Asset Pricing When T Equals One

Paolo Zaffaroni
Imperial College Business School

March 16, 2026

Abstract

This paper develops a methodology for conducting inference on the pricing ability of asset pricing models, and constructing the corresponding mean-variance portfolios, that uses one large cross-section of N asset returns corresponding to *one* time observation T. *No assumption* is required on the form of time variation of risk premia and risk exposures. My analysis is based on the novel asymptotics of Principal Component Analysis valid for $T = 1$ and large N. My methodology permits to study how the pricing and portfolio performance of asset pricing models changes across time period by period, as opposed to existing methodologies. Moreover, I can show *analytically* what affects the out-of-sample pricing ability and portfolio performance of the candidate asset pricing model. The methodology can be used for either cases of observed or latent risk factors, and a combination of the two, and is designed to exploit the rich information available in large cross-sections of individual stocks, although it can be used more widely. An empirical application testing the out-of-sample pricing ability *across time* of the zoo factors as well as the commonly used cross-sectional models, demonstrates the benefits of this methodology.

1 Introduction

This paper challenges the common-sense view that accurate estimation and testing of the pricing ability of asset pricing models requires a large dataset, and in particular a long time-series span T. While this view is grounded in standard statistical considerations, it comes at a substantial cost: it severely limits the ability of asset pricing models to explain the dynamics of asset prices or, more broadly, to quantify the interaction between asset prices and the wider economy as new information arrives. Examples include changes in monetary policy, geopolitical events, or the sudden onset of economic and financial crises, which are seen to dramatically affect asset prices. At the same time, widespread empirical evidence suggests that individual asset returns' distribution changes across time, again questioning the use of long time series span of data.

By contrast, if we could estimate asset pricing models, test their pricing ability, and construct efficient portfolios using only a single time observation, we would be able to exploit the well-known immediate response of asset prices to economic news, and allow for model-free time variation in the distribution of asset prices. More generally, restricting attention to one time period makes it possible to use conditional asset pricing models without imposing virtually any restriction on the functional form or the degree of time-variation in assets' risk premia, their volatility, and other parameters. The restrictions that are typically imposed when T is large–purely for tractability and statistical convenience–can be completely avoided when $T = 1$. These commonly adopted restrictions substantially increase the risk of model misspecification and significantly reduce the usefulness of conditional asset pricing models.

A further central objective of asset pricing is to explain cross-sectional differences in asset prices, that is, differences across individual assets. However, for both practical and statistical reasons, most empirical applications work with portfolio-level data. By construction, portfolios obscure and, in particular, substantially reduce cross-sectional heterogeneity in asset prices and returns.

Taken together, these considerations imply that the effective use of asset pricing models calls for a large cross-section of asset returns of size N observed over a very short time span T, ideally with $T = 1$, in stark contrast to common practice in empirical asset pricing. Resolving the methodological challenges associated with this setting is the main contribution of this paper. I develop a formal methodology that enables the estimation and testing of the out-of-sample pricing ability of (conditional) factor asset pricing models–the workhorses of asset pricing and related fields–using a single large cross-section of returns observed over a single period.

Because the pricing ability of an asset pricing model can be evaluated in two conceptually distinct ways–either through the ability of factor risk premia to explain the cross-section of expected returns, or through the ability of the stochastic discount factor (SDF) to price the cross-section of assets' payoffs–I construct two corresponding classes of pricing tests.[1] I also provide a formal treatment of how to estimate mean–variance efficient portfolios, and their implied returns, in environments where N is large but $T = 1$.

There are many advantages, but also important challenges, associated with using my $T = 1$ methodology. First, the methodology has the virtue of simplicity: setting $T = 1$ greatly simplifies the analytical derivations. Because I derive the asymptotic properties of both in-sample (IS) and out-of-sample (OOS)

[1] Although, in population, the pricing errors from fitting expected returns and those from SDF-based pricing are identical up to a scaling factor given by the inverse of the gross risk-free return (see Hansen and Jagannathan (1997)), in finite samples they will generally differ, and thus require distinct asymptotic analyses.

versions of my pricing tests, the analysis explains analytically how and why IS and OOS results differ. This goes beyond the conventional view that IS results are inconclusive because of overfitting, while OOS results are unstable and unreliable for the opposite reason. I develop this IS–OOS comparison both for the statistical properties of the pricing tests and for the performance of mean–variance portfolios. The analytical results provide clear guidance on how to mitigate, at least partially, some of the well-known pitfalls that affect OOS performance.

Second, when data are available over long time spans, my methodology permits the construction of time series of pricing tests and their corresponding p-values by applying the tests period by period. This allows one to investigate whether a model's pricing ability varies over time–for instance, across different phases of the business cycle. An important special case is the analysis of the time-varying pricing ability of the hundreds of risk factors (the "factor zoo") proposed in the empirical asset pricing literature. In this way, one can rank these factors by the fraction of periods in which they significantly price the cross-section of stock returns, either unconditionally or conditionally. For example, some factors may prove more useful during economic downturns, while others may be more relevant during expansions.

Third, I develop the methodology for models with both latent and observed risk factors, including settings in which both types of factors are present simultaneously. When risk factors are latent and estimated via Principal Component Analysis (PCA), I derive an asymptotic theory for PCA that is valid when $T = 1$ but N is large, which I refer to as *nano-PCA*. With these results in hand, I can analyze the costs and benefits of using factor models with latent (PCA-estimated) factors versus observed factors. On the one hand, PCA-estimated factors eliminate the risk of model misspecification arising from an incorrect choice of observed risk factors. On the other hand, they are affected by estimation noise, which leads to larger standard errors for estimators and test statistics. My methodology allows this trade-off to be evaluated analytically.

There are also many challenges that must be addressed in order to obtain a workable methodology. First, holding T fixed, and in particular setting $T = 1$, implies that model estimates will generally be subject to estimation bias, whose magnitude differs across parameters. For example, while PCA estimates of latent risk factors are accurately recovered (up to a rotation matrix) even when T is fixed, provided $N \to \infty$ the corresponding risk exposures remain biased regardless of how large N becomes. However, as my nano-PCA analysis shows, it is possible to estimate the moments of these noisy risk exposures—moments that are required for accurate estimation of standard errors for pricing tests and portfolio weights estimators—by applying suitable bias corrections.

Second, when $T = 1$ one can only work with a one-factor model. Moreover, whether the factor is latent or observed, the estimated model attains a perfect in-sample fit, which makes it impossible to estimate the variance of the error term. This is crucial because the error variance enters the bias adjustments and standard errors of my pricing test statistics and mean–variance portfolios. By contrast, OOS analysis requires at least two time observations–one for estimation and one for OOS evaluation. This allows both observations to be used for bias adjustment and for the estimation of standard errors, including, in particular, the error variance. Third, as already noted, when $T = 1$ I can only accommodate a single risk factor, whereas in reality the cross-section of returns may require a multi-factor structure. In such cases, my pricing tests and mean–variance portfolios are affected not only by the estimation bias associated with $T = 1$, but also by a misspecification bias arising from imposing a one-factor model when more factors are in fact needed. I show how to modify the methodology so as to mitigate this

misspecification bias. A further, well-known challenge in factor asset pricing is the identification of the number of priced risk factors. In my setting, this issue largely disappears: it is redundant when$T = 1$ and, more generally, when $T = r$, where r denotes the number of priced risk factors.[2]

My methodology relies on a number of key insights. Foremost, I build the analysis on the assumption of no-arbitrage. Although this is an economic assumption, no-arbitrage implies that the intercept term (i.e., alpha) is square-summable across assets and hence is diversified away when estimating the factor model. This means that I can avoid estimating alpha or, equivalently, that I do not need to de-mean the data. In my $T = 1$ setting, this is crucial: if the data were de-meaned, they would become identically equal to zero. As a consequence, sample variances and covariances are replaced by sample second moments (i.e., moments computed without de-meaning).

Next, as noted, I can allow only for a single risk factor when $T = 1$. More generally, I can allow for at most r risk factors when $T = r \geq 1$. In either case, the model attains perfect in-sample fit. Moreover, when $T = 1$ OOS pricing test statistics, model-implied expected returns, and mean–variance portfolios are identical regardless of whether the risk factor is PCA-estimated or observed, and they coincide with the corresponding formulae based solely on sample moments of the test assets. The same applies when $T = r \geq 1$. However, the necessary bias adjustments and standard errors do depend on whether one assumes a latent or an observed factor model, and they differ from the pure sample-moment formulae because at least two time periods are used—or, more generally, $r + 1$ periods when $r \geq 1$.

Lastly, the ability of nano-PCA to estimate accurately the moments of the noisy (i.e., inconsistent) risk exposures implies that I can still estimate the returns of mean–variance strategies with high accuracy, even though the corresponding portfolio weights—approximately linear in the risk exposures—remain severely biased when $T = 1$. This, in turn, reflects the fact that the ability of mean–variance portfolios to diversify away idiosyncratic risk requires only that N be large, regardless of the value of T.

The asymptotic theory for my methodology, including nano-PCA, rests on a single economic assumption—the absence of arbitrage—and on several statistical assumptions. No-arbitrage allows me to avoid de-meaning the data, which is essential for the asymptotic validity of the procedure when $T = 1$.

No assumptions are required regarding the time-variation of the factor asset pricing model for estimation under $T = 1$, and only a very limited degree of local smoothness in the time-variation of risk exposures is needed when two periods are considered. I also require the number of risk factors to remain locally constant over two periods. However, I do impose that the covariances of the error terms across assets "average out," which rules out time-variation in variances and covariances when averaged cross-sectionally, while still allowing for cross-sectional dependence in the idiosyncratic errors.

Some of these assumptions are consistent with empirical evidence: for example, stock returns tend to exhibit weak autocorrelation over time but strong cross-sectional dependence. Finally, the methodology is robust to mispricing; that is, it remains valid under an Arbitrage Pricing Theory (APT) setting in which expected returns can systematically deviate from those implied by the factor structure.

My paper advances three different streams of literature.

Cross-Sectional Asset Pricing. Principal Component Analysis (PCA) has become an increasingly popular tool in empirical asset pricing because, for portfolio construction and pricing, the risk factors need not be directly observed. As highlighted in the Arbitrage Pricing Theory (APT) of Ross (1976), this yields a reduced-form stochastic discount factor (SDF) with strong pricing ability and robustness to

[2]When there are $T > r$ periods available for estimation, the consistent estimator for r in Zaffaroni (2025) can be used.

omitted priced risk factors. From an asset pricing perspective, my work is closely related to Pelger (2019, 2020), Lettau and Pelger (2020a,b), Giglio and Xiu (2021), Giglio, Xiu, and Zhang (2025), Ma and Gagliardini (2020), and Zaffaroni (2025).

It differs from these contributions in that I develop an inferential theory for *conditional* asset pricing models based on my nano-PCA approach, whereas—with the exception of Pelger (2019, 2020)—these papers rely on constant-PCA or variants thereof. From an econometric standpoint, I extend the consistency results for constant PCA (i.e., under constant model parameters) established in Connor and Korajczyk (1986) and Bai (2003, Theorem 4), as well as the local-PCA analysis (i.e., when T remains fixed but exceeds r) in Zaffaroni (2025), by providing the asymptotic theory for nano-PCA in the case $T = r = 1$.

Conditional Asset Pricing. Several PCA-based procedures have been proposed for conditional asset pricing models, in which risk exposures, risk premia, and sometimes alphas are modeled as functions of observed instruments, such as firm characteristics; see, among others, Connor, M., and Linton (2012), Fan, Liao, and Wang (2016), Kelly, Pruitt, and Su (2019), Kim, Korajczyk, and Neuhierl (2021), Pelger and Xiong (2019), Haddad, Kozak, and Santosh (2020), and Ma and Gagliardini (2020). Further details are provided in Section OA2. Despite their appeal, these approaches may suffer from misspecification of the functional form (when parametrically imposed), from an incorrect choice of instruments, and often from the lack of a fully developed asymptotic theory. An alternative and extremely popular approach in empirical asset pricing is to exploit rolling windows of data, as in the influential methodology of Fama and MacBeth (1973), further refined in Andersen, Bollerselv, Diebold, and Wu (2006), Lewellen and Nagel (2006), and Ang and Kristensen (2012), among others.

My methodology advances this line of work by capturing fully model-free time variation in the parameters of the factor asset pricing model through time averages over short windows of fixed length T, particularly in the case $T = 1$.

Out-of-Sample Asset Pricing. Understanding the discrepancy between in-sample (IS) and out-of-sample (OOS) performance is central in empirical asset pricing. Among many important empirical contributions, Campbell and Thompson (2008) show that imposing simple economic constraints (e.g., a non-negative equity premium) improves OOS forecasts relative to unconstrained regressions, outperforming the historical mean both statistically and economically. In contrast, Goyal and Welch (2008) conduct a comprehensive horse race among popular equity premium predictors using real-time OOS tests, concluding that, for most sample periods and predictors, it is very difficult to beat the historical average excess return. Rapach, Strauss, and Zhou (2010) and Aiolfi and Timmermann (2006) advocate forecast combinations to enhance OOS performance. Kelly and Pruitt (2013) develop a framework that extracts market expectations of long-run cash flows and discount rates from the cross-section of asset prices via present-value identities, and show that these expectations have strong OOS predictive power for risk premia. In a Bayesian setting, Pettenuzzo, Timmermann, and Valkanov (2014) demonstrate OOS gains from imposing economic constraints. Gu, Kelly, and Xiu (2020) show that machine-learning-based models substantially improve OOS return prediction and deliver high-Sharpe long–short portfolios, while clarifying which predictors are most important. On the methodological front, Martin (2017) derives an option-implied measure of the expected equity return that does not rely on historical averages and shows that this measure outperforms conventional predictors in forecasting future market excess returns out of sample. Martin and Nagel (2019) examine whether the proliferation of signals and big-data-based

strategies is consistent with market efficiency, finding that many apparent IS profits largely disappear when evaluated OOS. Feng, Giglio, and Xiu (2020) construct a methodology that allows to sort (i.e., tame) the hundreds of observed risk factors proposed in the literature, in terms of their ability to explain the cross-section of asset returns. Bryzgalova, Huang, and Julliard (2023) use Bayesian Model Averaging to construct an SDF with superior OOS performance. Jacquier, Muhle-Karbe, and Mulligan (2025), extending Kan, Wang, and Zheng (2024), analyze theoretically how IS Sharpe ratios of linear predictive strategies translate into OOS performance.

My paper contributes to the OOS asset pricing literature by establishing *analytically* the differences between IS and OOS performance both for tests of pricing ability (based on risk premia or on the SDF) and for portfolio performance (measured by the Sharpe ratio).

I use my methodology to address the zoo problem, i.e., evaluating the contribution to asset pricing of the many individual observed risk factors, in terms of both OOS pricing ability and OOS portfolio performance. In contrast to existing methodologies, I am able to assess the ability of each risk factor across time, for example in correspondence to specific time periods of interest such as recessions and expansions. Finally, I am able to assess their pricing ability also adjusting for the possibility of misspecification bias, evincing whether some risk factor spans additional sources of risk to the core set of systematic risk factors identified by PCA. I will apply my methodology in terms of the risk premia pricing ability, and in terms of mean-variance portfolio performance, but not in terms of the SDF pricing ability, although I provide the complete asymptotic analysis of testing the SDF pricing ability.

In terms of pricing ability test, It emerges that core asset-pricing factors (market, size, value, profitability, investment, momentum, BAB/QMJ, HML_Devil) are systematically among the best. Profitability and investment characteristics (HXZ/Fama–French-style) and related variables (NOA/DNOA, GMA, ORGCAP, CINVEST) perform very well, confirming the importance of the q-factor style dimensions. Momentum and liquidity/trading variables (UMD, MOM6M/MOMS12M, IND MOM, ILL, LIQ_PS, BASPREAD, ZERO TRADE) are clearly useful but slightly below the very top. Working capital / balance-sheet composition factors (PCHSALEINV, DNCL, DNCO, SALECASH, SALEINV, SALEREC, QUICK, CURRAT) are surprisingly strong, often among the top-ranked factors. Accruals and distress/constraints factors (ACC, ABSACC, KZ, WW) appear moderate, not leading the ranking. Interestingly, to a certain extent, the high-performing risk factors from my analysis subsume the factors selected by Feng, Giglio, and Xiu (2020).

Second, comparing crisis-period means of the test statistics to their full-sample counterparts shows that most of the statistical power is concentrated in crises. For many factors, the average test statistic in crisis periods exceeds the full-sample average by several units. The market excess return (MktRf) has a crisis-period mean around 6 while its overall mean is slightly negative, and similar patterns arise for size (SMB), value (HML), investment (CMA), profitability (RMW), and many zoo factors, including investment-adjusted cash-flow and intermediary-capital measures. Group-wise, core priced factors from the modern asset-pricing canon (market, size, value, profitability, investment, momentum, and extensions such as BAB, QMJ, HML_Devil, and HXZ-style factors) show some of the largest and most significant crisis shifts, with crisis means well above their overall averages. Profitability, investment, working-capital, and valuation characteristics also become much more informative in crises, consistent with balance-sheet strength, investment policy, and operating efficiency mattering most when financial conditions are tight. Liquidity, volatility, and distress-related variables follow the same qualitative

pattern: illiquidity and volatility measures, turnover and zero-trade proxies, and financing-constraint indices (KZ, WW), as well as several accruals and distress-style signals, all exhibit stronger tests in crises than in normal times. Overall, financial crises appear as periods when risk premia, mispricing, and cross-sectional return spreads are amplified, making most well-known factors statistically more visible than in the rest of the sample.

Third, the comparison between the baseline model without PCA and specifications with PCA3 and PCA5/10 shows that principal-component adjustment substantially reshuffles factor rankings. In the no-PCA specification, headline factors such as BAB and illiquidity-type measures look crucial, but they become much less important once PCA absorbs broad common components. Conversely, a set of "second-tier" characteristics—including `nef`, `nop`, `saleinv`, `tang`, `ato`, beta, `op`, and `hire`—move into the upper ranks under PCA, suggesting that they carry priced information that was previously masked by collinearity and misspecification.

Rankings based on single-factor Sharpe ratios tell a qualitatively similar story when contrasted with pricing-based rankings: a small core of variables (RF, BAB, herf, HML, CMA, QMJ, roic, and, to some extent, ms and UMD) score well on both criteria, while many others display sharp trade-offs. For instance, some characteristics (dncl, cinvest, indmom, pchsaleinv, STR, and several investment-change measures) are excellent in terms of pricing tests but weak on Sharpe, whereas others (realestate_hxz, pm, and several microstructure and volatility proxies) look weak in pricing yet strong in Sharpe terms. These patterns highlight that factor "importance" depends critically on whether one evaluates pricing ability, portfolio performance, or both, and on how aggressively common components are removed.

I perform a second empirical exercise where I fit traditional multi-factor models, such as FF3, FF5, and FF5 plus momentum, and compare with latent multi-factor models estimated by PCA. In contrast to the previous empirical exercise, now I am assuming a multi-factor structure from the outset. Latent factor models appear more reactive to economic downturns (i.e., recessions), although their portfolio performance is comparable to the performance of the observed multi-factor models.

The paper proceeds as follows. Section 2 introduces the conditional asset pricing framework and embeds it in a factor asset pricing setting. Section 3 formalizes the nano-PCA estimator that is valid when $T = 1$. Section 4 develops tests of OOS pricing ability of the factor asset pricing model for $T = 1$, both in terms of risk premia and the SDF. It also presents the analysis of mean–variance portfolios and the corresponding performance measures in the $T = 1$ case. Section 5 generalizes the methodology to settings in which the one-factor model is misspecified, that is, when more than one risk factor is needed to span the cross-section of returns. My methodology is completed by the analysis of multi-factor models when they are assumed correctly specified, as exemplified in Section 6. Section 7 describes the empirical applications, which demonstrate the effectiveness of the proposed approach in assessing the ability of factor asset pricing models to price the cross-section over time. Section 8 concludes. The mathematical appendix states the technical assumptions and provides the proofs, with additional material relegated to the Online Appendix.

2 Asset Pricing Setting

I assume that x_{is}, the excess return on asset i at time s, satisfies a one-dimensional conditional factor model, in the sense that each x_{is} is a linear function of a single risk factor f_t and an idiosyncratic error

e_{it}:

$$x_{it} = \alpha_{it-1} + \lambda_{it-1} f_t + e_{it} \quad \text{for every } t \in \mathcal{Z} \text{ and } i \in \mathcal{N}, \tag{1}$$

where α_{it-1} is a time-varying intercept, γ_{t-1} is time-varying risk premium of the risk factor f_t satisfying $\mathcal{E}_{t-1}(f_t) = \gamma_{t-1}$, and $\mathcal{E}_{t-1}(\cdot)$ denotes the conditional expectation given all information available up to time $t-1$. The scalar coefficient λ_{it-1} is the time-varying risk exposure (i.e., the loading). The index sets are $\mathcal{N} = 1, 2, \ldots$ and $\mathcal{Z} = \ldots, -1, 0, 1, \ldots$. Finally, I assume that a risk-free asset with return r_{ft-1} is tradable for every $t \in \mathcal{Z}$.[3]

My results are initially developed for the case in which only one time observation, say t (i.e., $T = 1$), is available for estimation, relative to a generic time period $t \in \mathcal{Z}$, with one additional observation, say $t+1$, used for OOS evaluation (i.e., $T_{oos} = 1$). In this setting, I necessarily restrict (1) to a single risk factor and assume that the model is correctly specified.

Subsequently, I consider the case in which (1) is misspecified, namely when one still uses $T = 1$ for estimation but the cross-section of asset returns is actually driven by $r_t > 1$ risk factors. Note that the number of risk factors is allowed to vary over time. Although the model is estimated using one period ($T = 1$), I require at minimum $r_t + 1$ periods of data (i.e., $T + T + \text{oos} = r_t + 1$) to be able to make inference, i.e., construct the standard errors and the necessary bias-adjustment that mitigates the fixed T biases.

I will then analyze the correctly-specified multi-factor case with $r_t > 1$, in which the model from the outset allows for r_t risk factors. Estimation requires at least $T = r_t$ time periods, whereas inference requires at minimum $T + T_{\mathsf{oos}} = r_t + 1$ just like for the previous case of misspecification. Although in both cases of misspecified and correctly-specified models I use $r_t + 1$ periods of data, the latter case is more efficient, i.e., the test statistics will exhibit a smaller variance because it starts with the correctly-specified model but, of course, it might lead to incorrect inference should the posited model be wrong.

Therefore, I will present my methodology along these these three separate settings, namely, first, when the true and estimated model coincide with a one-factor model, second when the true model is multi-factor but one keeps estimating a one-factor model, and third when the true and estimated model coincide with a multi-factor model.

When $T = r_t = 1$, the conditional factor model (1) can, *without loss of generality*, be rewritten as a model with time-invariant parameters,

$$x_{it} = \alpha_i + \lambda_i f_t + e_{it},$$

since the time dimension plays no role for estimation in this case. I am initially assume the factor f_t to be *latent*, i.e., not observed. Therefore, I will initially develop the methodology under the latent factor specification and then show how it simplifies when the risk factor is observed, when for instance the risk factor is the excess return on a tradable portfolio (including a mimicking portfolio for a non-tradable risk factor).

The following assumption, which is uncommon in the statistical PCA literature but central in asset pricing, provides the structure on $\mathcal{E}_{t-1}(x_{it})$ needed for the methodology. Let $\text{var}_{t-1}(\cdot)$ denote the conditional variance given information up to $t-1$, and define the excess return on a portfolio p as $r^p_{N,t} \equiv \mathbf{w}^{p\prime}_{Nt-1} \mathbf{x}_{t.}$, where $\mathbf{w}^p_{Nt-1} = (w^p_{1t-1}, \ldots, w^p_{Nt-1})'$ are the portfolio weights and $\mathbf{x}_{t.} \equiv (x_{1t}, \ldots, x_{Nt})'$

[3]The case in which a risk-free asset is not tradable is analyzed in Section XXX.

is the N-vector of excess returns realized at the end of period t.

Assumption 1 (asymptotic no-arbitrage) *There is no sequence of portfolios p, along any subsequence N', such that $\text{var}_{t-1}(r^p_{N',t}) \to 0$ almost surely as $N' \to \infty$, while $\mathcal{E}_{t-1}(r^p_{N',t}) \geq C > 0$ almost surely for some $t \in \mathcal{Z}$ and some arbitrary positive constant C.*

Under Assumption 1, together with a limited degree of cross-sectional dependence in e_{is} and sufficient smoothness in the risk exposures $\boldsymbol{\lambda}_{is-1}$ (see Proposition OA.4 in Zaffaroni (2025)), the intercept terms α_{is-1} satisfy[4]

$$\sum_{i=1}^{\infty} \alpha^2_{is-1} < \infty \quad \text{a.s. for every } s \in \mathcal{T}_t, \tag{2}$$

where $\mathcal{T}_t \equiv \{t, t+1\}$ collects the two in-sample and OOS observations.

The square-summability condition (2) on the intercepts α_{is-1}, known in asset pricing as *pricing errors*, implies that I do *not* need to estimate the α_{is-1}. This allows me to avoid de-meaning the x_{is}. This is trivially *essential* when $T = 1$: de-meaning would yield $x_{it} - \bar{x}_i = x_{it} - x_{it} = 0$, since $\bar{x}_i \equiv T^{-1} \sum_{s=t}^{t} x_{is} = x_{it}$. Even when more than one period is used for estimation—as when both the IS and OOS observations are exploited so that $T + T_{oos} = 2$—de-meaning must be avoided in the latent factor case because it would compromise the asymptotic properties of the local-PCA estimator.[5]

Formulation (1), together with (2), represents a conditional version of the APT of Ross (1976).[6]

3 Nano-PCA: Principal Component Analysis When $T = 1$

I show how to implement PCA for model (1) when $T = 1$, which I refer to as *nano-PCA*, and I establish its asymptotic properties as $N \to \infty$. Because the analysis focuses on a single period, say period t, I do not need to impose *any* assumption on the time evolution of the parameters of the conditional asset pricing model (1). Likewise, since I work with a single risk factor, I do not need to estimate the number of factors.

However, I do require a smoothing assumption on the time behaviour of the variances and covariances of the idiosyncratic errors e_{is}, ensuring that they are only mildly temporally dependent for $s \in \{t, t+1\}$.

[4]Condition (2) is the main implication of the APT: $\mathcal{E}_{s-1}(x_{is})$ may differ from $\lambda_{is-1}\gamma_{s-1}$, but not too much in the sense of (2). In the unconditional setting, see Huberman (1982), Chamberlain (1983), and Ingersoll (1984), among others; for the conditional setting, see Stambaugh (1983) and Zaffaroni (2025)[Proposition OA.4]. In general, the APT bound involves the inverse of the conditional covariance matrix of e_{is} with elements $\sigma^{(-1)}_{ij,s-1,s-1}$ (see, e.g., Ingersoll (1984, Eq. (9))), that is, $\sum_{i,j=1}^{\infty} \alpha_{is-1}\alpha_{js-1}\sigma^{(-1)}_{ij,s-1,s-1} < \infty$ a.s. (almost surely); under our assumptions, this reduces to (2).

[5]For local-PCA to be valid, the time-series dependence of the data must be weak. De-meaning induces $e^*_{is} \equiv e_{is} - T^{*-1} \sum_{r=1}^{T^*} e_{ir}$, with $T^* = T + T_{oos} \geq 2$, and the e^*_{is} are serially correlated even if the e_{is} are temporally independent; see Zaffaroni (2025).

[6]The APT leads to (1) as follows. Start from $x_{is} = a_{is-1} + \lambda_{is-1}F_s + e_{is}$ for some initial factor F_s (e.g., a non-tradable risk factor) and time-varying intercept a_{is-1}. When $\mathcal{E}_{s-1}(F_s) = 0$ one has $a_{is-1} = \mathcal{E}_{s-1}(x_{is})$, although this is not required here. Under Assumption 1 (and suitable regularity conditions on e_{is} and λ_{is-1}; see Proposition OA.4 in Zaffaroni (2025)), there exists a constant γ_{s-1} (the risk premium) such that $\mathcal{E}_{s-1}(x_{is}) = \alpha_{is-1} + \lambda_{is-1}\gamma_{s-1}$, where the pricing errors α_{is-1} satisfy (2). Setting $f_s \equiv F_s + \gamma_{s-1} - \mathcal{E}_{s-1}(F_s)$ then yields (1). The risk premium satisfies $\gamma_{s-1} = -r_{fs-1}\big(\mathcal{E}_{s-1}(m_{s-1,s}F_s) - \mathcal{E}_{s-1}(F_s)/r_{fs-1}\big)$, as well as $\gamma_{s-1} = \mathcal{E}_{s-1}(f_s)$, where $m_{s-1,s}$ denotes the SDF associated with (1) (see Hansen and Richard (1987, Theorem 2.1)). Zaffaroni (2025) shows that f_s is asymptotically equivalent to the return on a mimicking portfolio as $N \to \infty$ (see Section OA5).

This builds on Connor and Korajczyk (1986) and Bai (2003), and rules out aggregate (i.e., cross-sectionally averaged) time-heteroskedasticity in the e_{is}.[7] In particular, I require that $N^{-1}\sum_{i=1}^{N} var(e_{is})$. converges, at an appropriate rate, to a locally constant limit σ_t^2 for each $s \in \{t, t+1\}$ (see Assumption 6). These and the additional regularity conditions needed for nano-PCA are stated in the final appendix (Assumptions 4–8), where I also distinguish clearly between the assumptions needed when $T = 1$ is used for estimation and the additional ones required when both periods $T + T_{oos} = 2$ are exploited for testing.

As discussed above, I estimate model (1) under a *locally* constant-parameter specification (w.l.o.g. because $T = 1$) and without intercepts (by Assumption 1), using the nano-PCA estimator for the risk exposures and the risk factor. This estimator is obtained by solving the following least-squares problem in $(N + 1)$ parameters:

$$\hat{f},\ \hat{\boldsymbol{\lambda}} = \text{argmin}_{f,\boldsymbol{\lambda}} \frac{1}{N}\sum_{i=1}^{N}\big(x_{it} - f\lambda_i\big)^2, \tag{3}$$

where $\boldsymbol{\lambda} \equiv (\lambda_1, \ldots, \lambda_N)'$. For notational simplicity, I omit the time index for estimated quantities and retain it only for the corresponding population objects.

The solution to the nano-PCA problem (3), together with its asymptotic properties as $N \to \infty$, is derived below. Throughout, $\to_p$ and $\to_d$ denote convergence in probability and in distribution, respectively. All limits in this paper are taken as $N \to \infty$ with T fixed.

Theorem 1 (Nano-PCA: Solution and Asymptotics)

(i) The solution to (3) *is given by* $\hat{f} = \pm 1$*, corresponding to the (scalar) eigenvector of* $\mathbf{x}_t'\mathbf{x}_t/N$ *under the normalization* $\hat{f}^2 = 1$*, and*

$$\hat{\lambda}_i = x_{it},$$

which coincides with the OLS estimate from regressing x_{it} *on* $\hat{f}$ *without an intercept, i.e.,* $\hat{\lambda}_i = \frac{x_{it}\hat{f}}{\hat{f}^2}$*, so that, without loss of generality, we set* $\hat{f} = 1$*.*

When Assumption 4 holds for the single period t*:*

(ii) The risk factor estimator $\hat{f}$ *is super-consistent:*

$$\hat{f} = H_t f_t,$$

where the rotation scalar is given by $H_t = f_t^{-1}$*.*

(iii) When Assumptions 1 and 4–7 hold for the single period t*, the risk exposure estimator is inconsistent:*

$$\hat{\lambda}_i = x_{it} = \lambda_{it-1}H_t^{-1} + \alpha_{it-1} + e_{it},$$

[7] My assumption allows for time-heteroskedasticity in e_{is} that evolves independently across stocks (or with mild deviations from independence), but rules out a common component in the second-moment dynamics of e_{is}, except when that component changes very slowly (e.g., under extremely persistent dynamics). Since $T + T_{oos} = 2$, the assumption of constant error variance (when averaged across assets) over the sample window is mild if the stochastic volatility of e_{is} is highly persistent (i.e., smooth) and if volatility jumps are excluded.

and satisfies

$$\hat{\mu}_\lambda \equiv \frac{\mathbf{1}_N'\hat{\boldsymbol{\lambda}}}{N} \;\to_p\; \mu_{\lambda_t} H_t^{-1},$$

$$\hat{\sigma}_\lambda \equiv \frac{\hat{\boldsymbol{\lambda}}'\hat{\boldsymbol{\lambda}}}{N} \;\to_p\; H_t^{-1}\sigma_{\lambda_t} H_t^{-1} + \sigma_t^2,$$

where μ_{λ_t}, σ_{λ_t}, and σ_t^2 are defined in Assumptions 5 and 6.

Remark. Result (i) is entirely driven by the rotation scalar, which is fully pinned down by the eigenvector normalization $\hat{f}'\hat{f}/T = \hat{f}^2 = 1$. This forces $\hat{f} = \pm 1$, and we take $\hat{f} > 0$ w.l.o.g., as discussed above.[8] In the multifactor case with $r > 1$, the rotation matrix is generally a function of *averages* over the $T \geq r$ time observations (see Zaffaroni (2025, Theorem 1)). When $T = 1$, these time averages *coincide* with the single available observation. As a result, the rotation scalar completely absorbs the latent factor, eliminating any time variation in it. This phenomenon occurs *only* in the knife-edge case $T = r = 1$.

Remark. Since $\hat{\lambda}_i\hat{f} = x_{it}$, the nano-PCA estimator achieves perfect in-sample fit, i.e.,

$$\hat{e}_{it} \equiv x_{it} - \hat{\lambda}_i\hat{f} = 0,$$

leading to $\sum_{i=1}^N \left(x_{it} - \hat{\lambda}_i\hat{f}\right)^2 = 0$, exactly as in linear regression when the number of parameters equals the number of observations. As a consequence, σ_t^2 (the cross-sectional average of $var(e_{it})$) is not identified and cannot be estimated when $T = 1$. Estimating σ_t^2 is not required for constructing expected returns, the SDF, or mean–variance portfolios, in particular their large-N approximations. However, when constructing test statistics, a consistent estimate of σ_t^2 is needed. By exploiting $T + T_{oos} = 2$ periods, I will be able to estimate σ_t^2 consistently.

Remark. When $T = 1$, the nano-PCA estimator of the risk factor has no sampling variability once the rotation scalar is taken into account; it is effectively infinitely precise, so standard errors and confidence intervals for the factor itself are irrelevant. In contrast, the nano-PCA estimator of the risk exposures remains noisy, although it is still possible to estimate accurately the first and second moments of λ_{it-1} when $T = 1$.[9] While the first moment of λ_{it-1} (up to rotation) can be estimated consistently, the second-moment estimator is upward biased due to the presence of the σ_t^2 term. This bias can be corrected once a consistent estimator of σ_t^2 is available, as shown in the next section.

Remark. When testing the OOS pricing ability and mean–variance OOS portfolio performance of the asset pricing model, I will exploit data over two periods: one for model estimation (say, period t) and one for OOS evaluation (say, period $t + 1$). Thus, I will use the $T + T_{oos} = 2$ observations jointly to re-estimate model (1) for the purpose of constructing asymptotically valid test statistics. This in turn requires bias adjustments based on consistent estimators of σ_t^2 and on the two consecutive realizations

[8] In the general multifactor case with $r \geq 1$, let $\hat{\mathbf{F}}$ denote the $T \times r$ matrix of eigenvectors of $\mathbf{X}\mathbf{X}'/(TN)$ associated with the r largest eigenvalues, where $\mathbf{X} \equiv (\mathbf{x}_{\cdot 1} \cdots \mathbf{x}_{\cdot N}) \equiv (\mathbf{x}_{t-T+1.} \cdots \mathbf{x}_{s.} \cdots \mathbf{x}_{t.})'$, $\mathbf{x}_{\cdot i} \equiv (x_{i,t-T+1} \cdots x_{is} \cdots x_{it})'$, and $\mathbf{x}_{s.}$ is defined as above. The normalization $\hat{\mathbf{F}}'\hat{\mathbf{F}}/T = \mathbf{I}_r$ imposes $r(r+1)/2$ constraints. The rotation matrix has r^2 free entries, but when $r = 1$ we have $r(r+1)/2 = r^2 = 1$, so the rotation matrix is fully determined (although still unobserved). Note that $r^2 > r(r+1)/2$ whenever $r \geq 2$.

[9] The lack of consistency of the risk exposures is *not* driven by the factor being latent (except for the presence of the rotation scalar): it remains an issue even when f_t is an observed risk factor.

of the risk factor (say, f_t and f_{t+1}).[10] As explained below, I rely on the limiting distribution theory for local-PCA in Zaffaroni (2025), which is valid for any fixed $T > r_t$, where r_t denotes the number of risk factors.

4 Cross Sectional Asset Pricing When $T = 1$ - Correct Specification

Cross sectional asset pricing is, generally speaking, about understanding how different assets earn different risk premia (see Zaffaroni and Zhou (2024)). Methodologically, it is about developing methodologies for estimating, testing, and using asset pricing models when a *large cross-section* of financial data, such as returns and firms' characteristics, is available, typically in the form of panel data. I am specifically focusing on methodologies for *conditonal* asset pricing. Although the range of applications is wide, I am going to focus on three problems that attracted the most interest, namely (i) estimation of risk premia and testing their pricing ability to fit the cross-section, (ii) mean-variance portfolios construction and their performance evaluation (i.e., Sharpe ratios), and (iii) estimation of the SDF, and testing its pricing ability.

In my case, I am going to construct methodologies using a single, possibly very large, cross-section of data, i.e., N large and T equal to one. However, for evaluating the OOS performance, I will necessarily use two observations (one for IS, one for OOS). Having estimated the asset pricing model with a $T = 1$ time observation, that is estimating expected returns, mean-variance portfolios, and the SDF, I will then use both observations to construct the OOS pricing tests and portfolio's Sharpe ratio, and the corresponding bias-adjustments and standard errors.

4.1 Pricing Ability of Risk Premia When $T = 1$ - Correct Specification

I now explain how to test the ability of the one-factor asset pricing model (1) to fit the cross-section of expected returns, i.e., the pricing ability of the risk premium associated with the factor f_t, distinguishing between IS and OOS performance.

I define the IS pricing errors from expected returns as

$$\hat{h}_{it} \equiv x_{it} - \hat{\lambda}_i \hat{f},$$

which, by Theorem 1(i), are identically zero, since $\hat{h}_{it} = x_{it} - x_{it} = 0$. Therefore, when $T = 1$, IS pricing performance cannot be assessed.[11] In contrast, the corresponding OOS pricing errors are

$$\begin{aligned} \hat{h}_{i,t+1|t} &\equiv x_{i,t+1} - \hat{\lambda}_i \hat{f} = x_{i,t+1} - x_{it} \\ &= \lambda_{it-1}(f_{t+1} - f_t) + e_{i,t+1} - e_{it} + \ell_{i,t+1} + \alpha_{it} - \alpha_{it-1}, \end{aligned} \tag{4}$$

where, w.l.o.g., I take t as the estimation period and $t + 1$ as the OOS period. Expression (4) follows from substituting (1) and the nano-PCA estimates into $\hat{h}_{i,t+1|t}$. I also impose *local* constancy of the risk exposures across the two periods t and $t + 1$, justified by Assumption 2 below. This assumption rules out erratic jump-type behavior and motivates the approximation term $\ell_{is} \equiv (\lambda_{is-1} - \lambda_{it-1})f_s$, which satisfies $\sum_{i=1}^{N} \ell_{is-1} = o(N^{1/2})$. I allow for time variation in the intercepts, $\alpha_{it} \neq \alpha_{it-1}$, because, by (2), their contribution vanishes asymptotically.

[10] Since f_t and f_{t+1} are now estimated jointly, care is needed in handling the rotation scalar, which will be common across the two estimates, unlike in the $T = 1$ case.

[11] Note that $\hat{h}_{it} = \hat{e}_{it}$, the model residuals, which were shown to be identically zero.

To test the OOS pricing ability of the factor model in the cross-section of expected returns, I consider a test statistic based on the cross-sectional average of the OOS pricing errors $\hat{h}_{i,t+1|t}$, denoted by $\tilde{L}_{t+1|t}$.[12]

Because of (4), which in turn reflects the $T = 1$ structure, a bias arises: the cross-sectional average of OOS pricing errors does not converge to zero,

$$\frac{1}{N}\sum_{i=1}^{N}\hat{h}_{i,t+1|t} \;\to_p\; \mu_{\lambda_t}(f_{t+1} - f_t),$$

even when the factor model (1) is correctly specified and exactly prices the cross-section. This contrasts with the large-T case, where such bias disappears as $T \to \infty$, but where conditional models are more prone to misspecification.

I therefore need to estimate the bias term $\mu_{\lambda_t}(f_{t+1} - f_t)$ and the standard errors of the *bias-adjusted* pricing test statistic. To this end, I re-estimate the conditional factor model (1) using the local-PCA method of Zaffaroni (2025) with $T + T_{oos} = 2$ observations. The corresponding risk exposures and factor realizations are obtained by solving the least-squares problem[13]

$$\tilde{f}_t,\ \tilde{f}_{t+1},\ \tilde{\boldsymbol{\lambda}} = \text{argmin}_{f_t, f_{t+1}, \boldsymbol{\lambda}} \frac{1}{2N}\sum_{s=t}^{t+1}\sum_{i=1}^{N}\big(x_{is} - f_s\lambda_i\big)^2, \tag{5}$$

where $\mathbf{X} \equiv (\mathbf{x}_{\cdot 1} \cdots \mathbf{x}_{\cdot N}) \equiv (\mathbf{x}_{t\cdot}\ \mathbf{x}_{t+1\cdot})'$, with $\mathbf{x}_{\cdot i} \equiv (x_{it}, x_{i,t+1})'$. [14]

As is well known, $\tilde{\mathbf{f}} \equiv (\tilde{f}_t\ \tilde{f}_{t+1})'$ equals $\sqrt{2}$ times the eigenvector associated with the largest eigenvalue of $\mathbf{X}\mathbf{X}'/(2N)$, and hence satisfies $\tilde{\mathbf{f}}'\tilde{\mathbf{f}}/2 = 1$. The factor loadings are given by

$$\tilde{\boldsymbol{\lambda}} \equiv \mathbf{X}'\tilde{\mathbf{f}}(\tilde{\mathbf{f}}'\tilde{\mathbf{f}})^{-1} = \frac{\mathbf{X}'\tilde{\mathbf{f}}}{2}.$$

In what follows, I use a hat ($\hat{\cdot}$) for nano-PCA estimates based on one observation ($T = 1$), and a tilde ($\tilde{\cdot}$) for local-PCA estimates based on more than one observation ($T + T_{oos} \geq 2$).

Estimation of (1) via local-PCA with two time observations requires additional technical conditions, listed in the final appendix (Section 9). In particular, I need the risk exposures to be *locally* constant across the two periods.

Hereafter, for a generic matrix $\mathbf{A}$, let $\|\mathbf{A}\| = [\text{tr}(\mathbf{A}'\mathbf{A})]^{1/2}$ denote the Frobenius norm. For deterministic sequences $\{a_N\}$ and $\{b_N\}$, the notation $a_N = O(b_N)$ and $a_N = o(b_N)$ means that $|a_N|/b_N < C$ (for some constant $C > 0$) and that $|a_N|/b_N \to 0$ as $N \to \infty$, respectively; the symbols $O_p(\cdot)$ and $o_p(\cdot)$ have the analogous meanings in probability.

Assumption 2 (Risk exposures – local smoothness)

$$\|\boldsymbol{\lambda}_{t-1} - \boldsymbol{\lambda}_t\|^2 = o(N^{1/2}).$$

For example, one may have $\boldsymbol{\lambda}_{s-1} = \boldsymbol{\lambda}_0 + \boldsymbol{\Lambda}_1\mathbf{z}_{s-1}$ for some parameter vector $\boldsymbol{\lambda}_0$ and parameter matrix $\boldsymbol{\Lambda}_1$, with $\boldsymbol{\Lambda}_1'\boldsymbol{\Lambda}_1 = o(N^{1/2})$, where $\mathbf{z}_{s-1}$ is a vector of observed instruments. No restriction is imposed

[12] Alternatively, one may construct a related test based on the cross-sectional average of the *squared* OOS pricing errors $\hat{h}^2_{i,t+1|t}$. Details are available upon request.

[13] See Magnus and Neudecker (2007, Chapter 17, Theorem 7) for an elegant proof. An asymptotically equivalent PCA estimator can be obtained by taking $\hat{\boldsymbol{\lambda}}$ as the eigenvector of $\mathbf{X}'\mathbf{X}/(2N)$ corresponding to its largest eigenvalue, multiplied by $\sqrt{N}$, and setting $\hat{\mathbf{f}} = \mathbf{X}\hat{\boldsymbol{\lambda}}/N$; see Bai and Ng (2002). Since N is typically much larger than $T = 2$, the approach based on $\mathbf{X}\mathbf{X}'$ is computationally more convenient. Recall that the nonzero eigenvalues of $\mathbf{X}'\mathbf{X}$ and $\mathbf{X}\mathbf{X}'$ coincide.

[14] With a slight abuse of notation, I use the same symbols for the feasible values in (5) as for the true factor realizations in the data-generating process (1), to avoid unnecessary proliferation of notation.

on $\boldsymbol{\lambda}_0$. The time variation allowed by this specification accommodates, for instance, the case in which $\mathbf{z}_{s-1}$ follows a (possibly multidimensional) random walk. A further smoothness condition on the second moments of e_{is}, analogous to the one used for nano-PCA, is also required; this and other regularity conditions for local-PCA are collected in the final appendix. Since the time window of length 2 is very short, these smoothness assumptions are mild.

The next theorem defines the bias-adjusted test statistic and characterizes its behavior under the null of correct specification and under the alternative.

Theorem 2 (Cross-Sectional Pricing Test – OOS, Correct Specification, $T = 1$**)**
Define

$$\tilde{L}_{t+1|t} \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,t+1|t} \; - \; \tilde{\mu}_\lambda(\tilde{f}_{t+1} - \tilde{f}_t), \quad \textit{with} \quad \tilde{\mu}_\lambda \equiv \frac{\mathbf{1}_N'\tilde{\boldsymbol{\lambda}}}{N}.$$

(i) Under Assumptions 1 and 2–8,

$$\sqrt{N}\,\tilde{L}_{t+1|t} \;\to_d\; N\big(0,\; \sigma_t^2(\boldsymbol{\iota}_{-1,1}'\mathbf{M}_f\boldsymbol{\iota}_{-1,1})\big(1 - \mu_{\lambda_t}^2/\sigma_{\lambda_t}\big) + \mathbf{A}_{L1}'\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}\big)\,,$$

where

$$\mathbf{A}_{L1} \equiv \left(-\mu_{\lambda_t}U_t^{-1}\frac{\mathbf{f}}{2} \otimes \mathbf{M}_f\boldsymbol{\iota}_{-1,1}\right),$$

with $\mathbf{f} \equiv (f_t, f_{t+1})'$, $\hat{\sigma}_f \equiv \mathbf{f}'\mathbf{f}/2$, $\boldsymbol{\iota}_{-1,1} \equiv (-1,1)'$, $\mathbf{M}_f \equiv \mathbf{I}_2 - \mathbf{f}\mathbf{f}'/(\mathbf{f}'\mathbf{f})$, $\mathbf{U}_{\epsilon,t}$ *being the* $(T+1)^2 \times (T+1)^2$ *matrix with* $T = 1$ *defined in Section 10, and* $U_t \equiv \sigma_{\lambda_t}\hat{\sigma}_f$.
(ii) Under the same assumptions, and imposing $\kappa_{4t} = 0$, *define* $\mathbf{M}_{\tilde{f}} \equiv \mathbf{I}_2 - \tilde{\mathbf{f}}\tilde{\mathbf{f}}'/2$, $\tilde{\mathbf{e}} \equiv \mathbf{X} - \tilde{\mathbf{f}}\tilde{\boldsymbol{\lambda}}'$, *and let* $\tilde{U}$ *be the largest eigenvalue of* $\mathbf{X}\mathbf{X}'/(2N) - (\tilde{\sigma}^2/2)\mathbf{I}_2$. *Then*

$$\tilde{\mathbf{A}}_{L1} \equiv \left(-\tilde{\mu}_\lambda\tilde{U}^{-1}\frac{\tilde{\mathbf{f}}}{2} \otimes \mathbf{M}_{\tilde{f}}\boldsymbol{\iota}_{-1,1}\right) \;\to_p\; \mathbf{A}_{L1},$$

$$\tilde{\sigma}^2 \equiv \mathrm{tr}(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \;\to_p\; \sigma_t^2, \quad \tilde{\mathbf{U}}_\epsilon \;\to_p\; \mathbf{U}_{\epsilon,t},$$

$$\tilde{\sigma}_\lambda \equiv \tilde{\boldsymbol{\lambda}}'\tilde{\boldsymbol{\lambda}}/N - \tilde{\sigma}^2/2 \;\to_p\; H_t^{-2}\sigma_{\lambda_t}, \quad \tilde{\mu}_\lambda \;\to_p\; H_t^{-1}\mu_{\lambda_t},$$

where $\tilde{\mathbf{U}}_\epsilon$ *is obtained by replacing* σ_t^2 *and* κ_{4t} *with* $\tilde{\sigma}^2$ *and* 0, *respectively, in* $\mathbf{U}_{\epsilon,t}$, *and*

$$H_t \equiv \sigma_{\lambda_t}^{1/2}U_t^{-1/2} = \hat{\sigma}_f^{-1/2}.$$

(iii) When (1) *does not hold and* $|\tilde{L}_{t+1|t}| \geq CN^{-1/2+c}$ *a.s. for some* $C, c > 0$, *we have*

$$\sqrt{N}\,\tilde{L}_{t+1|t} \;\to_p\; \pm\infty.$$

Remark. There are many alternative ways to construct feasible pricing tests. For instance, one may base a test on the *squared* pricing errors, satisfying,

$$\frac{\sum_{i=1}^N \hat{h}_{t+1|t}^2}{N} \to_p (f_{t+1} - f_t)^2\sigma_{\lambda_t} + 2\sigma_t^2,$$

suggesting the (bias-adjusted) statistic

$$\tilde{L}_{2,t+1|t} \equiv \frac{\sum_{i=1}^N \hat{h}_{t+1|t}^2}{N} - (\tilde{f}_{t+1} - \tilde{f}_t)^2\tilde{\sigma}_\lambda - 2\tilde{\sigma}^2.$$

Using arguments analogous to those employed for $\tilde{L}_{t+1|t}$, one can derive the asymptotic distribution of $\sqrt{N}\,\tilde{L}_{2,t+1|t}$.

4.1.1 Pricing Ability of Risk Premia When $T = 1$ - Correct Specification - Observed Risk Factor

.

An important, yet specific, case of the pricing test $\tilde{L}_{t+1|t}$ arises when an *observed* risk factor is used in model (1). I denote this factor by f_t^o, to distinguish it from the latent factor f_t.[15] This setting covers, for example, the case in which an analyst wishes to test whether a given portfolio return, such as the market excess return, prices the cross-section of expected returns.

In this case, the pricing errors $\hat{h}_{it}$ are *identical* to those in the latent factor case. The difference arises only in the construction of the bias adjustment and the standard errors. The observed-factor case can be derived from the latent-factor analysis by treating the factor as non-random from the sampling perspective (i.e., with no estimation error) and by setting the rotation matrix equal to one. To make this clear, I reconstruct the pricing errors and the test statistic, and then derive its limiting distribution.

Since f_t^o is observed in (1), the estimated risk exposures for $T = 1$ are the usual OLS estimators ("covariance over variance") computed on a single time observation:

$$\hat{\lambda}_i^o \equiv \frac{x_{it} f_t^o}{(f_t^o)^2} = \frac{x_{it}}{f_t^o}.$$

The model-implied expected return is then

$$\hat{\lambda}_i^o f_t^o = x_{it},$$

which is *identical* to the nano-PCA fitted value in (4), and also coincides with the time-series sample mean of x_{it} when only one observation is available. Hence the OOS pricing errors are exactly the same as in the latent factor case:

$$\hat{h}_{i,t+1|t}^o \equiv x_{i,t+1} - \hat{\lambda}_i^o f_t^o = x_{i,t+1} - x_{it} = \lambda_{it-1}(f_{t+1} - f_t) + e_{i,t+1} - e_{it} + \ell_{it},$$

so that $\hat{h}_{i,t+1|t}^o = \hat{h}_{i,t+1|t}$. As in the nano-PCA case, $\hat{\lambda}_i^o$ is a noisy estimator of the true risk exposure:

$$\hat{\lambda}_i^o = \frac{x_{it}}{f_t^o} = \lambda_{it-1} + \frac{\alpha_{it-1}}{f_t^o} + \frac{e_{it}}{f_t^o}.$$

Relative to nano-PCA, there is now no rotation term, and the observed factor f_t^o appears in the denominator because it is not constrained to equal 1.

The main difference with the latent factor case lies in the asymptotics of the pricing test statistic

$$\tilde{L}_{t+1|t}^o \equiv \frac{\mathbf{1}_N' \mathbf{h}_{t+1|t}^o}{N} - \tilde{\mu}_\lambda^o (f_{t+1}^o - f_t^o),$$

where the two-period estimator of $\mu_{\lambda_t}^o$ is

$$\tilde{\mu}_\lambda^o \equiv \frac{\mathbf{1}_N' \tilde{\boldsymbol{\lambda}}^o}{N},$$

with

$$\tilde{\boldsymbol{\lambda}}^o \equiv \frac{\sum_{s=t}^{t+1} \mathbf{x}_s f_s^o}{\sum_{s=t}^{t+1} (f_s^o)^2} = \boldsymbol{\lambda}_{t-1} + \frac{\sum_{s=t}^{t+1} \boldsymbol{\alpha}_{s-1} f_s^o}{\sum_{s=t}^{t+1} (f_s^o)^2} + \frac{\sum_{s=t}^{t+1} \mathbf{e}_s f_s^o}{\sum_{s=t}^{t+1} (f_s^o)^2} + \frac{\sum_{s=t}^{t+1} \ell_s f_s^o}{\sum_{s=t}^{t+1} (f_s^o)^2}. \tag{6}$$

It follows that

$$\frac{\mathbf{1}_N' \boldsymbol{\lambda}_{t-1}}{N} - \tilde{\mu}_\lambda^o = -\frac{\sum_{s=t}^{t+1} \bar{e}_s f_s^o}{\sum_{s=t}^{t+1} (f_s^o)^2} + o_p(N^{-1/2}),$$

[15] I use the superscript ^{o} for all quantities associated with the observed risk factor.

where $\bar{e}_s \equiv N^{-1}\sum_{i=1}^{N} e_{is}$. Substituting this expression into $\tilde{L}^o_{t+1|t}$ and rearranging yields

$$\tilde{L}^o_{t+1|t} = \boldsymbol{\iota}'_{-1,1}\mathbf{M}_{f^o}\mathbf{e}'\mathbf{1}_N/N + o_p(N^{-1/2}),$$

where $\mathbf{e} \equiv (\mathbf{e}_t, \mathbf{e}_{t+1})'$,

$$\boldsymbol{\iota}_{-1,1} \equiv (-1,1)', \quad \mathbf{f}^o \equiv (f^o_t, f^o_{t+1})', \quad \mathbf{M}_{f^o} \equiv \mathbf{I}_2 - \frac{\mathbf{f}^o\mathbf{f}^{o\prime}}{\mathbf{f}^{o\prime}\mathbf{f}^o}.$$

Therefore,

$$\sqrt{N}\,\tilde{L}^o_{t+1|t} \;\to_d\; N\big(0,\ \sigma_t^2\,\boldsymbol{\iota}'_{-1,1}\mathbf{M}_{f^o}\boldsymbol{\iota}_{-1,1}\big)\,.$$

The asymptotic variance

$$\sigma_t^2\,\boldsymbol{\iota}'_{-1,1}\mathbf{M}_{f^o}\boldsymbol{\iota}_{-1,1} = \sigma_t^2\left(2 - \frac{(f^o_{t+1} - f^o_t)^2}{\mathbf{f}^{o\prime}\mathbf{f}^o}\right)$$

coincides exactly with the variance in Theorem 2 once the sampling variability of the local-PCA factor estimate is switched off (see Appendix XXX for a proof). To construct the standard error for $\sqrt{N}\,\tilde{L}^o_{t+1|t}$, it suffices to estimate σ_t^2 by

$$\tilde{\sigma}_o^2 \equiv \frac{\mathrm{tr}(\tilde{\mathbf{e}}^o\tilde{\mathbf{e}}^{o\prime})}{N},$$

where $\tilde{\mathbf{e}}^o \equiv \mathbf{X} - \mathbf{f}^o\tilde{\boldsymbol{\lambda}}^{o\prime}$, and $\tilde{\sigma}_o^2 \to_p \sigma_t^2$ (see Appendix XXX).

Remark. One may wish to evaluate OOS performance using non-consecutive periods, e.g., using period t for estimation and period s for OOS analysis, with $s > t+1$, as in multi-step prediction settings.

Remark. In empirical applications, my test statistics can be applied to a *sequence* of short (possibly unbalanced) panels of dimension $2 \times N_t$, indexed by $t = 2, \ldots, T_0$, where T_0 is the overall time span of the data. This allows one to assess the ability of model (1) to price the cross-section *dynamically* over time. The two observations in each panel need not be consecutive. Thus one can compute $\hat{L}_{t|t-1}$ and $\hat{L}^o_{t|t-1}$ for any $2 \le t \le T_0$, where the quantities σ_t^2, μ_{λ_t}, and σ_{λ_t} will, in general, vary across panels. My empirical application illustrates the flexibility and usefulness of this *dynamic* approach: for example, one can revisit the "factor zoo" and examine which observed factors price the cross-section more frequently, or which are more effective during economic expansions versus recessions, and also compare latent and observed factor specifications. These issues are explored in detail in the empirical section.

Remark. Because the factor is observed, alternative bias adjustments are possible. For instance, one can use nano-PCA-based estimates of the risk exposures obtained from either period t or period $t+1$ to construct

$$\hat{L}^{o(t)}_{t+1|t} \equiv \frac{\mathbf{1}'_N\mathbf{h}^o_{t+1|t}}{N} - \hat{\mu}^{o(t)}_\lambda(f^o_{t+1} - f^o_t), \quad \hat{L}^{o(t+1)}_{t+1|t} \equiv \frac{\mathbf{1}'_N\mathbf{h}^o_{t+1|t}}{N} - \hat{\mu}^{o(t+1)}_\lambda(f^o_{t+1} - f^o_t),$$

where the superscript on $\hat{\mu}^o_\lambda$ indicates whether period t or $t+1$ is used for estimation. Under correct specification of (1), the limiting distributions are

$$\sqrt{N}\,\hat{L}^{o(t+1)}_{t+1|t} \;\to_d\; N\Big(0,\ \sigma_t^2\Big(1 + (f^o_{t+1}/f^o_t)^2\Big)\Big), \quad \sqrt{N}\,\hat{L}^{o(t)}_{t+1|t} \;\to_d\; N\Big(0,\ \sigma_t^2\Big(1 + (f^o_t/f^o_{t+1})^2\Big)\Big)\,.$$

4.2 Mean-Variance Portfolios When $T = 1$ - Correct Specification

An investor who posits model (1) and wishes to construct the mean–variance efficient portfolio must estimate the N-vector of (conditional) portfolio weights

$$\mathbf{w}_t = \frac{1}{\delta}\,\mathbf{V}_t^{-1}\boldsymbol{\lambda}_t\gamma_t,$$

where

$$\mathbf{V}_t = \sigma_{t,f}^2 \boldsymbol{\lambda}_t \boldsymbol{\lambda}_t' + \sigma_t^2 \mathbf{I}_N$$

is the (conditional) covariance matrix of the excess returns in (1), $\boldsymbol{\lambda}_t = (\lambda_{1t}, \ldots, \lambda_{Nt})'$, $\mathbf{I}_N$ is the $N \times N$ identity matrix, γ_t and $\sigma_{t,f}^2$ denote the conditional risk premium (equal to the conditional mean; see Footnote 6) and the conditional variance of the risk factor, respectively, and δ is the (constant) risk-aversion parameter.

Estimating $\mathbf{w}_t$ is challenging when $T = 1$ because both $\sigma_{t,f}^2$ and σ_t^2 must be estimated.[16] While $\sigma_{t,f}^2$ can be estimated by the sample second moment f_t^2 (since $T = 1$), this approach fails for σ_t^2 because the model achieves perfect in-sample fit when $T = 1$, implying residuals identically equal to zero. A way forward is to work with the large-N approximation of $\mathbf{w}_t$. Using the Sherman–Morrison formula,

$$\mathbf{w}_t = \frac{\boldsymbol{\lambda}_t \gamma_t}{\delta \sigma_t^2} \left(\frac{\sigma_{t,f}^{-2}}{\sigma_{t,f}^{-2} + \boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t / \sigma_t^2} \right) = \frac{\gamma_t}{\delta \sigma_{t,f}^2} \left(\frac{\boldsymbol{\lambda}_t}{\boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t} \right) + o(1), \tag{7}$$

where $o(1)$ denotes a remainder term that vanishes in the evaluation of $\mathbf{w}_t' \mathbf{x}_{t+1}$ as $N \to \infty$.[17] Inspecting (7), we see that σ_t^2 cancels out: the portfolio weights $\mathbf{w}_t$ become *functionally independent* of σ_t^2. Hence, estimation of σ_t^2 is not required for constructing the portfolio weights.

This leads to the nano-PCA estimator of $\mathbf{w}_t$:

$$\hat{\mathbf{w}} \equiv \frac{\hat{f}}{\delta \hat{f}^2} \left(\frac{\hat{\boldsymbol{\lambda}}}{\hat{\boldsymbol{\lambda}}' \hat{\boldsymbol{\lambda}}} \right) = \frac{1}{\delta} \left(\frac{\mathbf{x}_t}{\mathbf{x}_t' \mathbf{x}_t} \right), \tag{8}$$

where I used $\hat{f} = 1$ and $\hat{\boldsymbol{\lambda}} = \mathbf{x}_t$ from Theorem 1(i).

I can then define the corresponding portfolio returns, both in-sample and out-of-sample:

$$r_{t|t} \equiv \hat{\mathbf{w}}' \mathbf{x}_t, \qquad r_{t+1|t} \equiv \hat{\mathbf{w}}' \mathbf{x}_{t+1}.$$

To understand the difference between IS and OOS performance, I exploit the large cross-section and derive the limiting behavior of these portfolio returns as $N \to \infty$. Moreover, using both periods t and $t+1$, I can evaluate the Sharpe ratio of the portfolio (8), as I can now estimate its volatility, as well as f_t and f_{t+1}, via local-PCA with two observations:

$$\widehat{SR}_{t,t+1} \equiv \frac{\bar{\mathbf{x}}' \hat{\mathbf{w}}}{\sqrt{\hat{\mathbf{w}}' \mathbf{S} \hat{\mathbf{w}}}}, \tag{9}$$

where

$$\bar{\mathbf{x}} \equiv \frac{1}{2}\big(\mathbf{x}_t + \mathbf{x}_{t+1}\big), \quad \mathbf{S} \equiv \frac{1}{2}\Big[(\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_t - \bar{\mathbf{x}})' + (\mathbf{x}_{t+1} - \bar{\mathbf{x}})(\mathbf{x}_{t+1} - \bar{\mathbf{x}})'\Big].$$

Theorem 3 (Mean Variance Portfolio - IS and OOS - Correct Specification: $T = 1$) *When Assumptions 1 and 4-7 hold:*
(i)

$$r_{t|t} = \frac{1}{\delta} \frac{(\mathbf{x}_t' \mathbf{x}_t)}{(\mathbf{x}_t' \mathbf{x}_t)} = \frac{1}{\delta}. \tag{10}$$

$$r_{t+1|t} = \frac{1}{\delta} \frac{(\mathbf{x}_{t+1}' \mathbf{x}_t)}{(\mathbf{x}_t' \mathbf{x}_t)} \to_p \frac{1}{\delta} \frac{(\sigma_{\lambda_t} f_t f_{t+1})}{(\sigma_{\lambda_t} f_t^2 + \sigma_t^2)}. \tag{11}$$

[16] In my finite-T setting, the sample moments of the factor, computed from its true realizations, coincide with population moments.

[17] Expression (7) relies on f_t being a *strong* factor for the cross-section, i.e., $\boldsymbol{\lambda}_t' \boldsymbol{\lambda}_t / N \to \sigma_{\lambda_t} > 0$ (see Assumption 4).

(ii)

$$\widehat{SR}_{t,t+1} = \frac{(\mathbf{x}'_{t+1}\mathbf{x}_t) + (\mathbf{x}'_t\mathbf{x}_t)}{\sqrt{((\mathbf{x}'_{t+1}\mathbf{x}_t) - (\mathbf{x}'_t\mathbf{x}_t))^2}} + o_p(1) \to_p SR_{t,t+1} = \frac{(1 + \frac{(\sigma_{\lambda_t} f_t f_{t+1})}{(\sigma_{\lambda_t} f_t^2 + \sigma_t^2)})}{|(1 - \frac{(\sigma_{\lambda_t} f_t f_{t+1})}{(\sigma_{\lambda_t} f_t^2 + \sigma_t^2)}|}. \tag{12}$$

(iii) Under Assumptions 1 and 2-8 with $\kappa_{4t} = 0$, for any $0 < \alpha < 1/2$, setting $z_{\alpha/2}$ the (absolute value of the) $\alpha/2$-th quantile of the standard normal

$$Pr(\widehat{SR}_{t,t+1} - z_{\alpha/2} SE_{SR} \leq SR_{t,t+1} \leq \widehat{SR}_{t,t+1} + z_{\alpha/2} SE_{SR}) \to 1 - \alpha, \tag{13}$$

setting $SE^2_{SR} \equiv \frac{1}{N} \hat{\mathbf{a}}'_{SR} \tilde{\mathbf{B}}_{SR} \hat{\mathbf{a}}_{SR}$, with

$$\hat{\mathbf{a}}_{SR} \equiv \frac{1}{\sqrt{(\mathbf{x}'_{t+1}\mathbf{x}_t/N - \mathbf{x}'_t\mathbf{x}_t/N)^2}} \left(1 - \frac{1}{2}\frac{(\mathbf{x}'_{t+1}\mathbf{x}_t/N)^2 - (\mathbf{x}'_t\mathbf{x}_t/N)^2}{(\mathbf{x}'_{t+1}\mathbf{x}_t/N - \mathbf{x}'_t\mathbf{x}_t/N)^2}, 1 + \frac{1}{2}\frac{(\mathbf{x}'_{t+1}\mathbf{x}_t/N)^2 - (\mathbf{x}'_t\mathbf{x}_t/N)^2}{(\mathbf{x}'_{t+1}\mathbf{x}_t/N - \mathbf{x}'_t\mathbf{x}_t/N)^2}\right),$$

$$\tilde{\mathbf{B}}_{SR} \equiv \begin{bmatrix} 2\tilde{\sigma}^4 + 4\tilde{f}_t^2\tilde{\sigma}^2\tilde{\sigma}_\lambda & 2\tilde{f}_t\tilde{f}_{t+1}\tilde{\sigma}^2\tilde{\sigma}_\lambda \\ 2\tilde{f}_t\tilde{f}_{t+1}\tilde{\sigma}^2\tilde{\sigma}_\lambda & \tilde{\sigma}^4 + (\tilde{f}_t^2 + \tilde{f}_{t+1}^2)\tilde{\sigma}^2\tilde{\sigma}_\lambda \end{bmatrix},$$

referring to Theorem 2 for the definitions of $\tilde{\sigma}^2, \tilde{\sigma}_\lambda$, and $\tilde{f}_t$.

Remark. The OOS portfolio return is neither constant nor guaranteed to be positive; the conditions for positivity can be read off from (11). A sufficient condition is that f_t and f_{t+1} have the same sign, i.e., that the risk factor is aligned in-sample and out-of-sample. More generally, (11) highlights two key effects driving OOS performance. There is an *attenuation* effect, captured by the positive term

$$\frac{\sigma_{\lambda_t}}{\sigma_{\lambda_t} f_t^2 + \sigma_t^2},$$

and a *sign* effect, governed by the product $f_t f_{t+1}$. The attenuation effect can be fully mitigated ex post through a suitable bias adjustment, since it can be estimated accurately using both the IS and OOS observations. In contrast, the sign effect cannot be controlled ex ante, though it can be quantified and interpreted.

By comparison, the IS portfolio return is always positive and converges to the constant $1/\delta$, a direct consequence of overfitting under $T = 1$. The corresponding theoretical Sharpe ratio is unbounded.

4.2.1 Mean-Variance Portfolios When. $T = 1$: Observed Risk Factor

Given that $\hat{\boldsymbol{\lambda}}\hat{f} = \hat{\boldsymbol{\lambda}}^o f_t^o$, the estimated portfolio weights $\hat{\mathbf{w}}$ for $T = 1$ are unchanged, regardless of whether the factor is treated as latent and estimated via nano-PCA, or observed. Nonetheless, it is often of interest to compare, over time, the *economic value* of different observed risk factors in terms of their OOS portfolio performance.[18]

To make such comparisons meaningful, it is sufficient to use $T = 2$ periods for estimation, say t and $t + 1$, plus one additional period $t + 2$ for OOS evaluation. With two periods used for estimation, portfolio weights no longer remain invariant to the choice of the risk factor, unlike in the $T = 1$ case.

In this setting, the in-sample window consists of periods t and $t + 1$, and the $T = 2$ estimator of $\mathbf{w}_{t+1}$ follows from (7) and from the estimator

$$\tilde{\boldsymbol{\lambda}}^o \equiv \frac{\sum_{s=t}^{t+1} \mathbf{x}_s f_s^o}{\sum_{s=t}^{t+1} (f_s^o)^2}, \tag{14}$$

[18]This type of horse race evaluates the economic value of observed risk factors through Sharpe ratios, contributing to the *factor zoo* literature. It also helps identify whether certain factors are more valuable during expansions and others during recessions, as explored in the empirical section.

yielding

$$\tilde{\mathbf{w}}^o \equiv \frac{1}{\delta}\frac{\bar{f}^o}{\hat{\sigma}_f^o}\left(\frac{\tilde{\boldsymbol{\lambda}}^o}{\tilde{\boldsymbol{\lambda}}^{o\prime}\tilde{\boldsymbol{\lambda}}^o}\right),$$

where $\hat{\sigma}_f^o \equiv \mathbf{f}^{o\prime}\mathbf{f}^o/2$ and $\bar{f}^o \equiv (f_t^o + f_{t+1}^o)/2$.

The corresponding in-sample and OOS portfolio returns then satisfy

$$r_{s|t+1} \equiv \mathbf{x}_s'\tilde{\mathbf{w}}^o \;\to_p\; f_s^o\left(\frac{\bar{f}^o}{\hat{\sigma}_f^o}\right), \quad s = t, t+1,$$

$$r_{t+2|t+1} \equiv \mathbf{x}_{t+2}'\tilde{\mathbf{w}}^o \;\to_p\; f_{t+2}^o\left(\frac{\sigma_{\lambda_t}}{\sigma_{\lambda_t} + \sigma_t^2/(\mathbf{f}^{o\prime}\mathbf{f}^o)}\right)\left(\frac{\bar{f}^o}{\hat{\sigma}_f^o}\right).$$

As in the $T = 1$ case, both an *attenuation* effect (through the factor $\frac{\sigma_{\lambda_t}}{\sigma_{\lambda_t}+\sigma_t^2/(\mathbf{f}^{o\prime}\mathbf{f}^o)}$) and a *sign* effect (through f_{t+2}^o) shape the OOS portfolio return.

Using the three returns $r_{t|t+1}$, $r_{t+1|t+1}$, and $r_{t+2|t+1}$, one can construct a Sharpe ratio either from their sample moments or from the limiting expressions above. Define

$$\widetilde{SR}_{t,t+1,t+2} \equiv \frac{\bar{\mathbf{x}}'\tilde{\mathbf{w}}^o}{\sqrt{\tilde{\mathbf{w}}^{o\prime}\mathbf{S}\tilde{\mathbf{w}}^o}}, \tag{15}$$

where

$$\bar{\mathbf{x}} \equiv \frac{1}{3}(\mathbf{x}_t + \mathbf{x}_{t+1} + \mathbf{x}_{t+2}),$$

and

$$\mathbf{S} \equiv \frac{1}{3}\Big[(\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_t - \bar{\mathbf{x}})' + (\mathbf{x}_{t+1} - \bar{\mathbf{x}})(\mathbf{x}_{t+1} - \bar{\mathbf{x}})' + (\mathbf{x}_{t+2} - \bar{\mathbf{x}})(\mathbf{x}_{t+2} - \bar{\mathbf{x}})'\Big].$$

Asymptotically, $\widetilde{SR}_{t,t+1,t+2}$ is equivalent to the sample mean of the observed factor over $t, t+1, t+2$, divided by its sample standard deviation, with the OOS observation f_{t+2}^o effectively deflated by the factor

$$\frac{\sigma_{\lambda_t}}{\sigma_{\lambda_t} + \sigma_t^2/(\mathbf{f}^{o\prime}\mathbf{f}^o)}.$$

Alternatively, one may work with a modified OOS return that mitigates this deflation,

$$r_{t+2|t+1}^* \equiv \mathbf{x}_{t+2}'\tilde{\mathbf{w}}^o\left(\frac{\tilde{\sigma}_\lambda + \tilde{\sigma}^2/(\mathbf{f}^{o\prime}\mathbf{f}^o)}{\tilde{\sigma}_\lambda}\right) \;\to_p\; f_{t+2}^o\left(\frac{\bar{f}^o}{\hat{\sigma}_f^o}\right),$$

and then compute the Sharpe ratio using this adjusted OOS return. Standard errors and confidence intervals for these Sharpe ratios can be obtained using the asymptotic analysis developed in this paper.[19]

4.3 Pricing Ability of SDF When $T = 1$ - Correct Specification

The one-factor asset pricing model (1) implies the (ex-ante) SDF $m_{t,t+1} \equiv \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\gamma_t}{\sigma_{t,f}^2}\big(f_{t+1} - \gamma_t\big)$, such that $\mathcal{E}_t\big(m_{t,t+1}x_{i,t+1}\big) = \alpha_{it}/r_{ft} = O\big(N^{-1/2}\big)$, where the rate $O(N^{-1/2})$ follows from Assumption 1. Since the factor is latent, it is convenient to work with the *projected* SDF:

$$M_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\gamma_t\boldsymbol{\lambda}_t'\mathbf{V}_t^{-1}\big(\mathbf{x}_{t+1} - \boldsymbol{\lambda}_t\gamma_t\big), \tag{16}$$

[19] Details are available upon request.

which clearly coincides with $M_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\delta\, \mathbf{w}_t'\big(\mathbf{x}_{t+1} - \boldsymbol{\lambda}_t\gamma_t\big)$, by the definition of $\mathbf{w}_t$.

The natural OOS estimator of $M_{t,t+1}$ is then

$$\hat{M}_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\delta\hat{\mathbf{w}}'\mathbf{x}_{t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\mathbf{x}_t'\mathbf{x}_{t+1}}{\mathbf{x}_t'\mathbf{x}_t} + o_p(1),$$

recalling that the data are not de-meaned when constructing estimators.[20]

Using the large-N approximation, we have

$$\hat{M}_{t,t+1} \;\to_p\; \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\sigma_{\lambda_t} f_t f_{t+1}}{\sigma_{\lambda_t} f_t^2 + \sigma_t^2},$$

so a downward bias induced by σ_t^2 must be corrected. To this end, consider the bias-adjusted SDF estimators

$$\tilde{M}^*_{t,t} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\mathbf{x}_t'\mathbf{x}_t}{\mathbf{x}_t'\mathbf{x}_t} = 0 \text{ and } \tilde{M}^*_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\mathbf{x}_t'\mathbf{x}_{t+1}}{\mathbf{x}_t'\mathbf{x}_t - N\tilde{\sigma}^2} + o_p(1), \tag{17}$$

with $\tilde{M}^*_{t,t+1} \to_p \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{f_{t+1}}{f_t} = M^*_{t,t+1}$. The population (ex-post) OOS SDF $M^*_{t,t+1}$ is in the same spirit as $m_{t,t+1}$, but with f_{t+1} not de-meaned and with γ_t and $\sigma^2_{t,f}$ replaced by f_t and f_t^2, respectively.[21]

I use $\tilde{M}^*_{t,t}$ and $\tilde{M}^*_{t,t+1}$ to construct the sample pricing errors

$$\hat{\delta}_{i,t+1|t} \equiv \frac{1}{2}\big(\tilde{M}^*_{t,t}x_{it} + \tilde{M}^*_{t,t+1}x_{i,t+1}\big),$$

which satisfy $\mathbf{1}_N'\hat{\boldsymbol{\delta}}_{t+1|t}/N \;\to_p\; \frac{1}{r_{ft}}(f_t - f_{t+1})\frac{f_{t+1}}{f_t}\mu_{\lambda_t}$.

Theorem 4 (Pricing Test SDF - OOS - Correct Specification: $T = 1$)
Define

$$\tilde{S}_{t+1|t} \equiv \frac{1}{N}\sum_{i=1}^{N}\hat{\delta}_{it+1|t} - \frac{1}{r_{ft}}(1 - \frac{\tilde{f}_{t+1}}{\tilde{f}_t})\tilde{f}_{t+1}\tilde{\mu}_\lambda.$$

(i) Under Assumptions 1 and 2-8,

$$\sqrt{N}\tilde{S}_{t+1|t} \to_d N\left(0, a_{Lt}^2\left[\sigma_t^2(\boldsymbol{\iota}_{0,1}'\mathbf{M}_f\boldsymbol{\iota}_{0,1})(1 - \mu_{\lambda_t}^2/\sigma_{\lambda_t}) + \mathbf{A}_{L1}'\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}\right]\right), \tag{18}$$

setting

$$a_{Lt} \equiv \frac{1}{2r_{ft}}(1 - \frac{f_{t+1}}{f_t}), \tag{19}$$

$$\mathbf{A}_{L1} \equiv (-\mu_{\lambda_t}U_t^{-1}\frac{\mathbf{f}}{2} \otimes \mathbf{M}_f\boldsymbol{\iota}_{01}) \tag{20}$$

[20]The corresponding IS SDF estimator is

$$\hat{M}_{t,t} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\delta\hat{\mathbf{w}}'\mathbf{x}_t = \frac{0}{r_{ft}} = 0,$$

again reflecting the overfitting that arises in-sample when $T = 1$.

[21]Zaffaroni (2025) introduces the notion of an ex-post SDF, constructed from ex-post risk premia and the sample covariance matrix of the factors: $\frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\bar{f}}{\hat{\sigma}_f^2}(f_{t+1} - \bar{f})$, as opposed to the ex-ante SDF $m_{t,t+1}$. Under the $T = 1$ constraint (and hence no de-meaning), this expression simplifies to $M^*_{t,t+1}$. When T is fixed, only the ex-post SDF is identified and can be estimated. Despite the fact that the ex-post SDF differs from the ex-ante SDF by $O(T^{-1/2})$, its associated pricing errors are of order $O(T^{-1})$, making it a valuable object for empirical implementation.

where $\boldsymbol{\iota}_{0,1} \equiv (0,1)'$, $\mathbf{M}_f \equiv \mathbf{I}_2 - \mathbf{f}\mathbf{f}'/(\mathbf{f}'\mathbf{f})$, $\mathbf{U}_{\epsilon,t}$ *being the* $(T+1)^2 \times (T+1)^2$ *matrix with* $T=1$ *defined in Section 10,* $U_t \equiv \sigma_{\lambda_t}\hat{\sigma}_f$, *with* $\hat{\sigma}_f \equiv \mathbf{f}'\mathbf{f}/2$.
(ii) Under the same assumptions, with $\kappa_{4t}=0$, *setting* $\mathbf{M}_{\tilde{f}} \equiv \mathbf{I}_2 - \tilde{\mathbf{f}}\tilde{\mathbf{f}}'/2$, $\tilde{\mathbf{e}} \equiv \mathbf{X} - \tilde{\mathbf{f}}\tilde{\boldsymbol{\lambda}}'$, *and* $\tilde{U}$ *being the largest eigenvalue of* $(\mathbf{X}\mathbf{X}'/(2N) - (\tilde{\sigma}^2/2)\mathbf{I}_2)$,

$$\tilde{a}_{Lt} \equiv \frac{1}{2r_{ft}}(1 - \frac{\tilde{f}_{t+1}}{\tilde{f}_t}) \to_p a_{Lt} \tag{21}$$

$$\tilde{\mathbf{A}}_{L1} \equiv (-\tilde{\mu}_\lambda \tilde{U}^{-1}\frac{\tilde{\mathbf{f}}}{2} \otimes \mathbf{M}_{\tilde{f}}\boldsymbol{\iota}_{01}) \to_p \mathbf{A}_{L1}, \tag{22}$$

$$\tilde{\sigma}^2 \equiv tr(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \to_p \sigma^2,\ \tilde{\mathbf{U}}_\epsilon \to_p \mathbf{U}_{\epsilon,t}, \tag{23}$$

$$\tilde{\sigma}_\lambda \equiv \tilde{\boldsymbol{\lambda}}'\tilde{\boldsymbol{\lambda}}/N - \tilde{\sigma}^2/2 \to_p H_t^{-2}\sigma_{\lambda_t}, \tilde{\mu}_\lambda \to_p H_t^{-1}\mu_{\lambda_t}, \tag{24}$$

where $\tilde{\mathbf{U}}_\epsilon$ *is obtained replacing* σ_t^2 *and* $\kappa_{4,t}$ *with* $\tilde{\sigma}^2$ *and zero, respectively, in* $\mathbf{U}_{\epsilon,t}$, *and*

$$H_t \equiv \sigma_{\lambda_t}^{\frac{1}{2}} U_t^{-\frac{1}{2}} = \hat{\sigma}_f^{-\frac{1}{2}}.$$

(iii) When (1) *does not hold, and* $|\tilde{L}_{t+1|t}| \geq CN^{-1/2+c}$ *a.s. for some* $C, c > 0$, *we have*

$$\sqrt{N}\tilde{L}_{t+1|t} \to_p \pm\infty.$$

4.3.1 Pricing Ability of SDF When $T=1$ - Observed Risk Factor

Just as in the case of testing the pricing ability of risk premia, an important (though special) case of the SDF-based pricing test $\tilde{S}_{t+1|t}$ arises when an *observed* risk factor is used in model (1). As before, the pricing errors $\hat{\delta}_{it}$ are *identical* to those in the latent factor case, and the analysis for the observed factor can be obtained by freezing the sampling variability of the local-PCA factor estimator and setting the rotation matrix equal to one.

Define $\tilde{S}^o_{t+1|t} \equiv \frac{1}{N}\sum_{i=1}^N \hat{\delta}_{i,t+1|t} - \frac{1}{r_{ft}}\Big(1 - \frac{f^o_{t+1}}{f^o_t}\Big) f^o_{t+1}\tilde{\mu}^o_\lambda$, where the two-period estimator of μ_{λ_t} is $\tilde{\mu}^o_\lambda \equiv \frac{\mathbf{1}_N'\tilde{\boldsymbol{\lambda}}^o}{N}$, with $\tilde{\boldsymbol{\lambda}}^o$ as in (14), i.e., $\tilde{\boldsymbol{\lambda}}^o = \boldsymbol{\lambda}_{t-1} + \frac{\sum_{s=t}^{t+1}\boldsymbol{\alpha}_{s-1}f^o_s}{\sum_{s=t}^{t+1}(f^o_s)^2} + \frac{\sum_{s=t}^{t+1}\mathbf{e}_s f^o_s}{\sum_{s=t}^{t+1}(f^o_s)^2} + \frac{\sum_{s=t}^{t+1}\ell_s f^o_s}{\sum_{s=t}^{t+1}(f^o_s)^2}$. It follows that $\frac{\mathbf{1}_N'\boldsymbol{\lambda}_{t-1}}{N} - \tilde{\mu}^o_\lambda = -\frac{\sum_{s=t}^{t+1}\bar{e}_s f^o_s}{\sum_{s=t}^{t+1}(f^o_s)^2} + o_p(N^{-1/2})$, where $\bar{e}_s \equiv N^{-1}\sum_{i=1}^N e_{is}$. Substituting this into $\tilde{S}^o_{t+1|t}$ and rearranging yields $\tilde{S}^o_{t+1|t} = \boldsymbol{\iota}_{0,1}'\mathbf{M}_{f^o}\mathbf{e}'\mathbf{1}_N/N + o_p(N^{-1/2})$, where $\boldsymbol{\iota}_{0,1} \equiv (0,1)'$, $\mathbf{f}^o \equiv (f^o_t, f^o_{t+1})'$, $\mathbf{M}_{f^o} \equiv \mathbf{I}_2 - \frac{\mathbf{f}^o\mathbf{f}^{o\prime}}{\mathbf{f}^{o\prime}\mathbf{f}^o}$. Hence,

$$\sqrt{N}\,\tilde{S}^o_{t+1|t} \ \to_d\ N\big(0,\ \sigma_t^2\,\boldsymbol{\iota}_{0,1}'\mathbf{M}_{f^o}\boldsymbol{\iota}_{0,1}\big)\,.$$

The asymptotic variance

$$\sigma_t^2\,\boldsymbol{\iota}_{0,1}'\mathbf{M}_{f^o}\boldsymbol{\iota}_{0,1} = \sigma_t^2\left(1 - \frac{(f^o_{t+1})^2}{\mathbf{f}^{o\prime}\mathbf{f}^o}\right)$$

coincides exactly with the variance derived in Theorem 4 once the sampling variability of the local-PCA factor is removed (see Appendix XXX for a proof). To construct the standard error of $\sqrt{N}\,\tilde{S}^o_{t+1|t}$, it suffices to estimate σ_t^2 by $\tilde{\sigma}_o^2$, using the procedure described in Section 4.1.1.

5 Cross Sectional Asset Pricing When $T=1$ - Misspecification

My analysis so far has assumed that the factor asset pricing model (1) is correctly specified, i.e., that a single risk factor f_t suffices to price the cross-section of returns. In practice, however, empirical

evidence suggests that multiple risk factors are often needed to span the sources of risk driving returns. In my framework, I allow the number of risk factors to vary over time.

When a multi-factor structure is possible, two main approaches can be pursued.

First, one may continue to estimate the asset pricing model (1) with a single time observation ($T = 1$), maintaining (incorrectly) that there is only one factor. In this case, an additional bias arises on top of the small-sample bias due to $T = 1$, namely a *misspecification bias*. I show how to modify the methodology—in terms of both pricing tests and mean–variance portfolios—to mitigate this source of bias. When the goal is to assess the pricing ability of a given *observed* risk factor, I reduce misspecification bias by estimating the remaining $r_t - 1 > 0$ factors via local-PCA. In all such cases, although the estimation of the asset pricing model and the associated pricing tests and portfolios can still be based on a single period ($T = 1$), the construction of bias adjustments and standard errors requires $r_t + 1 > 1$ observations in total, that is, $T + T_{oos} = 1 + T_{oos} = r_t + 1$. Alternatively, one can treat all r_t risk factors as latent and estimate them jointly using local-PCA. Allowing for $r_t > 1$ introduces an additional challenge: determining the true number of priced factors on short samples of length $T + 1$. For this task, I rely on the local-PCA estimator of the number of factors proposed by Zaffaroni (2025).

A second, conceptually different route—developed in the next section—consists of allowing for multiple factors from the outset when estimating the asset pricing model, and modifying the pricing tests and mean–variance portfolio statistics accordingly. In this case there is no misspecification by construction. If $r_t > 1$ factors are included, at least $T = r_t$ time periods are required for estimation, plus at least one additional observation for OOS analysis. This completely departs from the $T = 1$ approach. Nonetheless, existing empirical work on both portfolio and individual stock data typically finds that the number of priced (strong) factors is quite small, which makes this approach still attractive in practice. It is particularly useful, for example, when one wishes to assess the pricing ability of a given multi-factor model based on observed risk factors—such as the three-factor model of Fama and French (1993)—from a *conditional* perspective, evaluating how its pricing performance varies over time.

5.1 Pricing Ability of Risk Premia When $T = 1$ - Misspecification

As discussed above, the asset pricing model (1) is now *misspecified* in the sense that a one-factor specification is estimated, even though the true DGP is multi-factor. Estimation remains based on a single period ($T = 1$), so the definition of the OOS pricing errors $\hat{h}_{i,t+1|t} = x_{i,t+1} - \hat{\lambda}_i \hat{f}$ does not change. However, under the true multi-factor structure, $\hat{h}_{i,t+1|t}$ satisfies

$$\hat{h}_{i,t+1|t} = \lambda_{it-1}(f^o_{t+1} - f^o_t) + \boldsymbol{\lambda}'_{it-1}(\mathbf{F}_{t+1} - \mathbf{F}_t) + e_{i,t+1} - e_{it} + \ell_{i,t+1} + \alpha_{it} - \alpha_{it-1},$$

because I now assume that the cross-section of returns follows the r_t-factor model

$$x_{it} = \alpha_{it-1} + \lambda_{it-1} f^o_t + \boldsymbol{\lambda}'_{it-1}\mathbf{F}_t + e_{it}, \quad \text{for all } t \in \mathcal{Z},\ i \in \mathcal{N}, \tag{25}$$

where f^o_t is an observed factor with exposure λ_{it-1}, $\mathbf{F}_t$ is an $(r_t - 1) \times 1$ vector of latent factors with exposures $\boldsymbol{\lambda}_{it-1}$, and $\ell_{i,t+1} \equiv (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{it-1})'\mathbf{F}_{t+1} + (\lambda_{it} - \lambda_{it-1}) f^o_{t+1}$. Two special cases are of particular interest: (i) the case in which no observed factor is posited and all r_t factors are latent, and (ii) the opposite case in which all r_t factors are observed. Both are treated below as special cases of (25).

In this multi-factor setting, $\hat{h}_{i,t+1|t}$ is contaminated not only by the small-sample bias arising from $T = 1$, but also by a *misspecification bias*. To construct a pricing test with correct asymptotic size and

power, this additional bias must be accounted for. Indeed,

$$\frac{\mathbf{1}_N'\hat{\mathbf{h}}_{t+1|t}}{N} \;\to_p\; \mu_{\lambda_t}(f^o_{t+1} - f^o_t) + \boldsymbol{\mu}_{\Lambda_t}'(\mathbf{F}_{t+1} - \mathbf{F}_t),$$

where $\boldsymbol{\Lambda}_{t-1}'\mathbf{1}_N/N \to \boldsymbol{\mu}_{\Lambda_t}$ for $\boldsymbol{\Lambda}_{t-1} \equiv (\boldsymbol{\lambda}_{1,t-1}, \ldots, \boldsymbol{\lambda}_{N,t-1})'$, and $\boldsymbol{\lambda}_{t-1}'\mathbf{1}_N/N \to \mu_{\lambda_t}$ for $\boldsymbol{\lambda}_{t-1} \equiv (\lambda_{1,t-1}, \ldots, \lambda_{N,t-1})'$

To proceed, I estimate the multi-factor version of (1) using the local-PCA methodology of Zaffaroni (2025) with $T + T_{oos} = r_t + 1 > 2$ observations. Let $\mathcal{T}_t \equiv \{t - T + 1, \ldots, t + 1\}$ denote the estimation window. Given the emphasis on evaluating the pricing ability of the observed factor f^o_t, it is convenient to estimate the remaining $r_t - 1$ latent factors $\mathbf{F}_t$ under the restriction that they are orthogonal to f^o_t.

Specifically, I first estimate the factor loadings on f^o_t as if it were the only regressor:

$$\tilde{\boldsymbol{\lambda}}^o \equiv \frac{\mathbf{X}'\mathbf{f}^o}{\mathbf{f}^{o\prime}\mathbf{f}^o},$$

where $\mathbf{X}$ is the $(T+1) \times N$ matrix of returns and $\mathbf{f}^o \equiv (f^o_{t-T+1}, \ldots, f^o_{t+1})'$. As before, I must use $T + 1$ periods; otherwise, with $T = 1 + (r_t - 1) = r_t$, the model would fit perfectly and σ^2_t could not be estimated.

Next, I orthogonalize the data with respect to f^o_t by applying $\mathbf{X}^* \equiv \mathbf{M}_{f^o}\mathbf{X}, \quad \mathbf{M}_{f^o} \equiv \mathbf{I}_{T+1} - \frac{\mathbf{f}^o\mathbf{f}^{o\prime}}{\mathbf{f}^{o\prime}\mathbf{f}^o}$, and then run local-PCA on $\mathbf{X}^*$. Formally,

$$\tilde{\mathbf{F}},\ \tilde{\boldsymbol{\Lambda}} = \text{argmin}_{\mathbf{F},\boldsymbol{\Lambda}} \frac{1}{(T+1)N} \sum_{s\in\mathcal{T}_t} \sum_{i=1}^{N} \left(x^*_{is} - \mathbf{F}_s'\boldsymbol{\lambda}_i\right)^2, \tag{26}$$

where $\boldsymbol{\Lambda} \equiv (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_N)'$ and x^*_{is} denotes the (i, s)-entry of $\mathbf{X}^*$. The matrix $\tilde{\mathbf{F}} \equiv (\tilde{\mathbf{F}}_{t-T+1}, \ldots, \tilde{\mathbf{F}}_{t+1})'$ is the $(T+1) \times (r_t - 1)$ matrix of estimated factors and consists of the $(r_t - 1)$ eigenvectors of $\mathbf{X}^*\mathbf{X}^{*\prime}/((T + 1)N)$ associated with its largest $r_t - 1$ eigenvalues, scaled by $\sqrt{T+1}$. Hence $\tilde{\mathbf{F}}'\tilde{\mathbf{F}}/(T + 1) = \mathbf{I}_{r_t-1}$, and $\tilde{\boldsymbol{\Lambda}} \equiv (\tilde{\boldsymbol{\lambda}}_1, \ldots, \tilde{\boldsymbol{\lambda}}_N)' = \mathbf{X}^{*\prime}\tilde{\mathbf{F}}(\tilde{\mathbf{F}}'\tilde{\mathbf{F}})^{-1} = \mathbf{X}^{*\prime}\tilde{\mathbf{F}}/(T + 1)$.[22]

The assumptions in Appendix XXX imply that all time-varying parameters are *locally constant* over the $(T + 1)$-period window. In particular, the factor loadings satisfy:

Assumption 3 (Risk exposures – smoothness, case $r_t > 1$) *For a given matrix $\boldsymbol{\Lambda}_{t-1}$,*

$$\|\boldsymbol{\Lambda}_s - \boldsymbol{\Lambda}_{t-1}\|^2 = o(N^{1/2}) \quad \textit{for every } s \in \mathcal{T}_t.$$

The number of factors r_t is allowed to vary over time, but not within the short window of length $T+1$. Many economically relevant forms of time variation in loadings are encompassed by Assumption 2 and its extension here.[23]

Since the time window $\mathcal{T}_t$ of length $T+1$ (with $T = r_t$), is still very short, the required smoothness of the loadings and other parameters is relatively mild. Moreover, by construction, the local-PCA factors are orthogonal to $\mathbf{f}^o$, and thus, without loss of generality, I assume the true latent factors $\mathbf{F}$ are also

[22] See Magnus and Neudecker (2007, Chapter 17, Theorem 7) for a proof. An asymptotically equivalent PCA estimator is obtained by taking $\tilde{\boldsymbol{\Lambda}}^\dagger$ as the $(r_t - 1)$ eigenvectors of $\mathbf{X}^{*\prime}\mathbf{X}^*/((T + 1)N)$ corresponding to the largest eigenvalues, scaled by $\sqrt{N}$, and setting $\tilde{\mathbf{F}} = \mathbf{X}^*\tilde{\boldsymbol{\Lambda}}^\dagger/N$; see Bai and Ng (2002). Since $N \gg T + 1$, the formulation based on $\mathbf{X}^*\mathbf{X}^{*\prime}$ is computationally more convenient. The nonzero eigenvalues of $\mathbf{X}^{*\prime}\mathbf{X}^*$ and $\mathbf{X}^*\mathbf{X}^{*\prime}$ coincide.

[23] Consistency of local-PCA only requires the weaker condition $\sup_{s\in\mathcal{T}_t} \|\boldsymbol{\Lambda}_s - \boldsymbol{\Lambda}_{t-1}\|^2 = o(N)$, whereas the limiting distribution results need the sharper $o(N^{1/2})$ rate. I do not distinguish between these for expositional simplicity. A related smoothness condition appears in Bates, Plagborg-Møller, Stock, and Watson (2013), although their rate depends on both N and T because they study constant-PCA.

orthogonal to $\mathbf{f}^o$.[24]

Although I present explicitly the case in which the true model is spanned by one observed factor and $(r_t - 1)$ latent factors, the case with no observed factor arises as a special case (by dropping f_t^o). Another special case arises when there are no latent risk factors. In all cases, I initially estimate (1) via nano-PCA, but for inference (pricing tests and portfolio analysis) I employ OLS for the observed factor f_t^o and local-PCA for the latent factors $\mathbf{F}_t$, as described above.

Theorem 5 (Pricing Test Cross Section - OOS - Misspecification: $T = r_t$)
Define

$$\tilde{L}_{t+1|t}^{miss} \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,t+1|t} - \tilde{\mu}_\lambda^o(f_{t+1}^o - f_t^o) - \tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}}'(\tilde{\mathbf{F}}_{t+1} - \tilde{\mathbf{F}}_t), \text{ with } \tilde{\mu}_\lambda^o \equiv \tilde{\boldsymbol{\lambda}}^{o\prime}\mathbf{1}_N/N \text{ and } \tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}'\mathbf{1}_N/N.$$

(i) Under Assumptions 1 and 2-8,

$$\sqrt{N}\tilde{L}_{t+1|t}^{miss} \to_d N\left(0, \sigma_t^2(\boldsymbol{\iota}_{\mathbf{0}-11}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}-11})(1 - \boldsymbol{\mu}_{\boldsymbol{\Lambda}_{t-1}}'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{t-1}}^{-1}\mu_{\boldsymbol{\Lambda}_{t-1}}) + \mathbf{A}_{L1}'\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}\right), \tag{27}$$

setting

$$\mathbf{A}_{L1} \equiv \frac{1}{(T+1)}(\mathbf{F}\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t} \otimes -\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}-11}). \tag{28}$$

where $\boldsymbol{\iota}_{\mathbf{0}-11} \equiv (\mathbf{0}_{T-1}', -1, 1)'$, $'\mathbf{M}_{f^o} \equiv \mathbf{I}_{T+1} - \mathbf{f}^o\mathbf{f}^{o\prime}/(\mathbf{f}^{o\prime}\mathbf{f}^o)$, $\mathbf{M}_{\mathbf{F}} \equiv \mathbf{I}_{T+1} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\mathbf{U}_{\epsilon,t}$ *being the* $(T+1)^2 \times (T+1)^2$ *matrix with* $T = r_t$ *defined in Section 10,* $\mathbf{U}_t$ *is the diagonal matrix of the* $(r_t - 1)$ *eigenvalues of* $\hat{\boldsymbol{\Sigma}}_{\mathbf{F}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}$*, with eigenvector matrix* $\boldsymbol{\Upsilon}_t$*.*
(ii) Under the same assumptions, with $\kappa_{4t} = 0$*, setting* $\mathbf{M}_{\tilde{F}} \equiv \mathbf{I}_{T+1} - \tilde{\mathbf{F}}(\tilde{\mathbf{F}}'\tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}'$*,* $\tilde{\mathbf{e}} \equiv \mathbf{X}^* - \tilde{\mathbf{F}}\tilde{\boldsymbol{\Lambda}}'$*, and* $\tilde{\mathbf{U}}$ *being the* $(r_t - 1)$ *largest eigenvalues of* $(\mathbf{X}^*\mathbf{X}^{*\prime}/((T+1)N) - (\tilde{\sigma}^2/(T+1))\mathbf{I}_{T+1})$*,*

$$\tilde{\mathbf{A}}_{L1} \equiv \frac{1}{(T+1)}(\tilde{\mathbf{F}}\tilde{\mathbf{U}}^{-1}\tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}} \otimes -\mathbf{M}_{\tilde{\mathbf{F}}}\mathbf{M}_{f^o}\boldsymbol{\iota}_{\mathbf{0}-11}) \to_p \mathbf{A}_{L1}, \tag{29}$$

$$\tilde{\sigma}^2 \equiv tr(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \to_p \sigma_t^2,\ \tilde{\mathbf{U}}_\epsilon \to_p \mathbf{U}_{\epsilon,t}, \tag{30}$$

$$\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}'\tilde{\boldsymbol{\Lambda}}/N - \tilde{\sigma}^2/(T+1)\mathbf{I}_{r_t-1} \to_p \mathbf{H}_t^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{H}_t^{-1\prime}, \tilde{\boldsymbol{\mu}}_\Lambda \to_p \mathbf{H}_t^{-1}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t}, \tag{31}$$

where $\tilde{\mathbf{U}}_\epsilon$ *is obtained replacing* σ_t^2 *and* $\kappa_{4,t}$ *with* $\tilde{\sigma}^2$ *and zero, respectively, in* $\mathbf{U}_{\epsilon,t}$*, and*

$$\mathbf{H}_t \equiv \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}^{\frac{1}{2}}\boldsymbol{\Upsilon}_t\mathbf{U}_t^{-\frac{1}{2}}.$$

(iii) When (1) *does not hold, and* $|\tilde{L}_{t+1|t}| \geq CN^{-1/2+c}$ *a.s. for some* $C, c > 0$*,*

$$\sqrt{N}\tilde{L}_{t+1|t} \to_p \pm\infty.$$

[24]Equivalently, one can always reparametrize the r_t-factor model via projections so that $\mathbf{F}$ and $\mathbf{f}^o$ are orthogonal. Let $\mathbf{P}_{f^o} = \mathbf{f}^o\mathbf{f}^{o\prime}/(\mathbf{f}^{o\prime}\mathbf{f}^o)$, with $\mathbf{P}_{f^o} + \mathbf{M}_{f^o} = \mathbf{I}_{T+1}$. Then

$$\mathbf{X} = \mathbf{f}^o\boldsymbol{\lambda}' + \mathbf{F}\boldsymbol{\Lambda} + \mathbf{e} = \mathbf{f}^o\boldsymbol{\lambda}' + [\mathbf{P}_{f^o}\mathbf{F} + \mathbf{M}_{f^o}\mathbf{F}]\boldsymbol{\Lambda} + \mathbf{e} = \mathbf{f}^o\boldsymbol{\lambda}^{*\prime} + \mathbf{F}^*\boldsymbol{\Lambda} + \mathbf{e},$$

with $\boldsymbol{\lambda}^* \equiv \boldsymbol{\lambda} + \boldsymbol{\Lambda}\mathbf{F}'\mathbf{f}^o/(\mathbf{f}^{o\prime}\mathbf{f}^o)$ and $\mathbf{F}^* \equiv \mathbf{M}_{f^o}\mathbf{F}$. This yields time-series orthogonality between the latent and observed factors, at the cost of inducing cross-sectional dependence between the corresponding loadings (unless $\mathbf{F}$ and $\mathbf{f}^o$ are already orthogonal).

When no observed risk factor is posited, I can derive the asymptotics of the pricing test as a special case of Theorem 5, in which all r_t factors are latent.[25] In this setting, $\hat{h}_{i,t+1|t}$ satisfies

$$\hat{h}_{i,t+1|t} = \boldsymbol{\lambda}'_{it-1}(\mathbf{F}_{t+1} - \mathbf{F}_t) + e_{i,t+1} - e_{it} + \ell_{i,t+1} + \alpha_{it} - \alpha_{it-1},$$

where $\mathbf{F}_t$ is an $r_t \times 1$ vector of latent factors, $\boldsymbol{\lambda}_{it-1}$ is the corresponding $r_t \times 1$ vector of factor loadings, and $\ell_{i,t+1} \equiv (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{it-1})'\mathbf{F}_{t+1}$, with $\sum_{i=1}^{N} \hat{h}_{i,t+1|t}/N \quad \to_p \quad \boldsymbol{\mu}'_{\Lambda_t}(\mathbf{F}_{t+1} - \mathbf{F}_t)$, and $\boldsymbol{\mu}_{\Lambda_t} \equiv \lim_{N\to\infty} \frac{\boldsymbol{\Lambda}'_{t-1}\mathbf{1}_N}{N}$, $\quad \boldsymbol{\Lambda}_{t-1} \equiv (\boldsymbol{\lambda}_{1,t-1}, \ldots, \boldsymbol{\lambda}_{N,t-1})'$.

Since there is no observed factor, I apply local-PCA directly to $\mathbf{X}$:[26]

$$\tilde{\mathbf{F}},\ \tilde{\boldsymbol{\Lambda}} = \text{argmin}_{\mathbf{F},\boldsymbol{\Lambda}} \frac{1}{(T+1)N} \sum_{s\in\mathcal{T}_t} \sum_{i=1}^{N} \big(x_{is} - \mathbf{F}'_s\boldsymbol{\lambda}_i\big)^2, \tag{32}$$

where $\tilde{\mathbf{F}} \equiv (\tilde{\mathbf{F}}_{t-T+1}, \ldots, \tilde{\mathbf{F}}_{t+1})'$ is the $(T+1) \times r_t$ matrix of estimated factors, consisting of the r_t eigenvectors of $\mathbf{XX}'/((T+1)N)$ corresponding to the largest r_t eigenvalues, scaled by $\sqrt{T+1}$, so that

$$\tilde{\mathbf{F}}'\tilde{\mathbf{F}}/(T+1) = \mathbf{I}_{r_t}, \quad \tilde{\boldsymbol{\Lambda}} \equiv \mathbf{X}'\tilde{\mathbf{F}}/(T+1).$$

The bias-adjusted pricing test statistic is now

$$\tilde{L}^{\texttt{miss}}_{t+1|t} \equiv \frac{1}{N}\sum_{i=1}^{N} \hat{h}_{i,t+1|t} - \tilde{\boldsymbol{\mu}}'_{\Lambda}(\tilde{\mathbf{F}}_{t+1} - \tilde{\mathbf{F}}_t),$$

where $\tilde{\boldsymbol{\mu}}_{\Lambda} \equiv \tilde{\boldsymbol{\Lambda}}'\mathbf{1}_N/N$, with limiting distribution

$$\sqrt{N}\,\tilde{L}^{\texttt{miss}}_{t+1|t} \ \to_d \ N\Big(0,\ \sigma_t^2\big(\boldsymbol{\iota}'_{\mathbf{0}-11}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}-11}\big)\Big(1 - \boldsymbol{\mu}'_{\boldsymbol{\Lambda}_{t-1}}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\Lambda}_{t-1}}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_{t-1}}\Big) + \mathbf{A}'_{L1}\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}\Big),$$

obtained from Theorem 5 after dropping $\mathbf{M}_{f^o}$ entirely. In particular,

$$\mathbf{A}_{L1} \equiv \frac{1}{T+1}\big(\mathbf{F}\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t} \ \otimes \ -\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}-11}\big),$$

where $\mathbf{F}$ collects the true factor realizations over $\mathcal{T}_t$, $\mathbf{M}_{\mathbf{F}} \equiv \mathbf{I}_{T+1} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\mathbf{U}_t$ is the diagonal matrix of the r_t eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\mathbf{F}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}$ (with eigenvector matrix $\boldsymbol{\Upsilon}$), and $\mathbf{U}_{\epsilon,t}$ is the $(T+1)^2 \times (T+1)^2$ matrix with $T = r_t$ defined in Section 10. Moreover, with $\tilde{\mathbf{e}} \equiv \mathbf{X} - \tilde{\mathbf{F}}\tilde{\boldsymbol{\Lambda}}'$ and $\tilde{\mathbf{U}}$ denoting the r_t largest eigenvalues of $\frac{\mathbf{XX}'}{(T+1)N} - \frac{\tilde{\sigma}^2}{T+1}\mathbf{I}_{T+1}$, we have

$$\begin{aligned}
\tilde{\mathbf{A}}_{L1} &\equiv \frac{1}{T+1}\big(\tilde{\mathbf{F}}\tilde{\mathbf{U}}^{-1}\tilde{\boldsymbol{\mu}}_{\Lambda} \ \otimes \ -\mathbf{M}_{\tilde{\mathbf{F}}}\boldsymbol{\iota}_{\mathbf{0}-11}\big) \ \to_p \ \mathbf{A}_{L1},\\
\tilde{\sigma}^2 &\equiv \text{tr}(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \ \to_p \ \sigma_t^2,\\
\tilde{\boldsymbol{\Sigma}}_{\Lambda} &\equiv \tilde{\boldsymbol{\Lambda}}'\tilde{\boldsymbol{\Lambda}}/N - \frac{\tilde{\sigma}^2}{T+1}\mathbf{I}_{r_t} \ \to_p \ \mathbf{H}_t^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{H}_t^{-1\prime},\\
\tilde{\boldsymbol{\mu}}_{\Lambda} &\ \to_p \ \mathbf{H}_t^{-1}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t},
\end{aligned}$$

where $\mathbf{H}_t$ is the (invertible) rotation matrix mapping the true factors $\mathbf{F}$ to their estimated counterparts $\tilde{\mathbf{F}}$, and all other notation is as in Theorem 5.

5.2 Mean-Variance Portfolios When $T = 1$ - Misspecification.

The investor continues to (incorrectly) believe that (1) is the true one-factor model and therefore constructs the mean–variance portfolio with $T = 1$ using the nano-PCA estimator (8), $\hat{\mathbf{w}} = \frac{1}{\delta}\left(\frac{\mathbf{x}_t}{\mathbf{x}'_t\mathbf{x}_t}\right)$.

[25] An analogous result can be derived for the case of r_t observed risk factors, but I do not pursue this here. Instead, I consider the case of r_t observed factors under a correctly specified multi-factor model in Section 6.

[26] See Footnote 14.

In reality, however, there are $r_t > 1$ risk factors spanning the cross-section of returns.

The construction of the IS and OOS portfolio returns and of the Sharpe ratio is unchanged, since it still relies on $T = 1$. Thus, $r_{t|t} \equiv \hat{\mathbf{w}}'\mathbf{x}_t$, $r_{t+1|t} \equiv \hat{\mathbf{w}}'\mathbf{x}_{t+1}$, and $\widehat{SR}_{t,t+1} \equiv \frac{\bar{\mathbf{x}}'\hat{\mathbf{w}}}{\sqrt{\hat{\mathbf{w}}'\mathbf{S}\hat{\mathbf{w}}}}$, where $\bar{\mathbf{x}} \equiv \frac{1}{2}(\mathbf{x}_t + \mathbf{x}_{t+1})$, $\mathbf{S} \equiv \frac{1}{2}\Big[(\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_t - \bar{\mathbf{x}})' + (\mathbf{x}_{t+1} - \bar{\mathbf{x}})(\mathbf{x}_{t+1} - \bar{\mathbf{x}})'\Big]$. The limiting behavior of these quantities, however, now reflects the true r_t-factor structure rather than the misspecified one-factor model.

Unlike the previous section, where I first develop the case of one observed risk factor with the remaining $r_t - 1$ latent factors, then generalizing to all latent or all observed risk factors, here I am considering only the case when the risk factors are all latent and estimated by local-PCA, as it appears more relevant in this context of misspecification. The results can be easily adapted to the other cases of one observed risk factor plus $r_t - 1$ latent risk factors, or when all factors are observed (see footnote 25).

Finally, in the mean–variance portfolio analysis I deliberately do not construct a formal test statistic (although this would be possible), so no additional bias adjustment is required. Instead, I provide asymptotically valid standard errors for the Sharpe ratio, which of course require $T + 1 > r_t$ periods of data, merging the IS and the OOS periods.

Theorem 6 (Mean Variance Portfolio - IS and OOS - Misspecification: $T = r_t$) *When Assumptions 1 and 4-7 hold:*
(i)

$$r_{t|t} = \frac{1}{\delta}\frac{(\mathbf{x}_t'\mathbf{x}_t)}{(\mathbf{x}_t'\mathbf{x}_t)} = \frac{1}{\delta}. \tag{33}$$

$$r_{t+1|t} = \frac{1}{\delta}\frac{(\mathbf{x}_{t+1}'\mathbf{x}_t)}{(\mathbf{x}_t'\mathbf{x}_t)} \to_p \frac{1}{\delta}\frac{(\mathbf{F}_{t+1}'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{F}_t)}{(\mathbf{F}_t'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{F}_t + \sigma_t^2)}. \tag{34}$$

(ii)

$$\widehat{SR}_{t,t+1} = \frac{(\mathbf{x}_{t+1}'\mathbf{x}_t) + (\mathbf{x}_t'\mathbf{x}_t)}{\sqrt{((\mathbf{x}_{t+1}'\mathbf{x}_t) - (\mathbf{x}_t'\mathbf{x}_t))^2}} + o_p(1) \to_p SR_{t,t+1} = \frac{(1 + \frac{(\mathbf{F}_{t+1}'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{F}_t)}{(\mathbf{F}_t'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{F}_t + \sigma_t^2)})}{|(1 - \frac{(\mathbf{F}_{t+1}'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{F}_t)}{(\mathbf{F}_t'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{F}_t + \sigma_t^2)}|}. \tag{35}$$

(iii) Assumptions 1 and 2-8 with $\kappa_{4t} = 0$, for any $0 < \alpha < 1/2$, setting $z_{\alpha/2}$ the (absolute value of the) $\alpha/2$-th quantile of the standard normal

$$Pr(\widehat{SR}_{t,t+1} - z_{\alpha/2}SE_{SR} \leq SR_{t,t+1} \leq \widehat{SR}_{t,t+1} + z_{\alpha/2}SE_{SR}) \to 1 - \alpha, \tag{36}$$

setting $SE_{SR}^2 \equiv \frac{1}{N}\hat{\mathbf{a}}_{1SR}'\tilde{\mathbf{A}}_{1SR}\hat{\mathbf{a}}_{1SR}$, with

$$\hat{\mathbf{a}}_{1SR} \equiv \frac{1}{\sqrt{(\mathbf{x}_{t+1}'\mathbf{x}_t/N - \mathbf{x}_t'\mathbf{x}_t/N)^2}}\left(1 - \frac{1}{2}\frac{(\mathbf{x}_{t+1}'\mathbf{x}_t/N)^2 - (\mathbf{x}_t'\mathbf{x}_t/N)^2}{(\mathbf{x}_{t+1}'\mathbf{x}_t/N - \mathbf{x}_t'\mathbf{x}_t/N)^2}, 1 + \frac{1}{2}\frac{(\mathbf{x}_{t+1}'\mathbf{x}_t/N)^2 - (\mathbf{x}_t'\mathbf{x}_t/N)^2}{(\mathbf{x}_{t+1}'\mathbf{x}_t/N - \mathbf{x}_t'\mathbf{x}_t/N)^2}\right),$$

$$\tilde{\mathbf{A}}_{1SR} \equiv \begin{bmatrix} 2\tilde{\sigma}^4 + 4\tilde{\sigma}^2\tilde{\mathbf{F}}_t'\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}}\tilde{\mathbf{F}}_t & 2\tilde{\sigma}^2\tilde{\mathbf{F}}_t'\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}}\tilde{\mathbf{F}}_{t+1} \\ 2\tilde{\sigma}^2\tilde{\mathbf{F}}_t'\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}}\tilde{\mathbf{F}}_{t+1} & \tilde{\sigma}^4 + \tilde{\sigma}^2(\tilde{\mathbf{F}}_t'\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}}\tilde{\mathbf{F}}_t + \tilde{\mathbf{F}}_{t+1}'\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}}\tilde{\mathbf{F}}_{t+1}) \end{bmatrix},$$

referring to Theorem 5 for the definitions of $\tilde{\sigma}^2, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}}$, and $\tilde{\mathbf{F}}_t$.

Remark. As in the $T = 1$ correctly specified case, the OOS portfolio return is neither constant nor guaranteed to be positive; the conditions for positivity can be derived from its limiting expression. A sufficient condition is that the factor realizations $\mathbf{F}_t$ and $\mathbf{F}_{t+1}$ are *aligned*, so that their contribution to the portfolio return has a non–negative sign. There remains an *attenuation* effect, captured by the

positive term

$$\mathbf{F}_t'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + \sigma_t^2,$$

which dampens the magnitude of the OOS portfolio return. This attenuation can be fully mitigated by an appropriate bias adjustment, since this term can be estimated accurately using both the in-sample and out-of-sample observations.

5.3 Pricing Ability of SDF When $T = 1$ - Misspecification.

.

As the investor continues to believe that the one-factor asset pricing model (1) holds, the implied estimator for the OOS $M_{t,t+1}$ is still given by $\hat{M}_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\delta\hat{\mathbf{w}}'(\mathbf{x}_{t+1}) = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\mathbf{x}_t'\mathbf{x}_{t+1}}{\mathbf{x}_t'\mathbf{x}_t} + o_p(1)$. although now, when the r_t-factor model holds, made up by the single observed risk factor f_t^o and $r_t - 1$ latent risk factors $\mathbf{F}_t$,

$$\hat{M}_{t,t+1} \to_p \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{(\mathbf{F}_{t+1}'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + f_t^o f_{t+1}^o\sigma_{\lambda_t})}{(\mathbf{F}_t'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + (f_t^o)^2\sigma_{\lambda_t} + \sigma_t^2)}, \tag{37}$$

where to simplify the exposition I am also assuming cross-sectional orthogonality between the λ_{it-1} and $\boldsymbol{\lambda}_{it-1}$, leading to the corrected OOS SDF estimator, using the local-PCA with $T + 1 = r_t + 1$ periods of data,

$$\tilde{M}_{t,t}^* = 0 \text{ and } \tilde{M}_{t,t+1}^* = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\frac{\mathbf{x}_t'\mathbf{x}_{t+1}}{\mathbf{x}_t'\mathbf{x}_t - N\tilde{\sigma}^2} + o_p(1), \tag{38}$$

using $\tilde{M}_{t,t}^*$ and $\tilde{M}_{t,t+1}^*$ to derive the sample pricing errors $\hat{\delta}_{it+1|t} \equiv \frac{1}{2}(\hat{M}_{t,t}^* x_{it} + \hat{M}_{t,t+1}^* x_{it+1})$, satisfying

$$\sum_{i=1}^{N}\hat{\delta}_{it+1|t}/N \to_p \frac{1}{2r_{ft}}\left(1 - \frac{(\mathbf{F}_{t+1}'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + f_t^o f_{t+1}^o\sigma_{\lambda_t})}{(\mathbf{F}_t'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + (f_t^o)^2\sigma_{\lambda_t})}\right)(f_{t+1}^o\mu_{\lambda_t} + \mathbf{F}_{t+1}'\boldsymbol{\mu}_{\mathbf{\Lambda}_t}).$$

Theorem 7 (Pricing Test Cross SDF- OOS - Misspecification: $T = r_t$)
Define

$$\tilde{S}_{t+1|t}^{miss} \equiv \frac{1}{N}\sum_{i=1}^{N}\hat{\delta}_{i,t+1|t} - \frac{1}{2r_{ft}}\left(1 - \frac{(\tilde{\mathbf{F}}_{t+1}'\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t + f_t^o f_{t+1}^o\tilde{\sigma}_\lambda)}{(\tilde{\mathbf{F}}_t'\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t + (f_t^o)^2\tilde{\sigma}_\lambda)}\right)(\tilde{\mu}_\lambda^o f_{t+1}^o + \tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}}'\tilde{\mathbf{F}}_{t+1}), \tag{39}$$

$$\text{with } \tilde{\mu}_\lambda^o \equiv \tilde{\boldsymbol{\lambda}}^{o\prime}\mathbf{1}_N/N \text{ and } \tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}} \equiv \tilde{\mathbf{\Lambda}}'\mathbf{1}_N/N. \tag{40}$$

(i) Under Assumptions 1 and 2-8,

$$\sqrt{N}\tilde{S}_{t+1|t}^{miss} \to_d N\left(0, a_{Lt}^2[\sigma_t^2(\boldsymbol{\iota}_{\mathbf{01}}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}})(1 - \boldsymbol{\mu}_{\mathbf{\Lambda}_{t-1}}'\mathbf{\Sigma}_{\mathbf{\Lambda}_{t-1}}^{-1}\mu_{\mathbf{\Lambda}_{t-1}}) + \mathbf{A}_{L1}'\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}]\right), \tag{41}$$

setting

$$a_{Lt} = \frac{1}{2r_{ft}}\left(1 - \frac{(\mathbf{F}_{t+1}'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + f_{t+1}f_t\sigma_{\lambda_t})}{(\mathbf{F}_t'\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{F}_t + f_t^2\sigma_{\lambda_t})}\right), \tag{42}$$

$$\mathbf{A}_{L1} \equiv \frac{1}{(T+1)}(\mathbf{F}\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t} \otimes -\mathbf{M}_f\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}). \tag{43}$$

where $\boldsymbol{\iota}_{\mathbf{01}} \equiv (\mathbf{0}_T', 1)'$, $\mathbf{M}_{\mathbf{F}} \equiv \mathbf{I}_{T+1} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\mathbf{U}_{\epsilon,t}$ *being the* $(T+1)^2 \times (T+1)^2$ *matrix with* $T = r_t$ *defined in Section 10,* $\mathbf{U}_t$ *is the diagonal matrix of the* $r_t - 1$ *eigenvalues of* $\hat{\Sigma}_{\mathbf{F}}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}$*, with eigenvector matrix* $\mathbf{\Upsilon}_t$*.*

(ii) Under the same assumptions, with $\kappa_{4t}=0$*, setting* $\mathbf{M}_{\tilde{F}}\equiv\mathbf{I}_{T+1}-\tilde{\mathbf{F}}(\tilde{\mathbf{F}}'\tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}'$*,* $\tilde{\mathbf{e}}\equiv\mathbf{X}^{*}-\tilde{\mathbf{F}}\tilde{\mathbf{\Lambda}}'$*, and* $\tilde{\mathbf{U}}$ *being the* r_t-1 *largest eigenvalues of* $(\mathbf{X}^{*}\mathbf{X}^{*\prime}/((T+1)N)-(\tilde{\sigma}^2/(T+1))\mathbf{I}_{T+1})$*,*

$$\tilde{a}_{Lt}=\frac{1}{r_{ft}}\left(1-\frac{(\tilde{\mathbf{F}}'_{t+1}\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t+f_{t+1}f_t\tilde{\sigma}_{\lambda})}{(\tilde{\mathbf{F}}'_t\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t+f_t^2\tilde{\sigma}_{\lambda})}\right)\rightarrow_p a_{Lt}, \tag{44}$$

$$\tilde{\mathbf{A}}_{L1}\equiv\frac{1}{(T+1)}(\tilde{\mathbf{F}}\tilde{\mathbf{U}}_t^{-1}\tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}}\otimes-\mathbf{M}_{\tilde{\mathbf{F}}}\mathbf{M}_f\boldsymbol{\iota}_{\mathbf{01}})\rightarrow_p\mathbf{A}_{L1}, \tag{45}$$

$$\tilde{\sigma}^2\equiv tr(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N\rightarrow_p\sigma_t^2,\ \tilde{\mathbf{U}}_{\epsilon}\rightarrow_p\mathbf{U}_{\epsilon,t}, \tag{46}$$

$$\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\equiv\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}/N-\tilde{\sigma}^2/(T+1)\mathbf{I}_{r_t-1}\rightarrow_p\mathbf{H}_t^{-1}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{H}_t^{-1\prime},\tilde{\boldsymbol{\mu}}_{\lambda}\rightarrow_p\mathbf{H}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t}, \tag{47}$$

where $\tilde{\mathbf{U}}_{\epsilon}$ *is obtained replacing* σ_t^2 *and* $\kappa_{4,t}$ *with* $\tilde{\sigma}^2$ *and zero, respectively, in* $\mathbf{U}_{\epsilon,t}$*, and*

$$\mathbf{H}_t\equiv\mathbf{\Sigma}_{\mathbf{\Lambda}_{t-1}}^{\frac{1}{2}}\mathbf{\Upsilon}_t\mathbf{U}_t^{-\frac{1}{2}}.$$

(iii) When (1) *does not hold, and* $|\tilde{S}_{t+1|t}|\geq CN^{-1/2+c}$ *a.s. for some* $C,c>0$*,*

$$\sqrt{N}\tilde{S}_{t+1|t}\rightarrow_p\pm\infty.$$

For the case of no observed risk factor being assumed, implying that the bias-adjustment only involves r_t latent risk factors,

$$x_{it}=\alpha_{it-1}+\boldsymbol{\lambda}'_{it-1}\mathbf{F}_t+e_{it}\text{ for every }t\in\mathcal{Z}\text{ and }i\in\mathcal{N}, \tag{48}$$

estimated local-PCA as explained above, the above formulae simplify to $\tilde{S}^{\texttt{miss}}_{t+1|t}\equiv\frac{1}{N}\sum_{i=1}^{N}\hat{\delta}_{i,t+1|t}-\frac{1}{2r_{ft}}\left(1-\frac{(\tilde{\mathbf{F}}'_{t+1}\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t)}{(\tilde{\mathbf{F}}'_t\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t)}\right)(\tilde{\boldsymbol{\mu}}'_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_{t+1})$, with limiting distribution

$$\sqrt{N}\tilde{S}^{\texttt{miss}}_{t+1|t}\rightarrow_d N\left(0,a_{Lt}^2[\sigma_t^2(\boldsymbol{\iota}'_{\mathbf{01}}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}})(1-\boldsymbol{\mu}'_{\mathbf{\Lambda}_{t-1}}\mathbf{\Sigma}^{-1}_{\mathbf{\Lambda}_{t-1}}\mu_{\mathbf{\Lambda}_{t-1}})+\mathbf{A}'_{L1}\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}]\right), \tag{49}$$

which is derived from Theorem 7 but omitting $\mathbf{M}_{f^o}$ entirely, where $a_{Lt}\equiv\frac{1}{2r_{ft}}\left(1-\frac{(\tilde{\mathbf{F}}'_{t+1}\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t)}{(\tilde{\mathbf{F}}'_t\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t)}\right)$, implying

$$\mathbf{A}_{L1}\equiv\frac{1}{(T+1)}(\mathbf{F}\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t}\otimes-\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}). \tag{50}$$

where now $\mathbf{U}_t$ is the diagonal matrix of the r_t eigenvalues of $\hat{\Sigma}_{\mathbf{F}}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}$, with eigenvector matrix $\mathbf{\Upsilon}_t$. Moreover, $\tilde{\mathbf{e}}\equiv\mathbf{X}-\tilde{\mathbf{F}}\tilde{\mathbf{\Lambda}}'$, $\tilde{\mathbf{U}}$ being the r_t largest eigenvalues of $(\mathbf{X}\mathbf{X}'/((T+1)N)-(\tilde{\sigma}^2/(T+1))\mathbf{I}_{T+1})$, and

$$\tilde{a}_{Lt}\equiv\frac{1}{2r_{ft}}\left(1-\frac{(\tilde{\mathbf{F}}'_{t+1}\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t)}{(\tilde{\mathbf{F}}'_t\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\tilde{\mathbf{F}}_t)}\right)\rightarrow_p a_{Lt}, \tag{51}$$

$$\tilde{\mathbf{A}}_{L1}\equiv\frac{1}{(T+1)}(\tilde{\mathbf{F}}\tilde{\mathbf{U}}_t^{-1}\tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}}\otimes-\mathbf{M}_{\tilde{\mathbf{F}}}\boldsymbol{\iota}_{\mathbf{01}})\rightarrow_p\mathbf{A}_{L1}, \tag{52}$$

$$\tilde{\sigma}^2\equiv tr(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N\rightarrow_p\sigma_t^2, \tag{53}$$

$$\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\equiv\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}/N-\tilde{\sigma}^2/(T+1)\mathbf{I}_{r_t}\rightarrow_p\mathbf{H}_t^{-1}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{H}_t^{-1\prime},\tilde{\boldsymbol{\mu}}_{\lambda}\rightarrow_p\mathbf{H}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t}, \tag{54}$$

with all the other definitions still standing.

6 Cross Sectional Asset Pricing When $T = r_t + 1$ - Correct Specification

In this section, I consider the case in which a multi-factor model is specified from the outset, replacing (1) with (25):

$$x_{it} = \alpha_{it-1} + \lambda_{it-1} f_t^o + \boldsymbol{\lambda}_{it-1}' \mathbf{F}_t + e_{it} \quad \text{for every } t \in \mathcal{Z} \text{ and } i \in \mathcal{N}.$$

This corresponds to the standard framework in empirical asset pricing—a correctly specified multi-factor model. The key difference with the common practice is that I estimate the model using a very small number of time observations ($T = r_t + 1$) in order to capture time variation in the parameters within a conditional asset pricing setting. In addition, at least one further time observation is required for out-of-sample (OOS) analysis.

As before, I begin with the case of one observed risk factor, f_t^o, and $r_t - 1$ latent risk factors, $\mathbf{F}_t$, and then consider the case in which all risk factors are latent. A third case is also of interest: a multi-factor model spanned exclusively by observed risk factors. This situation arises, for example, when assessing the pricing performance of the Fama and French (1993) three-factor model, among many other models proposed in the literature. Once again, however, the emphasis is on evaluating the performance of such well-known factor models from a conditional asset pricing perspective.

6.1 Pricing Ability of Risk Premia When $T = r_t + 1$ - Correct Specification

Given that I now assume the r_t-factor model (25) from the outset, at least r_t time observations are required for estimation. However, as I now explain, I use $T = r_t + 1$ periods for estimation. Indeed, when only r_t periods are employed, the estimated factor model attains a perfect fit, namely

$$x_{it} = \tilde{\lambda}_i^{o,\dagger} f_t^o + \tilde{\boldsymbol{\lambda}}_i^{\dagger\prime} \tilde{\mathbf{F}}_t^{\dagger},$$

where the dagger superscript $(.)^{\dagger}$ denotes estimates obtained using $T = r_t$ periods. Consequently, the pricing errors for model (25) satisfy

$$\tilde{h}_{it+1|t}^{\dagger} \equiv x_{it+1} - \tilde{\lambda}_i^{o,\dagger} f_t^o - \tilde{\boldsymbol{\lambda}}_i^{\dagger\prime} \tilde{\mathbf{F}}_t^{\dagger} = x_{it+1} - x_{it}, \tag{55}$$

which coincides with (4), i.e., the case $T = 1$ under the single-factor model (1).

By contrast, if pricing errors are constructed using $T = r_t + 1$ observations for estimation, yielding

$$\tilde{h}_{it+1|t} \equiv x_{it+1} - \tilde{\lambda}_i^o f_t^o - \tilde{\boldsymbol{\lambda}}_i' \tilde{\mathbf{F}}_t, \tag{56}$$

then $\tilde{h}_{it+1|t}$ now reflects the r_t-factor structure. Important special cases include: (i) all factors latent, obtained by omitting f_t^o; and (ii) all factors observed, again omitting f_t^o and replacing the local-PCA estimates $\tilde{\mathbf{F}}_t$ with the r_t-dimensional vector of observed risk factors $\mathbf{F}_t^o$, popular examples being the three- or five-factor specifications of Fama and French (1993, 2015). Setting $T = r_t > 1$ also introduces additional complications (which do not arise when $r_t = 1$) in the construction of mean–variance portfolios. For example, the sample covariance matrix of the r_t risk factors—required to be invertible for portfolio construction—is singular, as discussed below.

Since more than one period of data is available for IS estimation, one may also consider pricing errors based on estimated expected returns:

$$\tilde{h}_{it+1|t}^{ave} \equiv x_{it+1} - \tilde{\lambda}_i^o \bar{f}^o - \tilde{\boldsymbol{\lambda}}_i' \bar{\tilde{\mathbf{F}}}, \tag{57}$$

where $\bar{f}^o \equiv T^{-1} \sum_{s=t-T+1}^{t} f_s^o, \bar{\tilde{\mathbf{F}}} \equiv T^{-1} \sum_{s=t-T+1}^{t} f_s$ are the sample means computed over the T observations IS. This definition can be extended to the case where $T_{oos} > 1$ observations are used for OOS

evaluation, leading to the metric $\bar{x}_{i,OOS} - \tilde{\lambda}_i^o \bar{f}^o - \tilde{\boldsymbol{\lambda}}_i' \bar{\tilde{\mathbf{F}}}$, setting $\bar{x}_{i,OOS} \equiv T_{oos}^{-1} \sum_{s=T+1}^{T+T_{oos}} x_{is}$.

Again, it would be necessary to consider $T \geq r_t + 1$ because when $T = r_t$, the perfect fit implies that the model-implied estimator of expected returns coincides with the (local) sample mean of returns, i.e., $\tilde{\lambda}_i^o \bar{f}^{o,\dagger} + \tilde{\boldsymbol{\lambda}}_i^{\dagger\prime} \bar{\tilde{\mathbf{F}}}^\dagger = \bar{x}_i^\dagger \equiv T^{-1} \sum_{s=t-r_t+1}^{t} x_{is}$, so that $\tilde{h}_{it+1|t}^{ave} = x_{it+1} - \bar{x}_i^\dagger$. Hence, any combination of observed and latent risk factors yields identical pricing errors.[27] However, consistent with the overall analysis, I do not further develop the methodology based on $\tilde{h}_{it+1|t}^{ave}$, but instead focus on $\tilde{h}_{it+1|t}$, although admittily the former may also be of interest. The asymptotic analysis for $\tilde{h}_{it+1|t}^{ave}$ can be derived using the same arguments.[28]

Therefore, I focus on (56) except that I denote by $\tilde{\boldsymbol{\lambda}}_{-1,i}^o$ and $\tilde{\mathbf{F}}_{-1,t}$ the quantities estimated by OLS and local PCA, respectively, using $T = r_t + 1$ observations (the IS period), as in Section 5.1. Defining $l_{it+1} = (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{it-1})' \mathbf{F}_{t+1} + (\lambda_{it} - \lambda_{it-1}) f_{t+1}^o$, it follows that $\sum_{i=1}^N \tilde{h}_{it+1|t}/N \to_p \mu_{\lambda_t}(f_{t+1}^o - f_t^o) + \boldsymbol{\mu}_{\Lambda_t}'(\mathbf{F}_{t+1} - \mathbf{F}_t)$.

To construct the pricing test statistic, I re-estimate the model via local PCA using now $T+1 = r_t+2$ periods, so as to capture $\mathbf{F}_{t+1}$, obtaining $\tilde{\boldsymbol{\lambda}}_i$ and $\tilde{\mathbf{F}}_t$. Using both sets of local-PCA estimates—based on T and $T+1$ periods, respectively—to quantify the bias, I define the pricing test statistic as follows.

Theorem 8 (Pricing Test Cross Section - OOS - Correct Specification: $T = r_t + 1$**)**
Define

$$\tilde{L}_{t+1|t}^{corr} \equiv \sum_{i=1}^{N} h_{it+1|t}/N - (\tilde{\mu}_\lambda^o f_{t+1}^o - \tilde{\mu}_{-1,\lambda}^o f_t^o) - (\tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}}' \tilde{\mathbf{F}}_{t+1} - \tilde{\boldsymbol{\mu}}_{-1,\boldsymbol{\Lambda}}' \tilde{\mathbf{F}}_{-1,t}), \tag{58}$$

$$\text{with } \tilde{\mu}_\lambda^o \equiv \tilde{\boldsymbol{\lambda}}^{o\prime} \mathbf{1}_N/N \text{ and } \tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}' \mathbf{1}_N/N \text{ corresponding to } r_t + 2 \text{ observations}, \tag{59}$$

$$\text{and } \tilde{\mu}_{-1,\lambda}^o \equiv \tilde{\boldsymbol{\lambda}}_{-1}^{o\prime} \mathbf{1}_N/N \text{ and } \tilde{\boldsymbol{\mu}}_{-1,\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}_{-1}' \mathbf{1}_N/N \text{ corresponding to } r_t + 1 \text{ observations.} \tag{60}$$

(i) Under Assumptions 1 and 2-8,

$$\sqrt{N} \tilde{L}_{t+1|t}^{corr} \to_d N\left(0, \sigma_t^2 (\boldsymbol{\iota}_{\mathbf{01}}' \mathbf{M}_{f^o} \mathbf{M}_{\mathbf{F}} \boldsymbol{\iota}_{\mathbf{01}})(1 - \boldsymbol{\mu}_{\boldsymbol{\Lambda}_t}' \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}^{-1} \mu_{\boldsymbol{\Lambda}_t}) + \mathbf{A}_{L1}' \mathbf{U}_{\epsilon,t} \mathbf{A}_{L1}\right), \tag{61}$$

setting

$$\mathbf{A}_{L1} \equiv \frac{1}{(T+1)} (\mathbf{F} \mathbf{U}_t^{-1} \boldsymbol{\mu}_{\boldsymbol{\Lambda}_t} \otimes -\mathbf{M}_{f^o} \mathbf{M}_{\mathbf{F}} \boldsymbol{\iota}_{\mathbf{01}}). \tag{62}$$

where $\boldsymbol{\iota}_{\mathbf{01}} \equiv (\mathbf{0}_T', 1)'$ *with* $T = r_t + 1$, $\mathbf{M}_{f^o} \equiv \mathbf{I}_{T+1} - \mathbf{f}^o \mathbf{f}^{o\prime}/(\mathbf{f}^{o\prime} \mathbf{f}^o)$, $\mathbf{M}_{\mathbf{F}} \equiv \mathbf{I}_{T+1} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\mathbf{U}_{\epsilon,t}$ *being the* $(T+1)^2 \times (T+1)^2$ *matrix with* $T = r_t + 1$ *defined in Section 10,* $\mathbf{U}_t$ *is the diagonal matrix of the* $r_t - 1$ *eigenvalues of* $\hat{\Sigma}_{\mathbf{F}} \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}$*, with eigenvector matrix* $\boldsymbol{\Upsilon}_t$*.*
(ii) Under the same assumptions, with $\kappa_{4t} = 0$*, setting* $\mathbf{M}_{\tilde{F}} \equiv \mathbf{I}_{T+1} - \tilde{\mathbf{F}}(\tilde{\mathbf{F}}'\tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}'$*,* $\tilde{\mathbf{e}} \equiv \mathbf{X}^* - \tilde{\mathbf{F}}\tilde{\boldsymbol{\Lambda}}'$*, and* $\tilde{\mathbf{U}}$ *being the* $r_t - 1$ *largest eigenvalues of* $(\mathbf{X}^*\mathbf{X}^{*\prime}/((T+1)N) - (\tilde{\sigma}^2/(T+1))\mathbf{I}_{T+1})$*,*

$$\tilde{\mathbf{A}}_{L1} \equiv \frac{1}{(T+1)} (\tilde{\mathbf{F}} \tilde{\mathbf{U}}^{-1} \tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}} \otimes -\mathbf{M}_{f^o} \mathbf{M}_{\tilde{\mathbf{F}}} \boldsymbol{\iota}_{\mathbf{01}}) \to_p \mathbf{A}_{L1}, \tag{63}$$

$$\tilde{\sigma}^2 \equiv tr(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \to_p \sigma_t^2, \ \tilde{\mathbf{U}}_\epsilon \to_p \mathbf{U}_{\epsilon,t}, \tag{64}$$

$$\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}' \tilde{\boldsymbol{\Lambda}}/N - \tilde{\sigma}^2/(T+1)\mathbf{I}_{r_t-1} \to_p \mathbf{H}_t^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t} \mathbf{H}_t^{-1\prime}, \tilde{\boldsymbol{\mu}}_\lambda \to_p \mathbf{H}_t^{-1} \boldsymbol{\mu}_{\boldsymbol{\Lambda}_t}, \tag{65}$$

[27] Model (25) is estimated without an intercept (i.e., without demeaning), so the residuals do not necessarily have zero sample mean.

[28] Details are available upon request.

where $\tilde{\mathbf{U}}_\epsilon$ *is obtained replacing* σ_t^2 *and* $\kappa_{4,t}$ *with* $\tilde{\sigma}^2$ *and zero, respectively, in* $\mathbf{U}_{\epsilon,t}$*, and*

$$\mathbf{H}_t \equiv \mathbf{\Sigma}_{\mathbf{\Lambda}_t}^{\frac{1}{2}} \mathbf{\Upsilon}_t \mathbf{U}_t^{-\frac{1}{2}}.$$

(iii) When (1) *does not hold, and* $|\tilde{L}_{t+1|t}| \geq CN^{-1/2+c}$ *a.s. for some* $C, c > 0$*,*

$$\sqrt{N}\tilde{L}_{t+1|t} \to_p \pm\infty.$$

Comparing the pricing test statistic $\tilde{L}_{t+1|t}^{\texttt{corr}}$ with $\tilde{L}_{t+1|t}^{\texttt{miss}}$, corresponding to the misspecification case, one can see the gain in efficiency stemming from positing the r_t-factor model from the outset rather than using it only to quantify the bias adjustment, as for $\tilde{L}_{t+1|t}^{\texttt{miss}}$. In fact, the asymptotic covariance matrix of $\tilde{L}_{t+1|t}^{\texttt{corr}}$ is smaller than the one for $\tilde{L}_{t+1|t}^{\texttt{miss}}$. This can be seen by comparing the two covariance matrices term by term. In particular, $\big(\boldsymbol{\iota}_{\mathbf{01}}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}\big)$ is smaller than $\big(\boldsymbol{\iota}_{\mathbf{0}-11}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}-11}\big)$, the former being approximately half of the latter.[29] A similar argument holds for the second term of the asymptotic covariance matrix, where the divergence can be even stronger due to the $1/T$ term in the case $\tilde{L}_{t+1|t}^{\texttt{miss}}$, which is larger than the $1/(T+1)$ term that enters the corresponding quantity in $\tilde{L}_{t+1|t}^{\texttt{corr}}$.

When no observed risk factor is posited, I can derive the asymptotics of the pricing test as a special case of Theorem 8 for the case when the r_t risk factors are all assumed latent, where

$$\tilde{h}_{it+1|t} = \boldsymbol{\lambda}_{it-1}'\mathbf{F}_{t+1} + e_{it+1} + l_{it+1} + \alpha_{it} - \tilde{\boldsymbol{\lambda}}_{-1,i}'\tilde{\mathbf{F}}_{-1,t},$$

for an $r_t \times 1$ vector of latent risk factors $\mathbf{F}_t$, and a corresponding $r_t \times 1$ vector of risk exposures $\boldsymbol{\lambda}_{it-1}$, where now $l_{it+1} = (\boldsymbol{\lambda}_{it} - \boldsymbol{\lambda}_{it-1})'\mathbf{F}_{t+1}$, and $\sum_{i=1}^N \hat{h}_{it+1|t}/N \to_p \boldsymbol{\mu}_{\mathbf{\Lambda}_t}'(\mathbf{F}_{t+1} - \mathbf{F}_t)$.

As there is no observed risk factor, I can derive the local PCA for $\mathbf{X}$ directly from (32).[30] In particular, let $\tilde{\mathbf{F}} \equiv (\tilde{\mathbf{F}}_{t-T+1}, \ldots, \tilde{\mathbf{F}}_{t+1})'$ be the matrix of the r_t eigenvectors, multiplied by $\sqrt{T+1}$, corresponding to the largest r_t eigenvalues of $\mathbf{X}\mathbf{X}'/\big((T+1)N\big)$, satisfying $\tilde{\mathbf{F}}'\tilde{\mathbf{F}}/(T+1) = \mathbf{I}_{r_t}$, and define $\tilde{\mathbf{\Lambda}} \equiv \mathbf{X}'\tilde{\mathbf{F}}/(T+1)$.

The test statistic now takes the form

$$\tilde{L}_{t+1|t}^{\texttt{corr}} \equiv \frac{1}{N}\sum_{i=1}^N \tilde{h}_{i,t+1|t} - \big(\tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}}'\tilde{\mathbf{F}}_{t+1} - \tilde{\boldsymbol{\mu}}_{-1,\mathbf{\Lambda}}'\tilde{\mathbf{F}}_{-1,t}\big),$$

with limiting distribution

$$\sqrt{N}\,\tilde{L}_{t+1|t}^{\texttt{corr}} \to_d N\Big(0,\ \sigma_t^2\big(\boldsymbol{\iota}_{\mathbf{01}}'\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}\big)\big(1 - \boldsymbol{\mu}_{\mathbf{\Lambda}_{t-1}}'\mathbf{\Sigma}_{\mathbf{\Lambda}_{t-1}}^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_{t-1}}\big) + \mathbf{A}_{L1}'\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}\Big), \tag{66}$$

which is derived from Theorem 8 by omitting $\mathbf{M}_{f^o}$ entirely, implying

$$\mathbf{A}_{L1} \equiv \frac{1}{T+1}\big(\mathbf{F}\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t} \otimes -\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}\big), \tag{67}$$

[29] This can be seen by noting that $\big(\boldsymbol{\iota}_{\mathbf{01}}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}\big) = \sum_{t=1}^{T+1} v_t u_{t,T+1}^2$, where v_t and $\mathbf{u}_{T+1}$ are the eigenvalues and the $(T+1)$th eigenvector of the projection matrix $\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}$. As the eigenvalues satisfy $v_t = 1$ for $1 \leq t \leq T-1$, then $\sum_{t=1}^{T+1} v_t u_{t,T+1}^2 = \sum_{t=1}^{T-1} v_t u_{t,T+1}^2 \approx \sum_{t=1}^{T-1} u_{t,T+1}^2 \approx 1$, the approximation improving with T. On the other hand, $\big(\boldsymbol{\iota}_{\mathbf{0}-11}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}-11}\big) = \sum_{t=1}^{T+1} v_t\big(u_{t,T+1} - u_{t,T}\big)^2 = \sum_{t=1}^{T-1} u_{t,T+1}^2 + \sum_{t=1}^{T-1} u_{t,T}^2 - 2\sum_{t=1}^{T-1} u_{t,T+1}u_{t,T} \approx 2$, since $\sum_{t=1}^{T-1} u_{t,T+1}u_{t,T} \approx 0$ by the mutual orthogonality of the eigenvectors (it is identically zero when summing over $1 \leq t \leq T+1$), and the first two sums are each approximately equal to one. These approximations improve with T.

[30] See footnote 14.

where $\mathbf{U}_t$ is the diagonal matrix of the r_t eigenvalues of $\hat{\Sigma}_{\mathbf{F}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}$, with eigenvector matrix $\boldsymbol{\Upsilon}_t$. Moreover, let $\tilde{\mathbf{e}} \equiv \mathbf{X} - \tilde{\mathbf{F}}\tilde{\boldsymbol{\Lambda}}'$, and let $\tilde{\mathbf{U}}$ be the r_t largest eigenvalues of $\frac{\mathbf{X}\mathbf{X}'}{(T+1)N} - \frac{\tilde{\sigma}^2}{T+1}\mathbf{I}_{T+1}$, and define

$$\tilde{\mathbf{A}}_{L1} \equiv \frac{1}{T+1}\big(\tilde{\mathbf{F}}\tilde{\mathbf{U}}_t^{-1}\tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}} \otimes -\mathbf{M}_{\tilde{\mathbf{F}}}\boldsymbol{\iota}_{\mathbf{0}1}\big) \;\to_p\; \mathbf{A}_{L1}, \tag{68}$$

$$\tilde{\sigma}^2 \equiv \mathrm{tr}(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \;\to_p\; \sigma_t^2, \tag{69}$$

$$\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}'\tilde{\boldsymbol{\Lambda}}/N - \frac{\tilde{\sigma}^2}{T+1}\mathbf{I}_{r_t} \;\to_p\; \mathbf{H}_t^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\mathbf{H}_t^{-1\prime}, \tag{70}$$

$$\tilde{\boldsymbol{\mu}}_{\lambda} \to_p \mathbf{H}_t^{-1}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t}, \tag{71}$$

with all the other definitions still standing.

Finally, as another important special case of Theorem 8, I consider the case when the r_t risk factors $\mathbf{F}_t^o$ are assumed observed. In this case, the test statistic takes the form

$$\tilde{L}_{t+1|t}^{o\text{-}\mathtt{corr}} \equiv \frac{1}{N}\sum_{i=1}^{N}\tilde{h}_{i,t+1|t} - \big(\tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}}^{o\prime}\mathbf{F}_{t+1}^o - \tilde{\boldsymbol{\mu}}_{-1,\boldsymbol{\Lambda}}^{o\prime}\mathbf{F}_t^o\big),$$

with limiting distribution

$$\sqrt{N}\,\tilde{L}_{t+1|t}^{o\text{-}\mathtt{corr}} \;\to_d\; N\Big(0,\; \sigma_t^2\big(\boldsymbol{\iota}_{\mathbf{0}1}'\mathbf{M}_{\mathbf{F}^o}\boldsymbol{\iota}_{\mathbf{0}1}\big)\Big). \tag{72}$$

This is derived again from Theorem 8, but now omitting $\mathbf{M}_{f^o}$ and imposing no estimation uncertainty on the risk factors (and setting the rotation matrix equal to the identity matrix of size $r_t \times r_t$). This result is very useful, as one can now test the (time-varying) pricing ability of the various multifactor asset pricing models proposed in the empirical asset pricing literature for the case of observed risk factors, starting with the three-factor model of Fama and French (1993).[31]

6.2 Mean-Variance Portfolios When $T = r_t + 1$ - Correct Specification

.

The investor is aware that the multi-factor model (25) holds. As before, I first consider the case in which the factor space consists of a single observed factor f_t^o and $r_t - 1$ latent factors $\mathbf{F}_t$, and then discuss the special cases in which the r_t risk factors are either all latent or all observed.

I construct mean–variance portfolios with $T = r_t + 1 > 1$ using the local-PCA estimator, after

[31] Among many others XXX.

orthogonalizing with respect to the observed risk factor f_t^o. This yields[32]

$$\tilde{\mathbf{w}} \equiv \frac{1}{\tilde{\delta}} \tilde{\mathbf{B}} \left(\tilde{\mathbf{B}}' \tilde{\mathbf{B}} \right)^{-1} \begin{pmatrix} \bar{f}^o \\ \overline{s^2_{f^o}} \\ \tilde{\mathbf{S}}_F^{-1} \bar{\tilde{\mathbf{F}}} \end{pmatrix}, \tag{73}$$

where

$$\hat{\mathbf{B}} \equiv \left(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\Lambda}} \right) = \left(\frac{\mathbf{X}' \mathbf{f}^o}{\mathbf{f}^{o\prime} \mathbf{f}^o}, \frac{\mathbf{X}' \tilde{\mathbf{F}}}{T} \right), \quad \bar{f}^o \equiv \frac{\mathbf{1}_T' \mathbf{f}^o}{T}, \quad s^2_{f^o} \equiv \frac{\mathbf{f}^{o\prime} \mathbf{M}_{1_T} \mathbf{f}^o}{T},$$

$$\bar{\tilde{\mathbf{F}}} \equiv \frac{\tilde{\mathbf{F}}' \mathbf{1}_T}{T}, \quad \tilde{\mathbf{S}}_F \equiv \frac{\tilde{\mathbf{F}}' \mathbf{M}_{1_T} \tilde{\mathbf{F}}}{T} = \mathbf{I}_{r_t - 1} - \bar{\tilde{\mathbf{F}}} \, \bar{\tilde{\mathbf{F}}}'.$$

As periods from $t-T+1$ to t, are used for estimation, where $T = r_t+1$, perfect fit is not achieved, namely $\mathbf{X} \neq \mathbf{f} \tilde{\boldsymbol{\lambda}}' + \tilde{\mathbf{F}} \tilde{\boldsymbol{\Lambda}}'$, and I can now estimate σ_t^2. However, this is not relevant for portfolio construction as one does not need to estimate σ_t^2 using the large-N approximation. Moreover, as I now have $T > 1$ periods, I can afford de-meaning the data to construct the sample variance and covariance matrix of the observed and latent factors, i.e., $s^2_{f^o}$ and $\tilde{\mathbf{S}}_F$ use de-meaned data for portfolio construction. I thus obtain a simplified large-N representation for the portfolio returns:

$$\begin{pmatrix} r_{t-T+1} \\ \vdots \\ r_t \\ r_{t+1} \end{pmatrix} = \begin{pmatrix} \mathbf{X} \\ \mathbf{x}_{t+1}' \end{pmatrix} \tilde{\mathbf{w}} = \begin{pmatrix} \left(\mathbf{f}^o, \tilde{\mathbf{F}} \right) \mathbf{c} + o_p(1) \\ \left(f^o_{t+1}, \mathbf{F}_{t+1}' \right) \mathbf{C} \mathbf{c} + o_p(1) \end{pmatrix}, \tag{74}$$

where

$$\mathbf{C} \equiv \begin{pmatrix} \dfrac{\tilde{\boldsymbol{\lambda}}^{o\prime} \tilde{\boldsymbol{\lambda}}^o}{N} - \dfrac{\tilde{\sigma}^2}{\mathbf{f}^{o\prime} \mathbf{f}^o} & \dfrac{\tilde{\boldsymbol{\lambda}}^{o\prime} \tilde{\boldsymbol{\Lambda}}}{N} \\ \dfrac{\tilde{\boldsymbol{\Lambda}}' \tilde{\boldsymbol{\lambda}}^o}{N} & \dfrac{\tilde{\boldsymbol{\Lambda}}' \tilde{\boldsymbol{\Lambda}}}{N} - \dfrac{\tilde{\sigma}^2}{T} \mathbf{I}_{r_t - 1} \end{pmatrix} \begin{pmatrix} \dfrac{\tilde{\boldsymbol{\lambda}}^{o\prime} \tilde{\boldsymbol{\lambda}}^o}{N} & \dfrac{\tilde{\boldsymbol{\lambda}}^{o\prime} \tilde{\boldsymbol{\Lambda}}}{N} \\ \dfrac{\tilde{\boldsymbol{\Lambda}}' \tilde{\boldsymbol{\lambda}}^o}{N} & \dfrac{\tilde{\boldsymbol{\Lambda}}' \tilde{\boldsymbol{\Lambda}}}{N} \end{pmatrix}^{-1}, \quad \mathbf{c} \equiv \begin{pmatrix} \bar{f}^o \\ \overline{s^2_{f^o}} \\ \tilde{\mathbf{S}}_F^{-1} \bar{\tilde{\mathbf{F}}} \end{pmatrix}.$$

[32]Mean–variance theory implies that the (conditional) population portfolio weights under (25) are

$$\mathbf{w}_t = \frac{1}{\delta} \mathbf{V}_t^{-1} \mathbf{B}_t \boldsymbol{\gamma}_t = \frac{1}{\delta} \mathbf{B}_t \left(\mathbf{B}_t' \mathbf{B}_t \right)^{-1} \begin{pmatrix} \mathbb{E}_t \left(f^o_{t+1} \right) \\ s^2_f \\ \mathbf{S}_F^{-1} \mathbb{E}_t \left(\mathbf{F}_{t+1} \right) \end{pmatrix} + o(1),$$

where the approximation holds for large N (the $o_p(1)$ term indicating that the same portfolio return is obtained asymptotically in N). Here,

$$\mathbf{V}_t = \mathbf{B}_t \begin{pmatrix} \sigma^2_{f^o} & \mathbf{0}' \\ \mathbf{0} & \boldsymbol{\Sigma}_F \end{pmatrix} \mathbf{B}_t' + \sigma_t^2 \mathbf{I}_N$$

is the (conditional) population covariance matrix of excess returns in (25), $\mathbf{B}_t \equiv (\boldsymbol{\lambda}_t, \boldsymbol{\Lambda}_t)$, and $\boldsymbol{\gamma}_t$ is the r_t-dimensional vector of risk premia associated with $(f_t^o, \mathbf{F}_t')'$ (i.e., equal to their conditional means; see footnote 6). Moreover, $\sigma^2_{f^o}$ and $\boldsymbol{\Sigma}_F$ denote the population variance and covariance matrix of f_t^o and $\mathbf{F}_t$, respectively, and δ is the (constant) risk-aversion parameter.

When no observed risk factor is present, one obtains

$$\begin{pmatrix} r_{t-T+1} \\ \vdots \\ r_t \\ r_{t+1} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{F}}\,\tilde{\mathbf{S}}_F^{-1}\bar{\tilde{\mathbf{F}}} + o_p(1) \\ \mathbf{F}'_{t+1}\left(\frac{\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}}{N} - \frac{\tilde{\sigma}^2}{T}\mathbf{I}_{r_t}\right)^{-1}\left(\frac{\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}}{N}\right)\tilde{\mathbf{S}}_F^{-1}\bar{\tilde{\mathbf{F}}} + o_p(1) \end{pmatrix}, \tag{75}$$

whereas in the case of no latent factors, with $r_t > 1$ observed risk factors, the expression becomes

$$\begin{pmatrix} r_{t-T+1} \\ \vdots \\ r_t \\ r_{t+1} \end{pmatrix} = \begin{pmatrix} \mathbf{F}^o\mathbf{S}_{F^o}^{-1}\bar{\mathbf{F}}^o + o_p(1) \\ \mathbf{F}^{o\prime}_{t+1}\left(\frac{\tilde{\mathbf{\Lambda}}^{o\prime}\tilde{\mathbf{\Lambda}}^o}{N} - \frac{\tilde{\sigma}^2}{T}\mathbf{I}_{r_t}\right)^{-1}\left(\frac{\tilde{\mathbf{\Lambda}}^{o\prime}\tilde{\mathbf{\Lambda}}^o}{N}\right)\mathbf{S}_{F^o}^{-1}\bar{\mathbf{F}}^o + o_p(1) \end{pmatrix}, \tag{76}$$

where now all r_t factors are observed and $\tilde{\mathbf{\Lambda}}^o \equiv \mathbf{X}'\mathbf{F}^o\left(\mathbf{F}^{o\prime}\mathbf{F}^o\right)^{-1}$.[33]

For each of the above cases, one can compute the sample Sharpe ratio $\widetilde{SR}_{1:T,t+1} \equiv \bar{r}/s_r$, where $\bar{r} \equiv \frac{(r_{t-T+1},\ldots,r_t,r_{t+1})\mathbf{1}_T}{T}$, $s_r^2 \equiv \frac{(r_{t-T+1},\ldots,r_t,r_{t+1})\mathbf{M}_{1_T}(r_{t-T+1},\ldots,r_t,r_{t+1})'}{T}$. Its limiting distribution, along with corresponding standard errors, can be derived using the asymptotic analysis in Theorems 3 and 6.[34]

As expected, the IS return has always a positive valued SR, as it is equals to the square root of the sum of two quadratic form.[35]. More interestingly, as for the previous cases, one can deduce the conditions for positivity of the OOS returns (and SR), namely that the OOS return is aligned with the IS means. The attenuation bias also arises, given by the positive definite matrix $\tilde{\mathbf{C}}$, which can be shown to have eigenvalues all smaller than unity.

6.3 Pricing Ability of SDF When $T = r_t + 1$ - Correct Specification

As the investor now estimate the multi-factor asset pricing model (25) holds, I re-define the projected-version SDF $M_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\delta\mathbf{w}_t'(\mathbf{x}_{t+1} - \mathbf{\Lambda}_t\boldsymbol{\gamma}_t)$, i.e., by re-centering its mean to $1/r_{ft}$,unlike Section 5.3, as I can now afford to de-mean the data when estimating the SDF. Correspondingly, the implied estimator for $M_{t,t+1}$ would be $\tilde{M}_{s-1,s} = \frac{1}{r_{fs}} - \frac{1}{r_{fs}}\delta\tilde{\mathbf{w}}'(\mathbf{x}_s - \tilde{\boldsymbol{\lambda}}^o\bar{f}^o - \tilde{\mathbf{\Lambda}}\bar{\tilde{\mathbf{F}}})$, with $t-T+1 \leq s \leq t$ for the IS periods and $s = t+1$ for the OOS period, and $\tilde{\mathbf{w}}$ from (73), estimated using the T IS periods.

I can construct the SDF-induced pricing errors by averaging across the $(T+1)$ periods the vector of excess returns with , $\boldsymbol{\delta}_{t+1|t} = (T+1)^{-1}\sum_{s=t-T+1}^{t+1}\tilde{M}_{s-1,s}\mathbf{x}_s'$, leading to the average pricing errors $\mathbf{1}_N'\boldsymbol{\delta}_{t+1|t}/N$. It is useful to distinguish between the IS and the OOS part of the average pricing errors. For the IS portion of $\boldsymbol{\delta}_{t+1|t}$ I obtain, setting $r_{fs} = r_f$ constant across the T periods to simplify the analysis,

$$\frac{1}{T+1}\sum_{s=t-T+1}^{t}\tilde{M}_{s-1,s}\mathbf{x}_s' = -\frac{1}{T+1}\sum_{s=t-T+1}^{t}\frac{1}{r_f}(\bar{f}^o, \bar{\tilde{\mathbf{F}}}')\begin{bmatrix} 1/s_{f^o}^2 & \mathbf{0}_{r_t-1}' \\ \mathbf{0}_{r_t-1} & \tilde{\mathbf{S}}_F^{-1} \end{bmatrix}(\tilde{\mathbf{B}}'\tilde{\mathbf{B}})^{-1}\tilde{\mathbf{B}}'\tilde{\mathbf{e}}_s\tilde{\mathbf{e}}_s', \tag{77}$$

because $\sum_{s=t-T+1}^{t}\tilde{\mathbf{e}}_s = \mathbf{0}_N$ and $\sum_{s=t-T+1}^{t}\tilde{\mathbf{e}}_s[f_s^0, \tilde{\mathbf{F}}_s'] = \mathbf{0}_{N\times r_t}$. Finally, given $\tilde{\mathbf{e}}_s = \mathbf{B}[f_s^0, \mathbf{F}_s']' + \mathbf{e}_s -$

[33]This case can be viewed as a multi-factor generalization of the setting analyzed in Section 4.2.1.

[34]Detailed derivations are available upon request.

[35]Some care is needed as we impose orthogonality of the risk factors *before* de-meaning.

$\tilde{\mathbf{B}}[f_s^0, \tilde{\mathbf{F}}_s']'$ then $\tilde{\mathbf{B}}'\tilde{\mathbf{e}}_s/N = O_p(N^{-1/2})$ and $\tilde{\mathbf{e}}_s'\mathbf{1}_N/N = O_p(N^{-1/2})$, implying

$$\frac{1}{T+1}\sum_{s=t-T+1}^{t}\tilde{M}_{s-1,s}\mathbf{x}_s'\mathbf{1}_N/N = O_p(N^{-1}). \tag{78}$$

Therefore no bias adjustment is required for the IS part of the SDF average pricing errors, which in fact does not contribute to the limiting distribution of the SDF pricing test.

Next the OOS part of the average pricing errors is given by

$$\frac{1}{T+1}\tilde{M}_{t,t+1}\mathbf{x}_{t+1}' = \tag{79}$$

$$\frac{1}{T+1}\frac{1}{r_f}\left(1-(\bar{f}^o,\bar{\tilde{\mathbf{F}}}')\begin{bmatrix}1/s_{f^o}^2 & \mathbf{0}'\\ \mathbf{0} & \tilde{\mathbf{S}}_F^{-1}\end{bmatrix}(\tilde{\mathbf{B}}'\tilde{\mathbf{B}})^{-\mathbf{1}}\tilde{\mathbf{B}}'[\mathbf{B}\begin{bmatrix}f_{t+1}^0\\ \mathbf{F}_{t+1}\end{bmatrix}+\mathbf{e_{t+1}}-\tilde{\mathbf{B}}\begin{bmatrix}\bar{f}^o\\ \bar{\tilde{\mathbf{F}}}\end{bmatrix}]\right)([f_{t+1}^o,\mathbf{F}_{t+1}']\mathbf{B}'+\mathbf{e}_{t+1}') \tag{80}$$

$$\tag{81}$$

implying

$$\frac{1}{T+1}\tilde{M}_{t,t+1}\mathbf{x}_{t+1}'\mathbf{1}_N/N \to_p \frac{1}{T+1}M_{t,t+1}(\mu_{\lambda_t}f_{t+1}^o+\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t}'\mathbf{F}_{t+1}), \tag{82}$$

setting $M_{t,t+1} \equiv \frac{1}{r_f}\left(1-\mathbf{c}_t'\left\{\mathbf{C}_t'\begin{bmatrix}f_{t+1}^0\\ \mathbf{F}_{t+1}\end{bmatrix}-\begin{bmatrix}\bar{f}^0\\ \bar{\mathbf{F}}\end{bmatrix}\right\}\right)$, where

$$\mathbf{C}_t \equiv \begin{pmatrix}\sigma_{\lambda_t} & \boldsymbol{\Sigma}_{\lambda_t\boldsymbol{\Lambda}_t}'\\ \boldsymbol{\Sigma}_{\lambda_t\boldsymbol{\Lambda}_t} & \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}\end{pmatrix}\begin{pmatrix}\sigma_{\lambda_t}+\sigma_t^2/(\mathbf{f}^{o\prime}\mathbf{f}^o) & \boldsymbol{\Sigma}_{\lambda_t\boldsymbol{\Lambda}_t}'\\ \boldsymbol{\Sigma}_{\lambda_t\boldsymbol{\Lambda}_t} & \boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}+\frac{\sigma_t^2}{T}\mathbf{I}_{r_t-1}\end{pmatrix}^{-1} \text{ and } \mathbf{c}_t \equiv \begin{pmatrix}\frac{\bar{f}^o}{s_{f^o}^2}\\ \mathbf{S}_F^{-1}\bar{\mathbf{F}}\end{pmatrix}. \tag{83}$$

In summary, I can estimate the SDF $\tilde{M}_{s-1,s}$ using T periods, but use $T+1$ periods (adding the OOS observation) to estimate the test statistics and to conduct inference.

Theorem 9 (Pricing Test Cross SDF- OOS - Correct Specification: $T = r_t + 1$)
Define

$$\tilde{S}_{t+1|t}^{corr} \equiv \frac{1}{N}\sum_{i=1}^{N}\tilde{\delta}_{i,t+1|t} - \frac{\tilde{M}_{t,t+1}}{T+1}(\tilde{\mu}_\lambda^o f_{t+1}^o + \tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}}'\tilde{\mathbf{F}}_{t+1}), \text{ with } \tilde{\mu}_\lambda^o \equiv \tilde{\boldsymbol{\lambda}}^{o\prime}\mathbf{1}_N/N \text{ and } \tilde{\boldsymbol{\mu}}_{\boldsymbol{\Lambda}} \equiv \tilde{\boldsymbol{\Lambda}}'\mathbf{1}_N/N. \tag{84}$$

(i) Under Assumptions 1 and 2-8,

$$\sqrt{N}\tilde{S}_{t+1|t}^{corr} \to_d N\left(0, a_{Lt}^2[\sigma_t^2(\boldsymbol{\iota}_{\mathbf{0}1}'\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}1})(1-\boldsymbol{\mu}_{\boldsymbol{\Lambda}_{t-1}}'\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_{t-1}}^{-1}\mu_{\boldsymbol{\Lambda}_{t-1}}) + \mathbf{A}_{L1}'\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}]\right), \tag{85}$$

setting

$$a_{Lt} \equiv \frac{1}{(T+1)r_{ft}}\left(1-\mathbf{c}_t'\left\{\mathbf{C}_t\begin{bmatrix}f_{t+1}^0\\ \mathbf{F}_{t+1}\end{bmatrix}-\begin{bmatrix}\bar{f}^0\\ \bar{\mathbf{F}}\end{bmatrix}\right\}\right), \tag{86}$$

$$\mathbf{A}_{L1} \equiv \frac{1}{(T+1)}(\mathbf{F}\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\boldsymbol{\Lambda}_t}\otimes -\mathbf{M}_{f^o}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{0}1}). \tag{87}$$

where $\boldsymbol{\iota}_{\mathbf{0}1} \equiv (\mathbf{0}_T', 1)'$, $\mathbf{M}_{\mathbf{F}} \equiv \mathbf{I}_{T+1} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\mathbf{U}_{\epsilon,t}$ *being the* $(T+1)^2 \times (T+1)^2$ *matrix with* $T = r_t + 1$ *defined in Section 10,* $\mathbf{U}_t$ *is the diagonal matrix of the* $r_t - 1$ *eigenvalues of* $\hat{\Sigma}_{\mathbf{F}}\boldsymbol{\Sigma}_{\boldsymbol{\Lambda}_t}$*, with eigenvector matrix* $\boldsymbol{\Upsilon}_t$*, and* $\mathbf{C}_t, \mathbf{c}_t$ *are defined in* (83).

(ii) Under the same assumptions, with $\kappa_{4t} = 0$, setting $\mathbf{M}_{\tilde{F}} \equiv \mathbf{I}_{T+1} - \tilde{\mathbf{F}}(\tilde{\mathbf{F}}'\tilde{\mathbf{F}})^{-1}\tilde{\mathbf{F}}'$, $\tilde{\mathbf{e}} \equiv \mathbf{X}^* - \tilde{\mathbf{F}}\tilde{\mathbf{\Lambda}}'$, and $\tilde{\mathbf{U}}$ being the $r_t - 1$ largest eigenvalues of $(\mathbf{X}^*\mathbf{X}^{*\prime}/((T+1)N) - (\tilde{\sigma}^2/(T+1))\mathbf{I}_{T+1})$,

$$\tilde{a}_{Lt} \equiv \frac{1}{(T+1)r_{ft}}\left(1 - \tilde{\mathbf{c}}'\left\{\tilde{\mathbf{C}}'\begin{bmatrix} f^0_{t+1} \\ \tilde{\mathbf{F}}_{t+1}\end{bmatrix} - \begin{bmatrix}\bar{f}^0 \\ \bar{\tilde{\mathbf{F}}}\end{bmatrix}\right\}\right) \to_p a_{Lt}, \tag{88}$$

$$\tilde{\mathbf{A}}_{L1} \equiv \frac{1}{(T+1)}(\tilde{\mathbf{F}}\tilde{\mathbf{U}}^{-1}\tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}} \otimes -\mathbf{M}_{\tilde{\mathbf{F}}}\mathbf{M}_{f^o}\boldsymbol{\iota}_{\mathbf{01}}) \to_p \mathbf{A}_{L1}, \tag{89}$$

$$\tilde{\sigma}^2 \equiv tr(\tilde{\mathbf{e}}\tilde{\mathbf{e}}')/N \to_p \sigma_t^2,\ \tilde{\mathbf{U}}_\epsilon \to_p \mathbf{U}_{\epsilon,t}, \tag{90}$$

$$\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}} \equiv \tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}/N - (\tilde{\sigma}^2/T)\mathbf{I}_{r_t-1} \to_p \mathbf{H}_t^{-1}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\mathbf{H}_t^{-1\prime}, \tilde{\boldsymbol{\mu}}_\lambda \to_p \mathbf{H}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t}, \tag{91}$$

$$\tilde{\mathbf{\Sigma}}_{\lambda\mathbf{\Lambda}} \equiv \tilde{\mathbf{\Lambda}}'\tilde{\boldsymbol{\lambda}}/N \to_p \mathbf{\Sigma}_{\lambda_t\mathbf{\Lambda}_t}, \tilde{\sigma}_\lambda \equiv \tilde{\boldsymbol{\lambda}}'\tilde{\boldsymbol{\lambda}}/N - \tilde{\sigma}^2/(\mathbf{f}^{o\prime}\mathbf{f}^o) \to_p \sigma_{\lambda_t}, \tag{92}$$

where $\tilde{\mathbf{U}}_\epsilon$ is obtained replacing σ_t^2 and $\kappa_{4,t}$ with $\tilde{\sigma}^2$ and zero, respectively, in $\mathbf{U}_{\epsilon,t}$,

$$\tilde{\mathbf{C}} \equiv \begin{pmatrix} \tilde{\sigma}_\lambda & \tilde{\mathbf{\Sigma}}'_{\lambda\mathbf{\Lambda}} \\ \tilde{\mathbf{\Sigma}}_{\lambda\mathbf{\Lambda}} & \tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\end{pmatrix}\begin{pmatrix} \tilde{\sigma}_\lambda + \tilde{\sigma}^2/(\mathbf{f}^{o\prime}\mathbf{f}^o) & \tilde{\mathbf{\Sigma}}'_{\lambda\mathbf{\Lambda}} \\ \tilde{\mathbf{\Sigma}}_{\lambda\mathbf{\Lambda}} & \tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}} + \frac{\tilde{\sigma}^2}{T}\mathbf{I}_{r_t-1}\end{pmatrix}^{-1} \text{ and } \tilde{\mathbf{c}} \equiv \begin{pmatrix} \frac{\bar{f}^o}{s^2_{f^o}} \\ \tilde{\mathbf{S}}_F^{-1}\bar{\tilde{\mathbf{F}}}\end{pmatrix},$$

and

$$\mathbf{H}_t \equiv \mathbf{\Sigma}_{\mathbf{\Lambda}_t}^{\frac{1}{2}}\mathbf{\Upsilon}_t\mathbf{U}_t^{-\frac{1}{2}}.$$

(iii) When (1) does not hold, and $|\tilde{S}_{t+1|t}| \geq CN^{-1/2+c}$ a.s. for some $C, c > 0$,

$$\sqrt{N}\tilde{S}_{t+1|t} \to_p \pm\infty.$$

For the case in which no risk factors are observed, so that the bias adjustment only depends on r_t latent risk factors, the expressions above simplify by omitting all references to f_t^o—in particular $\mathbf{M}_{f^o}$—and by replacing $r_t - 1$ with r_t everywhere. This yields

$$\tilde{S}^{\texttt{corr}}_{t+1|t} \equiv \frac{1}{N}\sum_{i=1}^N \tilde{\delta}_{i,t+1|t} - \frac{\tilde{M}_{t,t+1}}{T+1}\,\tilde{\boldsymbol{\mu}}'_{\mathbf{\Lambda}}\bar{\tilde{\mathbf{F}}}_{t+1}, \tag{93}$$

with $\tilde{M}_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\,\delta\,\tilde{\mathbf{w}}'\big(\mathbf{x}_{t+1} - \tilde{\mathbf{\Lambda}}\bar{\tilde{\mathbf{F}}}\big)$, using $\tilde{\mathbf{w}} \equiv \frac{1}{\delta}\tilde{\mathbf{\Lambda}}\big(\tilde{\mathbf{\Lambda}}'\tilde{\mathbf{\Lambda}}\big)^{-1}\tilde{\mathbf{S}}_F^{-1}\bar{\tilde{\mathbf{F}}}$. I then obtain

$$\sqrt{N}\,\tilde{S}_{t+1|t} \to_d N\Big(0,\, a_{Lt}^2\big[\sigma_t^2\big(\boldsymbol{\iota}'_{\mathbf{01}}\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}\big)\big(1 - \boldsymbol{\mu}'_{\mathbf{\Lambda}_{t-1}}\mathbf{\Sigma}^{-1}_{\mathbf{\Lambda}_{t-1}}\mu_{\mathbf{\Lambda}_{t-1}}\big) + \mathbf{A}'_{L1}\mathbf{U}_{\epsilon,t}\mathbf{A}_{L1}\big]\Big), \tag{94}$$

where

$$a_{Lt} = \frac{1}{(T+1)\,r_{ft}}\left(1 - \mathbf{F}'_{t+1}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\big(\mathbf{\Sigma}_{\mathbf{\Lambda}_t} + \frac{\sigma_t^2}{T}\,\mathbf{I}_{r_t}\big)^{-1}\mathbf{S}_F^{-1}\bar{\mathbf{F}}\right), \tag{95}$$

$$\mathbf{A}_{L1} \equiv \frac{1}{T+1}\big(\mathbf{F}\,\mathbf{U}_t^{-1}\boldsymbol{\mu}_{\mathbf{\Lambda}_t} \otimes -\mathbf{M}_{\mathbf{F}}\boldsymbol{\iota}_{\mathbf{01}}\big). \tag{96}$$

Here $\boldsymbol{\iota}_{\mathbf{01}} \equiv (\mathbf{0}'_T, 1)'$, $\mathbf{M}_{\mathbf{F}} \equiv \mathbf{I}_{T+1} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$, $\mathbf{U}_{\epsilon,t}$ being the $(T+1)^2 \times (T+1)^2$ matrix with $T = r_t + 1$ defined in Section 10, and $\mathbf{U}_t$ is the diagonal matrix of the r_t eigenvalues of $\hat{\Sigma}_{\mathbf{F}}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}$, with eigenvector matrix $\mathbf{\Upsilon}_t$.

Moreover, defining $\mathbf{M}_{\tilde{F}} \equiv \mathbf{I}_{T+1} - \tilde{\mathbf{F}}\big(\tilde{\mathbf{F}}'\tilde{\mathbf{F}}\big)^{-1}\tilde{\mathbf{F}}'$, $\quad \tilde{\mathbf{e}} \equiv \mathbf{X} - \tilde{\mathbf{F}}\tilde{\mathbf{\Lambda}}'$, and letting $\tilde{\mathbf{U}}$ be the r_t largest

eigenvalues of $\left(\frac{\mathbf{X}\mathbf{X}'}{(T+1)N} - \frac{\tilde{\sigma}^2}{T+1}\,\mathbf{I}_{T+1}\right)$. Then

$$\tilde{a}_{Lt} = \frac{1}{(T+1)\,r_{ft}}\left(1 - \tilde{\mathbf{F}}'_{t+1}\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}}\big(\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}} + \frac{\tilde{\sigma}^2}{T}\,\mathbf{I}_{r_t}\big)^{-1}\right) = \frac{\tilde{M}_{t,t+1}}{T+1}. \to_p a_{Lt}, \tag{97}$$

$$\tilde{\mathbf{A}}_{L1} \equiv \frac{1}{T+1}\big(\tilde{\mathbf{F}}\,\tilde{\mathbf{U}}^{-1}\tilde{\boldsymbol{\mu}}_{\mathbf{\Lambda}} \otimes -\mathbf{M}_{\tilde{\mathbf{F}}}\boldsymbol{\iota}_{\mathbf{01}}\big) \to_p \mathbf{A}_{L1}, \tag{98}$$

where all other definitions remain unchanged.

As a final important special case, consider the situation in which all r_t risk factors are observed. In this case,

$$\tilde{S}^{\texttt{o-corr}}_{t+1|t} \equiv \frac{1}{N}\sum_{i=1}^{N}\tilde{\delta}_{i,t+1|t} - \frac{\tilde{M}_{t,t+1}}{T+1}\big(\tilde{\boldsymbol{\mu}}^{o\prime}_{\mathbf{\Lambda}}\mathbf{F}^{o}_{t+1}\big), \tag{99}$$

with $\tilde{M}_{t,t+1} = \frac{1}{r_{ft}} - \frac{1}{r_{ft}}\,\delta\,\tilde{\mathbf{w}}'\big(\mathbf{x}_{t+1} - \tilde{\mathbf{\Lambda}}^{o}\bar{\mathbf{F}}^{o}\big)$, using $\tilde{\mathbf{w}} \equiv \frac{1}{\delta}\tilde{\mathbf{\Lambda}}^{o}\big(\tilde{\mathbf{\Lambda}}^{o\prime}\tilde{\mathbf{\Lambda}}^{o}\big)^{-1}\mathbf{S}^{-1}_{F^o}\bar{\mathbf{F}}^{o}$, obtaining

$$\sqrt{N}\,\tilde{S}^{\texttt{corr}}_{t+1|t} \to_d N\Big(0,\, a^2_{Lt}\big[\sigma^2_t\big(\boldsymbol{\iota}'_{\mathbf{01}}\mathbf{M}_{\mathbf{F^o}}\boldsymbol{\iota}_{\mathbf{01}}\big)\big]\Big), \tag{100}$$

where

$$a_{Lt} = \frac{1}{(T+1)\,r_{ft}}\left(1 - \mathbf{F}^{o\prime}_{t+1}\mathbf{\Sigma}_{\mathbf{\Lambda}_t}\big(\mathbf{\Sigma}_{\mathbf{\Lambda}_t} + \sigma^2_t(\mathbf{F}^{o\prime}\mathbf{F}^{o})\big)^{-1}\mathbf{S}^{-1}_{F^o}\bar{\mathbf{F}}^{o}\right). \tag{101}$$

Here $\boldsymbol{\iota}_{\mathbf{01}} \equiv (\mathbf{0}'_T, 1)'$, $\mathbf{M}_{\mathbf{F^o}} \equiv \mathbf{I}_{T+1} - \mathbf{F}^{o}(\mathbf{F}^{o\prime}\mathbf{F}^{o})^{-1}\mathbf{F}^{o\prime}$, and $\tilde{\mathbf{e}} \equiv \mathbf{X} - \mathbf{F}^{o}\tilde{\mathbf{\Lambda}}^{o\prime}$, obtaining

$$\tilde{a}_{Lt} = \frac{1}{(T+1)\,r_{ft}}\left(1 - \mathbf{F}'_{t+1}\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}^o}\big(\tilde{\mathbf{\Sigma}}_{\mathbf{\Lambda}^o} + \tilde{\sigma}^2(\mathbf{F}^{o\prime}\mathbf{F}^{o})\big)^{-1}\right) \to_p a_{Lt}, \tag{102}$$

where all other definitions continue to apply.

7 Empirics

I consider several empirical examples to demonstrate the ability of my methodology.

7.1 Dissecting the factors zoo when $T = 1$

Empirical asset pricing seeks to explain the cross-section of expected returns using observable characteristics or risk factors. Since the 1990s, researchers have proposed hundreds of such factors—size, value, momentum, profitability, investment, quality, liquidity, betting-against-beta, and many more. This proliferation of proposed predictors is called the "factor zoo" (Cochrane, 2011), and a core challenge today is to distinguish truly robust factors from those that arise from data mining or weak identification.[36] As datasets and computational tools improved, researchers documented an ever-growing list of "significant" predictors. Harvey, Liu, and Zhu (2016) and related work catalog hundreds of factors and highlight the severe multiple-testing and p-hacking problems implied by this proliferation: many reported anomalies may not survive more stringent statistical thresholds or genuine out-of-sample tests. This recognition triggered a shift in emphasis from discovering new factors to disciplining and organizing the existing zoo and, at the same time, the development of novel methodologies to trying to resolve it. The machine-learning-based approach of Gu, Kelly, and Xiu (2020) treats firm characteristics as inputs in

[36]The early benchmark was the CAPM, with a single market factor, but it failed to explain systematic patterns such as the size and value effects. This led to multi-factor models, most prominently the Fama–French three-factor model (market, size, value), later extended by Carhart (adding momentum) and then by the Fama–French five/six-factor models (adding profitability, investment, and sometimes momentum). Parallel literatures introduced factors based on liquidity, volatility, idiosyncratic risk, leverage, and many firm-level characteristics, contributing to the rapid expansion of the zoo.

flexible prediction models, using regularization and rigorous out-of-sample evaluation to identify which characteristics truly matter once overfitting is controlled. Bryzgalova, Huang, and Julliard (Journal of Finance, 2023, "Bayesian Solutions for the Factor Zoo") develop a Bayesian framework to jointly evaluate a large set of candidate factors, addressing both weak identification and model uncertainty.

Overall, the field has evolved from factor proliferation to factor discipline with the aim of novel methodologies. For instance, the emerging consensus emphasizes (i) demanding out-of-sample and out-of-period robustness; (ii) adjusting rigorously for multiple testing and model misspecification (iii) grounding factors in coherent economic mechanisms, whether risk-based or behavioral, (iv) recognizing that many proposed factors are redundant once a core set of well-established dimensions is included, recently spanned by latent risk factors estimated by PCA. My methodology advances the factor zoo agenda across all these four dimensions.

In particular, my methodology allows to rank the ability of each factor of the zoo in terms of OOS pricing performance and portfolio performance for *each period of time*. I can then summarize each risk factors' performance according to its frequency of 'success' as well as characterize the empirical distribution of the perfomance metrics.

I use data from an unbalanced panel of monthly individual assets traded in the NYSE from January 1980 to December 2019, with N = 19,259 (sourced CRSP) as test assets, and the factor zoo dataset of Gu, Kelly, and Xiu (2020) for the set of candidate risk factors, omitting the risk factor coded as divi, divo, and mve_ia, because are not complete for the whole period. I am also omitting the risk free rate. I am adding the latent single-factor model estimated by nano-PCA as an additional risk factor, denoted by PCA1.

7.1.1 Dissecting the factors zoo when $T = 1$: Risk Premia Pricing Test

I apply the risk premium pricing test $\tilde{L}_{t+1|t}$ of Section 4.1 to assess the OOS ability of any given candidate risk factor to price the cross-section of expected return. Figure OA.1 report the time series of the t-ratios, given by dividing $\tilde{L}_{t+1|t}$ by its standard error, corresponding to the latent risk factor (red line - (see Theorem 2) and the market return (black line - see Section 4.1.1), together with the 95% confidence bands (green lines). Each t-ratio uses to consecutive observations, say t for estimating the pricing error $\hat{h}_{i,t+1|t}$ from (4), and then both t and $t+1$ for constructing the test statistic. In this latter case, I use local PCA for the latent risk factor case and OLS when considering the observed market risk factor.

Visually, one can see that although similar, the market risk factor is rejected more often than the latent factor estimated by the first PC. In the appendix, I report the time serie plots of the risk premium pricing test for other 14 observed risk factors, that corresponds to the most relevant risk factors identified by Gu, Kelly, and Xiu (2020)[Second Lasso - Table 1].

I then rank the 148 zoo factors, in Figure OA.2, according to the percentage of times that the t-test is significant at 5%, the smaller the value the more relevant being the corresponding risk factor.

It emerges that core asset-pricing factors (market, size, value, profitability, investment, momentum, BAB/QMJ, HML_Devil) are systematically among the best (low second-column scores). Profitability and investment characteristics (HXZ/Fama–French-style) and related variables (NOA/DNOA, GMA, ORGCAP, CINVEST) perform very well, confirming the importance of the q-factor style dimensions. Momentum and liquidity/trading variables (UMD, MOM6M/MOMS12M, IND MOM, ILL,

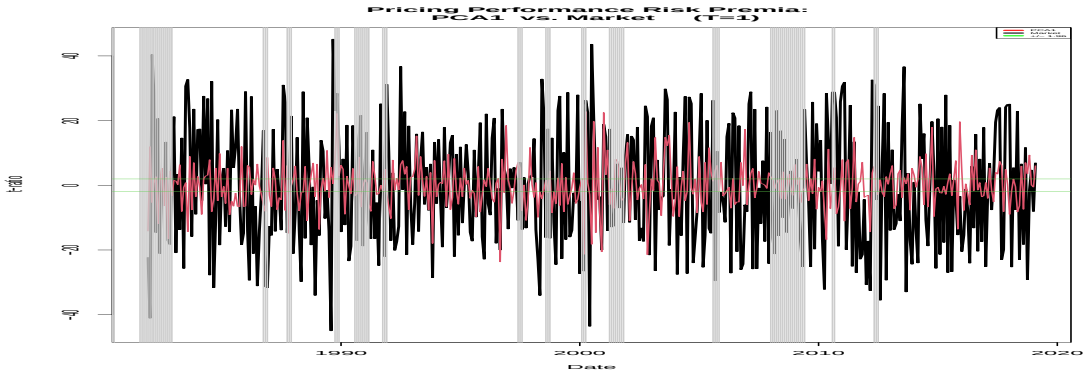


Figure OA.1: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and market risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

LIQ_PS, BASPREAD, ZERO TRADE) are clearly useful but slightly below the very top. Working capital / balance-sheet composition factors (PCHSALEINV, DNCL, DNCO, SALECASH, SALEINV, SALEREC, QUICK, CURRAT) are surprisingly strong, often among the top-ranked factors. Accruals and distress/constraints factors (ACC, ABSACC, KZ, WW) appear moderate, not leading the ranking.

In contrast, the worst performing risk factors appear related to volatility measures (RETVOL, IDIOVOL), profitability/portfolio efficiency proxies (OP, PM), trading/turnover (TURN, ZEROTRADE), constraints/distress (KZ, WW), some "quality" style factors (HIRE, RS, RSUP), and a subset of valuation/accrual variables.

I also perform a direct comparison between the ranking implied by the first metric (percentage of significant pricing test at 5%) and that implied by a second metric (average of absolute value of t-ratios). Not surprisingly, both metrics agree that a relatively small set of factors are "well behaved," but they differ in how they treat many core priced factors and several anomalies and volatility/trading variables.

Combining this with the previous "best factors" analysis, one getd a coherent picture, of consistently strong and consistently weak risk factors.

I now investigate the possibly different ability of the zoo factors to price the cross-section of expected returns in periods of economic and financial crises from normal periods. Figures OA.3-,OA.4 report the percentage of times that the t-test is significant at 5% corresponding to the expansion periods only and recession periods only, respectively.

The results are more accurately described in Tables 1-2-3-4. The tables report, for each risk factor, the regression coefficient of its test statistic on a financial-crisis dummy, the corresponding t-ratio, the average test statistic in crisis periods and overall, and the fraction of times the test exceeds standard significance thresholds. Taken together, these results paint a very consistent picture: financial crises are periods in which almost the entire factor zoo "lights up" much more strongly than in normal times. First, the coefficients on the crisis dummy are almost uniformly positive, and for most factors the associated t-ratios are clearly above conventional thresholds (frequently between 3 and 5). This implies that, on average, the test statistics are systematically higher in crisis periods. Impor-

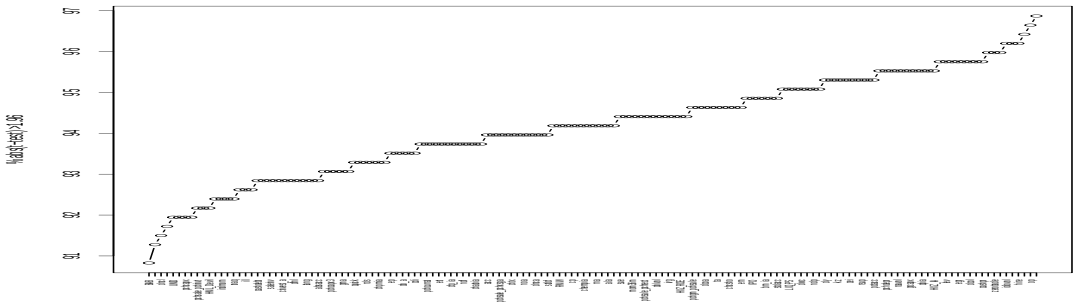


Figure OA.2: The plot reports the percentage of time the t-ratio for the risk premium pricing $\tilde{L}^o_{t+1|t}$, *sorted* across the 148 zoo risk factors of Feng, Giglio, and Xiu (2020), is significant at 5% level from Dec 1981 to Feb 2019.

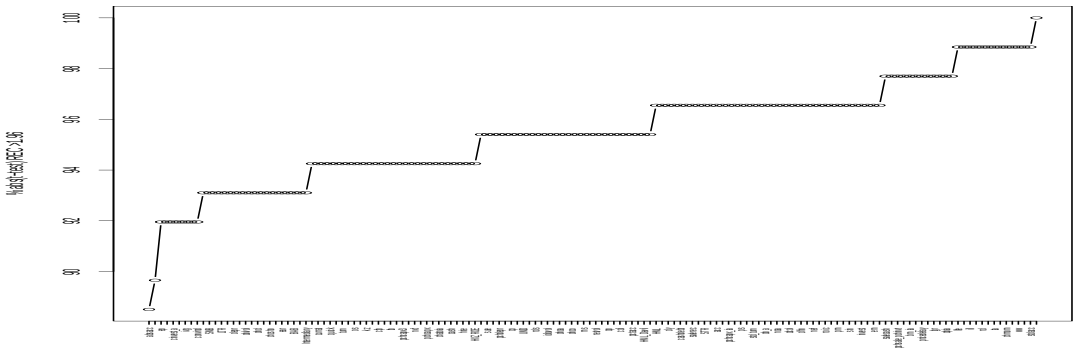


Figure OA.3: The plot reports the percentage of time the t-ratio for the risk premium pricing $\tilde{L}^o_{t+1|t}$, *sorted* across the 148 zoo risk factors of Feng, Giglio, and Xiu (2020), is significant at 5% level for the *recession periods* from Dec 1981 to Feb 2019.

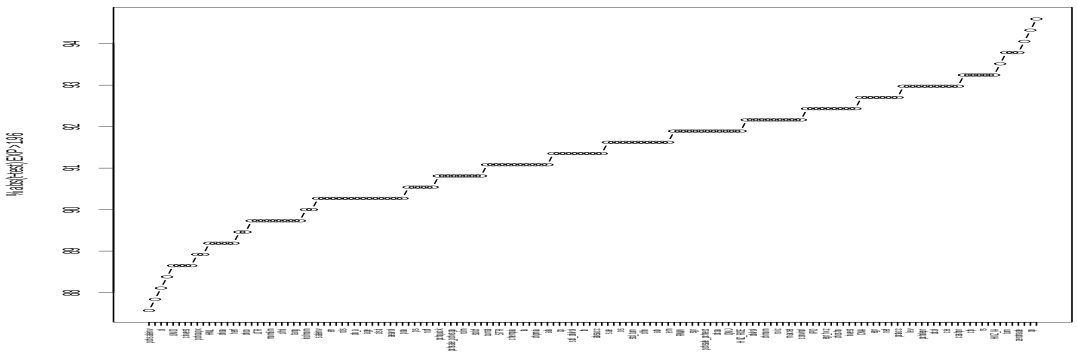


Figure OA.4: The plot reports the percentage of time the t-ratio for the risk premium pricing $\tilde{L}^o_{t+1|t}$, *sorted* across the 148 zoo risk factors of Feng, Giglio, and Xiu (2020), is significant at 5% level for the *no-recession periods* from Dec 1981 to Feb 2019.

tantly, this is not a marginal effect: the estimated shift is large and statistically precise, suggesting that crises are structurally different states of the world from the perspective of factor tests, not just rare noisy episodes. Second, comparing the crisis-period mean of the test statistics to the overall mean confirms this pattern. For many factors, the average test statistic in crisis periods is substantially larger than the full-sample average, often by several units. For example, the market excess return (MktRf) has a crisis mean around 6, while the overall mean is slightly negative; similar gaps appear for size (SMB), value (HML), investment (CMA), profitability (RMW), and for many "zoo" factors such as investment-adjusted cash-flow and intermediary-capital measures. In other words, most of the "action" in these tests is concentrated in crises. Third, the frequency measures show that the test statistics are very often "significant" in the classical sense. For the vast majority of factors, the test statistic exceeds 1.96 in over 90% of the periods, and even the more stringent thresholds (twice or five times 1.96) are crossed with surprisingly high frequency. This indicates that, conditional on your test design, these factors rarely look "inactive"; they are persistently strong, especially in crisis episodes. Fourth, the group-wise patterns (using the earlier classification of the factor zoo) are economically meaningful. Core priced factors from the modern asset-pricing canon-market, size, value, profitability, investment, momentum, and popular extensions such as BAB, QMJ, HML_Devil, and HXZ-style factors-show some of the largest and most significant crisis shifts, with crisis means far above their overall averages and high significance frequencies. Profitability, investment, and working-capital characteristics also become markedly more informative in crises, consistent with the idea that balance-sheet strength, investment policy, and operating efficiency matter most when financial conditions are tight. Valuation measures (e.g. earnings-price, cash-flow-price, investment-adjusted ratios) similarly display much stronger tests in crises than in normal times. Finally, liquidity, volatility, and distress-related variables follow the same qualitative pattern: volatility and illiquidity measures, turnover and zero-trade proxies, and financing-constraint indices (KZ, WW) all have positive and significant crisis coefficients and high frequencies of large test statistics. Some accruals, distress, and anomaly-style signals (such as sin stocks, max-return, short-interest proxies) also intensify in crises, though often less uniformly than core risk factors and investment/valuation signals. Overall, the evidence points to financial crises as periods in which risk premia, mispricing, and cross-sectional spreads in returns are generally amplified, making almost all well-known factors more visible and statistically powerful relative to the rest of the sample.

It appears interesting to relate the performance of the zoo risk factors from my analysis, based on Table 1 of Feng, Giglio, and Xiu (2020). Comparing the recession and expansion (rankings for the post-2012 factors reveals pronounced state dependence. Several of the new factors are clearly *recession-heavy*: RMW, the Intermediary Risk Factor, the convertible-debt factor (convind), Organizational Capital, and Employee growth (hire) all have relatively high positions (low ranks) in recessions but drop sharply in expansions, often by 60–100 places. This pattern is consistent with interpretations that emphasise downside risk and financial constraints: profitability (RMW), intermediaries' balance-sheet strength, capital structure (convertible debt), and certain intangible-capital or labour-adjustment proxies appear to matter disproportionately when conditions are stressed. By contrast, other factors are much more important in expansions. In particular, BAB moves from a decent but not exceptional rank in recessions to the very top of the list in expansions, and HML Devil similarly jumps from a low recession rank into the expansion top ten. In other specifications (e.g., PCA-based models), CMA also emerges as a top

expansion factor, reinforcing the view that some investment-style signals are rewarded more strongly in good times.

At the same time, there is a group of factors where the SDF evidence and the ranking-based importance do not fully align. QMJ, HXZ IA, and HXZ ROE exhibit strong and statistically significant prices of risk and average premia in the 2012–2016 SDF tests, yet in the recession/expansion rankings they occupy only middling positions, especially in expansions. This suggests that, once these factors are allowed to compete with a large set of characteristics, some of their priced variation is absorbed by other, more primitive signals, and they no longer dominate the cross-section. Organizational Capital and RMW sit somewhat in between: the ranks confirm that they are economically meaningful, particularly in recessions, but they are not uniformly top-tier across states. Overall, the comparison shows that the new factors are not simply "good" or "bad" in a time-averaged sense; rather, their relevance is conditional on the macroeconomic regime, with some factors acting primarily as recession hedges and others becoming central drivers of returns during expansions.

Risk Premia Pricing Test: Misspecification-Robust

The above results used the risk premia pricing test under the assumption that the cross-section is spanned by one single risk factor. I now examine the case when instead a multi-factor structure holds. This means that one now needs to bias-adjust the pricing test for both the $T = 1$-related estimation bias, in addition to the misspecification bias because the pricing errors are still being computed under the single-factor model. In particular, I consider three scenarios, namely when there are latent three, five, and 10 risk factors, in addition to the single risk factor (either latent or observed - one of the zoo factors), although I will only display the case of five latent risk factors.[37] See Section 5.1 - Theorem 5 - for details of the theory.

Therefore, I repeat the above analysis but with the salient difference of using a certain number of PCs to account for the misspecification bias. Figures OA.5, OA.6, and OA.7 report the time series of the market risk premium test statistics when using three, five, and 10 PCs, respectively, to account for misspecification. As expected, the pricing test appear not significant more often than when I did not allow for the latent factors. The same patters arises with respect to the other risk factors.

I then re-examine the rank of the 148 zoo factors, for the three cases of three, five, and 10 PCs, according to the percentage of times that the t-test is significant at 5%, with the smaller the value the more relevant being the corresponding risk factor, and compare with the previous case of no adjustment for misspecification was made. Figures OA.8-OA.9-OA.10 report the ranked percentages (significant at 5%), with the detailed results in Tables 7-8-9, where I also report the performances within the recession periods.

Most importantly, I evaluate how each single risk factor changes its performance once misspecification is taken into account.

A broad look at the rankings across specifications reveals a clear general pattern. Notice that in terms of rank, the performance is unchanged for PCA5 and PCA10. For a sizable group of characteristics, the move from the baseline model without principal components, discussed in the previous section, to models with PCA3, and PCA5/10, leaves their qualitative status largely unchanged: factors that are in the top tier tend to remain near the top, those in the middle stay roughly in the middle, and those in the bottom remain relatively unimportant. In other words, for many variables, the ordering is reasonably robust to the introduction of PCA-based controls for common variation. However, this stability is not universal. A non-trivial subset of factors exhibits substantial sensitivity to the inclusion and dimension of PCA: their ranks shift by on the order of 50 to 130 positions as one moves from the case without misspecification adjustment to the cases with the adjustment.

[37] Details are available upon request.

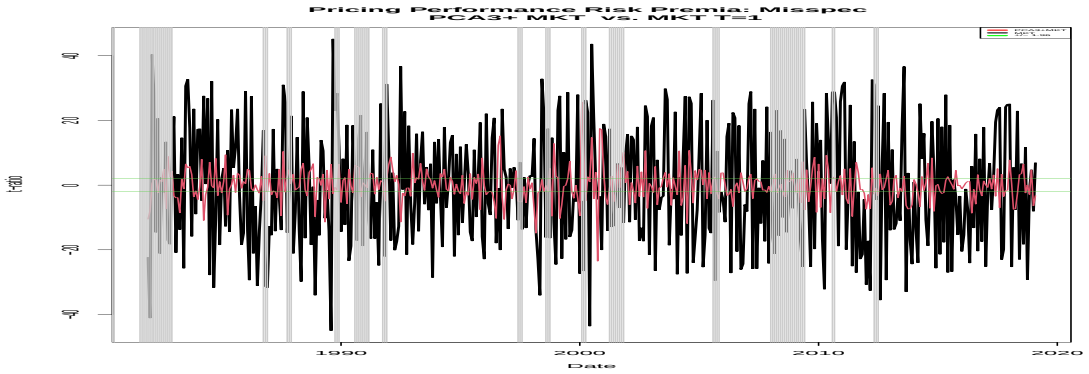


Figure OA.5: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (black line) corresponding to market risk factor and market risk factor+PCA3, respectively - see Section 5. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

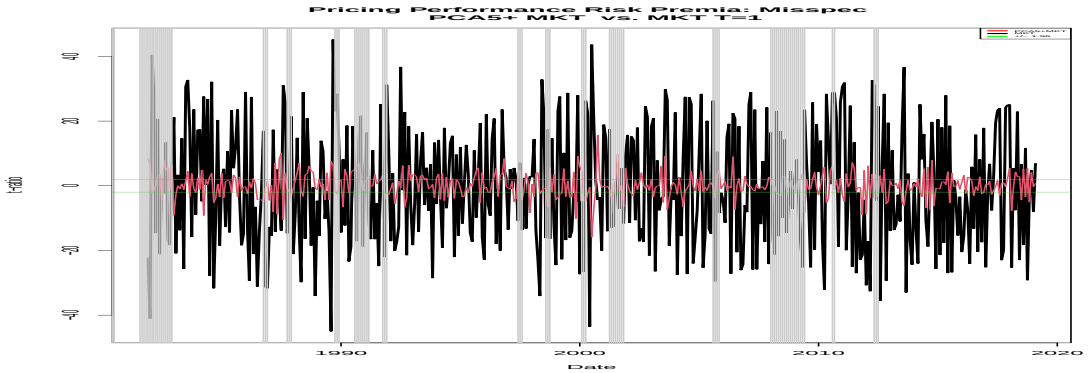


Figure OA.6: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (black line) corresponding to market risk factor and market risk factor+PCA5, respectively - see Section 5. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

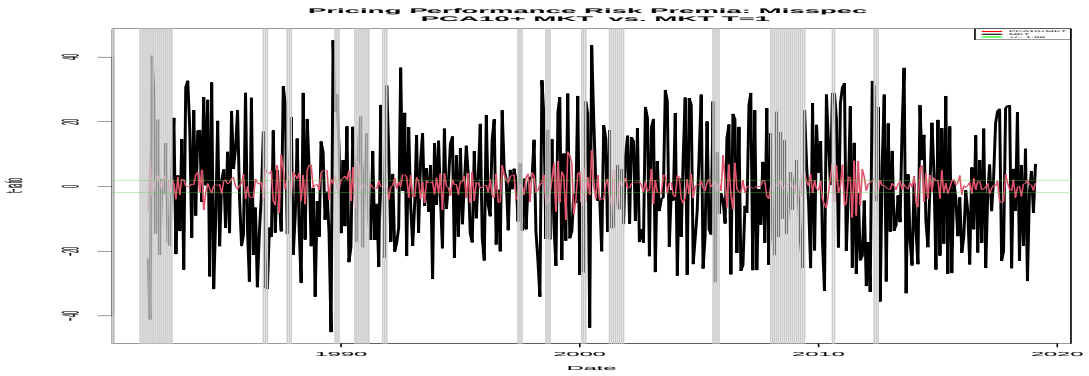


Figure OA.7: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (black line) corresponding to market risk factor and market risk factor+PCA10, respectively - see Section 5. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

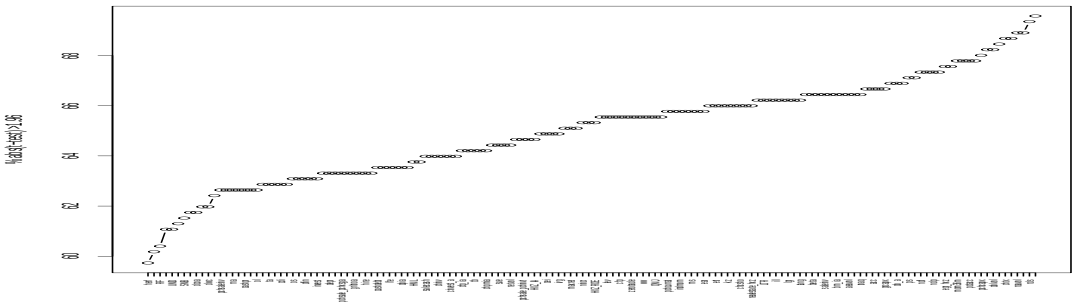


Figure OA.8: The plot reports the percentage of time the t-ratio for the risk premium pricing $\tilde{L}^{\texttt{miss}}_{t+1|t}$, *sorted* across the 148 zoo risk factors of Feng, Giglio, and Xiu (2020), is significant at 5% level from Dec 1981 to Feb 2019. The pricing test is bias-adjusted using the corresponding observed risk factor plus PCA3 - see Section 5.

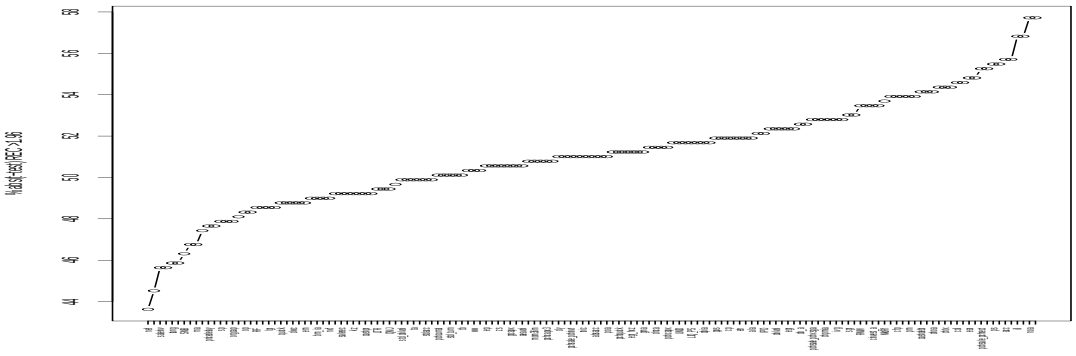


Figure OA.9: The plot reports the percentage of time the t-ratio for the risk premium pricing $\tilde{L}^{\texttt{miss}}_{t+1|t}$, *sorted* across the 148 zoo risk factors of Feng, Giglio, and Xiu (2020), is significant at 5% level from Dec 1981 to Feb 2019. The pricing test is bias-adjusted using the corresponding observed risk factor plus PCA5- see Section 5.

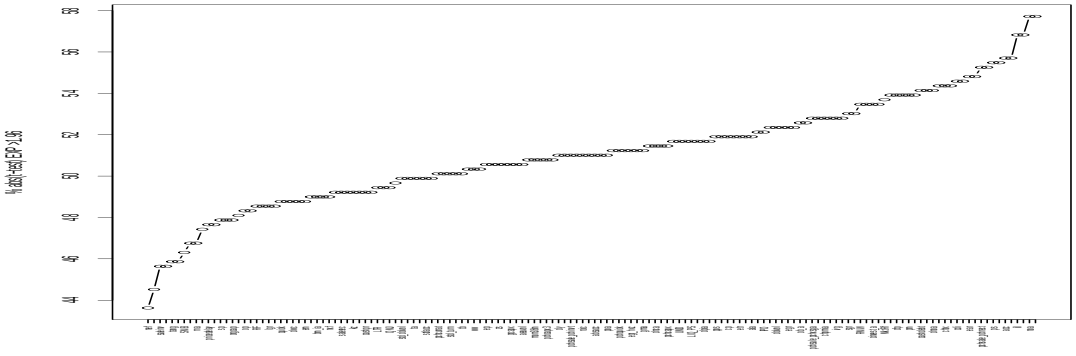


Figure OA.10: The plot reports the percentage of time the t-ratio for the risk premium pricing $\tilde{L}^{\texttt{miss}}_{t+1|t}$, *sorted* across the 148 zoo risk factors of Feng, Giglio, and Xiu (2020), is significant at 5% level from Dec 1981 to Feb 2019. The pricing test is bias-adjusted using the corresponding observed risk factor plus PCA10 - see Section 5.

These large rank swings strongly suggest that part of what the baseline specification is attributing to specific observed factors is in fact common variation that can be captured by a small number of latent components. The use of PCA5/PCA10, in particular, often re-orders the very top of the factor list relative to the other cases (no adjustment and PCA3), indicating that some factors which appear central in the baseline are acting as imperfect proxies for broader systematic forces that PCA later absorbs more efficiently.

The behaviour of individual factors makes this misspecification pattern particularly vivid. One striking group consists of extreme downgrades once PCA is introduced. These are factors that look very important in the no-adjustment, but lose their apparent status when PCA factors absorb the underlying common components. For instance, BAB is the top factor in the baseline, but plummets to rank 132 under PCA3 and remains low at rank 128 under PCA5/10. This is a huge collapse, from rank 1 to roughly rank 130, and it strongly indicates that BAB's "importance" in the baseline model is largely overlapping with the latent components captured by PCA; once those components are explicitly included, BAB no longer provides much incremental information. A similar pattern is visible for `herf`: it is high under no-adjustment (rank 6), becomes the single top factor under PCA3 (rank 1), and then collapses to rank 134 under PCA5/10. In a low-dimensional PCA (three components), `herf` seems to be used as a key factor alongside the PCs; but once more principal components are added, that role is absorbed by additional PCs, and `herf`'s standalone importance largely disappears. The investment-related factor `pchcapx` also illustrates strong sensitivity: it starts as a top-10 factor under no-adjustment (rank 7), drops dramatically to rank 139 in PCA3, and only partially recovers to rank 87 in PCA5/10. This suggests that its explanatory power is tightly tied to how common investment-related variation is treated. The illiquidity proxy `ill` moves from an important position under no-adjustment (rank 17) to rank 105 in PCA3 and further down to rank 145 in PCA5/10, making it essentially one of the least important factors once PCA is included. The Intermediary factor behaves similarly: it is mid-pack in the baseline (rank 90), but falls to rank 146 under PCA3 and to rank 148 under PCA5/10, becoming almost irrelevant in the PCA-augmented models. Finally, `rsup` exhibits a particularly wild pattern: it moves from rank 119 in M1 to a high rank 18 in M2, before collapsing to rank 146 in M3. This non-monotonic path underscores how strongly some variables' rankings depend on the exact PCA specification.

The flip side consists of large upgrades with PCA: factors that look weak or negligible in the baseline model, but become very important once common components are removed. The `nef` variable is the clearest example. It starts at rank 128 under no-adjustment, rises sharply to rank 22 in PCA3, and ultimately becomes the top factor under PCA5/10 (rank 1). When the generic common variation is stripped out by PCA, `nef` appears to carry unique, priced information that is not well captured by other characteristics. Similarly, `nop` begins at rank 142 under no-adjustment, moves to rank 64 in PCA3, and ends up as the number-two factor under PCA5/10 (rank 2). The factor `saleinv` (sales-to-investment) offers another instructive case: it is moderately important in the baseline (rank 21), seems to deteriorate under PCA3 (rank 113), and then becomes a top-3 factor in PCA5/10 (rank 3). The tangibility measure `tang` shows the same type of U-shaped pattern: it ranks 27 under no-adjustment, falls to 109 in PCA3, and then jumps to rank 5 under PCA5/10, with a similar pattern for the asset turnover ratio `ato` (ATO). Even the traditional market beta is affected: from rank 132 under no-adjustment and rank 111 in PCA3, beta rises to rank 16 in PCA5/10, suggesting that, once the main common components are removed, beta regains a non-trivial role in explaining the cross-section of returns. Operating profitability `op` offers another striking case: it is almost at the bottom under no-adjustment (rank 147) and still low in PCA3 (rank 136), but it jumps into the top 20 in PCA5/10 (rank 17). Likewise, the hiring or employee-growth proxy `hire` is nearly irrelevant in the baseline (rank 145), yet PCA3 already promotes it substantially (rank 37), and PCA5/10 lifts it even further to rank 18. In all of these cases, the introduction of PCA factors seems to strip away confounding common variation, allowing these "second-tier" factors to emerge as carrying more idiosyncratic priced risk than the baseline specification suggests.

Beyond extreme changes, there is a broad group of factors that experience moderate but still meaningful shifts in rank, typically in the range of 20–50 positions. These moves, while less dramatic than a leap of more than 100 ranks, still signal that PCA is relieving some misspecification and collinearity. Organizational Capital (`orgcap`) follows a steady upward trajectory: it moves from rank 34 under no-adjustment to rank 24 in PCA3 and then to rank 15 in PCA5/10, indicating a gradual recognition of its unique explanatory power once common components are controlled for more flexibly. By contrast, the standard HML factor exhibits a degrade-then-stabilize pattern: its rank worsens from 16 under no-adjustment to 45 in PCA3 and then remains roughly similar at 42 in PCA5/10, suggesting that part of HML's baseline prominence is taken over by the PCA factors, although it still retains some relevance. QMJ is clearly upgraded by PCA5/10: it moves from rank 88 under no-adjustment to 85 in PCA3 and then sharply improves to rank 41 in PCA5/10, consistent with the idea that quality-related effects are more distinct once overlapping exposures are netted out. Finally, HXZ_IA shows a substantial improvement: from a very low rank 131 under no-adjustment, it jumps to 65 in PCA3 and then to 56 in PCA5/10, indicating that investment-related effects may be more important than the baseline ranking suggests once collinearity with other characteristics and common components is mitigated. Taken together, these moderate shifts reinforce the intuition that PCA tends to clean up overlapping exposures and multicollinearity, allowing some factors to "shine" that were previously masked in the baseline model.

These patterns have direct implications for how we think about misspecification and the role of PCA in factor selection. First, PCA can effectively "steal" importance from some famous factors: variables such as BAB, the illiquidity proxy `ill`, and classic cross-sectional indicators like `herf` and `pchcapx` lose rank sharply once PCA factors are included. In the baseline M1 model, these characteristics appear to be central, but the PCA results suggest that they are, to a significant extent, capturing broad common variation that can be more parsimoniously represented by a few principal components. When those components are explicitly added, the incremental contribution of these well-known factors declines markedly. Second, PCA reveals hidden importance in other variables that were overshadowed by stronger, collinear signals under no-adjustment. Factors such as `nef`, `nop`, `saleinv`, `tang`, `ato`, beta, `op`, and `hire` move from weak or negligible positions in the baseline into the top ranks under PCA5/10. These variables likely represent more granular or idiosyncratic sources of priced risk that only become visible once the dominant common components are accounted for separately. Third, the dimension of PCA itself matters. PCA3 sometimes yields quite different rankings from PCA5/10.

In summary, comparing the baseline model under no-adjustment (no PCA) with PCA3 and PCA5/10 shows that many factors change ranking substantially when PCA is employed. Some headline factors, such as BAB and illiquidity, look crucial in the baseline but become much less important once PCA handles the underlying common components. At the same time, several "second-tier" factors—including `nef`, `nop`, `saleinv`, `tang`, `ato`, beta, `op`, and `hire`— move into the top ranks under PCA, suggesting that they carry unique, priced information that was previously obscured by collinearity and misspecification. The differences between PCA3 and PCA5/10 further demonstrate that the number of principal components is itself a crucial modelling choice: too few PCs can distort which observed characteristics appear most important, and only when a richer set of common components is included do the rankings begin to stabilise in a way that more plausibly reflects underlying economic structure.

7.1.2 Dissecting the factors zoo when $T = 1$: MV Portfolio

I now evaluate the ability of the zoo risk factors in terms of Sharpe ratio performance, using the methodology outlined in Section 4.2, and in particular the case of observed risk factor in Section 4.2.1. Figure OA.11 displays the time series of the Sharpe ratio $\widetilde{SR}_{t,t+1,t+2}$ corresponding to the market risk factor (red line), where I use two periods for estimation and one period for OOS analysis, together with the Sharpe ratio $\widehat{SR}_{t,t+1}$obtained using the first PC (black line). In the appendix, I report the corresponding histograms for the other 14 risk

factors examined in the previous section (see Figures OA.49-OA.61).

I also display the same data in terms of histograms (see Figure OA.12). Moreover, I compare the market factor Sharpe ratio with the one obtained *debiasing* the OOS effect (see Figure OA.13). Not surprisingly, the market and PCA1 appear to have a similar performance. Interestingly, debiasing appears to have a positive effect on performance.[38]

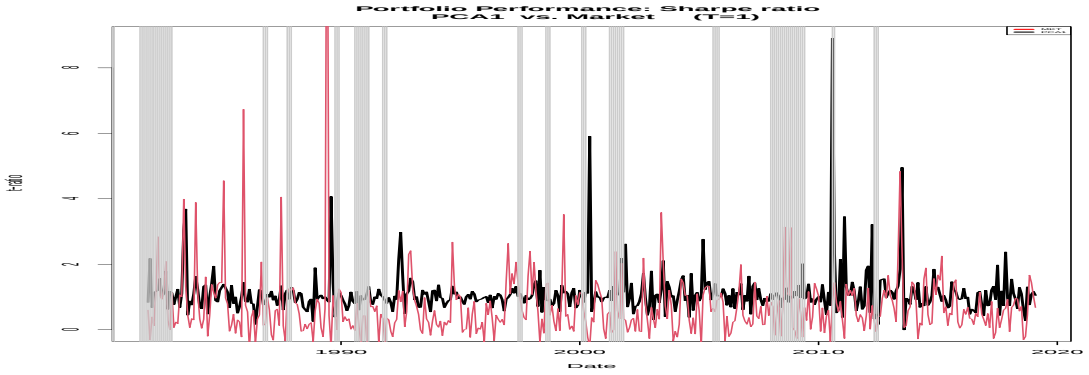


Figure OA.11: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (red line) and $\widetilde{SR}_{t,t+1,t+2}$ (black line) corresponding to PCA1 and the market risk factor respectively - see Section 4.2. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

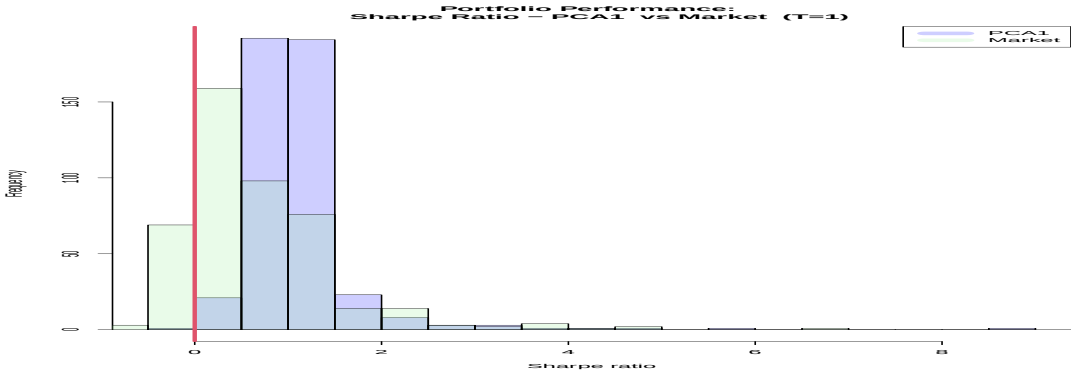


Figure OA.12: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the market risk factor respectively - see Section 4.2.

In Tables 11-13, I report the analysis of the behaviour of the SRs across risk factors during downturns, and across all periods, classified by groups of risk factors. Across the four groups of factors, the regression coefficients on the recession dummy and the Sharpe ratios paint a consistent picture of mildly pro-cyclical performance with some notable exceptions. Among core factors and standard characteristics, most coefficients

[38]For simplicity I am not adding the confidence bands for the Sharpe ratios. See Theorem 3.

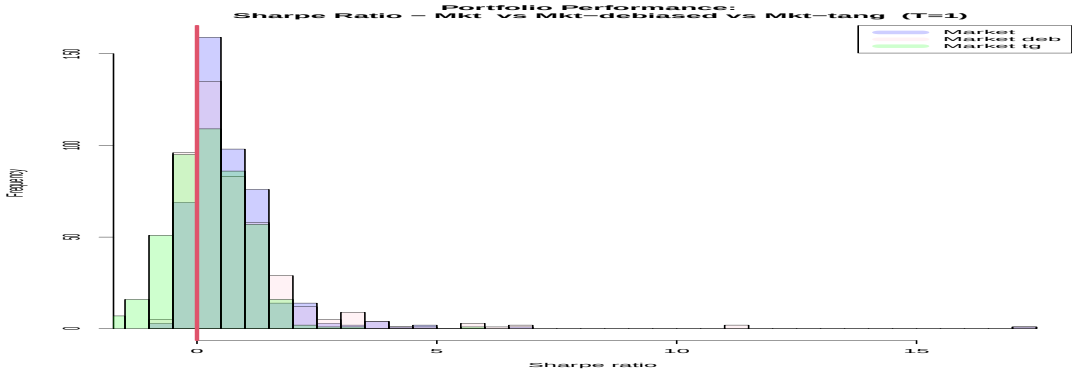


Figure OA.13: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widetilde{SR}_{t,t+1,t+2}$ (blue bars) corresponding to the market risk factor, the de-biased version based on $r^*_{t+2|t+1}$ (pink bars), and the corresponding tangency portfolio (green bars), respectively - see Section 4.2.

are small and statistically insignificant, indicating that their Sharpe ratios do not change dramatically between recessions and the full sample. In contrast, CMA, SMB, and dy have positive and sometimes statistically meaningful coefficients, coupled with elevated recession Sharpe ratios, suggesting that value–investment and size-like tilts can be relatively more attractive in downturns. BAB and QMJ exhibit high average Sharpe ratios overall, but their recession coefficients are negative and not statistically strong, indicating that their superior long-run risk-adjusted performance does not primarily come from recession states. For the core set, the main message is that classic pricing factors and standard characteristics maintain solid Sharpe ratios across regimes, with only moderate and often statistically weak recession-specific shifts, and with the risk-free rate standing out as behaving very differently across states.

The investment, profitability, balance-sheet, liquidity, and anomaly blocks add further nuance. Within investment and balance-sheet variables, several changes-in-investment measures (e.g., pchdepr, pchsaleinv, pchsale_pchinvt, STR, IPO, dnca) have economically sizeable and sometimes statistically significant coefficients, often with higher Sharpe ratios in recessions than in the full sample. This points to investment and financing frictions becoming more strongly rewarded when conditions are tight. Profitability- and quality-like variables such as chempia, orgcap, nop, HXZ_IA, op, gma, and rsup generally display high Sharpe ratios in recessions and across the full sample, with positive and occasionally significant coefficients, suggesting that profitability, quality, and balance-sheet strength are particularly valuable in downturns. In the liquidity and trading-activity group, several proxies (turn, baspread, roavol, retvol, zerotrade) have positive and statistically significant recession coefficients and elevated recession Sharpe ratios, indicating that exposures tied to illiquidity and volatility pay off especially well in bad times. Finally, the broad "remaining/anomalies" set shows that many classic accounting and constraint-related anomalies (e.g., absacc, ear, rds, grltnoa_hxz, dnca, rsup, realestate_hxz, HXZ_ROE) have non-trivial recession loadings and solid Sharpe ratios, but the t-statistics are mixed, underscoring that while a large number of signals deliver good performance on average, only a subset displays statistically robust, recession-specific enhancements in their Sharpe ratios.

Comparing the two rankings in Table 14, where a smaller number always indicates better performance, reveals both areas of agreement and sharp trade-offs between pricing ability and Sharpe performance. At the very top of the pricing test, BAB (rank test = 1) is among the best at explaining returns, while herf, ms, HML, and UMD also appear as strong pricing factors (ranks 5, 6, 26, 16, and 5 respectively). On the Sharpe side,

BAB remains extremely strong (rank SR = 2), and roic, QMJ, herf, and pm (ranks 3, 4, 5, and 8) also rise to the top. These patterns indicate a core set of variables—RF, BAB, herf, HML, CMA, QMJ, roic, and to some extent ms and UMD—that are jointly good at pricing the cross-section and at delivering attractive standalone Sharpe ratios. For example, CMA is relatively weak in the pricing test (rank test = 122) but quite good on Sharpe (rank SR = 16), while QMJ and roic, though only mid-pack in pricing (ranks 88 and 110), are among the very best in Sharpe terms (ranks 4 and 3). These are candidates for "dual-purpose" factors in the sense that they help with both pricing and portfolio performance, even if their strengths are not perfectly aligned.

At the same time, there is a substantial group of characteristics where pricing and Sharpe rankings diverge sharply, highlighting important trade-offs. Some variables are excellent pricing factors but weak Sharpe performers: dncl, cinvest, indmom, pchsaleinv, STR, and several investment-change measures sit in the top 20–30 for pricing but are mediocre to poor on Sharpe (rank_M1 often above 80–100). For instance, dncl has rank test = 3 but rank SR = 114, cinvest is 10 vs 140, and indmom is 13 vs 111, suggesting that these characteristics are valuable for spanning pricing errors in an SDF sense, but do not translate into especially high standalone Sharpe ratios. Conversely, a handful of variables look weak in pricing but strong in Sharpe: realestate_hxz (rank test = 130, rank SR = 7), pm (144, 8), and some microstructure or volatility proxies (e.g., dolvol at 85, 147; ill at 17, 148) exhibit much better Sharpe than pricing ranks, implying that their appeal lies more in return performance than in improving cross-sectional fit. Finally, many classic anomalies and accounting signals (e.g., accruals, earnings surprises, investment and profitability variants) occupy intermediate positions in both rankings, indicating that they provide a mix of pricing and performance benefits without dominating on either dimension. Overall, the joint ranking shows that a few core factors (RF, BAB, herf, QMJ, roic, HML, CMA) perform well along both criteria, while many others force a choice between superior pricing ability and superior Sharpe ratio.

MV Portfolio: Misspecification-Robust

Interpreting Table 15 with the convention that a smaller number corresponds to a better (higher) Sharpe ratio, the first column (single-factor Sharpe) is dominated by BAB, followed by roic, QMJ, herf, and ms (ranks 3–6). This indicates that, when each characteristic is evaluated in isolation, traditional benchmark factors and a small set of quality/profitability-type variables deliver the strongest risk-adjusted performance. Classic factors like HML, CMA, and UMD also appear relatively high in this single-factor ranking, while many illiquidity and microstructure proxies—ill, dolvol, baspread, zerotrade, tang—sit near the bottom (ranks above 140), suggesting weaker standalone Sharpe ratios. When we move to Sharpe with PCA3, the ordering shifts: pm now becomes the top-ranked characteristic (rank 1), and a number of anomalies or accounting variables (e.g., chempia, LTR, kz) rise toward the top. This reflects that with only three principal components, part of the systematic variation previously attributed to broad factors is now captured by the PCs, and some more idiosyncratic signals are rewarded in terms of residual Sharpe performance.

The PCA5 and PCA10 rankings further reshuffle the hierarchy and highlight strong model sensitivity. Under PCA5, BAB drops to the very bottom (148), while the best-performing characteristics become dolvol (rank 1), illiquidity-related and volatility measures (aeavol at 2, mom6m at 3, SMB at 4), and several investment or balance-sheet variables (ta, pchdepr, dnca, LTR) in the top decile. Under PCA10, zerotrade emerges as the single best factor (rank 1), followed closely by maxret (2), sp (3), stdcf (3rd-best at rank 3), and rdm and realestate_hxz also near the top, with BAB remaining at the bottom (ranks 148 and 147). Across the four specifications, the message is that single-factor Sharpe rankings favor RF, BAB, and certain quality/profitability signals, while PCA-augmented Sharpe rankings progressively shift Sharpe "credit" toward a mixture of volatility, liquidity, and more granular anomaly characteristics. The large rank swings for many variables (for example, ill from 148 to 71 to 23; BAB from 2 to 148–148) underscore that a factor's perceived importance in terms of Sharpe ratio is highly contingent on how aggressively one controls for latent common components.

7.2 Cross-sectional models when $T = r$

In this section I study the performance of multi-factor models, namely the case when a multi-factor model is posited from the outset, rather than just used to correct the biases arising from a misspecified single-factor model, as done in the previous sections. I will consider three latent factor models, PCA3, PCA5, and PCA10. I will also consider three observed factor models, FF3, FF5, FF5+momentum.

In terms of risk premia pricing performance, Figures OA.14-OA.16 display the corresponding time series of the test statistics corresponding to the PCA1 model (black line) with the multi-factor cases (red line). Clearly, increasing the number of latent risk factors improves the pricing performance.

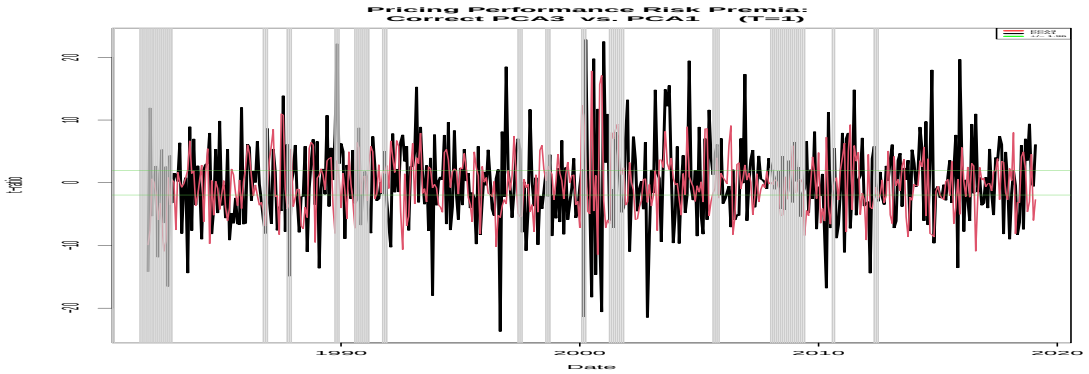


Figure OA.14: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (black line) and $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (red line) corresponding to PCA1 and PCA3, respectively - see Section 6. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

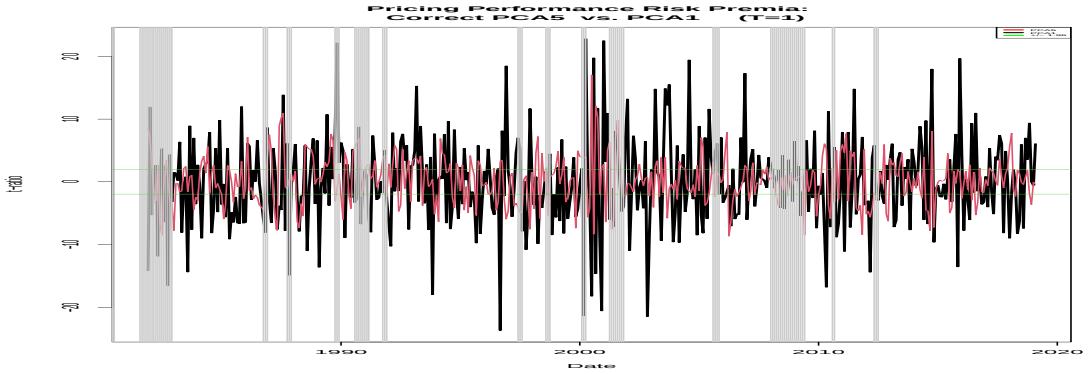


Figure OA.15: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (black line) and $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (red line) corresponding to PCA1 and PCA5, respectively - see Section 6. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

Figure OA.20 reports the histograms of the Sharpe ratios for the four latent factor models (I am including PCA1). As expected, OOS the performance of a model with a large number of factors becomes more erratic.

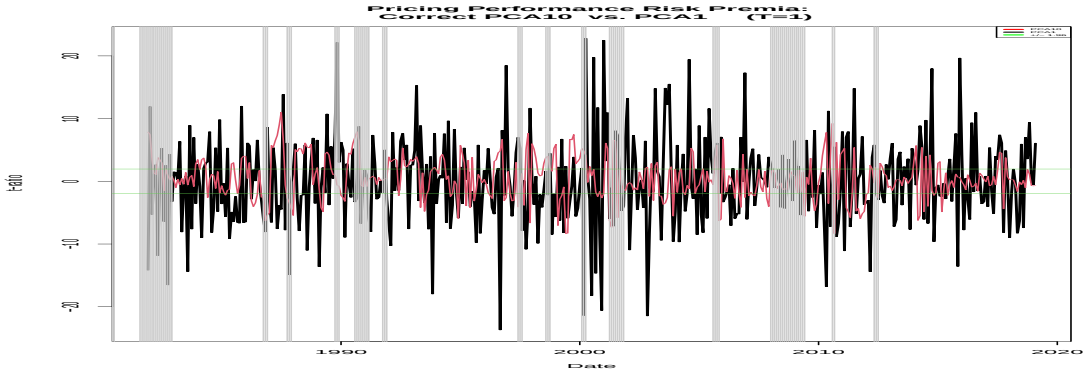


Figure OA.16: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (black line) and $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (red line) corresponding to PCA1 and PCA10, respectively - see Section 6. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

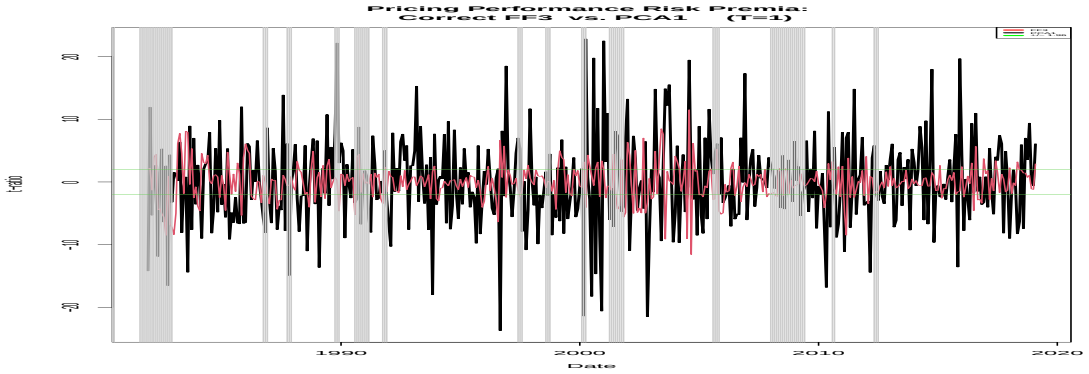


Figure OA.17: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (black line) and $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (red line) corresponding to PCA1 and FF3, respectively - see Section 6. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

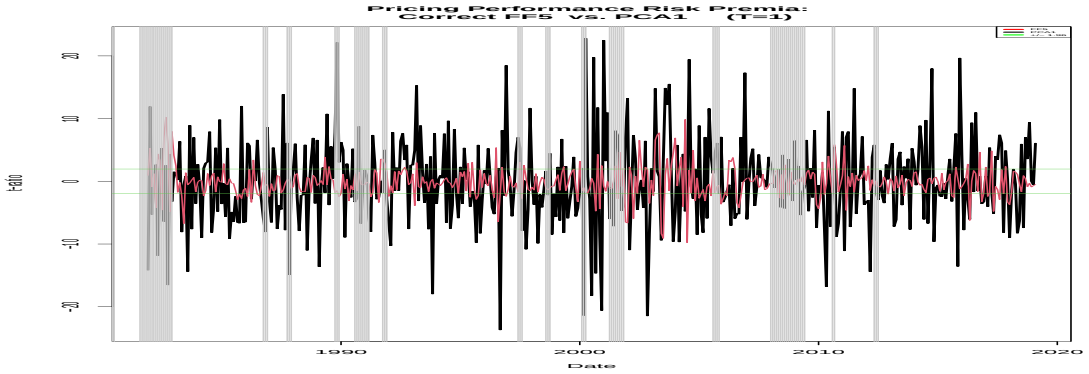


Figure OA.18: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (black line) and $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (red line) corresponding to PCA1 and FF5, respectively - see Section 6. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

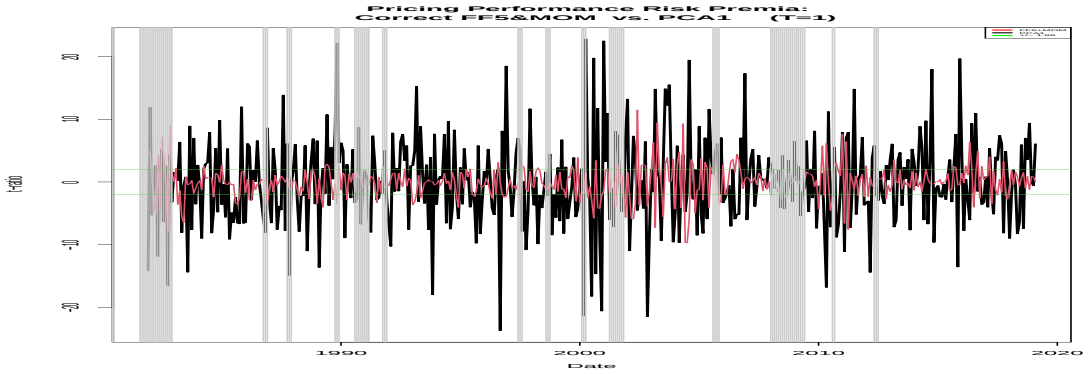


Figure OA.19: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (black line) and $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (red line) corresponding to PCA1 and FF5+Momentum, respectively - see Section 6. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

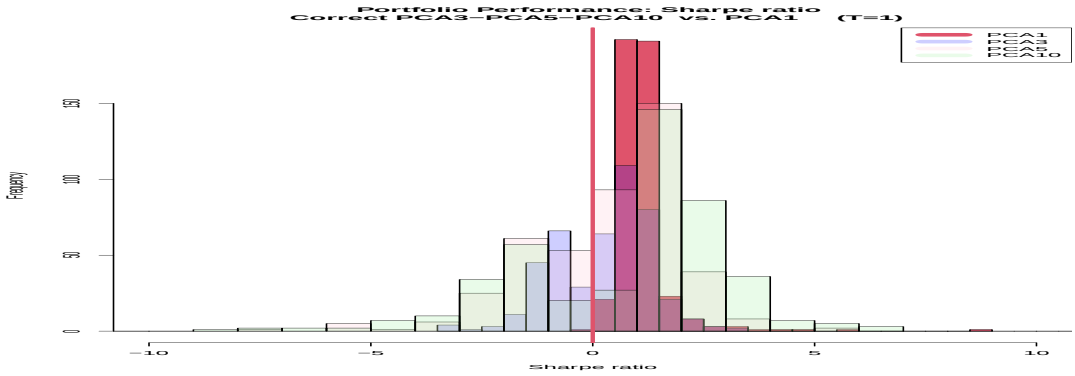


Figure OA.20: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widetilde{SR}_{t,t+1,t+2}$ (red bars) corresponding to PCA1, and three versions of $\widetilde{SR}_{1:T,t+1}$ corresponding to PCA3 (blue bars), PCA5 (pink bars), and PCA10 (green bars) - see Section 6.

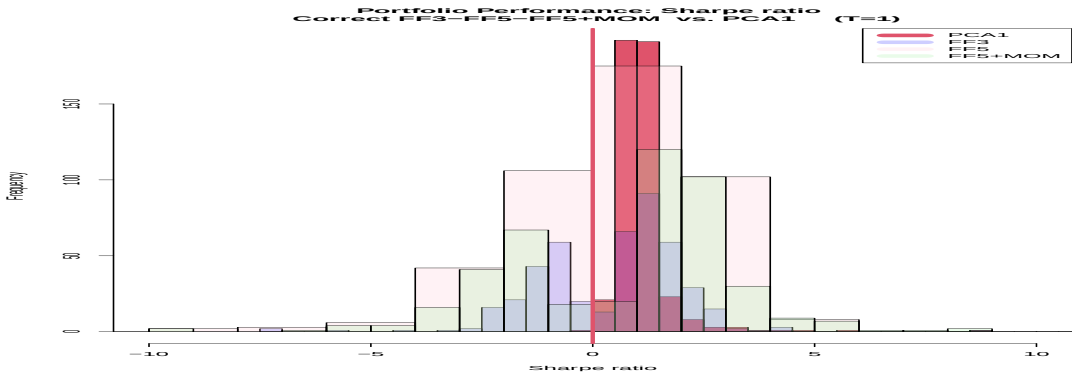


Figure OA.21: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widetilde{SR}_{t,t+1,t+2}$ (red bars) corresponding to PCA1, and three versions of $\widetilde{SR}_{1:T,t+1}$ corresponding to FF3 (blue bars), FF5 (pink bars), and FF5+Momentum (green bars) - see Section 6.

Tables 17 and 16 report the analysis in terms of behaviour of the pricing test and SR during the recession periods. The results demonstrate that the pricing performance deteriorates sharply during recessions, with slightly sharper results as the number of latent risk factors increase. This patterns is stronger in terms of Sharpe ratio performance, with the Sharpe ratios deteriorating sharply during recessions, both in terms of the (negative) magnitude and significance of the effect of the recession dummy on the Sharpe ratios, and in terms of the average Sharpe ratios (negative during recessions, positive otherwise), with again the strongest effect arising for PCA10.

Turning to the performance of the observed factor models, from Tables 19 and 18, one observes a similar pattern. Interestingly, in terms of Sharpe ratios, the observed risk factor models offer a slightly higher performance when averaged across time. On the other hand, the sensitivity of the pricing test to the recession dummy is weaker than for the latent models case.

Overall, the PCA models generate stronger raw cross-sectional "signal" (large absolute t's and decent Sharpe ratios, especially as the number of PCs grows), whereas the observed FF3/FF5/FF5+MOM models trade some of that raw signal for better risk adjustment and more efficient pricing, with FF5+MOM in particular offering a favourable combination of relatively high Sharpe ratios and reduced full-sample mispricing.

8 Conclusion

This paper develops a new methodology for evaluating the pricing ability and portfolio implications of conditional factor asset pricing models when only a single time-series observation, $T = 1$, is available but the cross-sectional dimension N is large. By exploiting the rich information contained in the cross-section of returns, I show that it is possible to conduct out-of-sample inference on risk premia, on the stochastic discount factor, and on mean–variance efficient portfolios without imposing tight parametric restrictions on the time-series dynamics of risk exposures or risk premia. The analysis rests on a novel asymptotic theory for Principal Component Analysis, "nano-PCA", which is valid when $T = 1$ and $N \to \infty$, and on pricing tests that remain informative even in this extreme short-panel setting.

The theoretical results clarify what can and cannot be learned about pricing ability when $T = 1$. First, I characterize the sources of sampling and misspecification bias that arise from relying on a one-factor representation and derive bias-adjusted statistics that deliver valid inference on pricing errors and Sharpe ratios. Second, I show that, despite the noisiness of estimated risk exposures, one can still recover the moments of these exposures that matter for mean–variance portfolio construction, and hence approximate the returns of efficient portfolios with high accuracy. Third, the framework accommodates both observed factors and latent factors extracted by nano-PCA, and it applies equally to pure factor models and to reduced-form stochastic discount factors.

The empirical applications illustrate the practical usefulness of the methodology for addressing the factor "zoo" problem. By constructing pricing tests and mean–variance portfolios at $T = 1$ across many hundreds of proposed risk factors, and by contrasting crisis and non-crisis periods as well as PCA-based and non-PCA specifications, I document that only a relatively small subset of factors displays robust pricing ability and economically meaningful portfolio performance. Moreover, the results reveal strong state dependence: many factors are particularly powerful during financial crises, when risk premia, mispricing, and cross-sectional dispersion are amplified. Finally, the comparison between specifications with and without PCA shows that common components play a central role in reshaping factor rankings, and that failure to account for them can lead to misleading inference about which characteristics are truly priced.

Overall, the paper demonstrates that meaningful asset pricing and portfolio analysis is feasible even in the extreme case where $T = 1$, provided that the cross-sectional dimension is large and that inference is built on appropriate asymptotics. The proposed methodology complements existing cross-sectional asset pricing

approaches and provides new tools for studying how the pricing ability of asset pricing models varies across factors and across economic states when data are short in the time dimension but rich in the cross-section.

9 Appendix

The following assumptions are assumed to hold for two generic periods t and $t+1$, that make the set $\mathcal{T}_t \equiv \{t, t+1\}$. For estimation with nano-PCA, I only need the assumptions below to hold for one period but for simplicity I am positing them for both periods of the set $\mathcal{T}_t$, needed for inference.

Remark: for esti and for test

Rem: nonrandom lambda. strong factor.

Remark. When $r_t = 1$ then scalars XXX.

Assumptions for $T = 1 = r_t = 1$.

Assumption 4 (factors) *$f_s^2 > 0$ a.s. for all $s \in \mathcal{T}_t$.*

Assumption 5 (risk exposures) *Setting $\boldsymbol{\lambda}_{s-1} = (\lambda_{1s-1}, \cdots, \lambda_{Ns-1})'$,*

$$\mathbf{1}_N'\boldsymbol{\lambda}_{s-1}/N \to \mu_{\lambda_{t-1}}, \ \ \boldsymbol{\lambda}_{s-1}'\boldsymbol{\lambda}_{s-1}/N \to \sigma_{\lambda_{t-1}} > 0 \text{ for every } s \in \mathcal{T}_t,$$

for some μ_{λ_t} and σ_{λ_t}, constant in $\mathcal{T}_t$, satisfying $0 < \inf_{t\in\mathcal{Z}} |\mu_{\lambda_t}|, \sigma_{\lambda_t} \leq \sup_{t\in\mathcal{Z}} |\mu_{\lambda_t}|, \sigma_{\lambda_t} < \infty$.

Assumption 6 (idiosyncratic error - temporal and cross-sectional dependence) *Set $\sigma_{ij,uv} = \mathcal{E}(e_{iu}e_{jv})$, with $\boldsymbol{\Sigma}_{et} \equiv [\sigma_{ij,tt}]_{i,j=1}^N$, and $\kappa_{h,i_1\ldots i_h,t_1\ldots t_h} = cumulant(e_{i_1t_1}, e_{i_2t_2}, \ldots, e_{i_ht_h})$. For for every $1 \leq i, j \leq N$, every $s, v \in \mathcal{T}_t$, and some positive constant C, where $g_1(\cdot)$ denotes the maximum eigenvalue operator and $\mathbb{1}_{(\cdot)}$ is the indicator function,*

$$\mathcal{E}(e_{is}) = 0, \mathcal{E}(e_{it}e_{is}e_{jv}) = 0, |\sigma_{ij,sv}| \leq C, |\kappa_{iijj,ttvs}| \leq C$$

$$\frac{1}{N}\sum_{i=1}^N (e_{is}e_{iv} - \sigma_{ii,sv}) \to_p 0 \text{ and } \frac{1}{N}\sum_{i=1}^N (e_{is}^4 - \kappa_{4,iiii,ssss} - 3\sigma_{ii,ss}^2) \to_p 0, \text{ with}$$

$$\frac{1}{N}\sum_{i=1}^N (\sigma_{ii,sv} - \mathbb{1}_{(s=v)}\sigma_t^2) = o(\frac{1}{\sqrt{N}}), \frac{1}{N}\sum_{i=1}^N \sum_{\substack{j=1 \\ j\neq i}}^N |\sigma_{ij,sv}| \to 0 \text{ and } \sup_{s\in\mathcal{T}_t} g_1(\boldsymbol{\Sigma}_{es}) = o(N^{\frac{1}{2}}),$$

$$\frac{1}{N}\sum_{i=1}^N (\sigma_{ii,ss}^2 - \sigma_{4t}) = o(\frac{1}{\sqrt{N}}), \frac{1}{N}\sum_{i=1}^N (\kappa_{4,iiii,ttsv} - \mathbb{1}_{(t=v=s)}\kappa_{4t}) = o(\frac{1}{\sqrt{N}}),$$

and $N^{-1}\sup_{i_1}\sum_{i_2,\ldots,i_h=1}^N |\kappa_{h,i_1\ldots i_h,t_1\ldots t_h}| \to 0$, for every $3 \leq h \leq 8$, every $t_j \in \mathcal{T}_t$ $(1 \leq j \leq h)$ and for at least one i_j $(2 \leq j \leq h)$ different from i_1, for some $\sigma_t^2, \sigma_{4t}, \kappa_{4t}$, constant in $\mathcal{T}_t$, satisfying

$$0 < \inf_{t\in\mathcal{Z}} \sigma_t^2 \leq \sup_{t\in\mathcal{Z}} \sigma_t^2 < \infty, 0 < \inf_{t\in\mathcal{Z}} \sigma_{4t} \leq \sup_{t\in\mathcal{Z}} \sigma_{4t} < \infty, 0 \leq \inf_{t\in\mathcal{Z}} |\kappa_{4t}| \leq \sup_{t\in\mathcal{Z}} |\kappa_{4t}| < \infty.$$

Assumption 7 (mixed moments) *For every $s \in \mathcal{T}_t$, setting $\mathbf{e}_{.i} \equiv (e_{it}, e_{it+1})'$,*

$$\frac{1}{N}\sum_{i=1}^N \lambda_{is-1}\mathbf{e}_{.i} \to_p \mathbf{0}_{2\times 1} \text{ and } \frac{1}{N}\sum_{i=1}^N \lambda_{is-1}^2 \mathcal{E}(\mathbf{e}_{.i}\mathbf{e}_{.i}')) \to \sigma_{\lambda_t}\sigma_t^2\mathbf{I}_2.$$

Assumption 8 (joint convergence in distribution) *For every $s \in \mathcal{T}_t$, setting ι_s the 2×1 vector with 1 in position $s-(t-1)$ and zero everywhere else,*

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \begin{bmatrix} (\mathbf{e}_{.i} e_{is} - \mathcal{E}(\mathbf{e}_{.i} e_{is})) \\ \lambda_{is-1} \mathbf{e}_{.i} \end{bmatrix} \to_d N\Big(\mathbf{0}_4, \begin{bmatrix} (\kappa_{4t} + \sigma_{4t}) \iota_s \iota_s' + \sigma_{4t} \mathbf{I}_2 & \mathbf{0}_{2\times 2} \\ \mathbf{0}_{2\times 2} & \sigma_{\lambda_t} \sigma_t^2 \mathbf{I}_2 \end{bmatrix}\Big).$$

Assumptions for $T > r_t \geq 1$.

The previous assumptions need to be generalized, except for Assumption 6. Now $\mathcal{T}_t \equiv \{t-T+1, \cdots, t+1\}$ and $\mathbf{F} \equiv (\mathbf{f}_{t-T+2}, \cdots, \mathbf{f}_{t+1})'$ with $\mathbf{f}_s \equiv (f_{1s}, \cdots, f_{r_t\, s})$ for every $s \in \mathcal{T}_t$.

Assumption 9 (number of factors) *The number of risk factors is locally constant: $r_s = r_t$ for every $s \in \mathcal{T}_t$.*

Assumption 10 (factors) $\hat{\mathbf{\Sigma}}_{\mathbf{F}} = \mathbf{F}'\mathbf{F}/T > 0$ *a.s. for any T such that $r_t \leq T < \infty$.*

Assumption 11 (risk exposures) $\mathbf{1}_N' \mathbf{\Lambda}_{s-1}/N \to \boldsymbol{\mu}_{\lambda_{t-1}}$, $\mathbf{\Lambda}_{s-1}' \mathbf{\Lambda}_{s-1}/N \to \mathbf{\Sigma}_{\Lambda_{t-1}} > 0$ *for every $s \in \mathcal{T}_t$.*

Assumption 12 (eigenvalues) *The eigenvalues of the $r_t \times r_t$ matrix $(\mathbf{\Sigma}_{\Lambda_{t-1}} \hat{\mathbf{\Sigma}}_{\mathbf{F}})$ are distinct a.s.*

Assumption 13 (mixed moments) *For every $s \in \mathcal{T}_t$, setting $\mathbf{e}_{.i} = (e_{it-T+2}, \cdots, , e_{it+1})'$,*

$$\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\lambda}_{is-1} \mathbf{e}_{.i}' \to_p \mathbf{0}_{r_t \times T} \text{ and } \frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{\lambda}_{is-1} \boldsymbol{\lambda}_{is-1}' \otimes \mathcal{E}(\mathbf{e}_{.i} \mathbf{e}_{.i}')) \to (\mathbf{\Sigma}_{\mathbf{\Lambda}_{t-1}} \otimes \sigma_t^2 \mathbf{I}_T).$$

Assumption 14 (joint convergence in distribution) *For every $s \in \mathcal{T}_t$, setting ι_s the $T \times 1$ vector with 1 in position $s-(t-T+1)$ and zero everywhere else,*

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \begin{bmatrix} (\mathbf{e}_{.i} e_{is} - \mathcal{E}(\mathbf{e}_{.i} e_{is})) \\ (\boldsymbol{\lambda}_{is-1} \otimes \mathbf{e}_{.i}) \end{bmatrix} \to_d N\Big(\mathbf{0}_{T(r_t+1)}, \begin{bmatrix} (\kappa_{4t} + \sigma_{4t}) \iota_s \iota_s' + \sigma_{4t} \mathbf{I}_T & \mathbf{0}_{T \times T r_t} \\ \mathbf{0}_{T r_t \times T} & (\mathbf{\Sigma}_{\mathbf{\Lambda}_{t-1}} \otimes \sigma_t^2 \mathbf{I}_T) \end{bmatrix}\Big).$$

10 The matrix $\mathbf{U}_\epsilon$

The expression for $\mathbf{U}_\epsilon$ can be derived in closed form. In particular, the $T^2 \times T^2$ matrix $\mathbf{U}_\epsilon$ has the following structure:

$$\mathbf{U}_\epsilon = \left[\begin{array}{ccc:c:c} U_{11} & \cdots & U_{1t} & \cdots & U_{1T} \\ \vdots & \ddots & \vdots & & \vdots \\ U_{t1} & \cdots & U_{tt} & \cdots & U_{tT} \\ \vdots & & \vdots & \ddots & \vdots \\ U_{T1} & \cdots & U_{Tt} & \cdots & U_{TT} \end{array}\right].$$

Each block of $\mathbf{U}_\epsilon$ is a $T \times T$ matrix. The blocks along the main diagonal, denoted U_{tt} for $t = 1, 2, \ldots, T$, are themselves diagonal matrices with $(\kappa_4 + 2\sigma_4)$ in the (t,t)-th position and σ_4 in the (s,s) position for every

$s \neq t$. The blocks outside the main diagonal, denoted U_{ts} for $s, t = 1, 2, \ldots, T$ with $s \neq t$, are all zero matrices except for the (s,t)-th element, which equals σ_4.

$$U_{tt} = \underset{\substack{t\text{-th}\\ \text{row}}}{\rightarrow} \begin{bmatrix} \sigma_4 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \sigma_4 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & (\kappa_4 + 2\sigma_4) & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & \sigma_4 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sigma_4 \end{bmatrix}, \quad U_{ts} = \underset{\substack{s\text{-th}\\ \text{row}}}{\rightarrow} \begin{bmatrix} 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & \sigma_4 & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix}.$$

$$\overset{\downarrow}{t\text{-th column}} \qquad \overset{\downarrow}{t\text{-th column}}$$

11 Time Series Plots of Zoo Factors

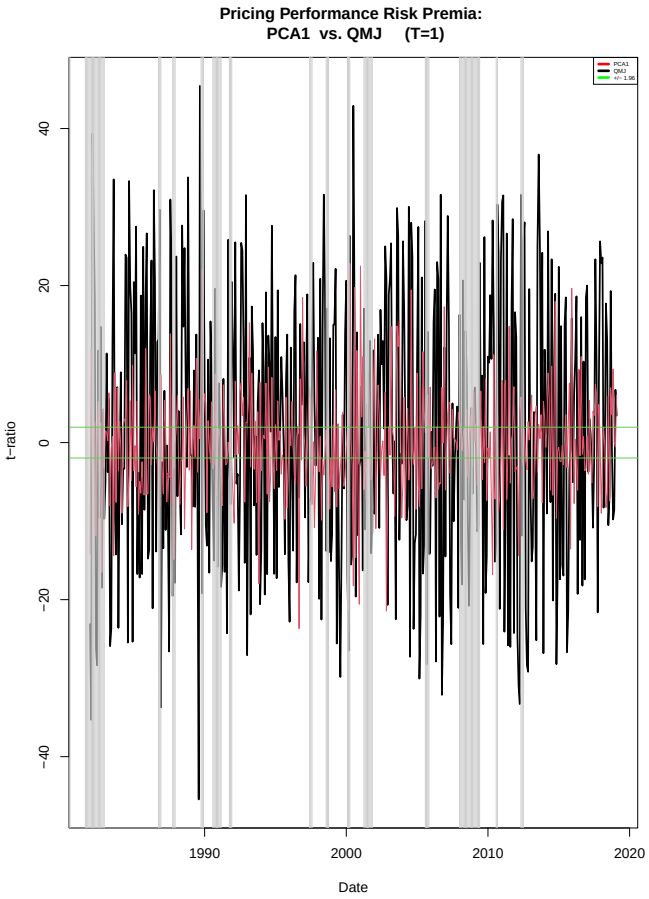


Figure OA.22: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and QMJ risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

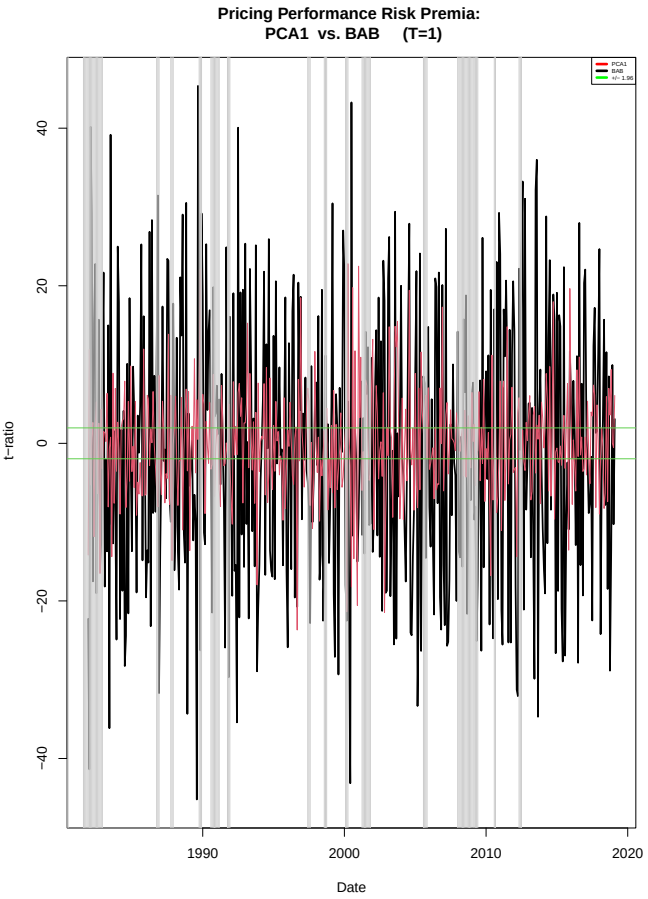
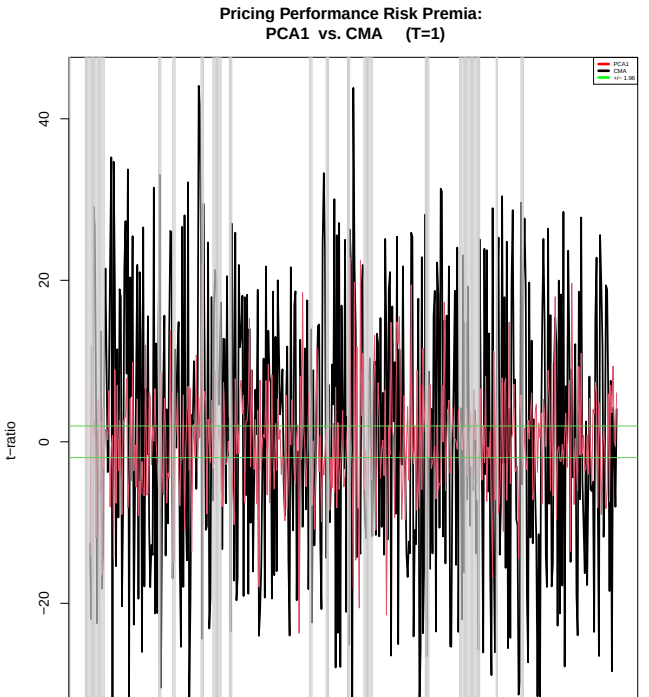
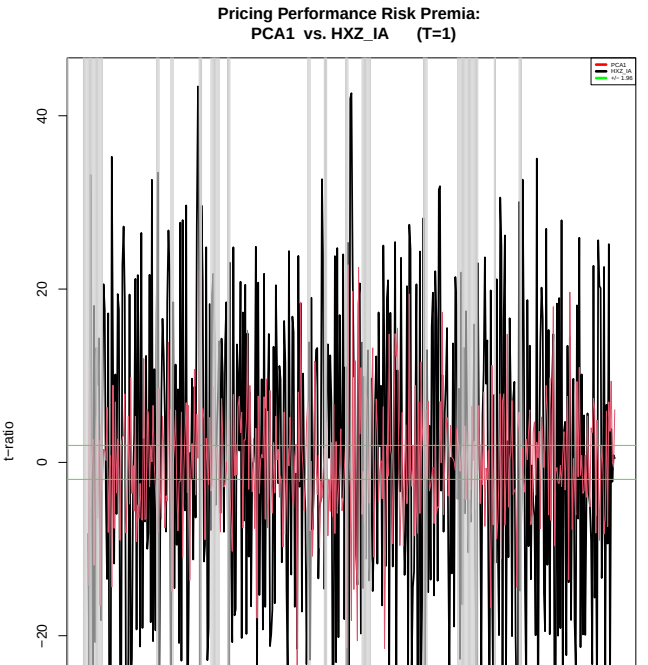


Figure OA.23: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and BAB risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.





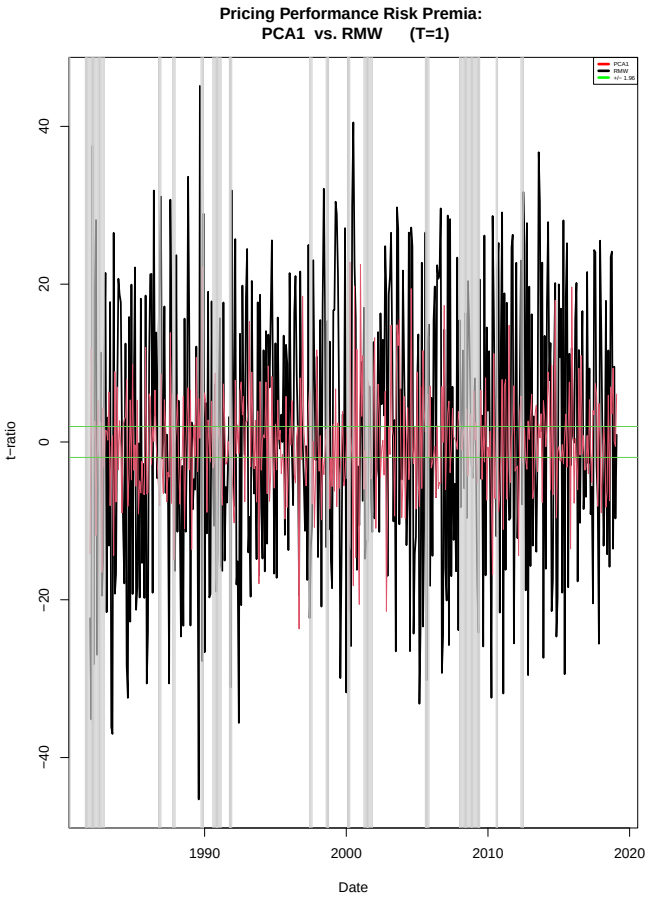


Figure OA.27: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and RMW risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

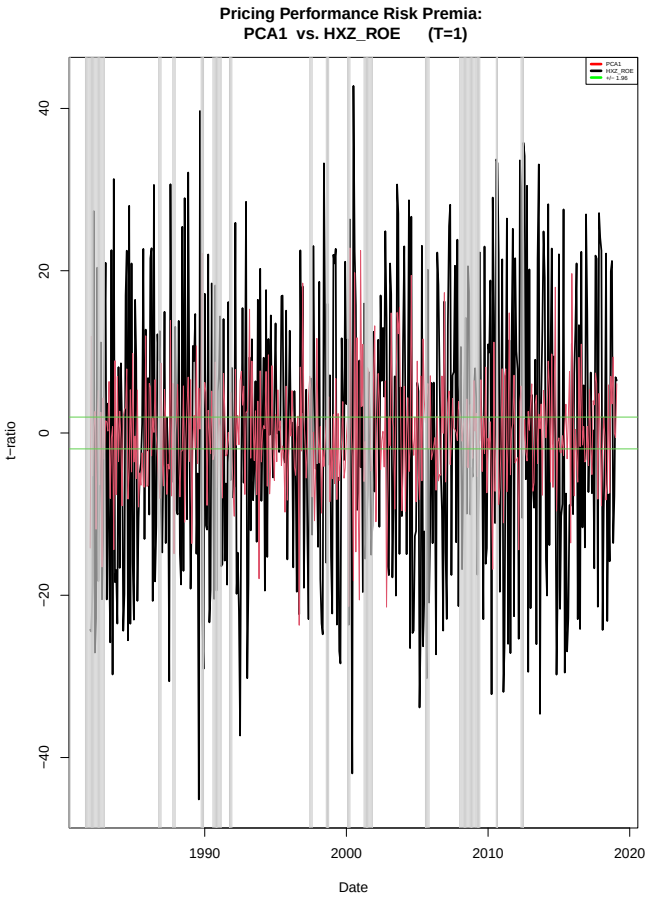
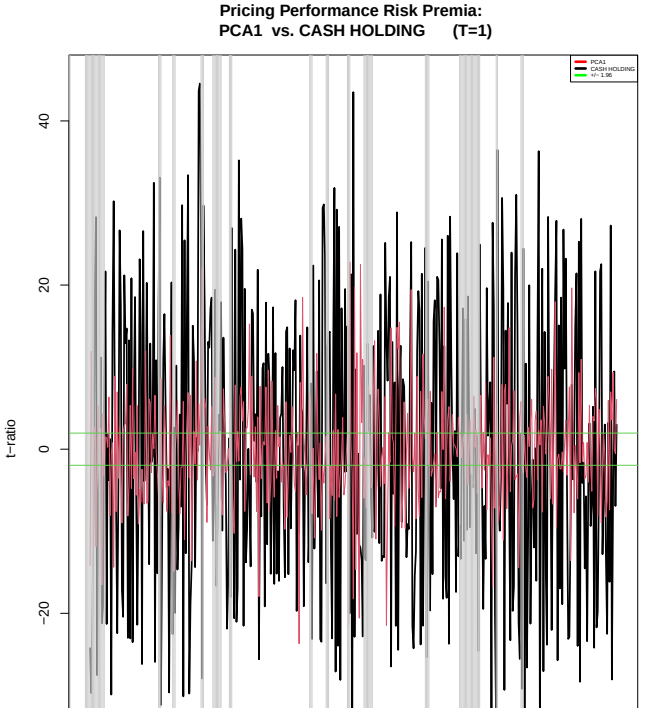
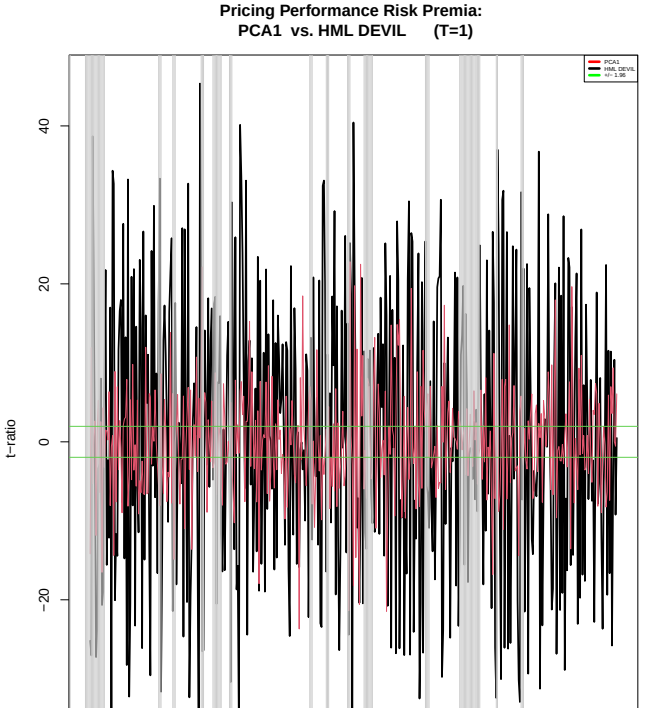


Figure OA.28: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and HXZ_ROE risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.





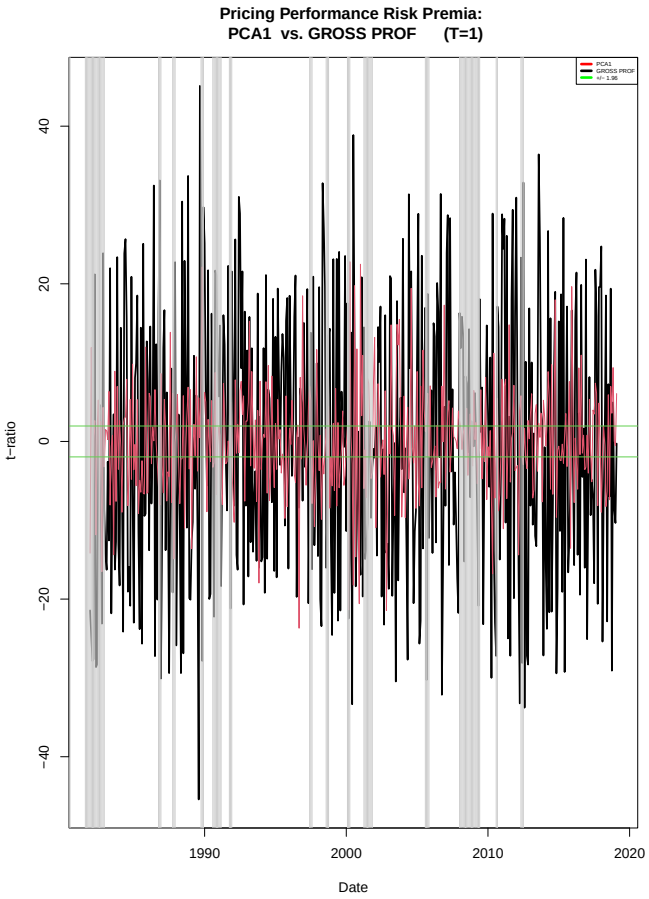


Figure OA.31: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and GROSS PROF risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

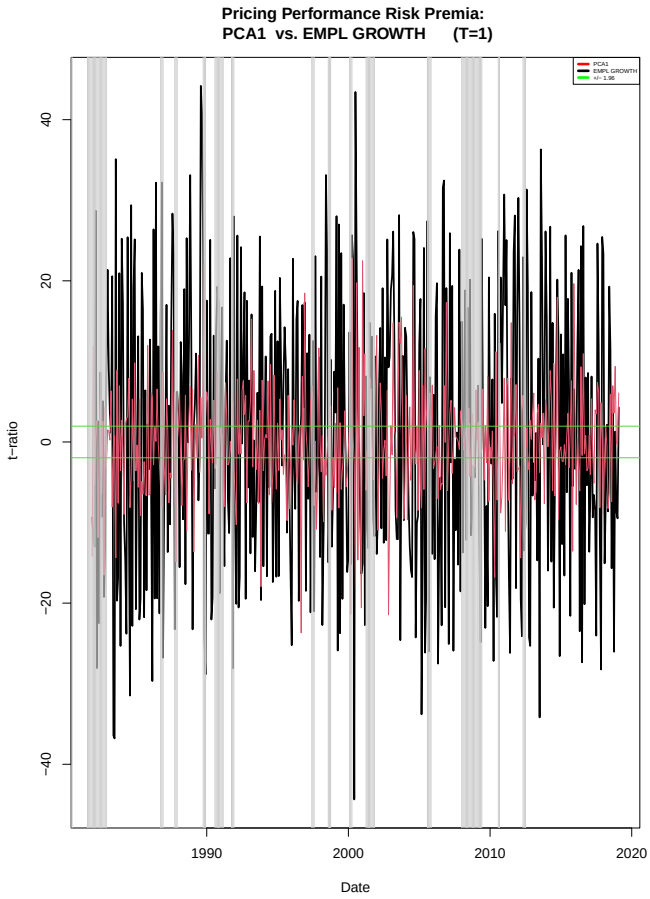
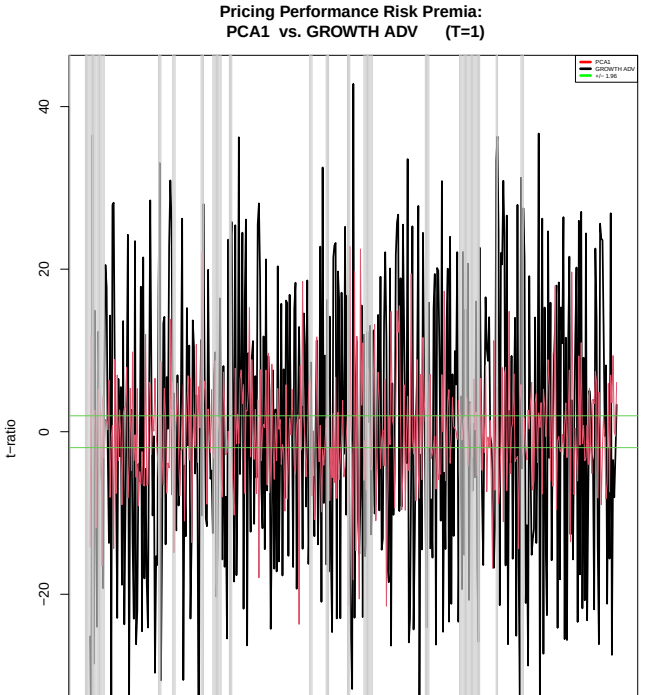
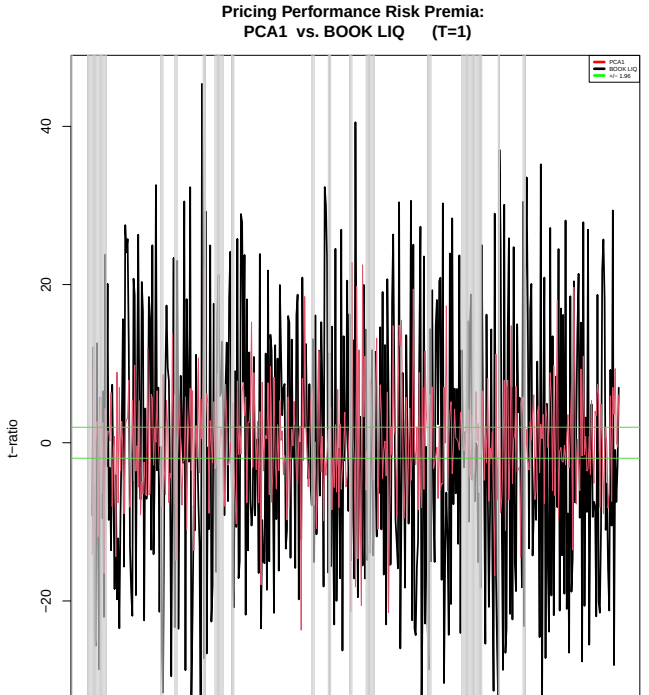


Figure OA.32: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and EMPL GROWTH risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.





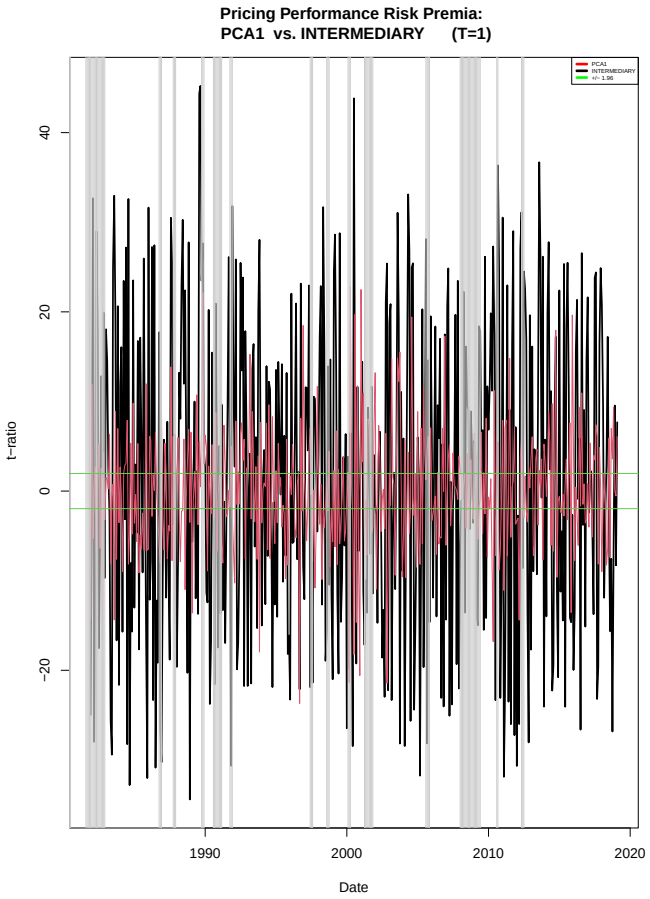


Figure OA.35: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to PCA1 and INTERMEDIARY risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

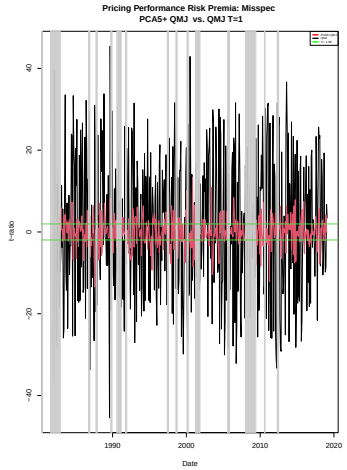


Figure OA.36: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to QMJ and QMJ+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

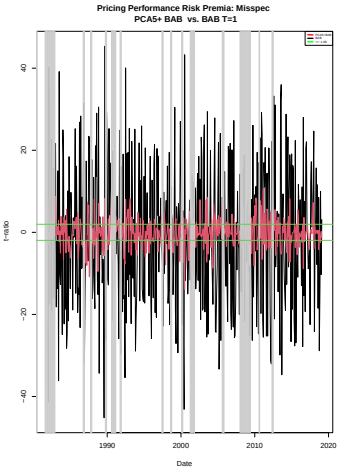


Figure OA.37: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to BAB and BAB+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

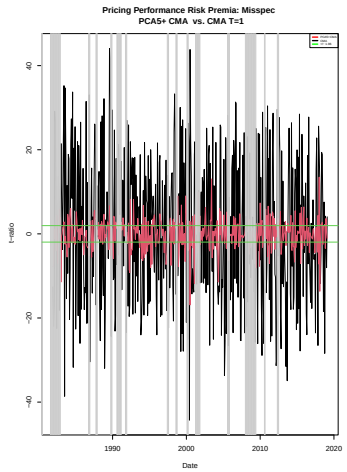


Figure OA.38: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to CMA and CMA+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to

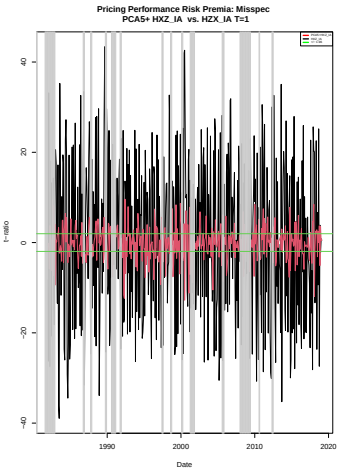


Figure OA.39: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to HXZ_IA and HXZ_IA +PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines corre-

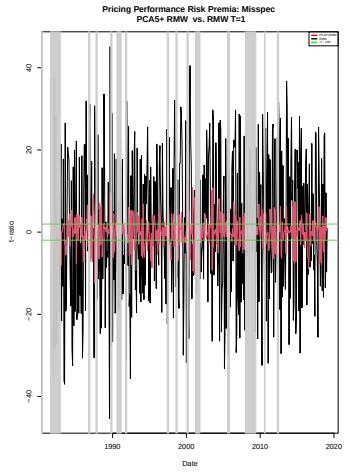


Figure OA.40: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to RMW and RMW+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

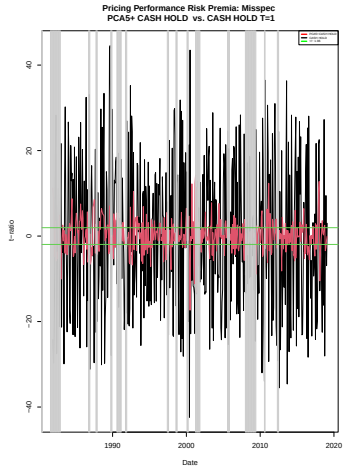


Figure OA.42: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to CASH HOLD and CASH HOLD+PCA5 risk factor, respectively - see Section 4.1. The hor-

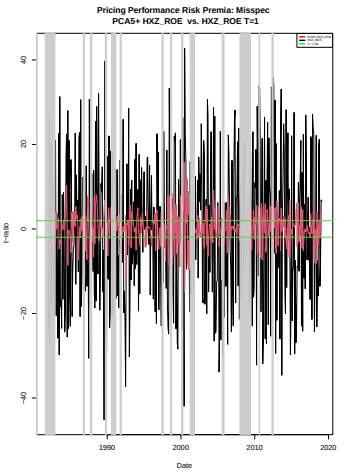


Figure OA.41: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to HXZ_ROE and HXZ_ROE+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

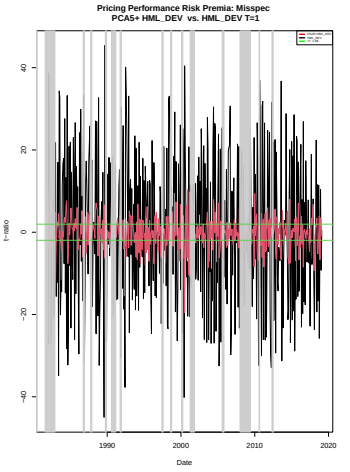


Figure OA.43: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to HML_DEV and HML_DEV+PCA5 risk factor, respectively - see Section 4.1. The horizontal

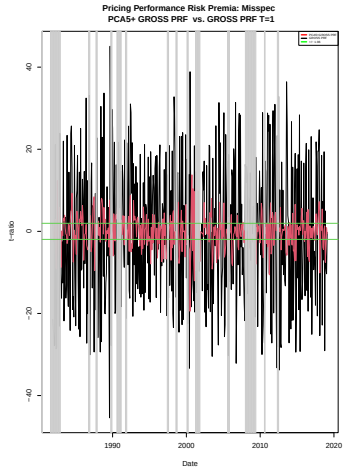


Figure OA.44: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to GROSS PROF and GROSS PROF+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

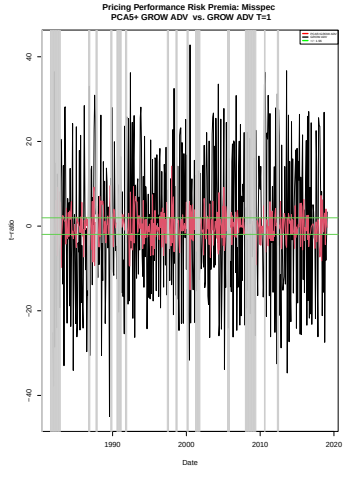


Figure OA.46: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to GROWTH AD and GROWTH AD+PCA5 risk factor, respectively -

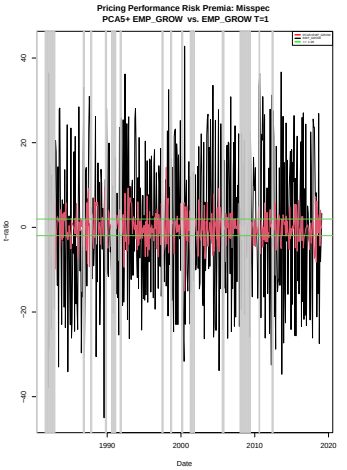


Figure OA.45: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to EMP GROW and EMP GROW+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

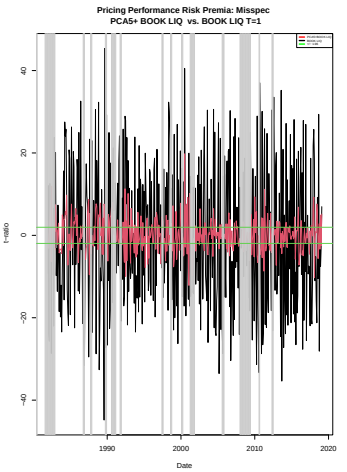


Figure OA.47: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to BOOK LIQ and BOOK LIQ+PCA5 risk factor, respectively - see Section 4.1.

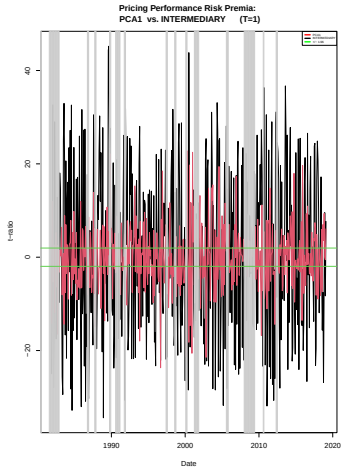


Figure OA.48: The plot reports the time series (from Dec 1981 to Feb 2019) of the t-ratio for the risk premium pricing test $\tilde{L}_{t+1|t}$ (red line) and $\tilde{L}^{o}_{t+1|t}$ (black line) corresponding to INTERM and INTERM+PCA5 risk factor, respectively - see Section 4.1. The horizontal green lines correspond to -1.96 and $+1.96$. The grey bars correspond to economic and financial crises, and the NBER recession indicator.

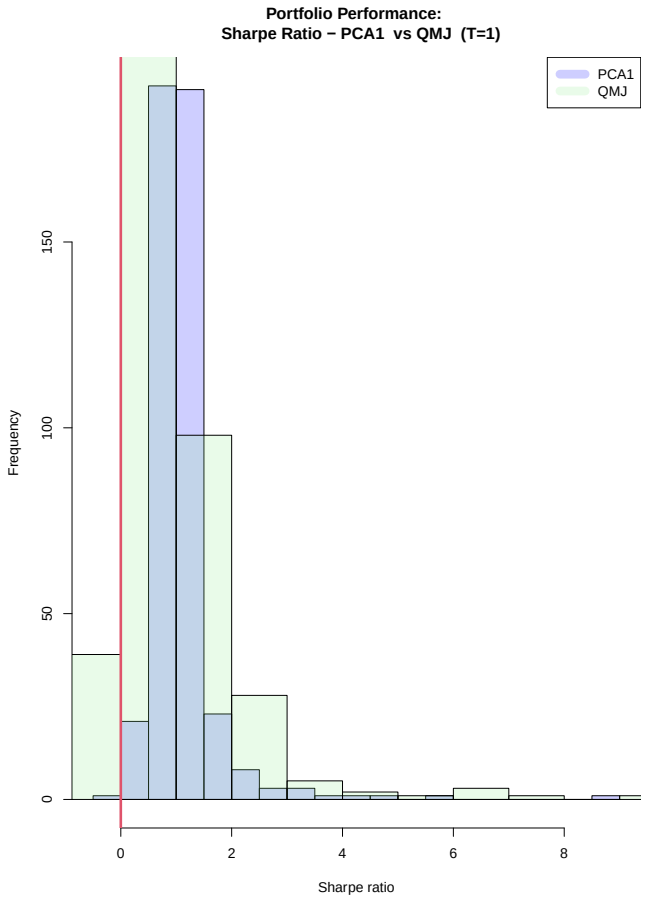


Figure OA.49: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the QMJ risk factor respectively - see Section 4.2.

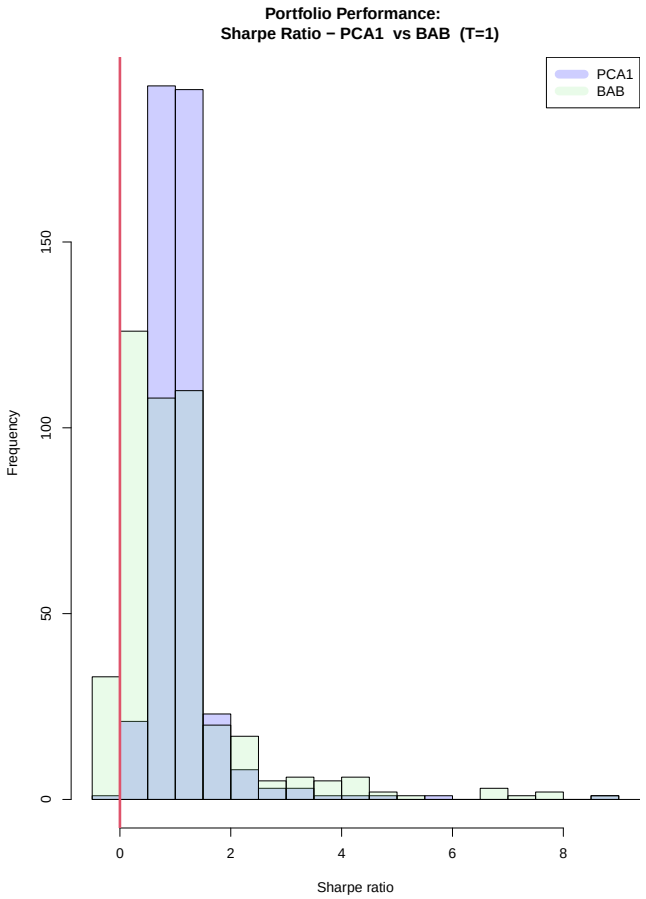
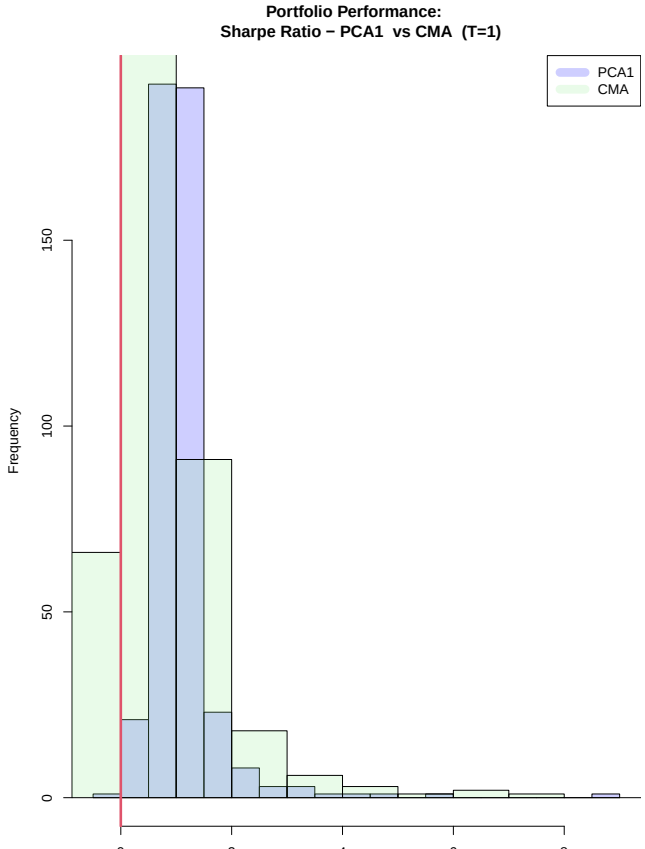
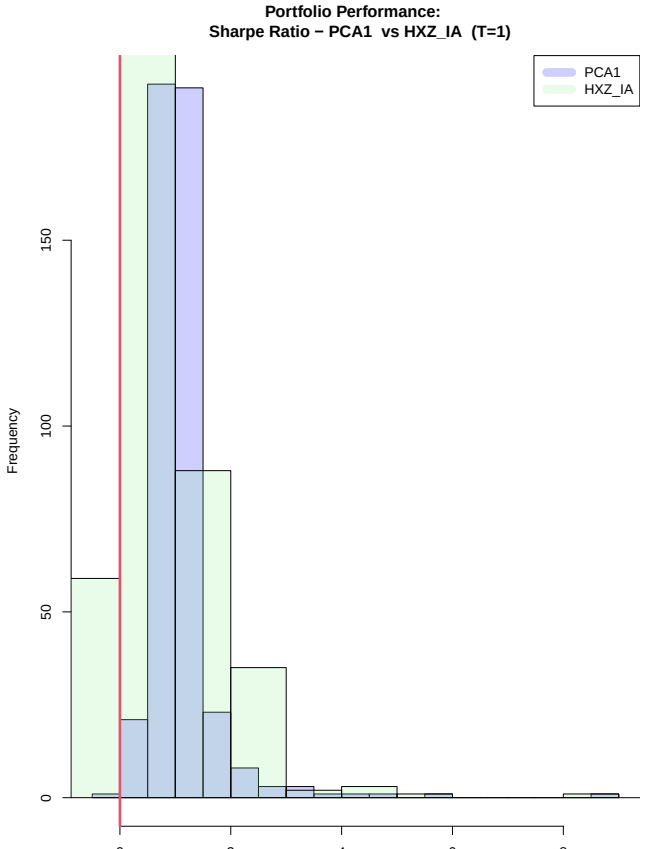


Figure OA.50: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the BAB risk factor respectively - see Section 4.2.





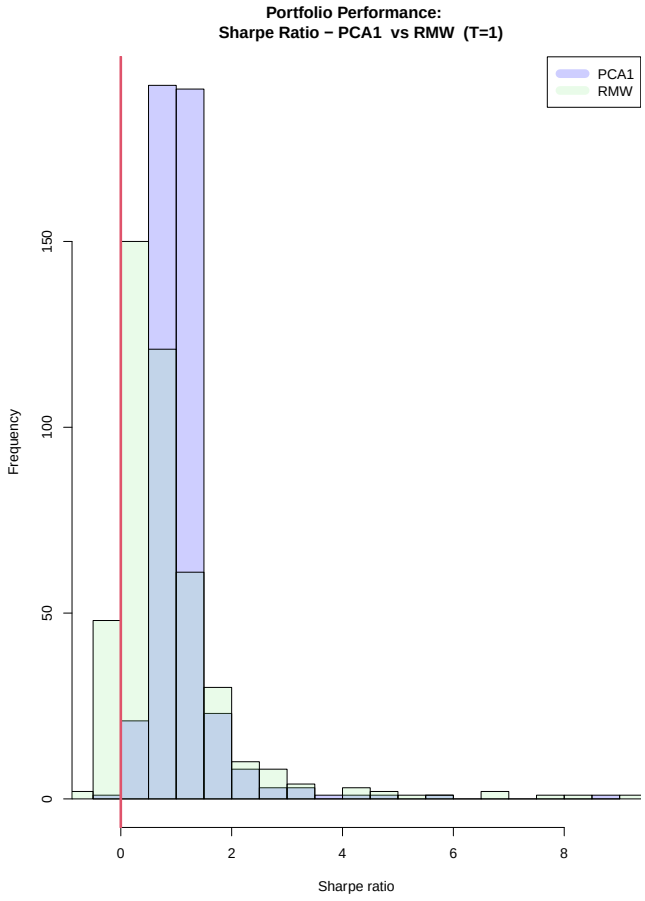


Figure OA.53: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the RMW risk factor respectively - see Section 4.2.

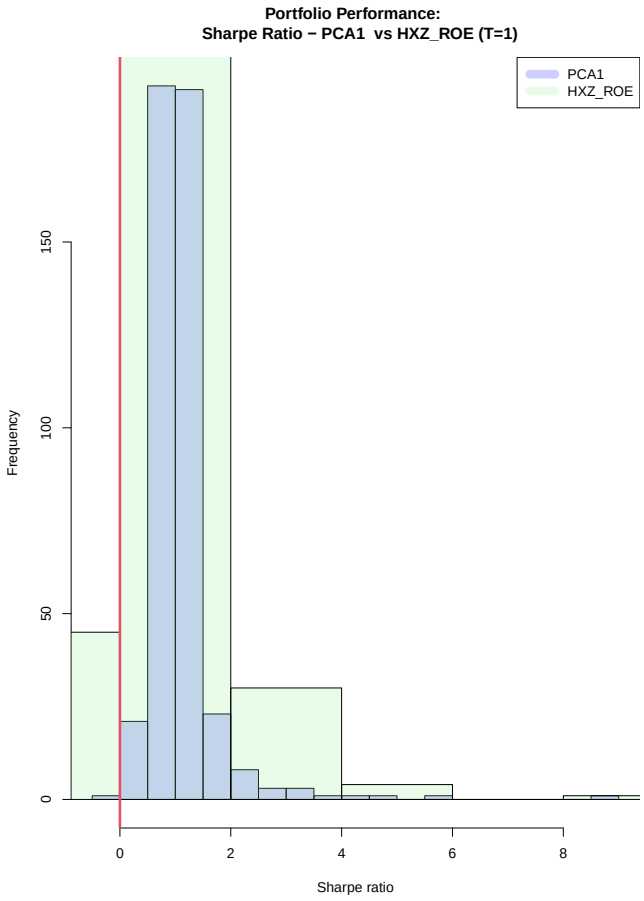
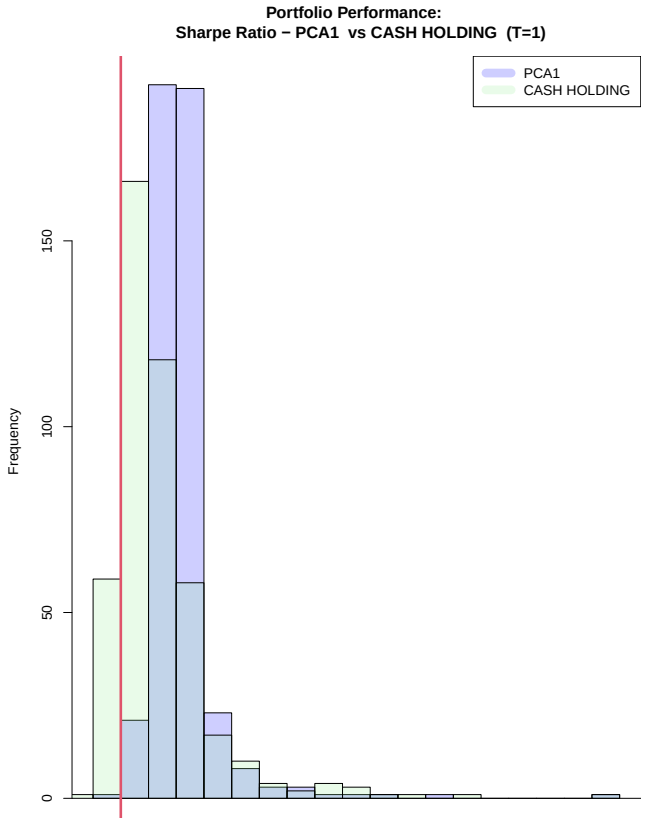
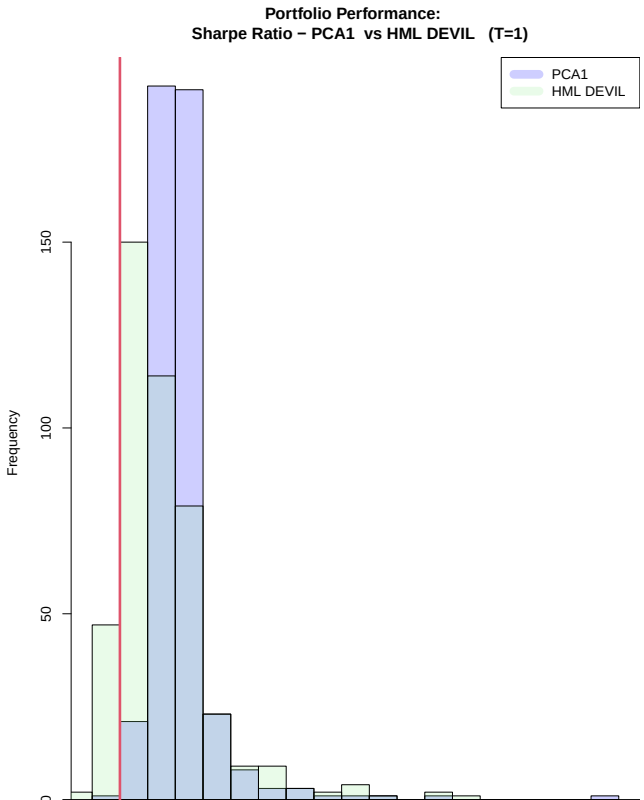


Figure OA.54: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the HXZ_ROE risk factor respectively - see Section 4.2.





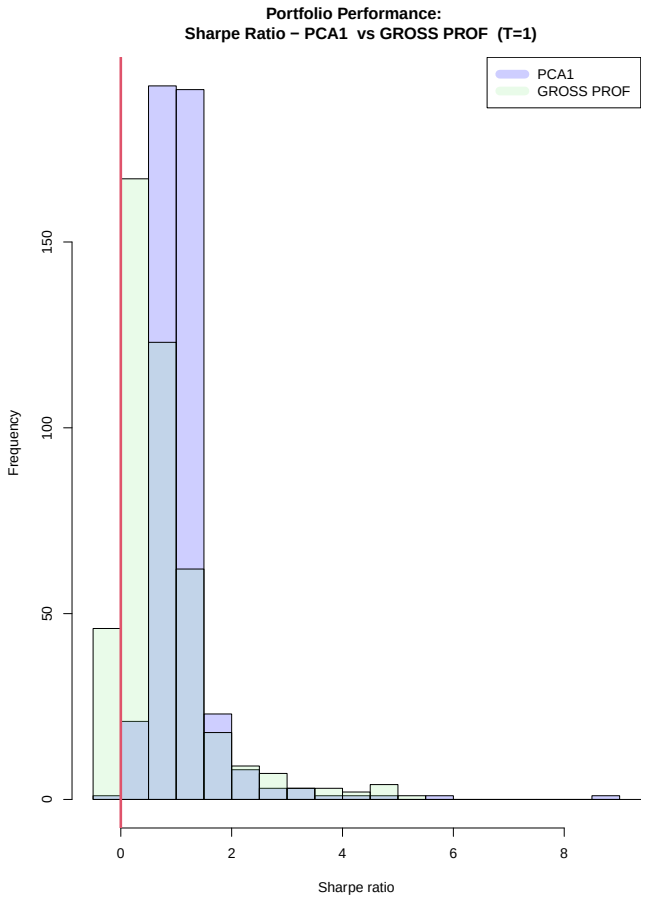


Figure OA.57: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the GROSS PROFIT risk factor respectively - see Section 4.2.

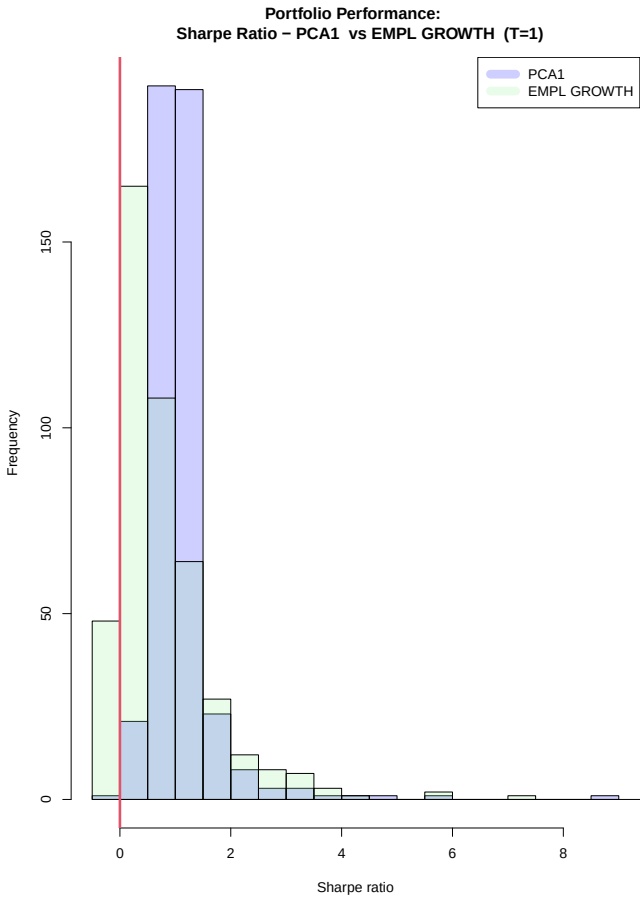
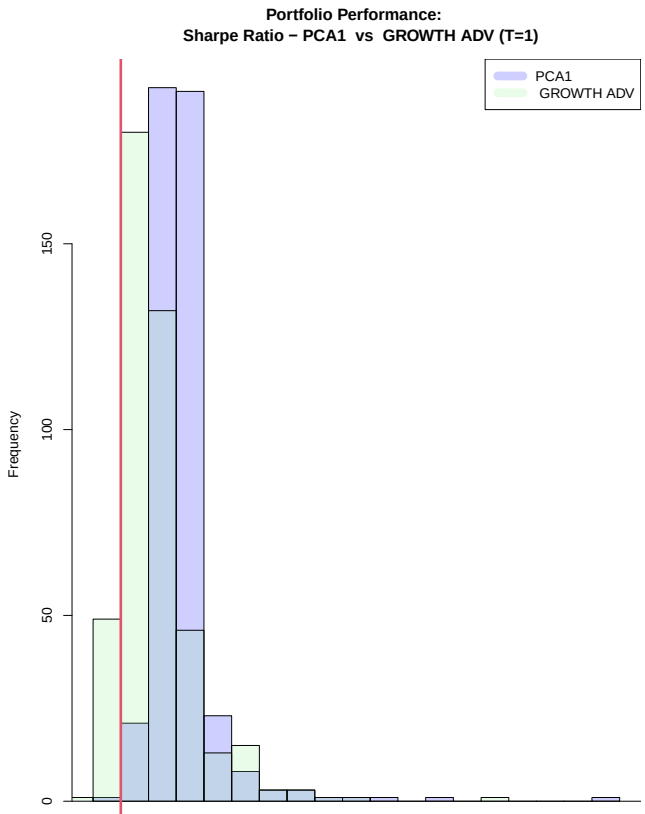
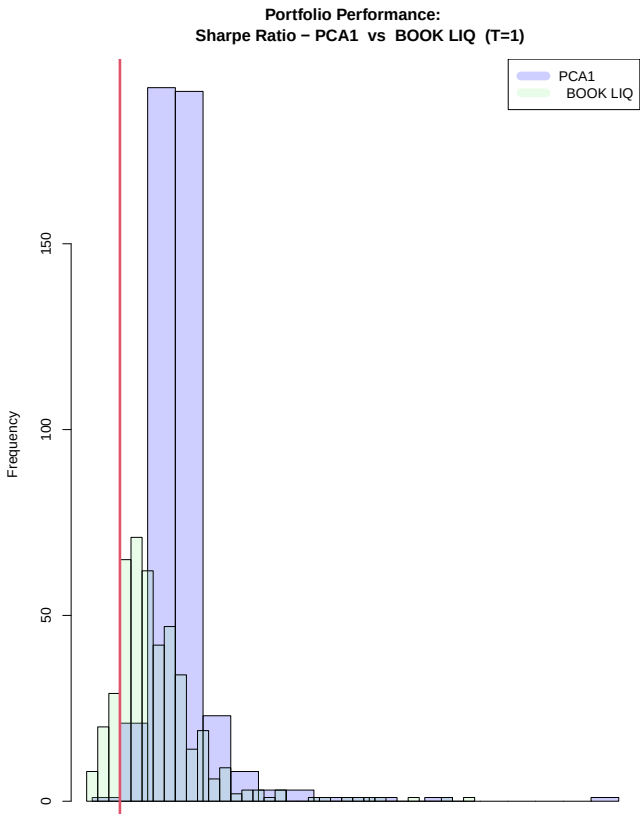


Figure OA.58: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the EMPL GROWTH risk factor respectively - see Section 4.2.





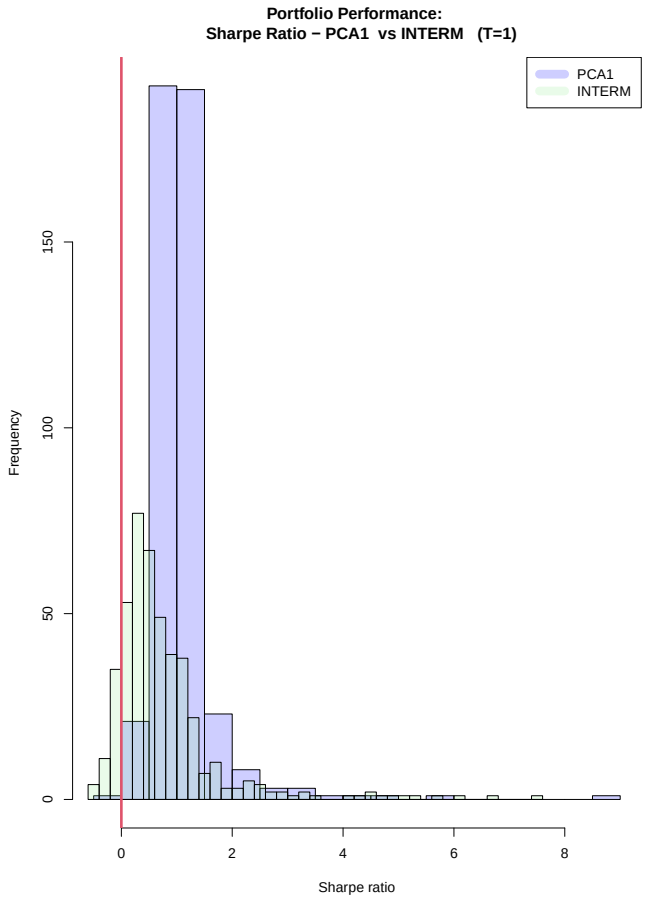


Figure OA.61: The plot reports the histograms (from Dec 1981 to Feb 2019) of the t-ratio for the Sharpe ratio $\widehat{SR}_{t+1|t}$ (blue bars) and $\widetilde{SR}_{t,t+1,t+2}$ (green bars) corresponding to PCA1 and the INTERMEDIARY risk factor respectively - see Section 4.2.

12 Tables

12.1 Risk Premia Test - $T = 1$ - Correct Specification

Table 1: Core Factors - Column 1: OLS estimator of recession dummy on pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test), namely the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, setting $\widetilde{test}_t \equiv \sqrt{N}|\tilde{L}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$ where $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 2) and d_t is a dummy variable that takes value one when t is one of the recession-financial crises periods, and zero otherwise; Column 2 t-ratio of OLS coefficient; Column 3 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over recession periods; Column 4 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over all periods; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 2×1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 5×1.96;

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times 1$
MktRf	3.490	3.141	6.294	−0.448	0.951	0.881	0.723
SMB	4.238	3.924	1.854	−0.128	0.942	0.866	0.658
HML	4.480	3.979	3.589	0.057	0.926	0.884	0.689
RMW	3.464	3.222	3.603	1.141	0.942	0.881	0.702
CMA	3.421	3.249	4.827	1.331	0.955	0.904	0.707
UMD	5.062	4.656	3.795	0.245	0.919	0.866	0.667
HML_Devil	5.794	5.208	3.793	0.552	0.922	0.864	0.689
BAB	4.414	3.963	1.658	−1.046	0.908	0.839	0.655
QMJ	2.618	2.433	5.605	1.550	0.944	0.864	0.694
HXZ_IA	3.762	3.639	2.385	0.754	0.955	0.899	0.714
HXZ_ROE	4.059	3.796	1.501	−0.409	0.944	0.886	0.673
Intermediary	4.537	4.172	5.793	1.370	0.944	0.861	0.658
convind	2.794	2.568	2.369	0.209	0.944	0.872	0.671
beta	4.080	3.870	3.020	−0.028	0.957	0.902	0.725
ep	3.124	2.806	4.317	0.940	0.935	0.879	0.723
dy	2.430	2.246	2.969	0.474	0.953	0.902	0.709
cp	3.579	3.269	4.059	0.685	0.942	0.879	0.698
pps	2.807	2.575	3.256	0.053	0.944	0.857	0.673
sue	3.508	3.283	3.637	0.554	0.944	0.872	0.671
LTR	4.538	4.211	2.453	−0.854	0.924	0.857	0.653
lev	4.325	3.913	6.165	0.806	0.957	0.868	0.689
cashdebt	5.578	5.069	4.262	0.498	0.928	0.868	0.689
cfp	3.114	2.933	4.540	0.461	0.955	0.888	0.707
cfp_ia	5.259	4.987	5.923	−0.209	0.937	0.877	0.671
bm_ia	3.408	3.214	2.829	0.945	0.949	0.895	0.707
ebp	3.440	3.214	5.110	1.158	0.953	0.893	0.705

Table 2: Liquidity, Volatility, and Trading Activity Factors - Column 1: OLS estimator of recession dummy on pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test), namely the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, setting $\widetilde{test}_t \equiv \sqrt{N}|\tilde{L}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$ where $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 2) and d_t is a dummy variable that takes value one when t is one of the recession-financial crises periods, and zero otherwise; Column 2 t-ratio of OLS coefficient; Column 3 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over recession periods; Column 4 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over all periods; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 2×1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 5×1.96;

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times 1$
turn	2.775	2.591	1.717	0.308	0.960	0.899	0.723
dolvol	3.222	3.030	3.164	0.329	0.944	0.893	0.687
std_dolvol	4.196	3.960	2.037	0.022	0.949	0.897	0.667
std_turn	3.452	3.156	2.151	0.408	0.946	0.872	0.702
ill	4.397	4.046	1.670	−0.579	0.926	0.850	0.660
LIQ_PS	2.719	2.460	4.672	−0.284	0.951	0.868	0.644
baspread	3.377	3.125	3.509	1.274	0.957	0.897	0.729
idiovol	2.589	2.458	2.677	0.105	0.962	0.893	0.723
retvol	3.450	3.230	3.506	0.751	0.969	0.913	0.707
zerotrade	2.884	2.640	2.304	0.372	0.960	0.893	0.716
pricedelay	3.702	3.428	5.029	−0.090	0.928	0.870	0.664
mom6m	4.789	4.470	0.434	−1.256	0.933	0.868	0.662
mom36m	4.620	4.359	2.026	−0.185	0.944	0.877	0.682
moms12m	4.917	4.610	0.220	0.145	0.931	0.859	0.655
chmom	4.759	4.622	1.536	−0.833	0.955	0.899	0.723
indmom	2.433	2.241	2.517	0.374	0.924	0.866	0.680

Table 4: Remaining Factors, Anomalies, and Constraints Factors - Column 1: OLS estimator of recession dummy on pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) namely the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, setting $\widetilde{test}_t \equiv \sqrt{N}|\tilde{L}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$ where $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 2) and d_t is a dummy variable that takes value one when t is one of the recession-financial crises periods, and zero otherwise; Column 2 t-ratio of OLS coefficient; Column 3 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over recession periods; Column 4 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over all periods; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 2×1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 5×1.96;

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
acc	1.857	1.741	2.907	−0.316	0.940	0.868	0.678
absacc	−0.226	−0.211	0.046	−0.401	0.928	0.881	0.723
stdacc	4.590	4.221	3.475	0.494	0.949	0.881	0.691
pctacc	2.002	1.932	1.885	−0.148	0.953	0.888	0.698
stdcf	3.669	3.377	4.154	0.780	0.940	0.884	0.691
gad	1.717	1.592	4.550	0.382	0.942	0.866	0.669
ala	4.232	4.064	0.300	−0.621	0.942	0.872	0.678
sin	4.724	4.375	0.564	−0.794	0.946	0.879	0.644
maxret	4.007	3.747	4.101	1.176	0.951	0.899	0.732
kz	3.827	3.573	3.538	0.269	0.953	0.906	0.729
ww	4.437	4.281	1.105	0.241	0.957	0.893	0.678
ms	3.222	2.833	4.662	1.157	0.928	0.843	0.620
rs	5.385	5.139	4.182	0.363	0.964	0.899	0.705
rsup	4.343	3.963	2.473	−0.286	0.953	0.881	0.685
herf	4.112	3.774	3.544	0.528	0.919	0.877	0.685
etr	5.197	4.793	4.388	−0.995	0.937	0.857	0.671
lfe	6.634	6.158	4.009	0.342	0.953	0.886	0.687
zs	3.540	3.181	4.157	0.751	0.949	0.881	0.694
rd	2.299	2.118	0.046	−1.053	0.951	0.904	0.676
adm	4.339	4.100	0.134	−0.523	0.942	0.879	0.678
rdm	2.512	2.426	1.929	−0.057	0.940	0.864	0.682
rds	2.367	2.299	0.290	−0.511	0.933	0.864	0.702
chinv	3.322	3.182	1.677	0.334	0.937	0.872	0.707
chtx	4.138	3.780	4.492	−0.131	0.940	0.890	0.673
grltnoa	4.453	4.309	4.684	0.068	0.933	0.861	0.689
grltnoa_hxz	3.716	3.500	2.897	−0.072	0.940	0.875	0.676
ob_a	4.859	4.582	0.623	−0.549	0.935	0.861	0.640
tb	3.324	2.979	3.188	0.172	0.940	0.857	0.664
age	3.379	3.093	4.794	1.334	0.935	0.868	0.687
egr	3.620	3.482	2.136	0.669	0.957	0.886	0.689
lgr	3.865	3.502	5.574	0.680	0.957	0.884	0.682

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
IPO	4.595	4.142	4.877	0.721	0.949	0.875	0.617
cei	4.734	4.481	1.066	−0.016	0.953	0.897	0.669
nef	4.122	3.826	6.357	1.270	0.955	0.893	0.727
ndf	4.363	4.151	4.264	0.267	0.937	0.888	0.658
nxf	3.716	3.406	7.363	1.605	0.946	0.879	0.680
dfin	4.058	3.969	3.977	0.010	0.946	0.897	0.694
dsti	4.960	4.639	−0.506	−1.049	0.940	0.864	0.662
dfnl	3.663	3.415	1.060	−0.719	0.928	0.857	0.669
egr_hxz	4.376	4.211	3.012	−0.298	0.953	0.904	0.685
ear	2.934	2.651	1.801	−0.433	0.937	0.868	0.664
chatoia	3.984	3.693	0.385	−0.405	0.937	0.886	0.698
chpmia	2.821	2.729	1.625	0.235	0.933	0.886	0.709
rna	3.409	3.113	0.129	−0.114	0.942	0.870	0.702
cdi	3.985	3.678	2.268	−0.113	0.935	0.881	0.667
chcsho	1.969	1.821	1.227	0.565	0.946	0.868	0.682
aeavol	3.868	3.559	2.640	−0.209	0.928	0.852	0.680
sp	4.417	4.054	4.991	0.049	0.944	0.866	0.687
cto	2.720	2.575	0.305	−0.463	0.940	0.875	0.709
STR	3.952	3.749	0.542	−0.589	0.942	0.886	0.671
sgr	3.536	3.378	1.622	−0.056	0.949	0.888	0.698
pchgm_pchsale	3.745	3.405	−0.273	−1.171	0.946	0.895	0.689
pchsale_pchinvt	4.836	4.438	1.031	−1.019	0.922	0.839	0.678
pchsale_pchrect	1.804	1.647	4.254	−0.793	0.944	0.879	0.669
pchsale_pchxsga	2.768	2.564	4.283	0.334	0.940	0.879	0.685

Table 3: Investment, Profitability, and Balance sheet Factors - Column 1: OLS estimator of recession dummy on pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) namely the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, setting $\widetilde{test}_t \equiv \sqrt{N}|\tilde{L}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$ where $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 2) and d_t is a dummy variable that takes value one when t is one of the recession-financial crises periods, and zero otherwise; Column 2 t-ratio of OLS coefficient; Column 3 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over recession periods; Column 4 average pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) over all periods; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 2×1.96; Column 5 frequency of pricing test $\tilde{L}_{t+1|t}$ (in terms of absolute value t-test) when bigger 5×1.96;

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times 1$
invest	5.124	4.929	3.658	0.788	0.953	0.897	0.678
cinvest	4.126	3.791	1.576	−0.734	0.922	0.857	0.673
cinvest_a	4.383	4.161	−1.710	−1.000	0.928	0.870	0.685
grcapx	4.085	3.862	3.921	0.513	0.955	0.886	0.696
pchcapx	4.049	3.797	2.672	0.355	0.919	0.859	0.649
pchcapx3	3.573	3.310	3.363	0.579	0.931	0.868	0.671
pchcapx_ia	2.758	2.561	3.173	0.082	0.944	0.877	0.694
noa	3.805	3.654	2.207	−0.499	0.940	0.866	0.631
dnoa	3.408	3.127	2.410	−0.382	0.924	0.861	0.680
gma	3.718	3.527	0.417	−0.385	0.931	0.864	0.694
orgcap	3.314	3.102	0.023	−0.637	0.931	0.868	0.676
roaq	4.965	4.585	4.639	−0.312	0.924	0.864	0.696
roavol	1.338	1.274	5.443	1.715	0.955	0.904	0.711
op	3.624	3.399	3.385	0.790	0.966	0.902	0.694
nop	3.317	3.094	2.149	0.684	0.960	0.897	0.716
pm	4.129	4.017	4.932	0.822	0.962	0.911	0.740
ato	4.066	3.727	5.052	0.616	0.944	0.861	0.678
cash	3.365	3.083	1.749	0.157	0.937	0.859	0.680
currat	2.160	2.056	1.356	0.942	0.937	0.877	0.736
pchcurrat	2.701	2.471	0.451	−0.382	0.937	0.861	0.669
quick	3.954	3.780	0.956	−0.383	0.933	0.879	0.727
pchquick	2.759	2.532	−2.057	−1.874	0.937	0.852	0.673
pchsaleinv	4.860	4.527	2.707	−0.498	0.913	0.846	0.669
saleinv	4.383	4.226	0.173	−0.443	0.928	0.881	0.680
salecash	4.105	3.849	5.468	0.770	0.928	0.881	0.680
salerec	4.181	3.793	4.227	0.813	0.931	0.843	0.685
dwc	4.057	4.051	4.028	0.424	0.951	0.893	0.691
dnca	3.988	3.723	4.475	0.338	0.940	0.879	0.664
dncl	3.548	3.222	1.112	0.280	0.915	0.837	0.647
dnco	3.781	3.512	5.406	1.573	0.926	0.861	0.664
dcoa	2.916	2.760	3.238	0.245	0.946	0.886	0.718
dcol	3.625	3.418	1.617	−0.350	0.957	0.897	0.723
ta	4.891	4.645	1.048	−0.434	0.946	0.864	0.653

Table 5: Ranking of zoo risk factors by the percentage of periods in which the corresponding risk premium pricing test is significant at the 5% level: Column 1 reports the ranking based on all periods, Column 2 on non-recession periods, and Column 3 on recession periods. A higher rank corresponds to a smaller percentage.

Zoo risk factor	rank_all	rank_exp	rank_rec
BAB	1	2	25
pchsaleinv	2	1	90
dncl	3	8	19
UMD	5	5	63
herf	6	15	20
pchcapx	7	9	47
ol	8	3	134
pchsale_pchinvt	9	4	125
cinvest	10	7	104
HML_Devil	11	10	83
LTR	12	19	13
indmom	13	27	4
dnoa	14	13	69
roaq	15	26	24
HML	16	11	85
ill	17	6	137
dnco	18	17	71
cashdebt	19	16	89
salecash	20	12	123
saleinv	21	29	14
depr	22	30	15
cinvest_a	23	46	5
pricedelay	24	14	129
dfnl	25	23	72
ms	26	24	73
tang	27	25	76
aeavol	28	41	22
absacc	29	75	1
salerec	30	20	91
pchcapx3	31	40	43
moms12m	32	42	46
gma	33	28	84
orgcap	34	56	8
quick	35	44	31
mom6m	36	21	124
rds	37	33	65
grltnoa	38	34	68
chpmia	39	65	6

Zoo risk factor	rank_all	rank_exp	rank_rec
poa	40	43	82
ep	41	68	3
nincr	42	32	98
ob_a	43	35	103
age	44	37	106
cdi	45	48	79
currat	46	57	29
pchcurrat	47	58	30
pchquick	48	49	58
etr	49	31	126
ps	50	45	99
cfp_ia	51	22	148
chinv	52	72	18
ndf	53	47	112
ear	54	74	23
chatoia	55	64	49
cash	56	66	51
acc	57	50	95
cto	58	60	62
pchsale_pchxsga	59	51	96
rdm	60	53	102
chtx	61	62	66
grltnoa_hxz	62	36	138
noa	63	54	105
tb	64	73	41
dnca	65	38	140
dsti	66	39	142
stdcf	67	55	118
SMB	68	88	11
RMW	69	89	12
STR	70	59	93
cp	71	69	61
os	72	79	35
chempia	73	61	100
adm	74	52	128
rna	75	84	48
gad	76	98	26
ala	77	67	122
convind	78	109	9
sue	79	77	57
pps	80	90	28
mom36m	81	78	60
sp	82	100	16

Zoo risk factor	rank_all	rank_exp	rank_rec
pchsale_pchrect	83	93	34
pchcapx_ia	84	70	97
dolvol	85	101	17
ato	86	85	80
ivg	87	118	7
QMJ	88	97	52
HXZ_ROE	89	99	55
Intermediary	90	108	27
pchgm_pchsale	91	92	64
std_turn	92	81	101
dcoa	93	95	70
dfin	94	83	109
ta	95	63	141
nxf	96	104	45
chcsho	97	115	21
sin	98	86	117
em	99	87	121
sgr	100	91	94
IPO	101	111	32
zs	102	112	36
bm_ia	103	80	127
std_dolvol	104	71	136
stdacc	105	76	147
MktRf	106	126	10
LIQ_PS	107	120	38
rd	108	82	139
dwc	109	102	108
roic	110	105	113
ndp	111	106	114
maxret	112	107	120
dy	113	110	87
lfe	114	94	135
kz	115	130	37
egr_hxz	116	113	110
cei	117	133	44
ebp	118	124	78
rsup	119	96	146
invest	120	117	119
pctacc	121	125	81
CMA	122	119	86
pchdepr	123	129	59
cfp	124	137	39
roavol	125	138	40

Zoo risk factor	rank_all	rank_exp	rank_rec
chmom	126	103	143
grcapx	127	132	74
nef	128	123	111
dpia	129	116	133
realestate_hxz	130	140	50
HXZ_IA	131	141	54
beta	132	136	56
lev	133	127	88
baspread	134	128	92
egr	135	121	130
lgr	136	122	131
dcol	137	131	107
ww	138	114	145
cashpr	139	135	116
turn	140	143	33
zerotrade	141	145	42
nop	142	134	132
idiovol	143	144	67
pm	144	142	115
hire	145	146	53
rs	146	139	144
op	147	147	77
retvol	148	148	75

12.2 Risk Premia Test - $T = 1$ - Misspecification

Table 6: Ranking of zoo risk factors by the percentage of periods in which the corresponding risk premium pricing test is significant at the 5% level: Column 1 reports the ranking based on all periods without bias adjustment; Columns 2, 3, and 4 report the rankings based on all periods with PCA3, PCA5, and PCA10 bias adjustments, respectively. A higher rank corresponds to a smaller percentage.

Zoo risk factor	rank_zoo	rank_zoo+PCA3	rank_zoo+PCA5	rank_zoo+PCA10
BAB	1	132	128	128
pchsaleinv	2	13	64	64
dncl	3	56	80	80
UMD	5	5	89	89
herf	6	1	134	134
pchcapx	7	139	87	87
ol	8	19	12	12
pchsale_pchinvt	9	63	71	71
cinvest	10	54	92	92

Zoo risk factor	rank_zoo	rank_zoo+PCA3	rank_zoo+PCA5	rank_zoo+PCA10
HML_Devil	11	72	68	68
LTR	12	103	39	39
indmom	13	89	104	104
dnoa	14	90	131	131
roaq	15	119	108	108
HML	16	45	42	42
ill	17	105	145	145
dnco	18	20	26	26
cashdebt	19	39	129	129
salecash	20	47	102	102
saleinv	21	113	3	3
depr	22	31	34	34
cinvest_a	23	51	121	121
pricedelay	24	42	11	11
dfnl	25	98	140	140
ms	26	91	142	142
tang	27	109	5	5
aeavol	28	117	63	63
absacc	29	70	75	75
salerec	30	46	33	33
pchcapx3	31	6	67	67
moms12m	32	68	74	74
gma	33	126	83	83
orgcap	34	24	15	15
quick	35	112	23	23
mom6m	36	48	96	96
rds	37	147	44	44
grltnoa	38	35	52	52
chpmia	39	57	113	113
poa	40	134	77	77
ep	41	76	57	57
nincr	42	73	50	50
ob_a	43	125	109	109
age	44	26	106	106
cdi	45	23	135	135
currat	46	140	78	78
pchcurrat	47	87	49	49
pchquick	48	30	79	79
etr	49	95	99	99
ps	50	127	141	141
cfp_ia	51	53	112	112
chinv	52	49	90	90
ndf	53	129	144	144

Zoo risk factor	rank_zoo	rank_zoo+PCA3	rank_zoo+PCA5	rank_zoo+PCA10
ear	54	93	137	137
chatoia	55	16	82	82
cash	56	128	48	48
acc	57	121	143	143
cto	58	2	14	14
pchsale_pchxsga	59	33	111	111
rdm	60	10	124	124
chtx	61	143	133	133
grltnoa_hxz	62	74	36	36
noa	63	144	147	147
tb	64	55	53	53
dnca	65	14	85	85
dsti	66	108	66	66
stdcf	67	44	118	118
SMB	68	7	7	7
RMW	69	110	119	119
STR	70	88	88	88
cp	71	104	97	97
os	72	25	130	130
chempia	73	122	120	120
adm	74	148	30	30
rna	75	15	9	9
gad	76	38	116	116
ala	77	142	101	101
convind	78	86	94	94
sue	79	59	70	70
pps	80	102	95	95
mom36m	81	135	65	65
sp	82	4	13	13
pchsale_pchrect	83	40	139	139
pchcapx_ia	84	34	132	132
dolvol	85	141	105	105
ato	86	28	6	6
ivg	87	69	115	115
QMJ	88	85	41	41
HXZ_ROE	89	75	110	110
Intermediary	90	146	148	148
pchgm_pchsale	91	114	54	54
std_turn	92	78	51	51
dcoa	93	9	8	8
dfin	94	27	10	10
ta	95	21	45	45
nxf	96	130	31	31

Zoo risk factor	rank_zoo	rank_zoo+PCA3	rank_zoo+PCA5	rank_zoo+PCA10
chcsho	97	99	4	4
sin	98	100	114	114
em	99	138	27	27
sgr	100	60	117	117
IPO	101	32	103	103
zs	102	96	59	59
bm_ia	103	115	29	29
std_dolvol	104	8	43	43
stdacc	105	36	47	47
MktRf	106	124	123	123
LIQ_PS	107	50	91	91
rd	108	116	60	60
dwc	109	11	25	25
roic	110	84	73	73
ndp	111	131	22	22
maxret	112	71	38	38
dy	113	120	69	69
lfe	114	41	58	58
kz	115	97	35	35
egr_hxz	116	133	81	81
cei	117	67	122	122
ebp	118	52	46	46
rsup	119	18	146	146
invest	120	29	76	76
pctacc	121	137	136	136
CMA	122	58	32	32
pchdepr	123	62	24	24
cfp	124	79	125	125
roavol	125	145	100	100
chmom	126	92	126	126
grcapx	127	123	61	61
nef	128	22	1	1
dpia	129	43	93	93
realestate_hxz	130	101	138	138
HXZ_IA	131	65	56	56
beta	132	111	16	16
lev	133	77	28	28
baspread	134	12	40	40
egr	135	106	107	107
lgr	136	107	21	21
dcol	137	80	84	84
ww	138	83	55	55
cashpr	139	17	37	37

Zoo risk factor	rank_zoo	rank_zoo+PCA3	rank_zoo+PCA5	rank_zoo+PCA10
turn	140	94	98	98
zerotrade	141	81	86	86
nop	142	64	2	2
idiovol	143	66	20	20
pm	144	118	127	127
hire	145	37	18	18
rs	146	82	62	62
op	147	136	17	17
retvol	148	61	72	72

Table 7: Zoo Factors with PCA3 bias adjustment. Column 1: OLS estimator of a recession dummy on pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (absolute-value t-test), that is, the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, where $\widetilde{test}_t \equiv \sqrt{N}\,|\tilde{L}^{\texttt{miss}}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$, and $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 5); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ over recession periods. Column 4: average pricing test over all periods. Columns 5–7: frequency with $|t| > 1.96$, $|t| > 2 \times 1.96$, and $|t| > 5 \times 1.96$, respectively.

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\lvert t\rvert > 1.96$)	Perc($\lvert t\rvert > 2 \times 1.96$)	Perc($\lvert t\rvert > 5 \times$
MktRf	0.649	1.614	−0.101	0.025	0.669	0.369	0.054
SMB	0.284	0.751	−0.106	−0.430	0.615	0.387	0.051
HML	−0.222	−0.594	0.357	−0.050	0.638	0.387	0.054
RMW	−0.207	−0.579	−0.220	−0.238	0.664	0.374	0.040
CMA	−0.172	−0.459	0.149	0.118	0.644	0.376	0.058
beta	0.242	0.677	−0.105	−0.094	0.664	0.412	0.049
ep	−0.017	−0.043	−0.226	−0.162	0.655	0.396	0.051
dy	0.191	0.518	−0.990	−0.168	0.667	0.396	0.051
sue	0.351	0.893	−0.850	−0.436	0.644	0.394	0.058
pps	0.088	0.239	−0.493	−0.456	0.662	0.412	0.047
LTR	0.378	0.964	−0.925	0.040	0.662	0.409	0.058
lev	−0.184	−0.473	−0.236	−0.107	0.655	0.365	0.058
cashdebt	−0.308	−0.775	−0.816	−0.144	0.635	0.412	0.063
currat	−0.218	−0.592	−0.375	−0.209	0.682	0.365	0.054
pchcurrat	0.352	0.964	−0.184	−0.134	0.658	0.389	0.038
pchquick	0.259	0.727	−0.388	−0.192	0.633	0.380	0.036
pchsaleinv	0.267	0.706	−0.433	−0.362	0.626	0.396	0.036
quick	−0.204	−0.545	−0.412	−0.116	0.664	0.362	0.058
salecash	−0.478	−1.221	−0.652	−0.307	0.640	0.365	0.056
saleinv	0.425	1.132	−0.018	−0.456	0.664	0.369	0.056

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
salerec	−0.214	−0.552	−0.300	−0.064	0.638	0.396	0.054
baspread	0.175	0.458	−0.266	−0.025	0.624	0.391	0.056
depr	−0.032	−0.079	−0.301	−0.272	0.633	0.396	0.069
pchdepr	−0.482	−1.190	−0.304	−0.183	0.647	0.349	0.058
STR	−0.348	−0.869	−0.323	0.008	0.658	0.416	0.060
mom6m	0.063	0.177	0.078	−0.108	0.640	0.369	0.040
mom36m	−0.169	−0.433	−0.250	−0.099	0.678	0.423	0.063
sgr	−0.243	−0.639	−0.703	−0.416	0.644	0.383	0.056
cp	0.090	0.232	−0.446	0.018	0.662	0.387	0.058
IPO	−0.048	−0.121	0.647	0.062	0.633	0.378	0.045
acc	−0.022	−0.057	−0.475	−0.190	0.667	0.378	0.067
sp	0.138	0.366	0.029	0.006	0.611	0.351	0.049
cto	−0.234	−0.573	−0.358	−0.018	0.602	0.371	0.058
UMD	0.024	0.063	0.444	0.174	0.611	0.367	0.051
turn	0.228	0.613	−0.239	0.068	0.660	0.412	0.060
pchgm_pchsale	0.066	0.161	−0.479	−0.387	0.664	0.416	0.063
pchsale_pchinvt	0.236	0.644	0.001	−0.251	0.647	0.398	0.036
pchsale_pchrect	0.017	0.044	−0.751	−0.437	0.635	0.376	0.045
pchsale_pchxsga	−0.227	−0.584	0.287	−0.067	0.633	0.407	0.054
etr	0.593	1.517	−0.510	−0.095	0.660	0.436	0.056
lfe	0.283	0.788	−0.366	−0.121	0.635	0.367	0.043
os	−0.193	−0.543	−0.306	−0.467	0.631	0.362	0.040
zs	−0.051	−0.131	−0.149	0.001	0.660	0.374	0.058
pchcapx_ia	−0.247	−0.622	0.359	−0.023	0.633	0.421	0.065
nincr	−0.033	−0.082	−0.734	0.113	0.653	0.403	0.065
indmom	0.112	0.312	−1.338	−0.333	0.658	0.423	0.040
ps	0.428	1.071	−0.220	0.090	0.671	0.414	0.063
bm_ia	−0.173	−0.429	−0.754	−0.333	0.664	0.418	0.067
cfp_ia	−0.189	−0.484	−0.072	−0.331	0.642	0.407	0.056
chempia	0.528	1.389	0.330	−0.097	0.667	0.421	0.056
dolvol	−0.064	−0.156	−0.508	−0.124	0.682	0.430	0.069
std_dolvol	−0.453	−1.242	−0.285	−0.187	0.617	0.351	0.043
std_turn	0.393	1.057	0.365	0.124	0.655	0.407	0.054
adm	−0.096	−0.267	−0.645	−0.331	0.696	0.414	0.051
rdm	0.140	0.384	0.532	−0.019	0.620	0.362	0.043
rds	0.234	0.629	0.166	−0.109	0.694	0.376	0.056
kz	−0.174	−0.459	−0.002	−0.207	0.660	0.378	0.054
chinv	0.498	1.299	−0.306	−0.204	0.640	0.358	0.043
chtx	−0.279	−0.765	0.748	−0.034	0.687	0.371	0.054
ill	−0.146	−0.404	0.041	−0.260	0.662	0.407	0.045
LIQ_PS	0.030	0.075	−0.949	−0.285	0.640	0.387	0.065

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
idiovol	0.311	0.838	−0.475	−0.184	0.649	0.409	0.054
grltnoa	−0.142	−0.368	−0.259	−0.386	0.633	0.360	0.049
ob_a	−0.886	−2.340	−0.249	0.147	0.669	0.416	0.051
grltnoa_hxz	−0.006	−0.017	0.003	−0.134	0.653	0.387	0.058
cfp	0.206	0.505	−0.629	−0.177	0.655	0.380	0.074
rd	0.130	0.328	−0.351	−0.364	0.664	0.427	0.054
cinvest	0.109	0.290	−0.927	−0.616	0.642	0.383	0.036
roavol	−0.033	−0.088	−0.685	−0.367	0.689	0.403	0.040
cinvest_a	0.047	0.122	−0.584	−0.447	0.640	0.391	0.058
noa	0.016	0.041	−0.611	−0.407	0.687	0.403	0.049
dnoa	0.289	0.749	−0.482	−0.314	0.658	0.398	0.058
tb	−0.250	−0.650	−0.433	−0.103	0.642	0.378	0.047
pricedelay	0.608	1.592	−0.894	−0.120	0.635	0.409	0.049
age	−0.180	−0.474	−0.241	0.197	0.631	0.398	0.058
egr	0.188	0.491	−0.579	−0.037	0.662	0.394	0.056
lgr	−0.231	−0.600	−0.180	−0.238	0.662	0.385	0.049
dcoa	−0.298	−0.816	0.466	0.126	0.617	0.367	0.034
dcol	−0.173	−0.447	0.476	0.048	0.655	0.405	0.056
dwc	−0.382	−1.062	−0.090	−0.275	0.620	0.353	0.040
dnca	0.161	0.446	0.042	−0.143	0.626	0.374	0.045
dncl	0.453	1.029	−0.756	−0.029	0.642	0.389	0.085
dnco	−0.030	−0.082	0.087	−0.099	0.629	0.360	0.040
dfin	−0.038	−0.102	−0.131	−0.183	0.631	0.378	0.045
ta	0.002	0.006	0.153	−0.099	0.629	0.383	0.049
dsti	−0.139	−0.357	−0.427	−0.232	0.662	0.409	0.054
dfnl	−0.292	−0.820	−0.604	−0.149	0.660	0.394	0.045
egr_hxz	0.104	0.284	−0.423	0.020	0.676	0.412	0.056
ms	−0.396	−1.046	−1.085	−0.160	0.658	0.383	0.054
chmom	0.299	0.801	−0.659	−0.051	0.658	0.398	0.040
grcapx	−0.109	−0.286	0.085	−0.057	0.667	0.398	0.051
retvol	0.146	0.390	−0.674	−0.180	0.644	0.398	0.056
zerotrade	0.341	0.942	−0.350	0.099	0.655	0.414	0.047
pchcapx3	0.325	0.844	−0.743	−0.048	0.613	0.362	0.054
cei	−0.272	−0.659	−0.114	−0.225	0.649	0.423	0.060
nef	0.195	0.515	−0.031	−0.060	0.629	0.385	0.043
ndf	−0.502	−1.330	−0.554	−0.319	0.673	0.416	0.047
nxf	0.400	1.097	−0.172	−0.006	0.673	0.409	0.043
rs	−0.212	−0.547	0.128	−0.079	0.655	0.394	0.054
herf	0.388	0.953	−1.291	−0.302	0.597	0.367	0.065
ww	0.144	0.397	0.084	0.255	0.655	0.365	0.043
roic	−0.189	−0.515	−1.163	−0.260	0.655	0.421	0.051

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
tang	−0.470	−1.299	−0.610	−0.163	0.662	0.400	0.040
op	0.248	0.653	−0.516	0.261	0.678	0.391	0.054
nop	−0.002	−0.005	−0.805	−0.015	0.647	0.385	0.060
ndp	−0.444	−1.140	−0.271	0.086	0.673	0.389	0.058
ebp	0.040	0.103	−0.139	−0.143	0.640	0.387	0.063
chcsho	−0.045	−0.119	−0.451	−0.043	0.660	0.421	0.038
aeavol	0.201	0.510	−0.110	−0.209	0.664	0.421	0.051
ear	−0.016	−0.037	−0.218	−0.497	0.658	0.416	0.072
moms12m	−0.030	−0.079	−0.370	−0.200	0.649	0.412	0.047
dpia	−0.353	−0.965	−0.203	−0.387	0.635	0.365	0.045
pchcapx	0.232	0.622	−0.312	−0.142	0.680	0.396	0.047
cdi	−0.103	−0.294	−0.465	−0.392	0.629	0.349	0.036
rna	0.200	0.533	−0.884	−0.220	0.626	0.371	0.045
pm	0.204	0.498	−0.791	−0.513	0.664	0.423	0.058
ato	0.457	1.184	−0.123	0.013	0.631	0.380	0.051
chatoia	−0.582	−1.630	0.507	−0.070	0.626	0.371	0.043
chpmia	−0.344	−0.891	−0.360	−0.236	0.642	0.380	0.051
cashpr	−0.293	−0.797	−0.005	−0.108	0.626	0.380	0.040
sin	0.278	0.763	−0.351	−0.179	0.660	0.412	0.038
rsup	−0.385	−1.007	−0.582	−0.272	0.626	0.338	0.060
stdcf	0.114	0.318	−0.506	−0.219	0.635	0.376	0.040
absacc	−0.395	−1.017	0.501	−0.047	0.651	0.380	0.054
invest	−0.303	−0.820	−0.288	−0.329	0.631	0.369	0.045
roaq	0.072	0.182	−0.459	−0.270	0.664	0.398	0.074
stdacc	0.250	0.707	−0.330	−0.089	0.633	0.371	0.045
realestate_hxz	0.030	0.079	0.135	0.099	0.660	0.436	0.054
pctacc	0.359	0.916	−0.233	0.169	0.678	0.391	0.065
maxret	−0.054	−0.147	−0.416	−0.049	0.651	0.389	0.051
ol	−0.554	−1.410	−0.538	−0.425	0.626	0.360	0.051
ivg	0.114	0.313	0.159	−0.034	0.649	0.362	0.045
poa	−0.321	−0.737	−0.267	−0.240	0.676	0.378	0.063
em	−0.401	−0.978	−0.420	−0.286	0.678	0.398	0.069
cash	−0.206	−0.506	−0.261	−0.139	0.671	0.414	0.063
HML_Devil	−0.107	−0.271	−0.408	−0.223	0.651	0.394	0.063
gma	0.007	0.015	−0.570	−0.163	0.669	0.412	0.076
orgcap	0.159	0.362	−0.244	0.000	0.629	0.371	0.078
BAB	−0.172	−0.457	−0.491	−0.200	0.673	0.405	0.038
QMJ	−0.087	−0.234	−0.378	−0.183	0.655	0.376	0.047
hire	−0.288	−0.743	0.266	−0.192	0.633	0.374	0.060
gad	0.009	0.027	−0.596	−0.147	0.633	0.380	0.043
ala	0.010	0.022	−0.804	−0.304	0.685	0.405	0.076

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
HXZ_IA	−0.063	−0.176	−0.207	0.079	0.647	0.383	0.047
HXZ_ROE	0.296	0.755	0.046	0.054	0.653	0.427	0.049
Intermediary	0.288	0.741	−0.309	−0.089	0.689	0.432	0.060
convind	−0.284	−0.743	0.004	−0.120	0.655	0.391	0.049

Table 8: Zoo Factors with PCA5 bias adjustment. Column 1: OLS estimator of a recession dummy on pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (absolute-value t-test), that is, the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, where $\widetilde{test}_t \equiv \sqrt{N}\,|\tilde{L}^{\texttt{miss}}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$, and $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 5); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ over recession periods. Column 4: average pricing test over all periods. Columns 5–7: frequency with $|t| > 1.96$, $|t| > 2 \times 1.96$, and $|t| > 5 \times 1.96$, respectively.

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
MktRf	−0.125	−0.424	0.156	0.156	0.559	0.280	0.011
SMB	0.409	1.321	0.312	0.015	0.555	0.268	0.022
HML	0.180	0.587	−0.289	−0.153	0.557	0.293	0.013
RMW	0.052	0.170	−0.002	−0.069	0.553	0.293	0.018
CMA	0.490	1.671	−0.011	−0.032	0.579	0.246	0.018
beta	0.768	2.581	−0.177	0.054	0.584	0.286	0.007
ep	−0.139	−0.447	−0.492	−0.069	0.602	0.327	0.018
dy	0.470	1.493	−0.342	−0.114	0.553	0.282	0.020
sue	0.633	2.088	0.137	0.032	0.570	0.253	0.025
pps	0.524	1.804	−0.619	−0.051	0.582	0.295	0.013
LTR	0.620	2.056	0.440	0.214	0.553	0.306	0.013
lev	0.147	0.474	0.144	−0.009	0.593	0.300	0.022
cashdebt	0.295	0.925	−0.225	−0.250	0.584	0.315	0.022
currat	−0.063	−0.191	−0.815	0.134	0.617	0.293	0.034
pchcurrat	0.115	0.391	−0.882	−0.036	0.575	0.298	0.007
pchquick	−0.029	−0.100	−0.741	−0.020	0.562	0.309	0.011
pchsaleinv	0.242	0.788	0.536	−0.042	0.617	0.300	0.020
quick	0.095	0.294	−0.262	0.139	0.588	0.266	0.029
salecash	−0.101	−0.329	−0.363	0.041	0.586	0.277	0.029
saleinv	0.342	1.146	0.383	0.072	0.544	0.289	0.011
salerec	0.154	0.522	−0.067	−0.109	0.573	0.320	0.009
baspread	0.544	1.760	−0.116	−0.118	0.575	0.286	0.011
depr	0.284	0.907	−0.278	0.022	0.593	0.302	0.022
pchdepr	−0.443	−1.357	0.328	−0.165	0.570	0.306	0.022
STR	0.164	0.516	0.041	0.193	0.570	0.331	0.016

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
mom6m	0.121	0.404	0.284	0.298	0.615	0.324	0.011
mom36m	0.576	1.839	−0.139	0.173	0.593	0.318	0.016
sgr	0.194	0.625	−0.330	−0.009	0.588	0.298	0.022
cp	−0.193	−0.634	−0.581	−0.047	0.600	0.275	0.022
IPO	0.161	0.508	0.071	0.117	0.553	0.300	0.011
acc	−0.140	−0.476	−0.338	0.133	0.591	0.327	0.011
sp	0.418	1.434	−0.047	0.011	0.555	0.275	0.013
cto	−0.135	−0.429	−0.001	0.032	0.539	0.275	0.013
UMD	0.208	0.656	0.134	0.098	0.573	0.345	0.013
turn	0.522	1.728	−0.133	0.082	0.591	0.320	0.009
pchgm_pchsale	0.031	0.099	0.131	−0.009	0.582	0.333	0.013
pchsale_pchinvt	0.082	0.276	0.669	−0.030	0.613	0.291	0.018
pchsale_pchrect	0.556	1.795	−0.381	0.062	0.597	0.293	0.025
pchsale_pchxsga	0.056	0.178	0.187	−0.067	0.584	0.313	0.016
etr	0.301	0.993	0.377	−0.020	0.604	0.333	0.013
lfe	0.145	0.492	0.349	0.284	0.615	0.306	0.011
os	0.002	0.006	−0.124	−0.004	0.557	0.313	0.009
zs	0.213	0.686	−0.076	−0.014	0.602	0.300	0.022
pchcapx_ia	0.216	0.691	0.517	0.167	0.617	0.342	0.016
nincr	−0.033	−0.100	0.030	0.119	0.593	0.320	0.034
indmom	0.591	1.927	0.020	−0.008	0.564	0.295	0.011
ps	0.818	2.671	−0.358	−0.048	0.553	0.304	0.011
bm_ia	0.112	0.355	−0.617	−0.147	0.593	0.318	0.020
cfp_ia	0.083	0.278	0.052	0.055	0.582	0.311	0.011
chempia	0.755	2.619	−0.213	−0.071	0.582	0.324	0.013
dolvol	0.178	0.560	−0.167	−0.020	0.591	0.331	0.013
std_dolvol	0.343	1.172	0.490	−0.032	0.600	0.286	0.016
std_turn	0.626	2.107	0.183	0.116	0.593	0.309	0.009
adm	0.261	0.933	−0.559	−0.155	0.562	0.318	0.007
rdm	0.584	2.064	−0.021	0.157	0.600	0.293	0.013
rds	0.535	1.854	0.065	0.176	0.575	0.284	0.013
kz	0.196	0.642	−0.308	−0.131	0.588	0.286	0.022
chinv	0.016	0.049	0.063	0.140	0.586	0.311	0.029
chtx	0.254	0.810	0.357	0.147	0.602	0.315	0.020
ill	0.392	1.327	0.459	−0.078	0.611	0.329	0.009
LIQ_PS	0.448	1.531	0.252	0.026	0.564	0.298	0.009
idiovol	0.780	2.479	0.138	0.117	0.575	0.313	0.011
grltnoa	0.357	1.132	0.090	0.029	0.584	0.304	0.018
ob_a	−0.047	−0.148	−0.108	−0.052	0.609	0.300	0.020
grltnoa_hxz	0.395	1.250	0.340	0.148	0.568	0.300	0.013
cfp	0.198	0.642	−0.169	−0.008	0.568	0.271	0.022

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
rd	0.454	1.532	−0.063	−0.008	0.575	0.293	0.009
cinvest	0.108	0.350	0.179	0.043	0.559	0.313	0.016
roavol	0.370	1.162	−0.110	−0.095	0.588	0.338	0.013
cinvest_a	0.242	0.772	0.156	0.020	0.615	0.300	0.020
noa	0.334	1.082	−0.252	−0.286	0.584	0.311	0.022
dnoa	0.243	0.779	0.017	−0.042	0.595	0.304	0.022
tb	−0.147	−0.469	0.719	0.367	0.579	0.311	0.022
pricedelay	0.530	1.839	−0.140	−0.174	0.570	0.277	0.004
age	0.188	0.593	−0.002	0.085	0.557	0.291	0.016
egr	0.384	1.137	−0.145	0.032	0.579	0.327	0.036
lgr	0.383	1.249	−0.510	0.044	0.604	0.306	0.016
dcoa	−0.141	−0.482	0.016	−0.058	0.595	0.306	0.020
dcol	−0.089	−0.291	−0.401	−0.164	0.606	0.304	0.022
dwc	0.676	2.377	−0.429	−0.072	0.521	0.257	0.011
dnca	0.387	1.301	−0.120	0.112	0.566	0.300	0.013
dncl	0.447	1.417	0.136	0.176	0.593	0.284	0.022
dnco	0.538	1.832	0.093	0.122	0.568	0.306	0.009
dfin	0.319	1.083	0.014	−0.231	0.535	0.304	0.009
ta	0.178	0.630	−0.522	−0.076	0.613	0.302	0.011
dsti	0.060	0.198	0.459	0.119	0.586	0.268	0.020
dfnl	−0.083	−0.270	−0.570	−0.023	0.597	0.336	0.016
egr_hxz	0.401	1.254	−0.072	0.053	0.588	0.340	0.022
ms	−0.321	−1.081	−0.175	0.083	0.570	0.275	0.027
chmom	0.358	1.145	0.349	0.158	0.620	0.338	0.018
grcapx	0.355	1.156	−0.535	0.037	0.570	0.315	0.013
retvol	0.470	1.497	0.187	0.103	0.562	0.304	0.016
zerotrade	0.372	1.243	−0.165	0.145	0.559	0.333	0.007
pchcapx3	0.189	0.598	−0.302	0.125	0.562	0.340	0.013
cei	0.198	0.633	0.266	0.192	0.629	0.331	0.025
nef	0.629	2.032	0.091	−0.014	0.553	0.275	0.020
ndf	−0.230	−0.778	0.445	0.243	0.602	0.342	0.018
nxf	0.622	1.998	0.261	0.261	0.577	0.313	0.029
rs	0.250	0.834	−0.422	−0.133	0.582	0.280	0.016
herf	−0.012	−0.040	−0.299	−0.099	0.604	0.331	0.013
ww	0.870	2.915	−0.106	0.037	0.582	0.286	0.013
roic	−0.072	−0.233	−0.331	0.014	0.555	0.286	0.022
tang	−0.029	−0.090	−0.160	0.169	0.555	0.306	0.022
op	0.690	2.140	−0.021	0.189	0.591	0.324	0.022
nop	0.657	2.049	−0.379	0.065	0.568	0.293	0.022
ndp	0.082	0.260	−0.307	−0.065	0.575	0.295	0.022
ebp	−0.066	−0.211	−0.164	−0.079	0.609	0.280	0.027

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
chcsho	0.463	1.361	−0.251	0.072	0.541	0.291	0.038
aeavol	−0.212	−0.656	−0.107	−0.009	0.564	0.277	0.016
ear	0.155	0.484	0.090	−0.246	0.584	0.327	0.018
moms12m	0.358	1.212	0.189	0.113	0.609	0.322	0.013
dpia	0.165	0.571	−0.075	−0.058	0.562	0.291	0.007
pchcapx	0.248	0.783	−0.501	−0.039	0.575	0.318	0.025
cdi	0.017	0.058	−0.329	−0.173	0.559	0.273	0.022
rna	0.298	1.032	0.039	0.398	0.544	0.271	0.011
pm	0.241	0.788	0.326	0.234	0.626	0.338	0.020
ato	0.505	1.741	−0.046	0.031	0.582	0.324	0.011
chatoia	−0.123	−0.448	0.352	0.122	0.591	0.284	0.009
chpmia	−0.042	−0.139	−0.689	−0.158	0.600	0.324	0.018
cashpr	0.024	0.081	0.042	−0.101	0.550	0.280	0.016
sin	0.646	2.051	0.087	−0.054	0.564	0.298	0.016
rsup	0.224	0.719	−0.244	−0.112	0.564	0.300	0.022
stdcf	0.208	0.680	0.001	−0.022	0.568	0.280	0.020
absacc	−0.135	−0.452	−0.336	0.065	0.582	0.298	0.013
invest	0.156	0.541	−0.077	−0.059	0.562	0.289	0.009
roaq	0.519	1.682	−0.390	−0.136	0.591	0.309	0.022
stdacc	0.201	0.655	−0.229	0.074	0.570	0.277	0.025
realestate_hxz	0.482	1.525	0.203	0.221	0.579	0.318	0.020
pctacc	−0.019	−0.063	0.223	0.165	0.613	0.291	0.016
maxret	0.255	0.844	0.286	0.021	0.586	0.304	0.011
ol	0.033	0.109	−0.278	0.064	0.521	0.289	0.013
ivg	0.220	0.730	−0.153	0.020	0.577	0.291	0.016
poa	−0.415	−1.297	0.361	0.143	0.577	0.273	0.022
em	−0.404	−1.278	0.164	−0.181	0.604	0.315	0.018
cash	−0.273	−0.863	−0.188	0.158	0.584	0.293	0.031
HML_Devil	0.145	0.465	−0.384	−0.160	0.586	0.306	0.016
gma	−0.333	−1.025	0.296	−0.178	0.577	0.320	0.018
orgcap	0.006	0.019	0.195	0.146	0.570	0.284	0.016
BAB	0.364	1.250	0.135	−0.101	0.595	0.275	0.007
QMJ	0.239	0.828	0.095	−0.070	0.559	0.262	0.013
hire	0.040	0.126	−0.664	−0.085	0.602	0.275	0.025
gad	0.233	0.785	0.120	0.001	0.537	0.293	0.009
ala	0.472	1.492	−0.574	0.220	0.611	0.318	0.025
HXZ_IA	0.286	0.975	−0.350	−0.125	0.573	0.282	0.013
HXZ_ROE	0.051	0.164	0.189	0.302	0.591	0.327	0.013
Intermediary	−0.232	−0.775	0.004	−0.157	0.624	0.333	0.016
convind	0.609	1.962	−0.638	0.065	0.597	0.342	0.022

Table 9: Zoo Factors with PCA10 bias adjustment. Column 1: OLS estimator of a recession dummy on pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (absolute-value t-test), that is, the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, where $\widetilde{test}_t \equiv \sqrt{N}\,|\tilde{L}^{\texttt{miss}}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$, and $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 5); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ over recession periods. Column 4: average pricing test over all periods. Columns 5–7: frequency with $|t| > 1.96$, $|t| > 2 \times 1.96$, and $|t| > 5 \times 1.96$, respectively.

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\lvert t\rvert > 1.96$)	Perc($\lvert t\rvert > 2 \times 1.96$)	Perc($\lvert t\rvert > 5 \times$
MktRf	−0.077	−0.283	0.472	0.317	0.537	0.242	0.007
SMB	0.419	1.599	0.257	0.212	0.463	0.221	0.002
HML	0.232	0.951	0.080	0.248	0.497	0.217	0.000
RMW	0.308	1.210	0.050	0.064	0.535	0.224	0.009
CMA	0.168	0.743	−0.195	0.071	0.492	0.204	0.000
beta	0.453	1.796	0.674	0.276	0.481	0.224	0.000
ep	0.521	2.148	0.638	0.393	0.506	0.228	0.000
dy	0.353	1.506	0.135	0.132	0.510	0.213	0.000
sue	0.562	2.171	0.338	0.073	0.510	0.239	0.004
pps	0.330	1.292	0.054	0.203	0.519	0.224	0.002
LTR	0.150	0.567	−0.133	−0.063	0.494	0.246	0.004
lev	0.409	1.590	0.276	0.208	0.490	0.242	0.002
cashdebt	0.191	0.677	−0.034	0.136	0.541	0.273	0.004
currat	0.525	2.193	0.452	0.258	0.512	0.217	0.004
pchcurrat	0.347	1.386	0.305	0.180	0.501	0.217	0.002
pchquick	0.251	1.018	0.060	0.195	0.512	0.221	0.007
pchsaleinv	−0.259	−0.919	0.630	0.344	0.508	0.277	0.007
quick	0.593	2.281	0.569	0.284	0.488	0.224	0.007
salecash	0.427	1.678	0.473	0.291	0.521	0.215	0.004
saleinv	0.538	2.359	0.118	0.046	0.456	0.228	0.000
salerec	0.210	0.823	1.005	0.360	0.492	0.228	0.004
baspread	0.425	1.711	0.164	0.190	0.494	0.244	0.000
depr	0.425	1.666	0.478	0.258	0.492	0.206	0.009
pchdepr	−0.273	−1.075	−0.135	−0.036	0.488	0.233	0.004
STR	−0.057	−0.214	−0.242	0.077	0.517	0.228	0.007
mom6m	0.129	0.471	−0.167	0.183	0.519	0.275	0.009
mom36m	0.234	0.894	0.428	0.135	0.508	0.262	0.000
sgr	0.358	1.525	0.190	0.195	0.530	0.226	0.000
cp	0.417	1.789	0.472	0.141	0.519	0.217	0.000
IPO	−0.016	−0.064	−0.016	0.297	0.521	0.242	0.002
acc	−0.022	−0.084	0.499	0.268	0.557	0.248	0.007
sp	0.434	1.804	0.657	0.301	0.479	0.215	0.002

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
cto	0.441	1.717	1.037	0.336	0.479	0.221	0.000
UMD	−0.159	−0.595	0.322	0.263	0.517	0.293	0.002
turn	0.284	1.176	0.506	0.259	0.519	0.215	0.002
pchgm_pchsale	−0.019	−0.071	0.242	0.284	0.503	0.244	0.004
pchsale_pchinvt	−0.500	−1.776	0.695	0.391	0.510	0.280	0.007
pchsale_pchrect	0.371	1.388	−0.180	0.019	0.553	0.244	0.004
pchsale_pchxsga	0.530	1.960	0.038	0.191	0.528	0.257	0.011
etr	0.067	0.270	0.338	0.114	0.519	0.242	0.004
lfe	0.375	1.404	0.410	0.217	0.506	0.248	0.007
os	−0.087	−0.296	0.377	0.180	0.541	0.239	0.011
zs	0.390	1.548	0.371	0.237	0.506	0.239	0.002
pchcapx_ia	0.498	1.897	0.231	0.275	0.544	0.266	0.004
nincr	0.235	0.917	0.099	0.114	0.501	0.208	0.002
indmom	0.061	0.224	0.297	0.186	0.523	0.228	0.009
ps	0.455	1.817	0.373	0.209	0.555	0.277	0.002
bm_ia	−0.007	−0.028	0.056	0.311	0.490	0.226	0.009
cfp_ia	0.465	1.816	0.373	0.246	0.528	0.233	0.004
chempia	0.478	1.975	0.320	0.274	0.535	0.224	0.002
dolvol	0.293	1.184	0.459	0.189	0.523	0.264	0.000
std_dolvol	−0.092	−0.361	0.046	0.224	0.499	0.221	0.002
std_turn	0.225	0.922	0.208	0.185	0.501	0.215	0.000
adm	0.009	0.034	0.490	0.284	0.490	0.255	0.002
rdm	0.843	3.061	0.537	0.280	0.539	0.260	0.009
rds	0.622	2.462	0.860	0.219	0.499	0.224	0.007
kz	0.425	1.676	−0.040	0.266	0.492	0.226	0.000
chinv	−0.183	−0.728	−0.036	0.090	0.517	0.235	0.000
chtx	0.344	1.422	0.262	0.197	0.544	0.260	0.000
ill	0.317	1.252	0.263	−0.033	0.568	0.271	0.004
LIQ_PS	−0.034	−0.130	0.182	0.134	0.517	0.260	0.004
idiovol	0.434	1.802	0.359	0.276	0.485	0.246	0.000
grltnoa	0.624	2.411	0.062	0.109	0.501	0.228	0.002
ob_a	0.034	0.130	0.676	0.375	0.526	0.244	0.002
grltnoa_hxz	0.574	2.410	0.023	0.154	0.492	0.201	0.002
cfp	0.253	1.005	0.366	0.368	0.539	0.235	0.002
rd	0.462	1.759	−0.173	−0.026	0.506	0.235	0.011
cinvest	0.266	1.050	0.202	0.050	0.517	0.268	0.000
roavol	0.736	3.059	0.173	0.183	0.519	0.242	0.000
cinvest_a	0.534	2.056	0.384	0.253	0.535	0.264	0.002
noa	0.224	0.906	−0.010	0.113	0.577	0.242	0.007
dnoa	0.168	0.635	0.497	0.304	0.541	0.246	0.007
tb	0.245	0.895	0.514	0.313	0.501	0.244	0.011

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
pricedelay	0.562	2.231	0.123	0.143	0.477	0.206	0.002
age	0.304	1.225	0.301	0.195	0.523	0.235	0.002
egr	0.017	0.067	0.023	0.071	0.523	0.226	0.004
lgr	0.307	1.334	−0.392	0.078	0.485	0.208	0.002
dcoa	0.210	0.894	0.287	0.164	0.468	0.201	0.000
dcol	0.186	0.779	0.318	0.311	0.515	0.219	0.002
dwc	0.744	2.985	0.389	0.136	0.488	0.244	0.002
dnca	0.347	1.401	0.270	0.091	0.515	0.197	0.000
dncl	0.264	1.092	−0.458	−0.092	0.512	0.204	0.002
dnco	0.381	1.637	0.324	0.097	0.488	0.192	0.000
dfin	0.209	0.791	0.297	0.106	0.474	0.257	0.000
ta	0.201	0.756	0.504	0.151	0.499	0.219	0.007
dsti	0.002	0.008	0.491	0.366	0.508	0.248	0.004
dfnl	0.032	0.119	0.275	0.146	0.553	0.284	0.007
egr_hxz	0.004	0.017	−0.021	0.112	0.512	0.237	0.000
ms	0.577	2.146	0.340	0.172	0.555	0.235	0.004
chmom	0.029	0.107	0.143	0.119	0.539	0.277	0.002
grcapx	0.209	0.800	0.633	0.138	0.506	0.257	0.002
retvol	0.353	1.422	0.365	0.207	0.510	0.251	0.000
zerotrade	0.214	0.890	0.301	0.212	0.515	0.219	0.000
pchcapx3	0.282	1.183	0.418	0.132	0.508	0.233	0.000
cei	−0.111	−0.415	0.477	0.344	0.535	0.246	0.002
nef	0.818	3.420	0.631	0.293	0.436	0.213	0.000
ndf	0.193	0.750	0.274	0.122	0.557	0.284	0.002
nxf	0.584	2.296	1.012	0.342	0.490	0.208	0.004
rs	0.160	0.635	0.240	0.320	0.506	0.237	0.002
herf	0.407	1.653	0.466	0.255	0.544	0.268	0.004
ww	−0.050	−0.199	0.146	0.101	0.503	0.235	0.002
roic	−0.066	−0.253	−0.012	0.162	0.510	0.255	0.002
tang	0.299	1.094	0.536	0.212	0.459	0.217	0.009
op	0.508	2.102	0.571	0.325	0.483	0.230	0.000
nop	0.655	2.803	0.153	0.177	0.445	0.206	0.000
ndp	0.229	0.915	0.083	0.292	0.485	0.239	0.002
ebp	0.262	1.051	0.326	0.246	0.499	0.208	0.000
chcsho	0.172	0.701	−0.165	0.190	0.456	0.219	0.000
aeavol	0.639	2.487	−0.053	0.030	0.506	0.235	0.002
ear	0.130	0.483	0.778	0.318	0.548	0.280	0.004
moms12m	−0.094	−0.376	−0.196	0.183	0.510	0.237	0.002
dpia	0.280	1.168	0.291	0.120	0.517	0.217	0.000
pchcapx	0.424	1.681	0.581	0.214	0.515	0.262	0.000
cdi	0.251	0.961	0.358	0.146	0.546	0.255	0.002

Continued on next

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\| > 5 \times$
rna	0.490	2.034	0.193	0.154	0.468	0.235	0.002
pm	0.347	1.291	0.159	0.226	0.539	0.273	0.004
ato	0.269	1.063	0.518	0.196	0.459	0.248	0.002
chatoia	−0.105	−0.398	0.242	0.217	0.512	0.228	0.002
chpmia	0.101	0.373	0.280	0.202	0.528	0.260	0.007
cashpr	0.111	0.437	0.297	0.364	0.492	0.226	0.002
sin	0.165	0.599	0.193	0.196	0.528	0.257	0.011
rsup	0.552	2.070	0.597	0.334	0.568	0.293	0.004
stdcf	0.059	0.217	0.108	0.136	0.530	0.253	0.007
absacc	0.423	1.591	0.267	0.456	0.510	0.260	0.004
invest	0.254	1.060	0.279	0.125	0.510	0.213	0.000
roaq	0.102	0.372	0.289	0.167	0.523	0.266	0.004
stdacc	−0.081	−0.307	0.123	0.152	0.499	0.228	0.007
realestate_hxz	−0.251	−0.949	0.290	0.225	0.548	0.246	0.004
pctacc	−0.051	−0.204	−0.124	0.201	0.546	0.206	0.002
maxret	0.355	1.497	0.301	0.254	0.492	0.210	0.000
ol	0.145	0.551	0.758	0.305	0.477	0.239	0.002
ivg	0.010	0.040	−0.072	0.225	0.528	0.233	0.004
poa	−0.011	−0.045	−0.098	0.107	0.510	0.217	0.000
em	0.378	1.514	0.579	0.146	0.488	0.233	0.000
cash	0.295	1.174	0.495	0.320	0.499	0.215	0.004
HML_Devil	−0.050	−0.205	0.233	0.135	0.508	0.235	0.000
gma	0.340	1.367	0.411	0.248	0.512	0.233	0.000
orgcap	0.180	0.718	0.130	0.119	0.479	0.221	0.002
BAB	0.487	1.937	0.367	0.227	0.539	0.266	0.002
QMJ	0.352	1.348	0.285	0.151	0.494	0.228	0.004
hire	−0.045	−0.198	0.367	0.204	0.483	0.206	0.000
gad	0.214	0.893	0.302	0.176	0.528	0.253	0.000
ala	0.374	1.482	0.565	0.172	0.519	0.248	0.000
HXZ_IA	0.342	1.456	0.180	0.105	0.503	0.206	0.000
HXZ_ROE	−0.006	−0.021	−0.290	0.042	0.526	0.289	0.011
Intermediary	0.198	0.781	0.502	0.172	0.577	0.286	0.004
convind	0.611	2.339	0.184	0.266	0.517	0.262	0.009

Zoo risk factor	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\|t\| > 1.96$)	Perc($\|t\| > 2 \times 1.96$)	Perc($\|t\|$
PCA3	0.376	1.110	2.369	-0.345	0.638	0.351	0.029
PCA5	0.430	1.538	2.369	-0.080	0.582	0.289	0.013
PCA10	0.303	1.232	2.369	0.278	0.521	0.262	0.002

Table 10: PCA1 factor with PCA3, PCA5, and PCA10 bias adjustment - Column 1: OLS estimator of recession dummy on pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (in terms of absolute value t-test), namely the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, setting $\widetilde{test}_t \equiv \sqrt{N}|\tilde{L}^{\texttt{miss}}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$ where $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 5) and d_t is a dummy variable that takes value one when t is one of the recession-financial crises periods, and zero otherwise; Column 2 t-ratio of OLS coefficient; Column 3 average pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (in terms of absolute value t-test) over recession periods; Column 4 average pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (in terms of absolute value t-test) over all periods; Column 5 frequency of pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (in terms of absolute value t-test) when bigger 1.96; Column 5 frequency of pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (in terms of absolute value t-test) when bigger 2×1.96; Column 5 frequency of pricing test $\tilde{L}^{\texttt{miss}}_{t+1|t}$ (in terms of absolute value t-test) when bigger 5×1.96.

12.3 MV Portfolio - $T = 1$ - Correct Specification

Table 11: Core Factors. Column 1: OLS estimator of a recession dummy on $\tilde{SR}_{t-2,t-1,t}$, that is, the estimated coefficient β in $\tilde{SR}_{t-2,t-1,t} = \alpha + \beta d_t + u_t$ (see Theorem 5); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average $\tilde{SR}_{t-2,t-1,t}$ over recession periods. Column 4: average $\tilde{SR}_{t-2,t-1,t}$ over all periods.

Zoo risk factor	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
MktRf	0.038	0.281	0.730	0.699
SMB	0.287	2.162	0.949	0.718
HML	−0.079	−0.757	0.687	0.758
RMW	−0.056	−0.416	0.785	0.831
CMA	0.636	2.670	1.417	0.905
UMD	−0.064	−0.434	0.756	0.807
beta	0.026	0.146	0.765	0.745
ep	−0.096	−0.706	0.692	0.770
dy	0.296	1.867	0.998	0.759
sue	−0.024	−0.110	0.888	0.907
ps	−0.097	−0.783	0.699	0.777
bm_ia	0.173	1.535	0.898	0.759
cfp_ia	−0.065	−0.492	0.719	0.772
cfp	−0.039	−0.124	0.901	0.932
acc	0.211	1.707	0.946	0.776
sgr	0.126	0.964	0.895	0.793
HML_Devil	−0.238	−1.966	0.615	0.807
BAB	−0.144	−0.934	0.936	1.052
QMJ	−0.171	−0.652	0.903	1.041
Intermediary	0.180	1.510	0.911	0.766
convind	−0.334	−1.969	0.542	0.811

Table 12: Investment, profitability, and balance-sheet factors. Column 1: OLS estimator of a recession dummy on $\tilde{SR}_{t-2,t-1,t}$, that is, the estimated coefficient β in $\tilde{SR}_{t-2,t-1,t} = \alpha + \beta d_t + u_t$ (see Theorem 5); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average $\tilde{SR}_{t-2,t-1,t}$ over recession periods. Column 4: average $\tilde{SR}_{t-2,t-1,t}$ over all periods.

Zoo risk factor	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
lev	−0.069	−0.473	0.739	0.794
cashdebt	−0.015	−0.130	0.741	0.752
currat	0.070	0.600	0.809	0.753

Continued on next page

Zoo risk factor	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
pchcurrat	−0.114	−0.235	0.701	0.793
pchquick	−0.024	−0.187	0.626	0.645
quick	−0.111	−0.455	0.686	0.775
salecash	−0.079	−0.817	0.633	0.697
saleinv	0.078	0.781	0.700	0.637
salerec	0.001	0.011	0.728	0.727
cp	0.012	0.105	0.761	0.751
cto	0.139	1.224	0.862	0.751
pchsaleinv	−0.327	−1.806	0.546	0.809
pchsale_pchinvt	−0.396	−2.098	0.486	0.805
pchsale_pchrect	0.096	0.990	0.754	0.677
pchsale_pchxsga	0.199	1.291	0.883	0.723
depr	0.139	1.214	0.826	0.714
pchdepr	0.531	2.861	1.200	0.772
STR	−0.283	−2.488	0.575	0.803
IPO	−0.274	−2.012	0.532	0.753
indmom	0.210	1.461	0.922	0.753
nincr	−0.035	−0.375	0.585	0.614
chempia	0.163	0.659	1.038	0.906
cinvest	0.095	0.640	0.784	0.708
cinvest_a	−0.157	−1.251	0.621	0.748
noa	−0.050	−0.316	0.793	0.833
dnoa	−0.177	−1.988	0.556	0.699
tb	0.245	2.497	0.867	0.669
age	0.129	1.202	0.814	0.710
egr	−0.037	−0.276	0.801	0.831
lgr	0.047	0.197	0.900	0.863
dcoa	−0.170	−1.148	0.613	0.750
dcol	−0.069	−0.648	0.684	0.740
dwc	0.118	0.696	0.921	0.826
dnca	−0.391	−2.169	0.548	0.863
dncl	0.141	1.435	0.779	0.665
dnco	0.267	2.019	1.050	0.835
dfin	−0.114	−0.938	0.613	0.705
ta	0.064	0.541	0.753	0.702
dfnl	0.262	2.146	0.980	0.769
grcapx	0.097	0.862	0.794	0.716
pchcapx3	0.081	0.498	0.874	0.809
invest	−0.233	−1.994	0.605	0.793
orgcap	0.098	0.686	0.833	0.753
tang	0.027	0.188	0.686	0.664

Continued on next page

Zoo risk factor	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
nop	0.112	0.958	0.860	0.769
HXZ_IA	0.121	0.619	0.947	0.849

Table 13: Remaining, anomalies, and constraints factors. Column 1: OLS estimator of a recession dummy on $\tilde{SR}_{t-2,t-1,t}$, that is, the estimated coefficient β in $\tilde{SR}_{t-2,t-1,t} = \alpha + \beta d_t + u_t$ (see Theorem 5); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average $\tilde{SR}_{t-2,t-1,t}$ over recession periods. Column 4: average $\tilde{SR}_{t-2,t-1,t}$ over all periods.

Zoo risk factor	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
pps	0.016	0.084	0.843	0.831
etr	−0.118	−0.819	0.657	0.752
lfe	0.007	0.041	0.710	0.704
os	−0.034	−0.201	0.821	0.848
zs	−0.200	−1.105	0.669	0.831
pchcapx_ia	0.016	0.152	0.697	0.684
kz	−0.059	−0.359	0.793	0.841
chinv	0.135	0.797	0.971	0.862
chtx	−0.182	−1.525	0.581	0.728
grltnoa	−0.139	−0.871	0.737	0.849
ob_a	−0.012	−0.124	0.724	0.734
grltnoa_hxz	−0.377	−2.126	0.564	0.867
rdm	−0.060	−0.244	0.909	0.958
rds	−0.239	−2.119	0.487	0.679
ms	0.094	0.573	0.984	0.908
cei	−0.361	−0.704	0.692	0.983
nef	0.161	0.724	0.928	0.798
ndf	−0.057	−0.477	0.614	0.660
nxf	−0.027	−0.299	0.669	0.691
rs	−0.093	−0.735	0.645	0.720
herf	0.127	1.120	0.865	0.763
ww	0.147	0.946	0.913	0.795
roic	−0.091	−0.844	0.735	0.808
op	0.318	2.156	1.049	0.793
ndp	−0.079	−0.526	0.711	0.774
ebp	−0.354	−0.720	0.683	0.968
chcsho	0.061	0.548	0.758	0.709
aeavol	0.020	0.149	0.704	0.688
ear	−0.219	−2.075	0.447	0.623
dpia	−0.223	−1.821	0.612	0.792

Continued on next page

Zoo risk factor	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
pchcapx	−0.011	−0.082	0.757	0.766
cdi	0.115	0.948	0.801	0.709
rna	0.011	0.087	0.768	0.759
pm	−0.155	−1.198	0.673	0.799
ato	−0.175	−1.154	0.610	0.751
chatoia	0.126	1.080	0.860	0.758
chpmia	−0.146	−1.044	0.598	0.715
cashpr	0.050	0.516	0.736	0.696
sin	−0.144	−0.800	0.712	0.828
rsup	0.345	2.526	1.108	0.830
stdcf	−0.107	−1.225	0.603	0.690
absacc	0.179	1.875	0.781	0.637
realestate_hxz	0.071	0.577	0.908	0.851
pctacc	0.069	0.706	0.746	0.690
maxret	0.059	0.413	0.742	0.694
ol	−0.061	−0.396	0.759	0.809
ivg	−0.020	−0.178	0.754	0.770
poa	−0.359	−0.729	0.572	0.861
em	0.193	1.331	0.911	0.755
cash	−0.026	−0.212	0.699	0.721
gma	0.253	2.004	0.973	0.770
hire	0.016	0.131	0.806	0.793
gad	−0.035	−0.269	0.639	0.667
ala	0.083	0.823	0.741	0.674
HXZ_ROE	−0.364	−1.108	0.629	0.921

Table 14: Joint ranking of zoo risk factors by pricing-test significance and Sharpe ratio. Column 2 reports the rank of each factor based on the percentage of periods in which the corresponding risk premium pricing test is significant at the 5% level (smaller rank = more often significant). Column 3 reports the rank of each factor based on its Sharpe ratio over all periods (smaller rank = higher Sharpe ratio).

Zoo risk factor	Rank (pricing test)	Rank (Sharpe ratio)
BAB	1	2
pchsaleinv	2	41
dncl	3	114
UMD	5	17
herf	6	5
pchcapx	7	91
ol	8	40

Continued on next page

Zoo risk factor	Rank (pricing test)	Rank (Sharpe ratio)
pchsale_pchinvt	9	92
cinvest	10	140
HML_Devil	11	32
LTR	12	75
indmom	13	111
dnoa	14	53
roaq	15	50
HML	16	9
ill	17	148
dnco	18	11
cashdebt	19	81
salecash	20	109
saleinv	21	144
depr	22	107
cinvest_a	23	61
pricedelay	24	100
dfnl	25	80
ms	26	6
tang	27	141
aeavol	28	134
absacc	29	137
salerec	30	65
pchcapx3	31	94
moms12m	32	72
gma	33	88
orgcap	34	67
quick	35	110
mom6m	36	83
rds	37	120
grltnoa	38	35
chpmia	39	103
poa	40	115
ep	41	84
nincr	42	122
ob_a	43	49
age	44	59
cdi	45	99
currat	46	96
pchcurrat	47	145
pchquick	48	139
etr	49	117

Continued on next page

Zoo risk factor	Rank (pricing test)	Rank (Sharpe ratio)
ps	50	25
cfp_ia	51	21
chinv	52	12
ndf	53	127
ear	54	129
chatoia	55	36
cash	56	128
acc	57	70
cto	58	27
pchsale_pchxsga	59	104
rdm	60	37
chtx	61	105
grltnoa_hxz	62	48
noa	63	14
tb	64	131
dnca	65	13
dsti	66	106
stdcf	67	26
SMB	68	102
RMW	69	46
STR	70	22
cp	71	55
os	72	44
chempia	73	23
adm	74	20
rna	75	86
gad	76	138
ala	77	101
convind	78	28
sue	79	119
pps	80	31
mom36m	81	56
sp	82	10
pchsale_pchrect	83	97
pchcapx_ia	84	116
dolvol	85	147
ato	86	113
ivg	87	24
QMJ	88	4
HXZ_ROE	89	54
Intermediary	90	85

Continued on next page

Zoo risk factor	Rank (pricing test)	Rank (Sharpe ratio)
pchgm_pchsale	91	132
std_turn	92	123
dcoa	93	60
dfin	94	126
ta	95	133
nxf	96	93
chcsho	97	121
sin	98	77
em	99	38
sgr	100	58
IPO	101	89
zs	102	51
bm_ia	103	79
std_dolvol	104	112
stdacc	105	108
MktRf	106	136
LIQ_PS	107	69
rd	108	87
dwc	109	15
roic	110	3
ndp	111	76
maxret	112	130
dy	113	73
lfe	114	146
kz	115	18
egr_hxz	116	33
cei	117	42
ebp	118	57
rsup	119	34
invest	120	64
pctacc	121	39
CMA	122	16
pchdepr	123	125
cfp	124	52
roavol	125	47
chmom	126	63
grcapx	127	62
nef	128	71
dpia	129	66
realestate_hxz	130	7
HXZ_IA	131	74

Continued on next page

Zoo risk factor	Rank (pricing test)	Rank (Sharpe ratio)
beta	132	118
lev	133	90
baspread	134	143
egr	135	43
lgr	136	29
dcol	137	19
ww	138	95
cashpr	139	78
turn	140	135
zerotrade	141	142
nop	142	45
idiovol	143	68
pm	144	8
hire	145	98
rs	146	82
op	147	30
retvol	148	124

12.4 MV Portfolio - $T = 1$ - Misspecification

Table 15: Ranking of Sharpe ratios for zoo risk factors with and without PCA bias adjustment. Column 2 reports the rank of each factor's Sharpe ratio over all periods using the baseline zoo specification. Columns 3–5 report the corresponding ranks when Sharpe ratios are computed after PCA3, PCA5, and PCA10 bias adjustments, respectively (larger rank = higher Sharpe ratio).

Zoo risk factor	Rank (SR, zoo)	Rank (SR, zoo+PCA3)	Rank (SR, zoo+PCA5)	Rank (SR, zoo+PCA10)
ill	148	140	71	23
dolvol	147	70	1	75
lfe	146	34	74	35
pchcurrat	145	89	92	130
saleinv	144	68	30	8
baspread	143	62	40	5
zerotrade	142	88	11	1
tang	141	77	26	7
cinvest	140	82	53	50
pchquick	139	107	84	110
gad	138	79	58	12
absacc	137	76	61	52
MktRf	136	84	43	138
turn	135	116	34	10
aeavol	134	71	2	28
ta	133	52	5	66
pchgm_pchsale	132	11	48	39
tb	131	135	138	63
maxret	130	113	12	2
ear	129	28	126	83
cash	128	22	16	30
ndf	127	83	50	47
dfin	126	86	73	37
pchdepr	125	96	6	97
retvol	124	64	52	11
std_turn	123	10	13	6
nincr	122	97	14	38
chcsho	121	131	119	46
rds	120	12	83	133
sue	119	91	63	36
beta	118	50	38	24
etr	117	25	17	16
pchcapx_ia	116	2	140	111
poa	115	24	41	88

Continued on next page

Zoo risk factor	Rank (SR, zoo)	Rank (SR, zoo+PCA3)	Rank (SR, zoo+PCA5)	Rank (SR, zoo+PCA10)
dncl	114	5	96	125
ato	113	133	47	31
std_dolvol	112	74	132	40
indmom	111	19	80	43
quick	110	54	116	29
salecash	109	110	25	90
stdacc	108	43	23	4
depr	107	134	36	56
dsti	106	81	109	78
chtx	105	127	136	59
pchsale_pchxsga	104	38	130	41
chpmia	103	75	137	74
SMB	102	27	4	27
ala	101	102	42	20
pricedelay	100	3	127	98
cdi	99	41	35	141
hire	98	20	128	68
pchsale_pchrect	97	121	55	21
currat	96	118	139	13
ww	95	17	59	48
pchcapx3	94	95	114	14
nxf	93	147	72	25
pchsale_pchinvt	92	44	135	112
pchcapx	91	72	97	107
lev	90	42	122	109
IPO	89	32	76	92
gma	88	142	62	99
rd	87	6	37	67
rna	86	143	91	113
Intermediary	85	93	8	9
ep	84	60	20	51
mom6m	83	15	3	22
rs	82	122	146	79
cashdebt	81	85	44	127
dfnl	80	115	103	108
bm_ia	79	33	32	84
cashpr	78	58	24	116
sin	77	69	27	93
ndp	76	129	66	64
LTR	75	8	9	32
HXZ_IA	74	53	104	70

Continued on next page

Zoo risk factor	Rank (SR, zoo)	Rank (SR, zoo+PCA3)	Rank (SR, zoo+PCA5)	Rank (SR, zoo+PCA10)
dy	73	117	7	15
moms12m	72	120	93	69
nef	71	94	95	19
acc	70	109	70	65
LIQ_PS	69	101	56	89
idiovol	68	146	45	55
orgcap	67	59	82	96
dpia	66	18	99	126
salerec	65	26	57	61
invest	64	47	106	124
chmom	63	45	75	76
grcapx	62	66	102	18
cinvest_a	61	100	19	49
dcoa	60	31	77	77
age	59	29	21	26
sgr	58	119	118	58
ebp	57	40	87	105
mom36m	56	55	105	128
cp	55	104	10	85
HXZ_ROE	54	106	29	142
dnoa	53	136	129	102
cfp	52	103	78	114
zs	51	56	81	82
roaq	50	46	51	100
ob_a	49	9	90	140
grltnoa_hxz	48	65	100	122
roavol	47	99	33	45
RMW	46	78	18	123
nop	45	137	64	44
os	44	144	94	91
egr	43	73	79	121
cei	42	90	111	134
pchsaleinv	41	13	121	103
ol	40	124	31	80
pctacc	39	49	133	73
em	38	35	120	86
rdm	37	14	112	145
chatoia	36	138	110	131
grltnoa	35	114	89	104
rsup	34	112	69	71
egr_hxz	33	67	124	54

Continued on next page

Zoo risk factor	Rank (SR, zoo)	Rank (SR, zoo+PCA3)	Rank (SR, zoo+PCA5)	Rank (SR, zoo+PCA10)
HML_Devil	32	61	28	62
pps	31	130	22	42
op	30	145	67	33
lgr	29	16	88	81
convind	28	23	107	53
cto	27	139	60	132
stdcf	26	98	15	3
ps	25	87	123	57
ivg	24	92	131	95
chempia	23	7	145	101
STR	22	51	134	136
cfp_ia	21	37	49	17
adm	20	123	113	118
dcol	19	80	108	34
kz	18	141	142	117
UMD	17	105	86	143
CMA	16	36	85	120
dwc	15	57	98	60
noa	14	111	141	94
dnca	13	48	144	137
chinv	12	39	39	115
dnco	11	108	101	106
sp	10	21	143	146
HML	9	125	68	72
pm	8	1	117	129
realestate_hxz	7	63	115	144
ms	6	30	65	119
herf	5	4	125	87
QMJ	4	128	54	135
roic	3	132	46	139
BAB	2	126	148	148
RF	1	148	147	147

12.5 Cross-Sectional Models When $T = r$

12.5.1 Cross-Sectional Models When $T = r$: Latent

Table 16: Latent Multi-Factor Models: Column 1: OLS estimator of a recession dummy on $\widetilde{SR}_{1:T,t}$, that is, the estimated coefficient β in $\widetilde{SR}_{1:T,t} = \alpha + \beta d_t + u_t$ (see Section 6.2); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average $\widetilde{SR}_{1:T,t}$ over recession periods. Column 4: average $\widetilde{SR}_{1:T,t}$ over all periods.

Multi-factor model	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
PCA3	−1.838	−5.429	−0.113	0.287
PCA5	−0.817	−3.288	−0.303	0.361
PCA10	−1.776	−5.961	−0.730	0.698

Table 17: Latent Multi-Factor Models: Column 1: OLS estimator of a recession dummy on pricin test $\tilde{L}^{\texttt{corr}}_{t+1|t}$ (absolute-value t-test), that is, the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, when $\widetilde{test}_t \equiv \sqrt{N}\,|\tilde{L}^{\texttt{corr}}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$, and $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 8); d_t is a dumm equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OL coefficient. Column 3: average pricing test $\tilde{L}^{\texttt{corr}}_{t+1|t}$ over recession periods. Column 4: average pricing tes over all periods. Columns 5–7: frequency with $|t| > 1.96$, $|t| > 2 \times 1.96$, and $|t| > 5 \times 1.96$, respectively

Multi-factor model	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\lvert t\rvert > 1.96$)	Perc($\lvert t\rvert > 2 \times 1.96$)	Perc($\lvert t\rvert > 5$
PCA3	0.270	0.794	0.310	−0.314	0.638	0.351	0.02
PCA5	−0.326	−1.289	0.317	0.212	0.582	0.289	0.01
PCA10	−0.326	−1.289	0.329	0.546	0.521	0.262	0.00

12.5.2 Cross-Sectional Models When $T = r$: Observed

Table 18: Observed multi-factor models. Column 1: OLS estimator of a recession dummy on $\widetilde{SR}_{1:T,t+1}$, that is, the estimated coefficient β in $\widetilde{SR}_{1:T,t+1} = \alpha + \beta d_t + u_t$ (see Section 6.2); d_t is a dummy equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OLS coefficient. Column 3: average $\widetilde{SR}_{1:T,t+1}$ over recession periods. Column 4: average $\widetilde{SR}_{1:T,t+1}$ over all periods.

Multi-factor model	Coeff	t-ratio	Mean(SR)-crises	Mean(SR)-overall
FF3	−0.595	−3.006	−0.009	0.471
FF5	−1.099	−3.054	−0.456	0.436
FF5+MOM	−1.464	−4.410	−0.442	0.750

Table 19: Observed Multi-Factor Models: Column 1: OLS estimator of a recession dummy on pricin test $\tilde{L}^{\texttt{o-corr}}_{t+1|t}$ (absolute-value t-test), that is, the estimated coefficient β in $|\widetilde{test}_t| = \alpha + \beta d_t + u_t$, wher $\widetilde{test}_t \equiv \sqrt{N}\,|\tilde{L}^{\texttt{o-corr}}_{t+1|t}|/SE_{\tilde{L}_{t+1|t}}$, and $SE_{\tilde{L}_{t+1|t}}$ denotes the standard error (see Theorem 8); d_t is a dumm equal to one in recession/financial-crisis periods and zero otherwise. Column 2: t-ratio of the OL coefficient. Column 3: average pricing test $\tilde{L}^{\texttt{o-corr}}_{t+1|t}$ over recession periods. Column 4: average pricin test over all periods. Columns 5–7: frequency with $|t| > 1.96$, $|t| > 2 \times 1.96$, and $|t| > 5 \times 1.96$ respectively.

Multi-factor model	Coeff	t-ratio	Mean(t)-crises	Mean(t)-overall	Perc($\lvert t\rvert > 1.96$)	Perc($\lvert t\rvert > 2 \times 1.96$)	Perc($\lvert t\rvert > 5$
FF3	0.354	1.528	0.317	−0.049	0.400	0.161	0.00
FF5	0.004	0.021	0.325	−0.129	0.300	0.094	0.00
FF5+MOM	−0.001	−0.002	0.329	−0.079	0.347	0.105	0.00

References

Aiolfi, M., and A. Timmermann (2006): "Persistence in forecasting performance and conditional combination strategies," *Journal of Econometrics*, 135, 31–53.

Andersen, T., T. Bollerselv, F. Diebold, and J. Wu (2006): "Realized beta: Persistence and predictability," in *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger, vol. B*, ed. by T. Fomby, and D. Terrell, pp. 1–40. Elsevier, Amsterdam, Netherlands.

Ang, A., and D. Kristensen (2012): "Testing conditional factor models," *Journal of Financial Economics*, 106, 132–156.

Bai, J. (2003): "Inferential theory for factor models of large dimensions," *Econometrica*, 71, 135–171.

Bai, J., and S. Ng (2002): "Determining the number of factors in approximate factor models," *Econometrica*, 70, 191–221.

Bates, J. B., M. Plagborg-Møller, J. Stock, and M. Watson (2013): "Consistent factor estimation in dynamic factor models with structural instability," *Journal of Econometrics*, 177, 289–304.

Bryzgalova, S., J. Huang, and C. Julliard (2023): "Bayesian solutions for the factor zoo: We just ran two quadrillion models," *The Journal of Finance*, 78, 487–557.

Campbell, J. Y., and S. B. Thompson (2008): "Predicting excess stock returns Out of Sample: Can anything beat the historical average?," *Review of Financial Studies*, 21, 1509–1531.

Chamberlain, G. (1983): "Funds, factors, and diversification in arbitrage pricing models," *Econometrica*, 51, 1305–1323.

Connor, G., and R. A. Korajczyk (1986): "Performance measurement with the Arbitrage Pricing Theory: A new framework for analysis," *Journal of Financial Economics*, 15, 373–394.

Connor, G., H. M., and O. Linton (2012): "Efficient estimation of a semiparametric characteristic-based factor model of security returns," *Econometrica*, 80, 713–754.

Fama, E. F., and K. R. French (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

——— (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116, 1–22.

Fama, E. F., and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607–636.

Fan, J., Y. Liao, and W. Wang (2016): "Projected principal component analysis in factor models," *The Annals of Statistics*, 44, 219–254.

Feng, G., S. Giglio, and D. Xiu (2020): "Taming the Factor Zoo: A Test of New Factors," *Journal of Finance*, 75, 1327–1370.

Giglio, S., and D. Xiu (2021): "Asset pricing with omitted factors," *Journal of Political Economy*, 129, 1947–1990.

Giglio, S., D. Xiu, and D. Zhang (2025): "Test Assets and Weak Factors," *Journal of Finance*, 80, 259–319.

Goyal, A., and I. Welch (2008): "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies*, 21, 1455–1508.

Gu, S., B. Kelly, and D. Xiu (2020): "Empirical asset pricing via machine learning," *Review of Financial Studies*, 33, 2223–2273.

Haddad, V., S. Kozak, and S. Santosh (2020): "Factor timing," *The Review of Financial Studies*, 33, 1980–2018.

Hansen, L., and R. Jagannathan (1997): "Assessing specification errors in stochastic discount factor models," *Journal of Finance*, 52(2), 557–590.

Hansen, L. P., and S. F. Richard (1987): "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," *Econometrica*, 55, 587–613.

Huberman, G. (1982): "A simple approach to Arbitrage Pricing Theory," *Journal of Financial Economics*, 28, 183–191.

Ingersoll, J. E. (1984): "Some results in the theory of arbitrage pricing," *The Journal of Finance*, 39, 1021–1039.

Jacquier, A., J. Muhle-Karbe, and J. Mulligan (2025): "In-Sample and Out-of-Sample Sharpe ratios for linear predictive models," *Quantitative Finance*, Forthcoming.

Kan, R., X. Wang, and X. Zheng (2024): "In-Sample and out-of-Sample Sharpe Ratios of Multi-Factor Asset Pricing Models," *Journal of Financial Economics*, 155, 103837.

Kelly, B., and S. Pruitt (2013): "Market Expectations in the Cross-Section of Present Values," *Journal of Finance*, 68, 1721–1756.

Kelly, B. T., S. Pruitt, and Y. Su (2019): "Characteristics are covariances: A unified model of risk and return.," *Journal of Financial Economics*, 134, 501–524.

Kim, K., R. A. Korajczyk, and A. Neuhierl (2021): "Arbitrage portfolios," *The Review of Financial Studies*, 34, 2813–2856.

Lettau, M., and M. Pelger (2020a): "Estimating latent asset-pricing factors," *Journal of Econometrics*, 218, 1–31.

——— (2020b): "Factors that fit the time-series and cross-section of stock returns," *The Review of Financial Studies*, 33, 2274–2325.

Lewellen, J., and S. Nagel (2006): "The conditional CAPM does not explain asset-pricing anomalies," *Journal of Financial Economics*, 82, 289–314.

Ma, H., and P. Gagliardini (2020): "Extracting statistical factors when betas are time-varying," *Working Paper, Swiss Finance Institute*.

Magnus, J., and R. H. Neudecker (2007): *Matrix Differential Calculus with Applications in Statistics and Econometrics*. J. Wiley & Sons, Revised Edition, Chichester.

Martin, I. (2017): "What is the expected return on the market?," *The Quarterly Journal of Economics*, 132, 367–433.

Martin, I., and S. Nagel (2019): "Market efficiency in the age of big data," *Journal of Financial Economics*, 135, 1–29.

Pelger, M. (2019): "Large-dimensional factor modeling based on high-frequency observations," *Journal of Econometrics*, 208, 23–42.

——— (2020): "Understanding systematic risk: A high-frequency approach," *Journal of Finance*, 75, 2179–2220.

Pelger, M., and R. Xiong (2019): "State-varying factor models of large dimensions," *Journal of Business & Economic Statistics*, 40, 1315–1333.

Pettenuzzo, D., A. Timmermann, and R. Valkanov (2014): "Forecasting stock returns under economic constraints," *Journal of Financial Economics*, 114, 517–553.

Rapach, D. E., J. Strauss, and G. Zhou (2010): "Out-of-Sample equity premium prediction: Combination forecasts and links to the real economy," *Review of Financial Studies*, 23, 821–862.

Ross, S. A. (1976): "The Arbitrage Theory of capital asset pricing," *Journal of Economic Theory*, 13, 341–360.

Stambaugh, R. F. (1983): "Arbitrage pricing with information," *Journal of Financial Economics*, 12, 357–369.

Zaffaroni, P. (2025): "Factor models for conditional asset pricing," *Journal of Political Economy, forthcoming*.

Zaffaroni, P., and G. Zhou (2024): *Asset Pricing: Cross-Section Predictability*. Oxford University Press.