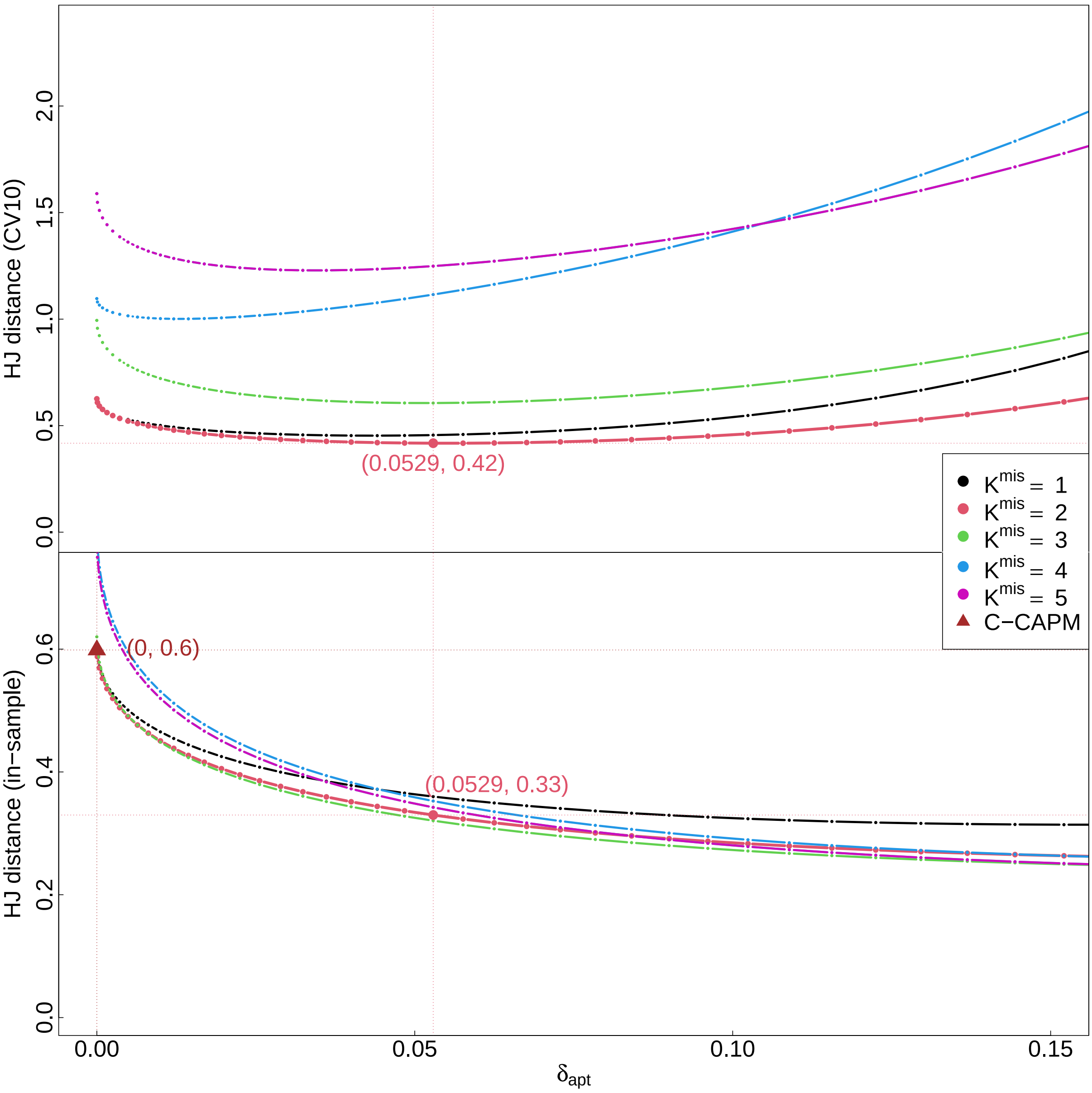


Figure IA.7: Time-series behavior of the corrected SDF and its components when the candidate model is the C-CAPM

This figure has four panels, which show the dynamics of the corrected SDF $\hat{M}_{\exp,t+1}$ and its three components: the unsystematic component $\hat{M}^{a}_{\exp,t+1}$, the component $\hat{M}^{\beta,\text{can}}_{\exp,t+1}$ corresponding to the candidate model with the consumption mimicking portfolio as the sole factor, and the missing systematic component $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$. Gray bars indicate NBER recession periods.

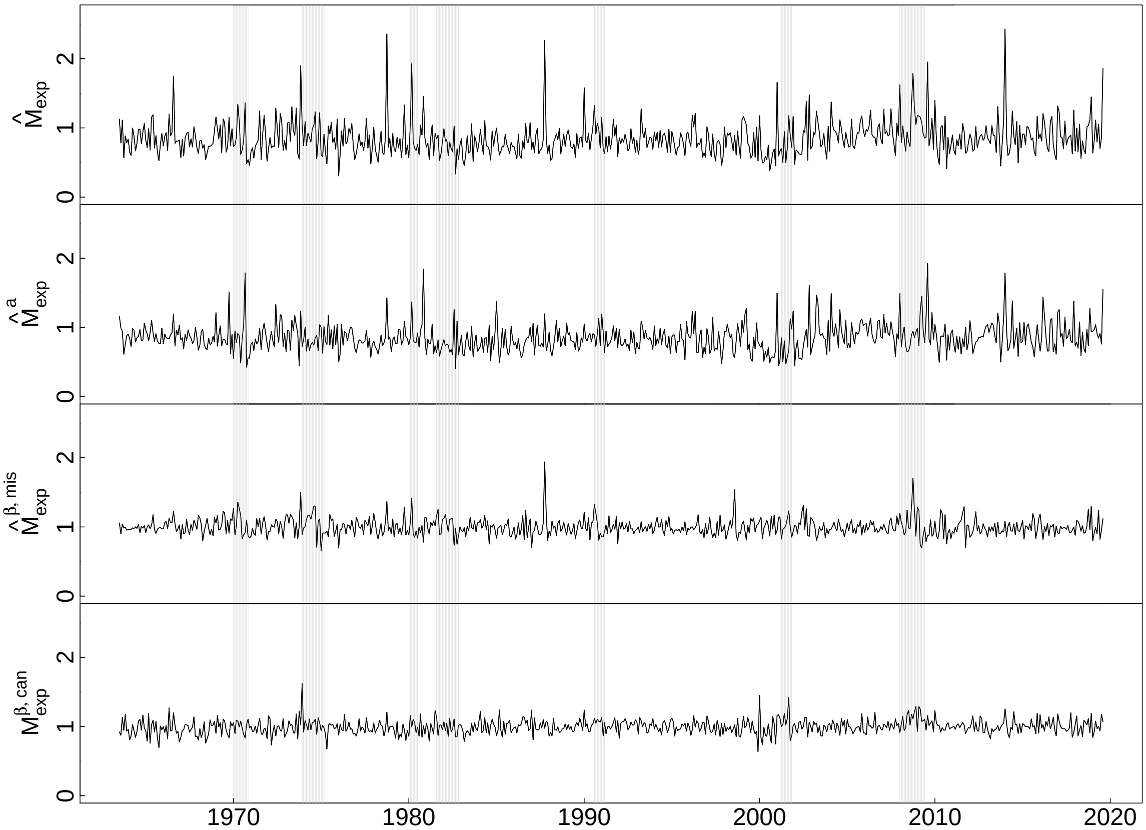


Figure IA.8: Correcting the FF3

This figure illustrates how the HJ distance changes with K^{mis} and δ_{apt} when we correct for misspecification the three-factor model of Fama and French (1993). The top panel shows the results of the cross-validation exercise for different combinations of K^{mis} and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are (δ_{apt}, HJ distance). The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors, W is the inverse of the second-order moment matrix of asset returns, and $\hat{M}_{\exp,t+1}$ is the SDF implied by the FF3 model (brown triangle in the bottom panel) or the corrected SDFs with different values of K^{mis} and δ_{apt}.

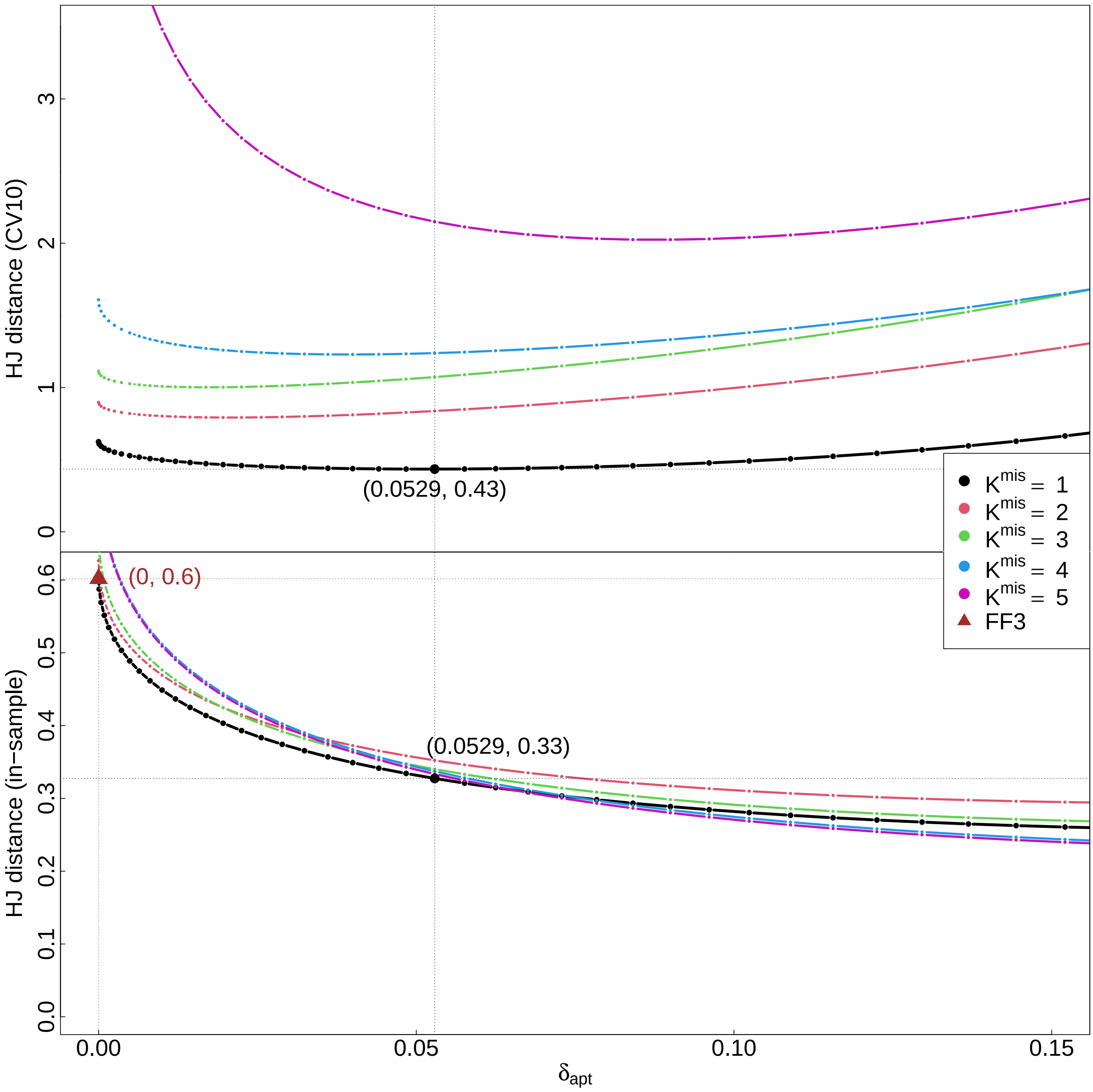


Figure IA.9: Time-series behavior of the corrected SDF and its components when the candidate model is FF3

This figure has four panels, which show the dynamics of the corrected SDF, $\hat{M}_{\exp,t+1}$ and its three components: the unsystematic component $\hat{M}^{a}_{\exp,t+1}$, the component $\hat{M}^{\beta,\text{can}}_{\exp,t+1}$ corresponding to the candidate FF3 model, and the missing systematic component $\hat{M}^{\beta,\text{mis}}_{\exp,t+1}$. Gray bars indicate NBER recession periods.

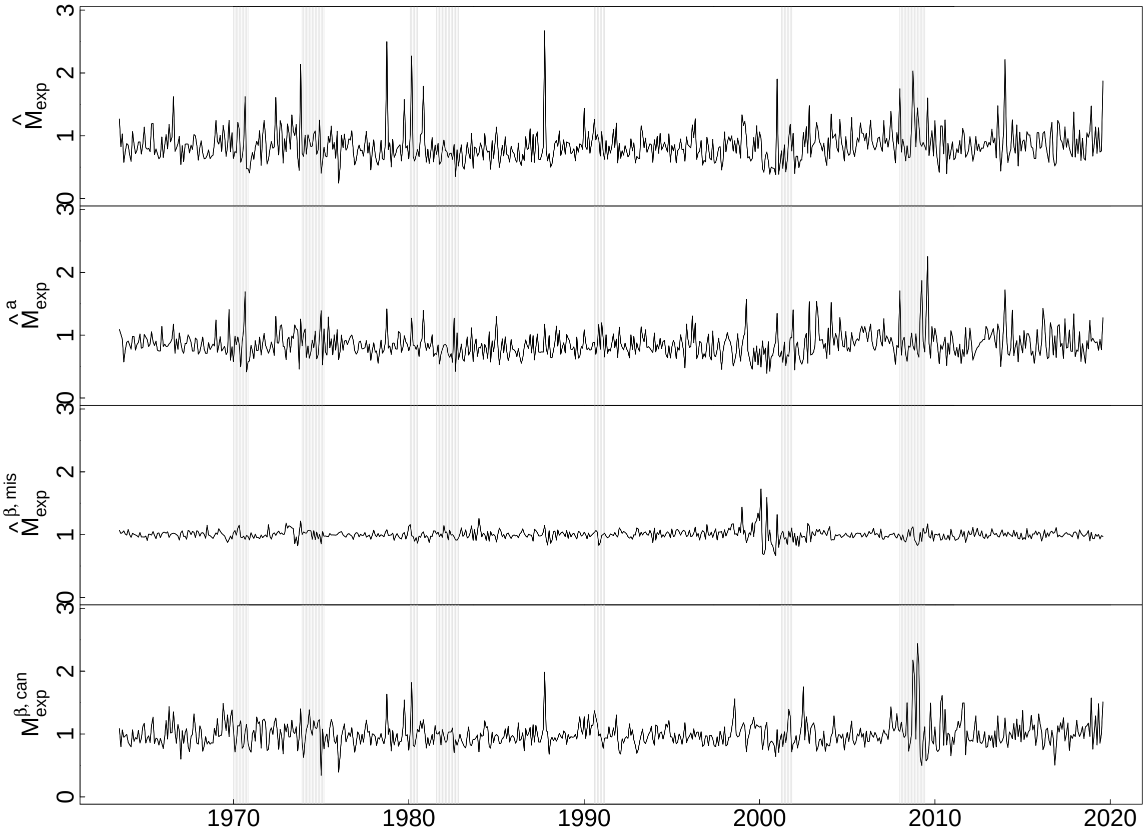


Figure IA.10: Pricing errors in the candidate and corrected FF3 model

The red dots in this figure indicate the annualized pricing errors for the 202 basis assets when using the candidate FF3 model. The blue dots indicate the annualized pricing errors using the *corrected* FF3 model.

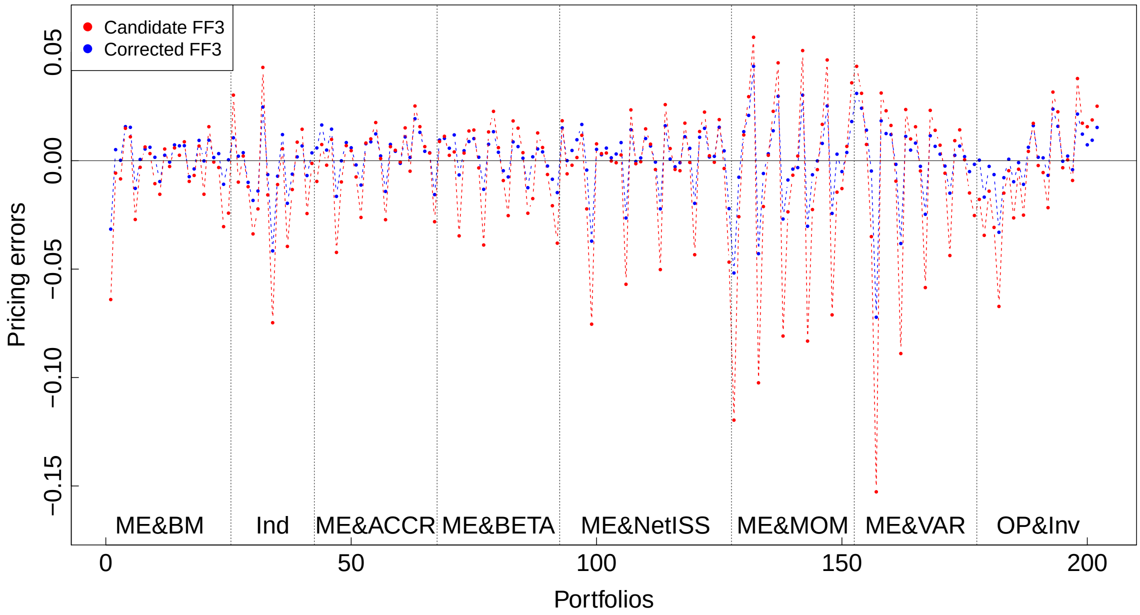


Figure IA.11: APT model selection on the KNS daily dataset

This figure illustrates how the HJ distance changes with the key parameters of the APT model, K and δ_{apt}. The top panel shows the results of the cross-validation exercise for different combinations of K and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are $(\delta_{\text{apt}}, \text{HJ distance})$. The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors and W is the inverse of the second-order moment matrix of asset returns.

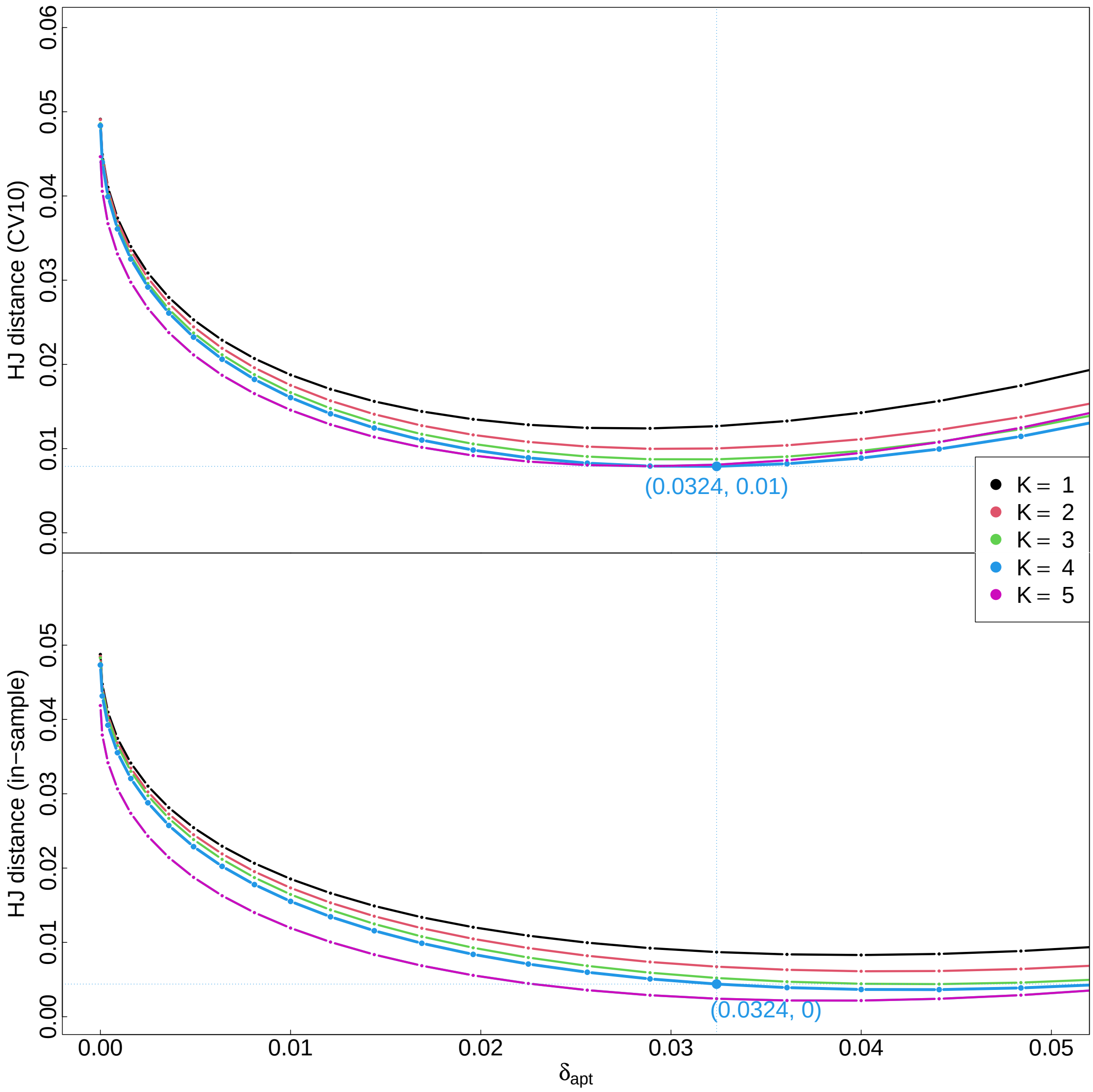


Figure IA.12: APT model selection on the KNS monthly dataset

This figure illustrates how the HJ distance changes with the key parameters of the APT model, K and δ_{apt}. The top panel shows the results of the cross-validation exercise for different combinations of K and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are $(\delta_{\text{apt}}, \text{HJ distance})$. The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors and W is the inverse of the second-order moment matrix of asset returns.

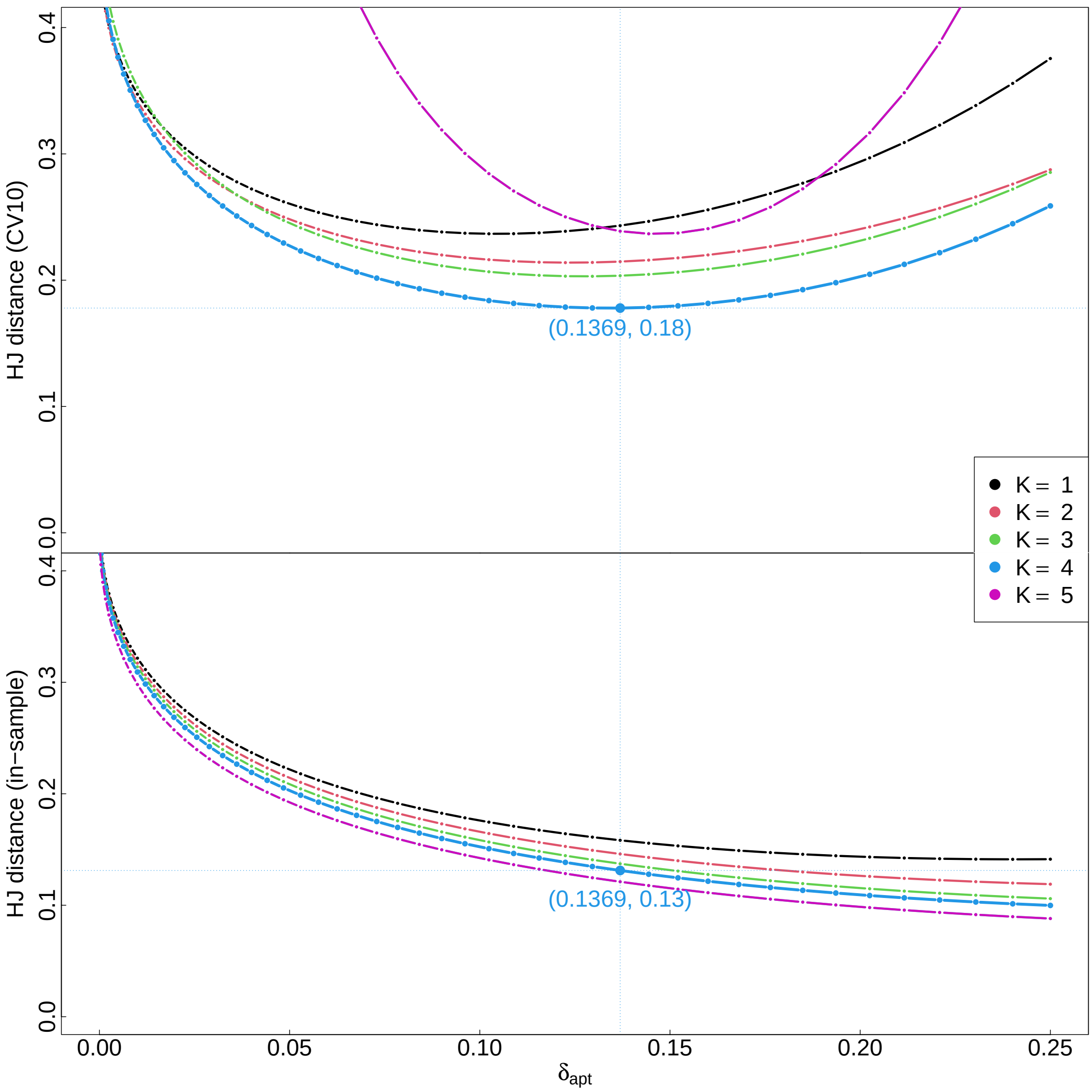


Figure IA.13: APT model selection on the LP monthly dataset

This figure illustrates how the HJ distance changes with the key parameters of the APT model, K and δ_{apt}. The top panel shows the results of the cross-validation exercise for different combinations of K and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are $(\delta_{\text{apt}}, \text{HJ distance})$. The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors and W is the inverse of the second-order moment matrix of asset returns.

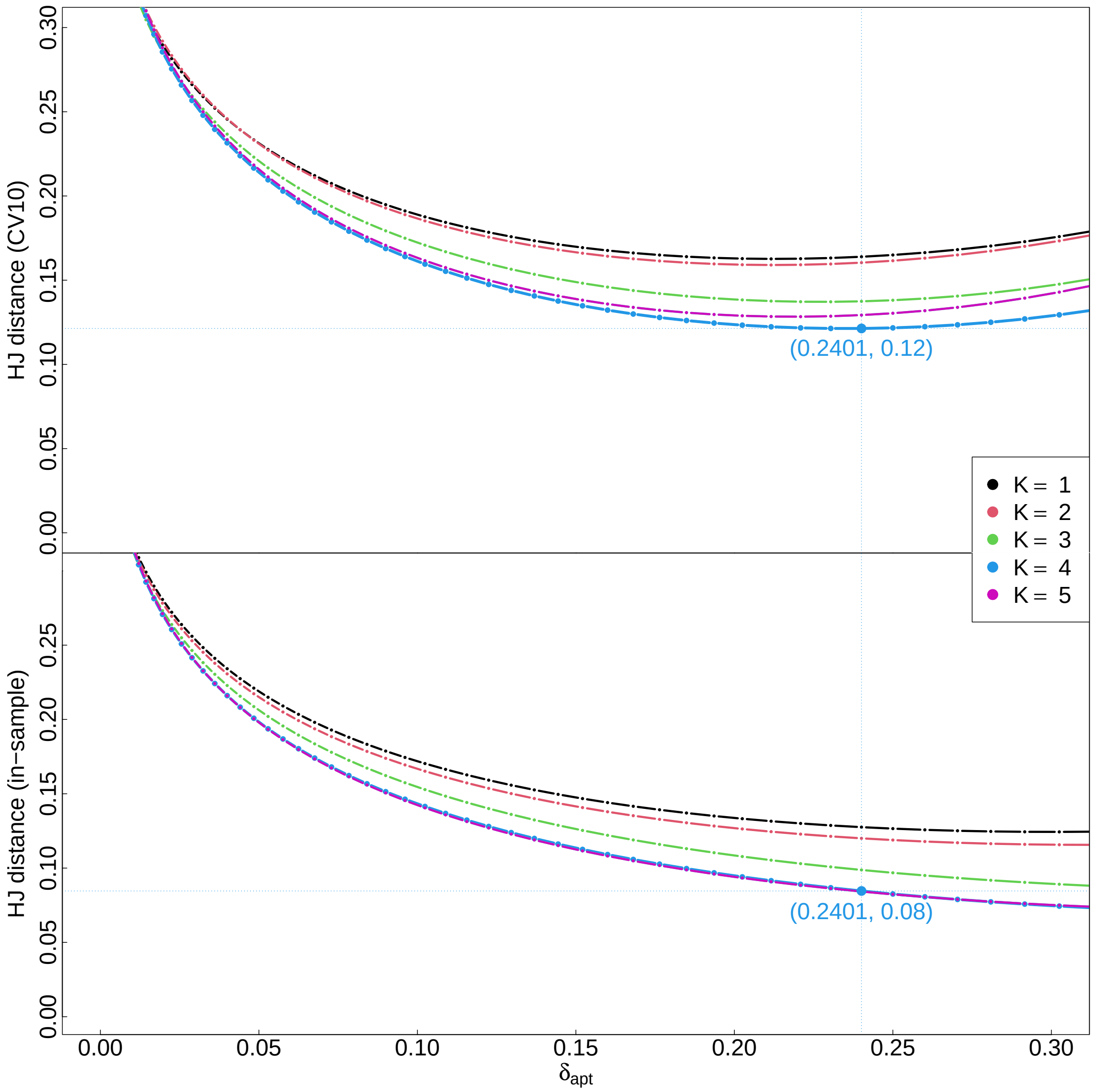


Figure IA.14: Correcting the SDF based on Kozak et al. (2020)

This figure illustrates how the HJ distance changes with K^{mis} and δ_{apt} when we correct for misspecification the model of Kozak et al. (2020). The top panel shows the results of the cross-validation exercise for different combinations of K^{mis} and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are (δ_{apt}, HJ distance). The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors, W is the inverse of the second-order moment matrix of asset returns, and $\hat{M}_{\exp,t+1}$ is the SDF based on Kozak et al. (2020) (brown triangle in the bottom panel) or the corrected SDFs with different values of K^{mis} and δ_{apt}. The SDFs are estimated on the original daily data used in Kozak et al. (2020).

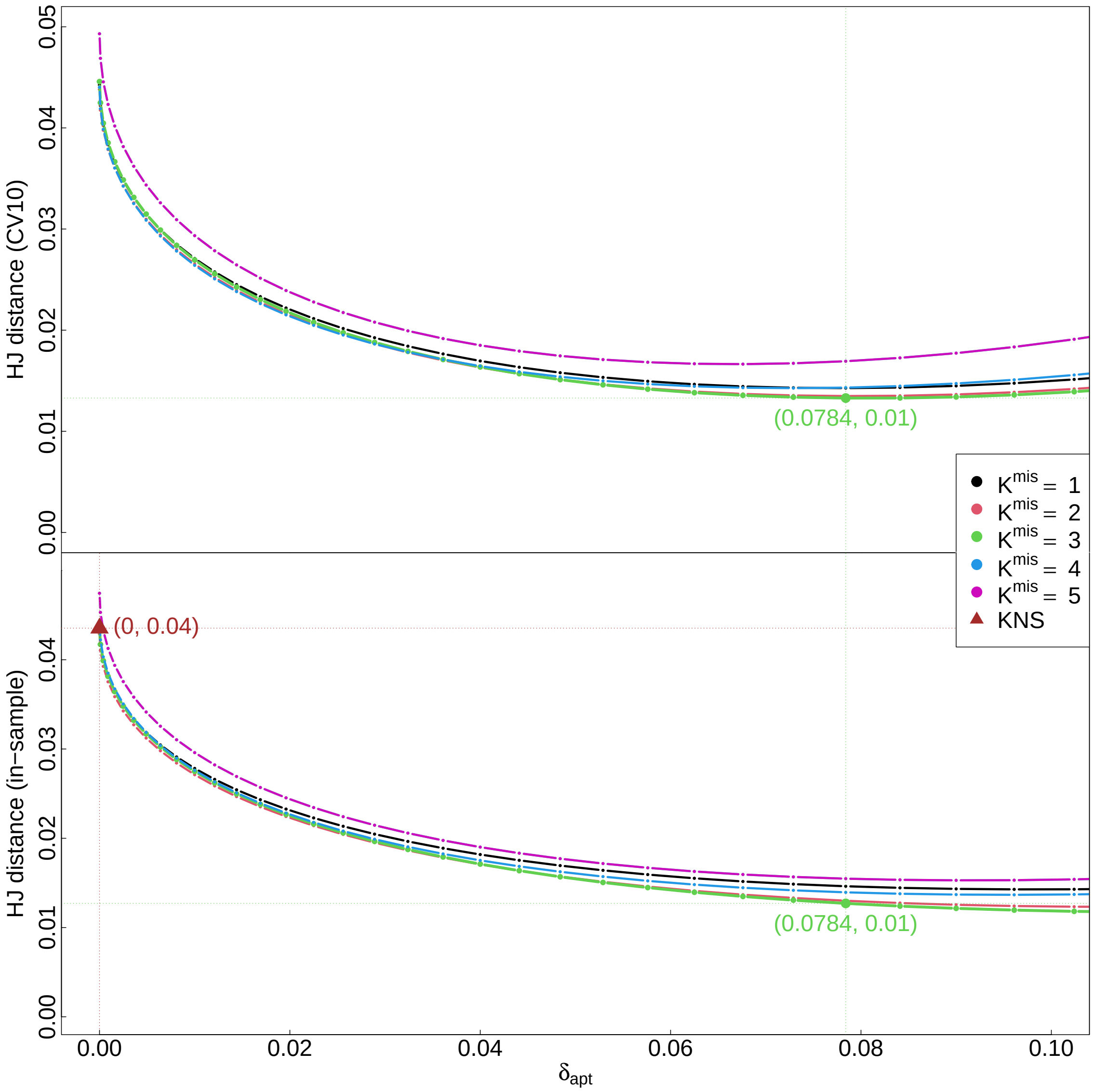


Figure IA.15: Correcting the RP-PCA SDF based on Lettau and Pelger (2020)

This figure illustrates how the HJ distance changes with K^{mis} and δ_{apt} when we correct for misspecification the model of Lettau and Pelger (2020). The top panel shows the results of the cross-validation exercise for different combinations of K^{mis} and δ_{apt}. We split the sample into ten folds and estimate the model on all but one fold, which we use for validation. We repeat this procedure ten times and compute the HJ distance on the validation folds. The bottom panel shows the results for the full sample. The numbers reported in the two plots are (δ_{apt}, HJ distance). The HJ distance is $p_e' W p_e$, where $p_e = \mathbb{E}(\hat{M}_{\exp,t+1}(R_{t+1} - R_f 1_N))$ is a vector of pricing errors, W is the inverse of the second-order moment matrix of asset returns, and $M_{\exp,t+1}$ is the SDF based on Lettau and Pelger (2020) (brown triangle in the bottom panel) or the corrected SDFs with different values of K^{mis} and δ_{apt}. The SDFs are estimated on the daily data used in Lettau and Pelger (2020).