Linear Factor Models and the Estimation of Expected Returns

Cisil Sarisoy, Peter de Goeij, and Bas J.M. Werker

2024-014

Linear Factor Models and the Estimation of Expected Returns

Cisil Sarisoy[*]

Federal Reserve Board

Peter de Goeij

Tilburg University

Bas J.M. Werker

Tilburg University

January 2024

Abstract

This paper analyzes the properties of expected return estimators on individual assets implied by the linear factor models of asset pricing, i.e., the product of β and λ. We provide the asymptotic properties of factor–model–based expected return estimators, which yield the standard errors for risk premium estimators for individual assets. We show that using factor-model-based risk premium estimates leads to sizable precision gains compared to using historical averages. Finally, inference about expected returns does not suffer from a small–beta bias when factors are traded. The more precise factor–model–based estimates of expected returns translate into sizable improvements in out–of–sample performance of optimal portfolios.

Keywords: Cross Section of Expected Returns, Risk Premium, Small β's.

1 Introduction

Estimating expected returns on individual assets or portfolios is perhaps one of the longest standing challenges in asset pricing. One standard approach at hand is to use historical averages. However, it is known that these estimates are generally very noisy. Even using daily data does not help much, if at all. There is a long history of papers trying to improve estimates of expected returns by using asset pricing models, in which expected excess returns on individual assets are linear in their exposures to the risk factors imposed (β). The coefficients in this linear relationship are the prices of risk for the factors (λ). Examples include Sharpe (1964)'s CAPM, Merton (1973)'s ICAPM, Breeden (1979)'s CCAPM, Ross (1976)' APT and Lettau and Ludvigson (2001)'s conditional CCAPM, among many others.

The literature on inference based on factor models mainly concentrates, in a frequentist setting, on the econometric properties of the prices of risk, λ, and evaluating the ability of the models in explaining the cross section of expected returns. In this paper, the focus is different: we analyze the estimation of the expected (excess) returns on *individual assets* or *portfolios* based on linear factor models, i.e., the product of exposures β and risk prices λ. In order to have an estimate of the expected (excess) return on an individual asset, both β and λ have to be estimated, and the dependence between these estimators introduce a nontrivial noise structure in the standard errors of the expected (excess) return estimators.

Jorion (1991) compares CAPM—based estimators with classical sample averages of past returns finding the former outperforming the latter in estimating expected stock returns. Pástor and Stambaugh (1999) investigate, in a Bayesian setting, the impact of prior uncertainty about mispricing in a factor model on the posterior estimates of the cost of equity. Similarly, Pástor (2000) develops Bayesian approaches to examine the role of prior mispricing in portfolio allocation decisions. Our paper complements this earlier work by providing the first asymptotic analysis for the expected (excess) return estimators for sev-

eral often–used factor models. Such limiting distributions yield the frequentist standard errors and, accordingly, confidence bounds for the expected (excess) return of individual assets or portfolios. Moreover, we evaluate the implications of weakly correlated factors on the estimation of expected (excess) returns. We examine the inference under various settings where the factors are traded, non-traded or their mimicking portfolios are used in the estimation.

First, we derive the asymptotic properties of expected (excess) return –risk premium– estimators based on factor models. These limiting distributions yield the standard errors for individual assets or portfolios. We thereby assess the precision gains from using factor–model based risk–premium estimators vis–à–vis the historical averages approach. In particular, we provide closed-form asymptotic expressions for these precision gains. We show in Theorems 4.2, 4.3, and 4.4 that exploiting the linear relationship implied by linear factor models indeed leads to more precise estimates of risk premiums as compared to historical averages. In an empirical analysis of the estimation of risk-premiums on 25 Fama and French (1992) size and book–to–market sorted portfolios, we document reductions in estimated variances of up to 24% for individual portfolios.

Second, we analyze the estimation of risk premiums in the presence of weakly correlated and spurious factors. When factors are weakly correlated with assets, i.e., β's are small, the standard confidence intervals of the price of risk estimates are known to be erroneous (see, e.g., Kleibergen, 2009). This effect may be severe in empirical research, as these confidence intervals may be unbounded as documented for the case of consumption CAPM of Lettau and Ludvigson (2001)[1]. This is a relevant issue in practice because macroeconomic variables are typically weakly related to individual asset/portfolio returns. We demonstrate that such issues do not exist if the object of interest is the risk premiums on

---

[1]See also Kan and Zhang (1999) Gospodinov, Kan, and Robotti (2014), Bryzgalova (2015), Burnside (2015), Gospodinov, Kan, and Robotti (2017, 2019), Giglio, Xiu, and Zhang (2021) on the role of spurious or weakly identified factors for inference about the prices of risk.

individual assets, but only in case factors are traded. In that case, the limiting variances of the risk–premium estimators are not affected by the β's being small, see Corollary 5.1-2. Monte Carlo simulation results document that those limiting variances provide reliable approximations of the finite-sample variances of the factor-model-based estimators of risk premiums.

Third, we explore the implications of the precision gains from using factor model based estimates of expected returns in Markowitz (1952)'s setting. The implementation of the mean–variance framework of Markowitz (1952) in practice requires the estimation of the first two moments of asset returns. Constructing optimal portfolios with the imprecise estimates of expected returns, using historical averages, and the sample covariance matrix generally lead to poor out–of–sample performance.[2] In the far end, this has led to simply abandoning the application of theoretically optimal decisions and using naive techniques such as the 1/N strategy or the global minimum variance (GMV) portfolio as these are not subject to estimation risk on expected returns (DeMiguel, Garlappi, and Uppal, 2009).[3] Our Monte Carlo simulations document strong improvements in the out–of–sample Sharpe ratios of optimal portfolios when constructed with factor–model–based estimates of risk premiums as compared to when constructed with the historical averages. Moreover, optimal portfolios constructed with the factor–model–based risk–premium estimates perform better than both the GMV portfolio and the 1/N strategy portfolio.

The remainder of the paper is organized as follows. Section 2 introduces our set–up and presents the linear factor model with the assumptions that form the basis of our statistical analysis. Next, we introduce factor–mimicking portfolios and clarify the link between the

---

[2]See, for example, Frost and Savarino (1988), Michaud (1989), Jobson and Korkie (1980), and Best and Grauer (1991).

[3]Several studies provide solutions on improving the covariance matrix estimates (see, e.g., Ledoit and Wolf, 2003, DeMiguel et al., 2009 among others). However, the estimation error in asset return means is more severe than error in covariance estimates (see Merton, 1980, Chopra and Ziemba, 1993) and the imprecision in estimates of the expected returns affects the optimal portfolio weights more drastically compared to the imprecision in covariance estimates (see DeMiguel et al., 2009).

expected returns obtained with non–traded factors and with factor–mimicking portfolios. Section 3 discusses in detail the standard GMM estimators we consider. In particular, we recall the different sets of moment conditions for various cases such as all factors being traded or using factor–mimicking portfolios. Section 4 derives the asymptotic properties of these induced GMM estimators, and we derive the efficiency gains over and above the risk–premium estimator based on historical averages. Section 5 presents the analysis for the small βs. Section 6 reports results from a Monte Carlo simulation experiment to study the finite–sample properties of the factor–model based estimators of expected (excess) returns. Section 7 presents our simulation analysis for portfolio optimization, and Section 8 concludes. All proofs are gathered in the appendix.

2   Model and Assumptions

Let M be a candidate stochastic discount factor such that for any traded asset $i = 1, 2, \ldots, N$ with excess return R_i^e

$$\mathrm{E}\left[MR_i^e\right] = 0. \tag{2.1}$$

Linear factor models additionally specify $M = a + b'F$, where $F = (F_1, ..., F_K)'$ is a vector of K factors. Note that (2.1) can be written in matrix notation using the vector of excess returns $R^e = (R_1^e, ..., R_N^e)'$. Throughout we impose the following.

Assumption 1. *The N–vector of excess asset returns R^e and the K–vector of factors F with $K < N$ satisfy the following conditions:*

1. *The covariance matrix of excess returns $\Sigma_{R^e R^e}$ has full rank N,*

2. *The covariance matrix of factors Σ_{FF} has full rank K,*

3. *The covariance between excess returns and factors, $Cov\left[R^e, F'\right]$, has full rank K.*

The linear asset-pricing model can be alternatively expressed using the beta representation

$$\mathrm{E}\left[R^e\right] = \beta\lambda, \tag{2.2}$$

where $\beta = \mathrm{Cov}\left[R^e, F'\right]\Sigma_{FF}^{-1}$, and $\lambda = -\frac{1}{\mathrm{E}[M]}\Sigma_{FF}b$.

Thus, (2.2) specifies a linear relationship between risk premiums on individual assets, $\mathrm{E}\left[R^e\right]$, and their exposures β to the risk factors, F. The vector λ denoted the so–called prices of risk of the factors.[4] The primary focus of our analysis is on inference about (2.2). For our main results, the following assumptions are needed.

Assumption 2. *Assume that $[R_t^{e\prime}, F_t']'$ is a jointly stationary and ergodic process with a finite fourth moment.*

Assumption 3. *Let $\varepsilon_t = R_t^e - \alpha - \beta F_t$. Assume that $E[\varepsilon_t|F_t] = 0$ and $Var[\varepsilon_t|F_t] = \Sigma_{\varepsilon\varepsilon}$.*

Assumption 2 provides primitive conditions for central limit theorem approximations for returns and factors. This assumption is sufficient to obtain limiting distributions for the GMM estimators that we focus in this paper. But to obtain explicit limiting results, we make further assumptions on the data. Assumption 3 is made for that purpose and can further be relaxed at the cost of a more cumbersome notation.

2.1   Factor–Mimicking Portfolios

A large number of studies in the asset pricing literature suggest "macroeconomic" factors that capture systematic risk. Examples include the C-CAPM of Breeden (1979), the I-CAPM of Merton (1973), and the conditional C-CAPM of Lettau and Ludvigson (2001). In order to assess the validity of macroeconomic risk factors being priced or not, it has been

---

[4]We focus on constant parameter factor models. Gagliardini, Ossola, and Scaillet (2016) and Kelly, Pruitt, and Su (2019) allow for time varying risk exposures and time varying factor risk premia by incorporating information from stock characteristics and macroeconomic variables. Our results can be extended to such a setting, at the cost of additional assumptions. As constant parameter models are still widely used in empirical applications, we've decided to focus on that setting.

suggested to refer to alternative formulations of such factor models replacing the factors by their projections on the linear span of the excess returns. This is commonly referred to as factor–mimicking portfolios and early references go back to Huberman, Kandel, and Stambaugh (1987) (see also, e.g., Fama, 1998, Lamont, 2001, and Balduzzi and Robotti, 2008). In this paper, we analyze the effect of such formulations on the estimation of risk premiums and we show, in Section 4, that there are efficiency gains from the information in mimicking portfolios when estimating risk premiums.

It is important to understand that, the prices of risk of (non–traded) factors generally differ from the risk–premiums on their factor–mimicking portfolios. However, using factor–mimicking portfolios leads to identical risk premiums on individual assets. This is shown in Theorem 2.1 below.

To be precise, we project the factors F_t onto the space of excess asset returns, augmented with a constant. In particular, given Assumption 3, there exists a K–vector Φ_0 and a $K \times N$ matrix Φ of constants and a K–vector of random variables u_t satisfying

$$F_t = \Phi_0 + \Phi R_t^e + u_t, \tag{2.3}$$

$$\mathrm{E}\left[u_t\right] = 0_{K \times 1}, \quad \text{and} \quad Eu_t R_t^{e\prime} = 0_{K \times N}.$$

We then define the factor–mimicking portfolios by

$$F_t^m = \Phi R_t^e. \tag{2.4}$$

Now, we obtain an alternative formulation of the linear factor model by replacing the original factors by their factor–mimicking portfolios

$$R_t^e = \alpha^m + \beta^m F_t^m + \varepsilon_t^m, \quad t = 1, 2, \ldots, T. \tag{2.5}$$

Recall that, using the projection results, we have

$$\Phi = \Sigma_{FF}\beta'\Sigma_{R^eR^e}^{-1}, \quad \text{and} \quad \beta^m = \beta\left(\beta'\Sigma_{R^eR^e}^{-1}\beta\right)^{-1}\Sigma_{FF}^{-1}. \tag{2.6}$$

The following theorem recalls that, while factor loadings and prices of risk change when using factor mimicking portfolios, expected (excess) returns on individual assets, i.e., their product, are not affected. For completeness we provide a proof in the appendix.

Theorem 2.1. *Under Assumption 1, we have $\beta\lambda = \beta^m\lambda^m$ with $\lambda^m = E[F_t^m]$.*

Note that since the factor–mimicking portfolio is an excess return itself, asset pricing theory implies that the price of risk attached to it, λ^m, equals its expectation. This additional information can be imposed in the estimation of expected (excess) returns and one may hope that the expected (excess) return estimators obtained with factor–mimicking portfolios are more precise than the expected (excess) return estimators obtained with the non-traded factors themselves. We study this question in Section 4.

3   Estimation

We concentrate on Hansen (1982)'s GMM estimation technique. The GMM approach is particularly useful in our paper as it avoids the use of two-step estimators and the resulting "errors-in-variables" problem when calculating limiting distributions. In addition, we immediately obtain the joint limiting distribution of estimates for β and λ which is needed as we are interested in their product.

In the following sections, we study the asymptotics of the expected (excess) return estimators by specifying different sets of moment conditions. In Section 3.1, we study a set of moment conditions which generally holds, i.e., both when factors are traded and when they are non-traded. In Section 3.2, we study the case where all factors are traded. We then incorporate the moment condition that factor prices equal expected factor values.

In Section 3.3, we consider expected (excess) return estimates based on factor–mimicking portfolios.

3.1 Moment Conditions - General Case

We first provide the moment conditions for a general case, i.e., where factors may represent excess returns themselves, but not necessarily. In that case, the standard moment conditions to estimate both factor loadings β and factor prices λ are

$$
\mathrm{E}\left[h_t(\alpha, \beta, \lambda)\right] = \mathrm{E}\left[\begin{array}{c} \begin{bmatrix} 1 \\ F_t \end{bmatrix} \otimes [R_t^e - \alpha - \beta F_t] \\ \\ R_t^e - \beta\lambda \end{array}\right] = 0. \tag{3.1}
$$

The first set of moment conditions identifies α and β as regression coefficients, while the last set of conditions represents the pricing restrictions. Note that there are $N(1 + K + 1)$ moment conditions although there are $N(1 + K) + K$ parameters, which implies that the system is overidentified. We set a linear combination of the given moment conditions to zero, that is, we set $A\mathrm{E}\left[h_t(\alpha, \beta, \lambda)\right] = 0$ with

$$
A = \begin{bmatrix} I_{N(1+K)} & 0_{N(1+K)\times N} \\ \\ 0_{K\times N(1+K)} & \Theta_{K\times N} \end{bmatrix}.
$$

Note that the matrix A specified above combines the last N moment conditions into K moment conditions so that the system becomes exactly identified. We take $\Theta = \beta^T \Sigma_{\varepsilon\varepsilon}^{-1}$. The advantage of this particular choice is that the resulting λ estimates coincide with the GLS cross–sectional estimates.

3.2 Moment Conditions - Traded Factor Case

Asset pricing theory provides an additional restriction on the prices of risk when factors are traded, meaning that they are excess returns themselves. If a factor is an excess return, we have $\lambda = \mathrm{E}[F_t]$. For example, the price of market risk is equal to the expected excess market return, and the prices of size and book–to–market risks, as captured by Fama-French's SMB and HML portfolio movements, are equal to the expected SMB and HML excess returns. Note that we use the term "excess return" for any difference of gross returns, that is, not only in excess of the risk-free rate. Prices of excess returns are zero, i.e., excess returns are zero investment portfolios.

The standard two–pass estimation procedure commonly found in the finance literature may not give reliable estimates of risk prices when factors are traded. Hou and Kimmel (2006) provide an interesting example to point out this issue. They generate standard two–pass expected (excess) return estimates (both OLS and GLS) in the three factor Fama–French model by using 25 size and book–to–market porfolios as test assets. As shown in their Table 1, both OLS and GLS risk price estimates of the market are significantly different from the sample average of the excess market return. It is important to point out that the two–pass procedure ignores the fact that the Fama–French factors are traded factors and it treats them in the same way as non–traded factors.

Consequently, when factors are traded we usually replace the second set of moment conditions with the condition that their expectation of the vector of factors equals λ. Then, the relevant moment conditions are given by

$$
\mathrm{E}\left[h_t(\alpha, \beta, \lambda)\right] = \mathrm{E}\left[
\begin{array}{c}
\begin{bmatrix} 1 \\ F_t \end{bmatrix} \otimes \left[R_t^e - \alpha - \beta F_t\right] \\
\\
F_t^e - \lambda
\end{array}
\right] = 0, \tag{3.2}
$$

where F_t is the $K \times 1$ vector of factor (excess) returns.

In this case, estimates are obtained by an exactly identified system, i.e., the number of parameters equals the number of moment conditions. Note that if the factor is traded, but we do not add the moment condition that the factor averages equal λ , then the results are just those of the non-traded case in Section 3.1.

Alternatively, we could incorporate the theoretical restriction on factor prices into the estimation by adding the factor portfolios as test assets in the linear pricing equation, $R^e - \beta\lambda$. This set of moment conditions would be similar to the general case, with the only difference being that the linear pricing restriction incorporates the factors as test assets in addition to the original set of test assets, i.e., we define $R_t^F = \begin{bmatrix} R_t^e \\ F_t \end{bmatrix}$. Under this setting, the moment conditions would be given by

$$\mathrm{E}\left[h_t(\alpha, \beta, \lambda)\right] = \mathrm{E}\left[\begin{bmatrix} \begin{bmatrix} 1 \\ F_t \end{bmatrix} \otimes [R_t^e - \alpha - \beta F_t] \\ R_t^F - \beta_{F,R}\lambda \end{bmatrix} \right] = 0, \tag{3.3}$$

with $\beta_{F,R} = \begin{bmatrix} \beta \\ I_K \end{bmatrix}$. Following the same procedure as in Section 3.1, we specify the A matrix and set $\Theta = \beta_{F,R}^T \Sigma_{R^F R^F}^{-1}$. Because we find that GMM based on (3.3) leads to the same asymptotic variance covariance matrices for risk premiums as GMM based on (3.2), we omit the conditions (3.3) in the rest of the paper.

3.3   Moment Conditions - Factor–Mimicking Portfolios

Following Balduzzi and Robotti (2008), we also consider the case where risk prices are equal to expected returns of factor–mimicking portfolios. Then, the moment conditions used are

$$
\mathrm{E}\left[h_t(\alpha^m, \beta^m, \Phi_0, \Phi, \lambda^m)\right] = \mathrm{E}
\begin{bmatrix}
\begin{bmatrix} 1 \\ R_t^e \end{bmatrix} \otimes [F_t - \Phi_0 - \Phi R_t^e] \\
\begin{bmatrix} 1 \\ F_t^m \end{bmatrix} \otimes [R_t^e - \alpha^m - \beta^m F_t^m] \\
\Phi R_t^e - \lambda^m
\end{bmatrix} = 0, \qquad (3.4)
$$

with $F_t^m = \Phi R_t^e$. In this case, there are $K(1+N) + N(1+K) + K$ moment conditions and parameters, which makes the system exactly identified.

4   Precision of Risk–Premium Estimators

As mentioned in the introduction, our focus is on estimating risk premiums of individual assets or portfolios. Much of the literature on multi–factor asset pricing models has primarily focused on the issue of a factor being priced or not. Formally, this is a test on (a component of) λ being zero or not and, accordingly, the properties of risk–price estimates for λ have been studied and compared.[5]

In this paper, since our focus is on analyzing the possible efficiency gains based on linear factor models in estimating expected (excess) returns, we first derive the joint distribution of estimates for β and λ for the three cases discussed in Sections 3.1 to 3.3.[6] Then, we derive the asymptotic distributions of the implied expected (excess) return estimators given by the product $\hat{\beta}\hat{\lambda}$. Moreover, we illustrate the empirical relevance of our asymptotic results

---

[5]Examples include Shanken (1992), Jagannathan and Wang (1998), Kleibergen (2009), Lewellen, Nagel, and Shanken (2010), Kan and Robotti (2011), and Kan, Robotti, and Shanken (2013).

[6]We also provide results for inference on risk premiums when the model is misspecified in the appendix.

using the Fama–French three factor model with 25 Fama–French size and book–to–market portfolios as test assets. In particular, we provide the (asymptotic) variances of the various risk–premium estimators with empirically reasonable parameter values and evaluate the benefits of using linear factor models in estimating risk premiums, see Table 1 below.

The asset data used in this paper consists of 25 portfolios formed by Fama-French (1992, 1993), downloaded from Kenneth French's website. The factors are the three factors of Fama and French (1992) (market, book–to–market, and size). Our analysis is based on monthly data from January 1963 until August 2020, i.e., we have 692 time–series observations.

The following theorem provides the limiting distribution of the historical averages estimator. It's classical and provided for comparison with the three GMM–based estimators in Sections 3.1-3.3.

Theorem 4.1. *Suppose that Assumptions 1 and 2 holds and that $R_t^e - E[R^e]$ forms a martingale difference sequence.*[7] *Then, we have $\sqrt{T}\left(\bar{R}^e - E[R^e]\right) \xrightarrow{d} \mathcal{N}(0, \Sigma_{R^e R^e})$.*

Note that Theorem 4.1 assumes no factor structure. We will, next, provide the asymptotic distributions of expected (excess) return estimators given the linear factor structure implied by asset pricing models. Note that the joint distributions of λ and β are different for each set of moment conditions, which leads to different asymptotic distributions for the risk premiums $\beta\lambda$ as well. Hence, we derive the asymptotic distributions of expected (excess) return estimators for the three set of moment conditions introduced in Sections 3.1, 3.2, and 3.3 separately.

4.1 Precision with General Moment Conditions

The following theorem provides the asymptotic variances of the risk–premium estimators based on the general moment conditions in Section 3.1. This result is valid for both traded

---

[7]We throughout assume that standard assumptions for martingale CLTs hold.

and non-traded factors.

Theorem 4.2. *Suppose that Assumptions 1-3 hold and that $h_t(\alpha, \beta, \lambda)$ forms a martingale difference sequence. Consider the moment conditions (3.1).*

Then, the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is given by

$$\Sigma_{R^e R^e} - \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\left(\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\right). \tag{4.1}$$

The proof is provided in the appendix. Theorem 4.2 provides the asymptotic covariance matrix of the factor–model based risk–premium estimators with the general moment conditions as in Section 3.1. This formula is useful mainly for two reasons. First, it can be used to compute the standard errors of these risk–premium estimates and, accordingly, the related t–statistics and p-values can be obtained. Second, it allows us to study the precision gains for estimating the risk premiums from incorporating information about the factor model.

In case of a one–factor model and a single test asset, the (asymptotic) variances of both the naive risk–premium estimator and the factor–model based risk–premium estimator are the same. When more assets/portfolios are available ($N > 1$), observe that the magnitude of the asymptotic variances of the risk–premium estimators depends on the prices of risk λ, the exposures β, and the idiosyncratic variances $\Sigma_{\varepsilon\varepsilon}$. Note that the difference between the asymptotic covariance matrix of the naive estimator, $\bar{R}^e$, and the factor–based risk–premium estimator is $\left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\left(\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\right)$. The following corollary formalizes this relation.

Corollary 4.1. *Suppose that Assumptions 1-3 hold and that $h_t(\alpha, \beta, \lambda)$ forms a martingale difference sequence. If $\lambda'\Sigma_{FF}^{-1}\lambda < 1$, then the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is at most $\Sigma_{R^e R^e}$.*

Corollary 4.1 shows that there may be precision gains for estimating risk premiums

from the added information about the factor model if $\lambda'\Sigma_{FF}^{-1}\lambda$ is smaller than one. Observe that, although $\lambda'\Sigma_{FF}^{-1}\lambda$ can be larger than one mathematically, in the one–factor case with a traded factor, $\lambda'\Sigma_{FF}^{-1}\lambda$ is the squared Sharpe ratio of that factor. This squared Sharpe ratio is, for stocks and stock portfolios, generally (much) smaller than 1. Moreover, plugging in the estimates from the Fama–French three factor model (based on GMM with moment conditions (3.1)) gives $\lambda'\Sigma_{FF}^{-1}\lambda = 0.034$. Note that the smaller the value for $\lambda'\Sigma_{FF}^{-1}\lambda$, the larger the efficiency gains from imposing a factor model. It may be surprising that, theoretically, the GMM estimator can behave worse than historical averages. Note, however, that we are, in line with the literature, not using optimal GMM weights.

We calculate the (asymptotic) variances of the factor–model based risk–premium estimates for all 25 FF portfolios by plugging the parameter estimates into (4.1). Table 1 presents the results. Comparing the asymptotic variances of the factor–model based risk–premium estimators to those of the naive estimators, we see that the factor–model based risk–premium estimators are more precise than the naive estimators for all 25 Fama–French portfolios. In particular, using the 3–factor model in estimating risk premiums of 25 FF portfolios leads to considerable gains in variances of up to 23%. Note that a 23% gain in variances means that the same (statistical) precision can be obtained with 23% less observations.

4.2   Precision with Moment Conditions for Traded Factors

When the risk factors are traded, meaning that the factor itself is an excess return, additional restrictions on the prices of risk can be incorporated into the estimation. With the availability of such information, one could again expect efficiency gains in estimating both the prices of risk and the expected (excess) returns. In this section, we consider this case and the following theorem gives the asymptotic variances of the expected (excess) return estimators with the moment conditions (3.2) for the case all factors are traded.

Theorem 4.3. *Suppose that all factors are excess returns. In addition, suppose that Assumptions 1-3 hold and that $h_t(\alpha, \beta, \lambda)$ forms a martingale difference sequence. Consider the moment conditions (3.2).*

Then, the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is given by

$$\Sigma_{R^e R^e} - \left(1 - \lambda' \Sigma_{FF}^{-1} \lambda\right) \Sigma_{\varepsilon\varepsilon}. \tag{4.2}$$

Theorem 4.3 allows us to study the efficiency gains for estimating risk premiums from a model where the factors are traded compared to historical averages. Comparing the asymptotic covariance matrix of the factor–based risk–premium estimators from GMM based on (3.2) to the one of the naive estimator, we observe that the difference is given by $\left(1 - \lambda' \Sigma_{FF}^{-1} \lambda\right) \Sigma_{\varepsilon\varepsilon}$. As before, the factor $1 - \lambda' \Sigma_{FF}^{-1} \lambda$ is empirically generally found to be positive implying an efficiency gain. Moreover, observe that asymptotic covariance matrix of the risk–premium estimator based on GMM with (3.2) can be different from the ones of the risk–premium estimator based on GMM with (3.1). This indicates that there may be efficiency gains even within the GMM framework from that the information that the factors are traded. The following corollary formalizes these issues.

Corollary 4.2. *Suppose that all factors are traded. In addition, suppose that Assumptions 1-3 hold and that $h_t(\alpha, \beta, \lambda)$ forms a martingale difference sequence. Consider the GMM estimator based on the moment conditons (3.2). Then, we have the following.*

1. *If $\lambda' \Sigma_{FF}^{-1} \lambda < 1$, then the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is at most $\Sigma_{R^e R^e}$.*

2. *The limiting variance of this expected (excess) return estimator is at most the limiting variance of the estimator based on the moment conditions (3.1).*

Plugging in the parameter estimates from the analysis of the Fama–French model gives $\lambda' \Sigma_{FF}^{-1} \lambda < 1 = 0.033$. Comparing the variances of the risk–premium estimates based on

GMM with (3.2) to those of the naive estimators in Table 1, we see that the risk–premium estimates based on GMM with (3.2) have smaller asymptotic variances than the naive estimators. In particular, the magnitude of efficiency gains goes up to 24%. Moreover, consistent with Corollary 4.2, the asymptotic variances of the risk–premium estimates based on GMM with (3.1) exceed those of the risk–premium estimators based on GMM with (3.2). Overall, these precision gains from estimating risk premiums based on factor models stem from two sources. First, the linear relation implied by asset pricing models is valuable information in the estimation of risk premiums. Second, when the factors are traded, the additional information that the prices of risk equal the expected factor returns leads to higher precision of risk–premium estimates.

4.3 Precision with Moment Conditions Using Factor–Mimicking Portfolios

One may hope that replacing factors by factor–mimicking portfolios may also bring efficiency gains compared to (4.1) since the additional restriction on the price of the factor risk can be incorporated into the estimation. In this section, we derive the asymptotic variances of expected (excess) return estimators obtained with factor–mimicking portfolios.

Theorem 4.4. *Suppose that Assumptions 1-3 hold and that $h_t(\alpha^m, \beta^m, \Phi_0, \Phi, \lambda^m)$ forms a martingale difference sequence. Consider the GMM estimator based on the moment conditions (3.4).*

Then, the limiting variance of the expected (excess) return estimator, $\hat{\beta}^m \hat{\lambda}^m$, is given by

$$
\Sigma_{R^e R^e} - \left(\mu'_{R^e} \left\{ \Sigma_{R^e R^e}^{-1} - \Sigma_{R^e R^e}^{-1} \beta \left(\beta' \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \beta' \Sigma_{R^e R^e}^{-1} \right\} \mu_{R^e} \right) \tag{4.3}
$$
$$
\times \left(\Sigma_{R^e R^e} - \beta \left(\beta' \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \Sigma_{FF}^{-1} \left(\beta \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \beta' \right)
$$
$$
- \left(1 - \mu'_{R^e} \Sigma_{R^e R^e}^{-1} \mu_{R^e} \right) \left(\Sigma_{R^e R^e} - \beta \left(\beta' \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \beta' \right),
$$

with $\mu_{R^e} = \mathrm{E}\left[R_t^e\right]$.

Theorem 4.4 enables us to study the efficiency gains in risk premiums using factor–mimicking portfolios. Observe that the difference between the asymptotic covariance matrices of the naive estimator and the factor–model based GMM risk–premium estimator with (3.4) is given by

$$\mu_{R^e}' \left\{ \Sigma_{R^e R^e}^{-1} - \Sigma_{R^e R^e}^{-1} \beta \left(\beta' \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \beta' \Sigma_{R^e R^e}^{-1} \right\} \mu_{R^e} \tag{4.4}$$
$$\times \left(\Sigma_{R^e R^e} - \beta \left(\beta' \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \Sigma_{FF}^{-1} \left(\beta \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \beta' \right)$$
$$+ \left(1 - \mu_{R^e}' \Sigma_{R^e R^e}^{-1} \mu_{R^e} \right) \left(\Sigma_{R^e R^e} - \beta \left(\beta' \Sigma_{R^e R^e}^{-1} \beta \right)^{-1} \beta' \right).$$

Efficiency gains with respect to the historical averages estimator are dependent on (4.4) being positive semi–definite or not. The results from our empirical analysis with FF-3 factor model illustrates that there are considerable efficiency gains over the naive estimation for all 25 Fama–French 25 portfolios (see Table 1). In particular, estimating risk premiums with GMM (3.4) leads to, of up to 22%, smaller variances than estimating them with the naive estimator. Moreover, we find that estimating risk premiums by making use of the mimicking portfolios leads to small efficiency losses over the estimation based on the general case, i.e, GMM based on (3.1) for all assets.

Note that one important difference between Theorem 4.2 and Theorem 4.4 may potentially come from the estimation of the mimicking portfolio weights. The estimation of the weights of the factor–mimicking portfolio potentially leads to different (intuitively higher) asymptotic variances for the betas of the mimicking factors as well as for the mimicking factor prices of risk, and the risk premiums, which are essentially a multiplication of β^m and λ^m. Such issue is similar to errors–in–variables type of corrections in two step Fama–Macbeth estimation, i.e., the Shanken (1992) correction in asymptotic variances for generated regressors. We recall here that GMM standard errors automatically account for

such effects as the system of moment conditions is solved simultaneously. In particular, in our setting with moments conditions (3.4), GMM treats the moments producing Φ simultaneously with the moments generating β^m and λ^m. Hence, the long–run covariance matrix captures the effects of estimation of Φ on the standard errors of the β^m and λ^m, hence the risk premiums.

If we consider the Fama–French three factor model with the 25 FF–portfolios, we can also intuitively gain insights about the difference between the inferences about risk premiums based on GMM with the two sets of moment conditions (3.2) and (3.4). In fact, since the factors are traded, meaning that they are excess returns themselves, we can estimate the risk premiums via the second set of moment conditions (3.2). Moreover, we can also estimate such system via the third set of moment conditions (3.4), which has the additional burden of estimating the coefficients for the construction of the mimicking portfolio. Accordingly, GMM estimation via the second set and the third set of moment conditions may lead to different precisions for the risk–premium estimates. The last column in Table 1 documents the efficiency comparisons in estimating risk premiums of 25 FF portfolios employing factor mimicking portfolios over risk premium estimation with moment conditions (3.2). Efficiency losses are present for all 25 Fama–French portfolios, meaning that risk–premium estimates employing factor mimicking portfolios, i.e., based on (3.4), are less precise than risk–premium estimates based on (3.2).

5   Inference about Risk Premiums when the β's are small

Several studies document inference issues regarding the prices of risk when the factors are weakly correlated with the asset returns (see, e.g., Kleibergen (2009), Gospodinov et al. (2014), Bryzgalova (2015), Burnside (2015), Kleibergen and Zhan (2015), and Gospodinov

et al. (2017, 2019)). When β's are close to zero and/or when the β matrix is almost of reduced rank, the confidence intervals of the prices of risk estimates, $\hat{\lambda}$, are erroneous, which leads to unreliable statistical inference. The effects may be severe in empirical research, as the confidence intervals of the risk price estimates may even be unbounded, as documented in Kleibergen (2009) for the case of the conditional consumption CAPM of Lettau and Ludvigson (2001).[8] Accordingly, Kleibergen (2009) provides identification-robust statistics and confidence sets for the risk price estimates when the β's are small.

However, once the interest is in estimating risk premiums on *individual assets* or portfolios, a natural question is whether similar issues exist in the presence of weakly correlated factors. Recall that the focus of this study is on risk premiums, the product of β and λ, rather than on λ alone.

In the rest of this section, we will focus on the specification where β has small but non-zero values. Following the literature on weak instruments (see, e.g., Staiger and Stock (1997)), and Kleibergen (2009), we consider a sequence of β's getting smaller as the sample size increases. [9]

Corollary 5.1. *Suppose Assumption 2 and Assumption 3 hold and consider the small β case where $\beta = \frac{1}{\sqrt{T}}B$ for a fixed full rank $N \times K$ matrix B. Then, the behavior of the risk premium estimators in large samples can be characterized as follows.*

1. Consider the moment conditions (3.1), then

$$\left(\hat{\beta}\hat{\lambda} - \beta\lambda\right) \xrightarrow{d} \left(B + Y\Sigma_{FF}^{-1}\right)\left[\left(B + Y\Sigma_{FF}^{-1}\right)'\Sigma_{\epsilon\epsilon}^{-1}\left(B + Y\Sigma_{FF}^{-1}\right)\right]^{-1}$$
$$\times \left(B + Y\Sigma_{FF}^{-1}\right)'\Sigma_{\epsilon\epsilon}^{-1}[B\lambda + U] - \beta\lambda, \tag{5.5}$$

where $U \sim N(0, \Sigma_{\epsilon\epsilon})$ and $vec(Y) \sim N(0, \Sigma_{FF} \otimes \Sigma_{\epsilon\epsilon})$ are independently distributed. [10]

---

[8]Kleibergen (2009) documents that %95 percent confidence bounds of the prices of risk on the scaled consumption growth coincides with the whole real line.

[9]The results in this section have been revised in the current version of the paper.

[10]The limiting variance of the risk price estimator in the presence of weak–factors is unbounded for Fama

2. Consider moment conditions (3.2), then

$$\left(\hat{\beta}\hat{\lambda} - \beta\lambda\right) \xrightarrow{d} N\left(0, \lambda'\Sigma_{FF}^{-1}\lambda\Sigma_{\epsilon\epsilon}\right). \tag{5.6}$$

The proof is provided in the appendix. Corollary 5.1 documents several important findings of our analysis regarding the issue of small but non-zero factor loadings. First, if the parameters of the linear factor model are estimated with GMM based on the moment conditions (3.1), the limiting behaviour of the risk premium estimator is non-standard and differs from normality. This result is in line with the literature documenting unreliable statistical inference about the prices of risk based on the Fama-Macbeth and GLS two-pass estimation and their unbounded confidence sets.

Corollary 5.1-2 shed light on the issue of small β's when all factors in the linear factor model of interest are traded. In this case, if one estimates the parameters of the model with GMM based on (3.2), then the limiting variances of the risk–premium estimators are not affected by the β having a small value or not. Observe that, the limiting variances in Corollary 5.1-2 is equal to (4.2) under $\beta = 0$.

6 Monte Carlo Simulations

In this section, we conduct a Monte Carlo experiment to study the finite-sample properties of the various-factor model based estimators of expected returns. We consider correctly specified models, and models with spurious and weak factors.

We calibrate the parameters of the true models by using the monthly returns on 25 Fama-French portfolios from January 1963 until August 2020 (available on Kenneth French's website). We use the nominal 1–month Treasury bill rate as a proxy for risk–free rate. In our simulations we consider three different time-series sample sizes: are $T = 240$,

---

and MacBeth (1973) and generalized least squares (GLS) regression methods (see, e.g. Kleibergen (2009), and Kleibergen and Zhan (2015).

$T = 480$, and $T = 960$. These choices cover the sample sizes typically encountered in empirical research with quarterly and monthly data. Varying T is useful for understanding the small-sample behaviour of the estimators and we use $T = 960$ to assess the quality of our asymptotic approximations. All results are based on $20,000$ replications.

6.1 Under the null that asset pricing restrictions hold

In this subsection, we assume that the asset pricing restrictions are true, i.e. Eqn. (2.2) holds for returns of $N = 25$ test assets. We present results for a one-factor model setting.[11] In this case, the model is calibrated to mimic the CAPM, estimated over the sample period from January 1963 until August 2020. All factors and the error terms are generated from multivariate normal distribution.[12]

Figure 1 reports the simulation results for the standard errors of risk premium estimates on 25 test assets. Panel (a), Panel (b), and Panel (c) show the results for $T = 240, T = 480$, and $T = 960$, respectively. Top exhibits provide the root–mean–square errors (RMSE) for risk premium estimators across 25 test assets (x-axis values). The middle exhibits illustrate the simulation averages of the estimated asymptotic standard errors (AEST). The lower exhibits present the percentage errors (PE) of AESTs as compared to the RMSEs.

There are several important observations to be made about the standard errors. First, AESTs for all 25 assets are very close to their corresponding RMSEs for all three factor-model based estimators of expected returns (GMM–Gen, GMM–Tr, GMM–Mim). In particular, absolute percentage errors across all assets are less than 0.5% for T=960, and they are less than 1% for T=480, and $T = 240$. Second, the RMSEs of all three factor–model based estimators of expected returns (GMM–Gen, GMM–Tr, GMM–Mim) are smaller than the RMSEs of the historical averages estimator, which holds true for all 25 individual assets

---

[11]Results for three-factor model is available upon request.
[12]In the three-factor model setting, the parameters of the model are calibrated to mimic the Fama–French three-factor model.

and for all sample sizes $T = 240$, $T = 480$, and $T = 960$. We observe a similar pattern for the estimated asymptotic standard errors. Those patterns are manifestations of the efficiency gains from using the factor–model based risk premium estimates.

These simulation results show that the limiting variances for factor–model based estimators of risk premiums that we provide in our paper provide accurate approximations to the finite sample behavior of standard errors of the estimates. Furthermore, comparing the RMSEs (and AESTs) of factor–model–based estimators of risk premiums to RMSEs (and AESTs) of the historical averages, we observe that the efficiency gains from using the factor–model based estimators of risk–premiums are present for all individual assets.

6.2 Presence of weak factors

In this section, we consider the case for the presence of a weak factor. The parameters are calibrated by using the monthly data from from January 1963 until August 2020 on 25 Fama French (1992) portfolios sorted by size and book–to–market and the corresponding market factor, $R_{m,t}^e$. The data generating process is given by

$$R_t^e = \alpha + \frac{B}{\sqrt{T}} R_{m,t}^e + \varepsilon_t, \quad t = 1, 2, \ldots, T, \tag{6.7}$$

with normally distributed $R_{m,t}^e$ and multivariate normal ε_t under the null of $\mathrm{E}\left[R_t^e\right] = \frac{B}{\sqrt{T}}\lambda$. We use 20000 replications for each estimator considered.

Figure 2 present the simulation results for the standard errors of the risk premium estimates in the presence of weak factor. Panel (a), Panel (b), and Panel (c) show the results for $T = 240$, $T = 480$, and $T = 960$, respectively. For all three factor–model based risk premium estimators, the estimated asymptotic standard errors are close to their corresponding RMSEs, with absolute percentage errors being lower than 3% for all assets for $T = 240, 480, 960$. Moreover, comparing the RMSEs (and the standard errors of the historical averages) to the RMSEs (and the standard errors of the factor–model–based

risk–premium estimators), we observe that there are substantial efficiency gains from using the factor–model–based risk–premium estimates. These efficiency gains are present for all individual assets.

To sum up, if one is interested in making econometric inference about the prices of risk λ, small but non-zero β's may may cause spurious inference. However, once the interest is in estimating risk premiums, i.e., expected (excess) returns on individual assets or portfolios, the limiting variances based on GMM with (3.1), (3.2) or (3.4) provide good approximations to the finite–sample variances of the factor model based estimators when the β's are weak.

7 Portfolio Choice with Parameter Uncertainty

In the previous sections, we provided an asymptotic analysis of the three factor–model based risk–premium estimators and analyzed the efficiency gains with respect to naive historical averages. In this section, we analyze the economic significance of these gains in a portfolio allocation problem à la Markowitz (1952).

The implementation of the mean–variance framework requires the estimation of first two moments of the asset returns. Although, in this setting, the optimal portfolios are supposed to achieve the best performance, in practice, the estimation error in the estimated moments leads to large deterioration of the out–of–sample performance of the optimal portfolios (see, e.g., DeMiguel et al. (2009)). In this section, we analyze the out–of–sample performances of the optimal portfolios based on factor–based risk–premium estimates as well as the historical averages, $1/N$ portfolio and global minimum variance portfolio in a simulation analysis.

We consider the following well known optimization problem: Suppose a risk–free asset exists and w is the vector of relative portfolio allocations of wealth to N risky assets. The investor has preferences that are characterized by the expected return and variance of her selected portfolio. The investor maximizes her expected utility, by choosing the vector of

portfolio weights w such that

$$\mathrm{E}\left[U\right] = w'\mu^e - \frac{\gamma}{2}w'\Sigma_{RR}w, \qquad (7.8)$$

is maximized, where γ measures the investor's risk aversion level and μ^e and Σ_{RR} denote the expected excess returns on the assets and covariance matrix of returns. The solution to this maximization problem is well known and given by

$$w_{opt} = \frac{1}{\gamma}\Sigma_{RR}^{-1}\mu^e. \qquad (7.9)$$

In the optimization problem above, since the true risk–premium vector, μ^e, and the true covariance matrix of asset returns, Σ_{RR}, are unknown, in empirical work, one needs to estimate them.

We consider four portfolios constructed with different risk–premium estimators: the optimal portfolio constructed with historical averages, and the optimal portfolios constructed with the three factor model–based GMM risk–premium estimates with moment conditions (3.1), (3.2) and (3.4), respectively. Note that the covariance matrix is estimated using the traditional sample counterpart.[13] We also consider the global minimum variance (GMV thereafter) portfolio to which we compare the performance of the portfolios based on the risk–premium estimates. Note that the implementation of this portfolio only requires estimation of the covariance matrix, for which we again use the sample counterpart, and completely ignores the estimation of expected returns.[14] Morover, we analyze the performance of the $1/N$ portfolio.

We compare performances of the portfolios by using their out-of-sample Sharpe Ratios.[15]

---

[13]$1/(T-1)\sum_1^T (R_t - \bar{R}_t)(R_t - \bar{R}_t)'$, where $\bar{R}_t$ is the sample average of returns.

[14]This portfolio is obtained by minimizing the portfolio variance with respect to the weights with the only constraint that weights sum up to 1 and the N–vector of portfolio weights is given by $w_{gmv} = \frac{\Sigma_{RR}\iota_N}{\iota_N'\Sigma_{RR}\iota_N}$

[15]See Peñaranda and Sentana (2011) for an analysis examining the improvements in the estimation of in–sample mean variance frontiers based on asset pricing model restrictions, tangency or spanning constraints.

We set the initial window length at 120 data points, corresponding to 10 years of monthly data. The parameters for the return–generating process are calibrated to mimic the Fama-French 3–factor model with 25 FF portfolios. All factors and the error terms are generated from multivariate normal distribution. We simulate independent sets of $Z = 20,000$ return samples.

Table 2 provides the simulation results for out–of–sample Sharpe ratios of different portfolios. For each portfolio, we present the average estimate over simulations, $\overline{SR}$ (first line), the bias as the percentage of the population Sharpe ratios, $(\overline{SR} - SR)/SR$ (second line) and the root–mean–square error (RMSE), the square root of $\sum_{s=1}^{Z}(\hat{SR}_s - SR)/Z$, (third line) , where $Z = 20,000$.

In order to isolate the effect of the error in risk–premium estimates, we present our results with both true and estimated Σ_{RR}. First, note that the true Sharpe ratio of the optimal portfolio is superior to the portfolios based on estimated risk–premiums or covariance matrix of asset returns. Comparing the average Sharpe ratio of the optimal portfolio based on historical averages to the true Sharpe ratio of optimal portfolio for enlarging samples, we see that the bias is strikingly large and negative with -42% and -43%, depending on whether the covariance matrix of asset returns is the true one or the estimated one. However, using the factor–models to estimate the risk–premiums reduces the bias in Sharpe ratios substantially to about -12% when the true coviariance matrix is used, and to about -9% when the covariance matrix is estimated. In particular, with GMM–Gen estimates, average Sharpe ratio of the optimal portfolio is 0.157 in case of true covariance matrix (with an improvement of about 50% over the average Sharpe ratios with the historical averages) and 0.161 in case of an estimated covariance matrix (with an improvement of about 60% over the average Sharpe ratios with the historical averages). Among the optimal portfolios constructed with factor–model based risk–premium estimates, the portfolio based on GMM–Tr estimates perform the best with 0.162. However, the differences in biases

26

are minimal for all optimal portfolios constructed with factor–model based risk–premium estimators.

Next, we analyse the RMSEs of the various portfolios. Out–of–sample Sharpe ratio of the optimal portfolios based on historical averages is extremely volatile across simulations. That is, for the case of enlarging samples, it has a RMSE of 0.09 (given the average estimate 0.10) if the covariance matrix is estimated. However, using factor-model based risk–premium estimators decreases the RMSEs substantially. The differences in RMSEs are minor for all optimal portfolios constructed with factor–model based risk–premium estimators.

Comparing the average Sharpe ratios of the optimal portfolios the factor model–based risk–premium estimates with GMV and 1/N, we see that optimal portfolios based on the naive estimator performs worse than both the 1/N strategy and the GMV portfolio. Moreover, both GMV and 1/N have substantially lower RMSEs. This result is consistent with the findings in the literature that GMV portfolio as well as 1/N strategy has better out–of–sample performance than the optimal portfolios based on sample moments.[16] However, the average Sharpe ratios for all optimal portfolios based on factor model–based risk–premium estimates are larger than both the GMV and 1/N porfolios. Moreover, their out of Sharpe ratios across simulations are almost as stable as the GMV portfolio as well as the 1/N strategy.

Overall, using the factor–model based risk–premium estimators improves the performance of optimal portfolios substantially over the optimal portfolios based on the plug in estimates of historical averages in terms of both bias and RMSEs. Moreover, in contrast to the optimal portfolios with historical averages, these portfolios perform considerably better than the global minimum variance portfolio. In practice, while there is little hope in knowing the universally true factor model, we know that particular factor models work

---

[16]See, e.g., DeMiguel et al. (2009), Jagannathan and Ma (2003), and Jorion (1985, 1991)).

well empirically in explaining expected returns of particular test assets (e.g. the Fama-French three factor model explains the returns of 25 size/book-to-market portfolios). Our results suggests that using factor-model-based expected return estimates likely improves the performance of optimal portfolios over the portfolios based on the historical averages estimators of expected returns, minimum variance portfolio, and $1/N$ portfolio.

8   Conclusions

One traditional technique in the literature is to use average historical returns as estimates of expected excess returns, that is risk premiums, on individual assets or portfolios. These estimators are usually noisy. This translates into the need for very long, in practice, mostly infeasible, samples of data in order to obtain some precision. However, the finance literature provides a wide variety of risk–return models which imply a linear relationship between the expected excess returns and their exposures.

In this paper, we show that, when correctly specified, such parametric specifications on the functional form of risk premiums lead to significant inference gains for estimating expected (excess) returns. In the standard Fama–French three factor model using MKT, SMB, HML as factors with 25 FF portfolios, the efficiency gains are sizeable and go up to about 25% for individual portfolios. For applications, this translates into the benefit of using only about 75% of the data with factor–model based risk–premium estimates to obtain the same precision as with the historical averages estimator. Moreover, we show that the presence of weakly identified factors, the confidence bounds of factor model based risk–premium estimators are not affected in case of traded factors. Finally, we show that out–of–sample performance of optimal portfolios significantly improves if factor–model based estimates of risk premium are used in portfolio weights instead of the classical historical averages.

References

Balduzzi, P., and Robotti, C. (2008). Mimicking portfolios, economic risk premia, and tests of multi-beta models. *Journal of Business & Economic Statistics*, *26*(3), 354–368.

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

Boyd, S., and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, *7*(3), 265–296.

Bryzgalova, S. (2015). Spurious factors in linear asset pricing models. *Working Paper*.

Burnside, C. (2015). Identification and Inference in Linear Stochastic Discount Factor Models with Excess Returns. *Journal of Financial Econometrics*, *14*(2), 295–330.

Chopra, V. K., and Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, *19*(2), 6–11.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, *22*(5), 1915–1953.

Fama, E. F. (1998). Determining the number of priced state variables in the icapm. *The Journal of Financial and Quantitative Analysis*, *33*(2), 217–231.

Fama, E. F., and French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, *47*(2), 427–465.

Fama, E. F., and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

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

Frost, P. A., and Savarino, J. E. (1988). For better performance. *The Journal of Portfolio Management*, *15*(1), 29–34.

Gagliardini, P., Ossola, E., and Scaillet, O. (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica*, *84*(3), 985–1046.

Giglio, S., Xiu, D., and Zhang, D. (2021). Test assets and weak factors. *Chicago Booth Research Paper*, *21*(4).

Gospodinov, N., Kan, R., and Robotti, C. (2014). Misspecification-Robust Inference in Linear Asset-Pricing Models with Irrelevant Risk Factors. *The Review of Financial Studies*, *27*(7), 2139–2170.

Gospodinov, N., Kan, R., and Robotti, C. (2017). Spurious inference in reduced-rank asset-pricing models. *Econometrica*, *85*(5), 1613–1628.

Gospodinov, N., Kan, R., and Robotti, C. (2019). Too good to be true? fallacies in evaluating risk factor models. *Journal of Financial Economics*, *132*(2), 451–471.

Hall, A. (2005). *Generalized Method of Moments*. Advanced texts in econometrics. Oxford University Press.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, (pp. 1029–1054).

Hou, K., and Kimmel, R. (2006). On the estimation of risk premia in linear factor models. *Unpublished working paper. Ohio State University*.

Huberman, G., Kandel, S., and Stambaugh, R. F. (1987). Mimicking portfolios and exact arbitrage pricing. *The Journal of Finance*, *42*(1), 1–9.

Jagannathan, R., and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance*, *58*(4), 1651–1683.

Jagannathan, R., and Wang, Z. (1998). An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *The Journal of Finance*, *53*(4), 1285–1309.

Jobson, J. D., and Korkie, B. (1980). Estimation for markowitz efficient portfolios. *Journal of the American Statistical Association*, *75*(371), 544–554.

Jorion, P. (1985). International portfolio diversification with estimation risk. *The Journal of Business*, *58*(3), 259–278.

Jorion, P. (1991). Bayesian and capm estimators of the means: Implications for portfolio selection. *Journal of Banking & Finance*, *15*(3), 717–727.

Kan, R., and Robotti, C. (2011). On the estimation of asset pricing models using univariate betas. *Economics Letters*, *110*(2), 117–121.

Kan, R., Robotti, C., and Shanken, J. (2013). Pricing model performance and the two-pass cross-sectional regression methodology. *The Journal of Finance*, *68*(6), 2617–2649.

Kan, R., and Zhang, C. (1999). Two-pass tests of asset pricing models with useless factors. *The Journal of Finance*, *54*(1), 203–235.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*(3), 501–524.

Kleibergen, F. (2009). Tests of risk premia in linear factor models. *Journal of Econometrics*, *149*(2), 149–173.

Kleibergen, F., and Zhan, Z. (2015). Unexplained factors and their effects on second pass r-squared's. *Journal of Econometrics*, *189*(1), 101–116.

Lamont, O. A. (2001). Economic tracking portfolios. *Journal of Econometrics*, *105*(1), 161–184.

Ledoit, O., and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, *10*(5), 603–621.

Lettau, M., and Ludvigson, S. (2001). Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, *109*(6), 1238–1287.

Lewellen, J., Nagel, S., and Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, *96*(2), 175–194.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, (pp. 867–887).

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, *8*(4), 323–361.

Michaud, R. O. (1989). The markowitz optimization enigma: Is 'optimized'optimal? *Financial analysts journal*, *45*(1), 31–42.

Pástor, L. (2000). Portfolio selection and asset pricing models. *The Journal of Finance*, *55*(1), 179–223.

Pástor, L., and Stambaugh, R. F. (1999). Costs of equity capital and model mispricing. *The Journal of Finance*, *54*(1), 67–121.

Peñaranda, F., and Sentana, E. (2011). Inferences about portfolio and stochastic discount factor mean variance frontiers. *Working paper*.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, *13*(3), 341–60.

Shanken, J. (1992). On the estimation of beta-pricing models. *The review of financial studies*, *5*(1), 1–33.

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

Staiger, D., and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica: journal of the Econometric Society*, (pp. 557–586).

A Proofs

In the rest of the paper, the covariance matrix of the factor–mimicking portfolios is denoted by $\Sigma_{F^m F^m}$.

A.1 Equivalence of factor pricing using mimicking portfolios

Proof of Theorem 2.1. Define M^m as the projection of M onto the augmented span of excess returns,

$$M^m = \mathbb{P}(M|1, R^e) \tag{A.1}$$

so that

$$= \mathrm{E}\left[M^m\right], \tag{A.2}$$

$$\mathrm{Cov}\left[M, R^e\right] = \mathrm{Cov}\left[M^m, R^e\right]. \tag{A.3}$$

Thus, we have

$$\begin{aligned}
\beta\lambda &= \mathrm{Cov}\left[R^e, F'\right]\Sigma_{FF}^{-1}\left(-\frac{1}{\mathrm{E}\left[M\right]}\Sigma_{FF}b\right) \tag{A.4}\\
&= -\frac{1}{\mathrm{E}\left[M\right]}\mathrm{Cov}\left[R^e, F'\right]b \\
&= -\frac{1}{\mathrm{E}\left[M^m\right]}\mathrm{Cov}\left[R^e, F^{m\prime}\right]b \\
&= -\frac{1}{\mathrm{E}\left[M^m\right]}\mathrm{Cov}\left[R^e, F^{m\prime}\right]\Sigma_{F^m F^m}^{-1}\Sigma_{F^m F^m}b \\
&= \beta^m\lambda^m,
\end{aligned}$$

which completes the proof. $\qquad\square$

A.2 Precision of Parameter Estimators Given a Factor Model

This section provides the proofs for asymptotic properties of the parameter estimators under the specified linear factor model. The lemma A.1 below illustrates the asymptotic distribution of the GMM estimators with a given set of moment conditions provided that a pre–specified matrix A, that essentially determines the weigths of the overidentifying moments, is introduced. Thereafter, these results will be used to calculate the variance covariance matrix for the moment conditions (3.1), (3.2) and (3.4), respectively.

Under appropriate regularity conditions, see, e.g., Hall (2005), Chapter 3.4, we have the following result.

Lemma A.1. *Let $\theta \in \mathbb{R}^p$ be a vector of parameters and the moment conditions are given by $E[h_t(\theta)] = 0$ where $h_t(\theta) \in \mathbb{R}^q$, stationary and ergodic process with finite fourth moment. Given a prespecified matrix $A \in \mathbb{R}^{p \times q}$, its consistent estimator $\hat{A}$ and $\hat{A}\frac{1}{T}\sum_{t=1}^{T} h_t(\hat{\theta}) = 0$,*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}\left(0, [AJ]^{-1}ASA'[J'A']^{-1}\right), \tag{A.5}$$

where,

$$J = E\left[\frac{\partial h_t(\theta)}{\partial \theta'}\right], \tag{A.6}$$

$$S = E[h_t(\theta)h_t(\theta)']. \tag{A.7}$$

The above lemma presents the asymptotic distribution of the parameters in a general GMM context. In the subsequent lemmas, we provide the limiting distributions for the parameter estimators based on the moment conditions (3.1), (3.2) and (3.4), respectively.

Lemma A.2. *Under Assumptions 1-3 and the moment conditions (3.1) with parameter vector $\theta = (\alpha', vec\,(\beta)', \lambda')'$, we have*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, V), \tag{A.8}$$

with

$$V = \begin{bmatrix} \begin{bmatrix} 1 + \mu_F'\Sigma_{FF}^{-1}\mu_F & -\mu_F'\Sigma_{FF}^{-1} \\ -\Sigma_{FF}^{-1}\mu_F & \Sigma_{FF}^{-1} \end{bmatrix} \otimes \Sigma_{\varepsilon\varepsilon} & V_c \\ \\ V_c' & (1 + \lambda'\Sigma_{FF}^{-1}\lambda)(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1} + \Sigma_{FF} \end{bmatrix}$$

where $\mu_F = E[F_t]$ and $V_c = \begin{bmatrix} 1 + \mu'_F \Sigma_{FF}^{-1} \lambda \\ -\Sigma_{FF}^{-1} \lambda \end{bmatrix} \otimes \beta(\beta' \Sigma_{\varepsilon\varepsilon}^{-1} \beta)^{-1}.$

Proof. The proof follows from plugging the appropriate matrices for the moment conditions provided in Section 3.1 into the variance covariance formula in (A.5) and performing the matrix multiplications. Below, we provide the limiting variance covariance matrix (S) and the Jacobian (J) for this specific set of moment conditions,

$$S = \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & \mu'_F \otimes \Sigma_{\varepsilon\varepsilon} & \Sigma_{\varepsilon\varepsilon} \\ \mu_F \otimes \Sigma_{\varepsilon\varepsilon} & [\Sigma_{FF} + \mu_F \mu'_F] \otimes \Sigma_{\varepsilon\varepsilon} & \mu_F \otimes \Sigma_{\varepsilon\varepsilon} \\ \Sigma_{\varepsilon\varepsilon} & \mu'_F \otimes \Sigma_{\varepsilon\varepsilon} & \beta \Sigma_{FF} \beta' + \Sigma_{\varepsilon\varepsilon} \end{bmatrix}.$$

$$J(\theta) = \mathrm{E}\left[\frac{\partial h_t(\theta)}{\partial \theta'}\right] = \begin{bmatrix} -\begin{bmatrix} 1 & \mu'_F \\ \mu_F & \Sigma_{FF} + \mu_F \mu'_F \end{bmatrix} \otimes I_N & 0_{N(K+1)\times K} \\ \begin{bmatrix} 0_{N\times N} & -\lambda' \otimes I_N \end{bmatrix} & -\beta \end{bmatrix}.$$

Furthermore

$$A = \begin{bmatrix} I_{N(K+1)} & 0_{N(K+1)\times N} \\ 0_{K\times N(K+1)} & \beta' \Sigma_{\varepsilon\varepsilon}^{-1} \end{bmatrix}.$$

so that the limiting variance of GMM estimator for θ is obtained by performing the matrix multiplications $[AJ]^{-1} ASA'[J'A']^{-1}$. $\qquad\square$

Lemma A.3. *Suppose that all factors are traded. Then, under Assumptions 1-3 and the moment conditions (3.2) with parameter vector $\theta = (\alpha', vec(\beta)', \lambda')'$, we have*

$$\sqrt{T}(\hat\theta - \theta) \xrightarrow{d} \mathcal{N}(0, V), \tag{A.9}$$

with

$$V = \left[\begin{array}{cc} \left[\begin{array}{cc} 1 + \mu_F' \Sigma_{FF}^{-1} \mu_F & -\mu_F' \Sigma_{FF}^{-1} \\ -\Sigma_{FF}^{-1} \mu_F & \Sigma_{FF}^{-1} \end{array} \right] \otimes \Sigma_{\varepsilon\varepsilon} & 0_{N(K+1)\times K} \\ 0_{K \times N(K+1)} & \Sigma_{FF} \end{array} \right].$$

Proof. The proof follows from plugging the appropriate matrices for the moment conditions (3.2) into the variance covariance formula in (A.5) and performing the matrix multiplications. Below, we provide the limiting variance covariance matrix (S), Jacobian (J) for this specific set of moment conditions, In this case,

$$S = \left[\begin{array}{ccc} \Sigma_{\varepsilon\varepsilon} & \mu_F' \otimes \Sigma_{\varepsilon\varepsilon} & 0_{N \times K} \\ \mu_F \otimes \Sigma_{\varepsilon\varepsilon} & [\Sigma_{FF} + \mu_F \mu_F'] \otimes \Sigma_{\varepsilon\varepsilon} & 0_{NK \times K} \\ 0_{K \times N} & 0_{K \times NK} & \Sigma_{FF} \end{array} \right],$$

and

$$J(\theta) = \left[\begin{array}{cc} -\left[\begin{array}{cc} 1 & \mu_F' \\ \mu_F & \Sigma_{FF} + \mu_F \mu_F' \end{array} \right] \otimes I_N & 0_{N(K+1)\times K} \\ 0_{K \times N(K+1)} & I_K \end{array} \right].$$

Thus, the limiting variance of the GMM estimator for θ is obtained by performing the matrix multiplications $J^{-1} S [J']^{-1}$ since $A = I_{N(K+1)+K}$. □

The next lemma provides the asymptotic properties of the GMM estimatior with factor–mimicking portfolios.

Lemma A.4. *Given that Assumptions 1-3 are satisfied and that (2.3) hold, then under the moment conditions (3.4), for $\theta = (vec(\beta^m)', \lambda^{m'})'$, we have*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, V), \tag{A.10}$$

with

$$
V = \left[\begin{array}{cc} \Sigma^{-1}_{F^m F^m} \Phi \Sigma_{R^e R^e} \Phi' \Sigma^{-1}_{F^m F^m} \otimes \beta^m \Sigma_{uu} \beta^{m\prime} + \Sigma^{-1}_{F^m F^m} \otimes \Sigma_{\varepsilon^m \varepsilon^m} & -\Sigma^{-1}_{F^m F^m} \mu_{F^m} \otimes \beta^m \Sigma_{uu} \\ \\ -\mu'_{F^m} \Sigma^{-1}_{F^m F^m} \otimes \Sigma_{uu} \beta^{m\prime} & \mu'_{R^e} \Sigma^{-1}_{R^e R^e} \mu_{R^e} \Sigma_{uu} + \Sigma_{F^m F^m} \end{array} \right].
$$

Proof. The proof follows again from plugging the appropriate matrices for the moment conditions (3.4) into the variance covariance formula in (A.5) and performing the matrix multiplications. Now, observe that from (A.7), we have

$$
S = \left[\begin{array}{ccc} \left[\begin{array}{cc} 1 & \mu'_{R^e} \\ \mu_{R^e} & \Sigma_{R^e R^e} + \mu_{R^e} \mu'_{R^e} \end{array} \right] \otimes \Sigma_{uu} & 0_{K(1+N) \times N(K+1)} & 0_{K(1+N) \times K} \\ \\ 0_{N(K+1) \times K(1+N)} & \left[\begin{array}{cc} 1 & \mu'_{F^m} \\ \mu_{F^m} & \Sigma_{F^m F^m} + \mu_{F^m} \mu_{F^m}{}' \end{array} \right] \otimes \Sigma_{\varepsilon^m \varepsilon^m} & 0_{N(K+1) \times K} \\ \\ 0_{K \times K(1+N)} & 0_{K \times N(K+1)} & \Sigma_{F^m F^m} \end{array} \right],
$$

and from (A.6), we have

$$
J(\theta) = \mathrm{E} \left[\begin{array}{ccc} - \left[\begin{array}{cc} 1 & R^{e\prime}_t \\ R^e_t & R^e_t R^{e\prime}_t \end{array} \right] \otimes I_K & 0_{K(1+N) \times N(K+1)} & 0_{K(1+N) \times K} \\ \\ - \left[\begin{array}{cc} 0 & R^{e\prime}_t \\ 0_{K \times 1} & \Phi(R^e_t R^{e\prime}_t) \end{array} \right] \otimes \beta^m & - \left[\begin{array}{cc} 1 & F^{m\prime}_t \\ F^m_t & F^m_t F^{m\prime}_t \end{array} \right] \otimes I_N & 0_{N(K+1) \times K} \\ \\ 0_K & R^{e\prime}_t \otimes I_K & 0_{K \times N(K+1)} & -I_K \end{array} \right],
$$

with $A = I_{K(1+N)+N(K+1)+K}$. Thus, the limiting variance of the GMM estimator for $\theta = (\mathrm{vec}\,(\beta^m)', \lambda^{m\prime})'$ is obtained by performing the matrix multiplications $J^{-1} S [J']^{-1}$.

Here, it is worth stressing that the limiting variance covariance matrix obtained by performing the matrix multiplications corresponds to the parameter vector

$$
(\Phi_0{}', \mathrm{vec}\,(\Phi)', \alpha^{m\prime}, \mathrm{vec}\,(\beta^m)', \lambda^{m\prime})'.
$$

Therefore, the asymptotic variance covariance matrix for $\theta = (\text{vec}\,(\beta^m)', \lambda^{m\prime})'$ is the lower-right $KN + K$ by $KN + K$ sub-matrix of the larger variance covariance matrix. $\square$

Lemmas A.2–A.4 allow us to study the asymptotic properties of the obtained risk premium estimators. It is worth mentioning that the lower–left $NK + K$ dimensional square matrices of the variance covariance matrices in Lemma A.2 and A.3 give the variance covariance matrices corresponding to parameters $(\text{vec}\,(\beta)', \lambda')'$. We will use these results to derive the variance covariance matrices of risk premium estimators in the following section.

Proof of Theorem 4.1. This follows from a direct application of the Central Limit Theorem. $\square$

Proofs of Theorems 4.2 and 4.3. We are interested in the asymptotic distribution of $g(\beta, \lambda) = \beta\lambda$. Given

$$\left(\text{vec}\,\left(\hat{\beta}\right)', \hat{\lambda}'\right)' - \left(\text{vec}\,(\beta)', \lambda'\right)' \xrightarrow{d} \mathcal{N}(0, V_{\beta,\lambda}),$$

we have, by applying the delta method, that

$$\sqrt{T}\left(g(\hat{\beta}, \hat{\lambda}) - g(\beta, \lambda)\right) \xrightarrow{d} \mathcal{N}(0, \dot{g}'V_{\beta,\lambda}\dot{g}),$$

with

$$\dot{g} = \begin{bmatrix} \lambda' \otimes I_N & \beta \end{bmatrix}.$$

Remember that Lemma A.2 and A.3 give the asymptotic distributions of $\sqrt{T}(\hat{\theta} - \theta)$ where $\theta = (\alpha', \text{vec}\,(\beta)', \lambda')'$ for the moment conditions (3.1) and (3.2). Observe that $V_{\beta,\lambda}$ is the lower $NK + K$ block diagonal matrix of the variance covariance matrices provided in Lemma A.2 and A.3. Hence, the asymptotic variances of the risk premium estimators in Theorems 4.2 and 4.3 follow from plugging in the limiting variance covariance matrices of $(\text{vec}\,(\beta)', \lambda')'$ and calculating $\dot{g}'V_{\beta,\lambda}\dot{g}$. $\square$

Proof of Theorem 4.4. We are interested in $g(\beta^m, \lambda^m) = \beta^m\lambda^m$. Given

$$\left(\text{vec}\,\left(\hat{\beta}^m\right)', \hat{\lambda}^{m\prime}\right)' - \left(\text{vec}\,(\beta^m)', \lambda^{m\prime}\right)' \xrightarrow{d} \mathcal{N}(0, V_{\beta^m,\lambda^m}),$$

Then, by applying the delta method, we have

$$\sqrt{T}(g(\hat{\beta}^m, \hat{\lambda}^m) - g(\beta^m, \lambda^m)) \xrightarrow{d} \mathcal{N}(0, \dot{g}'V_{\beta^m, \lambda^m}\dot{g})$$

and note that here

$$\dot{g} = \begin{bmatrix} \lambda^{m\prime} \otimes I_N & \beta^m \end{bmatrix}$$

Then, we have

$$\dot{g}'V_{\beta^m, \lambda^m}\dot{g} = \lambda^{m\prime}\Sigma_{F^m F^m}^{-1}\lambda^m \Sigma_{\varepsilon^m \varepsilon^m} + \beta^m \Sigma_{F^m F^m}\beta^{m\prime} \tag{A.11}$$
$$+ \left(\mu'_{R^e}\Sigma_{RR}^{-1}\mu_{R^e} - \lambda^{m\prime}\Sigma_{F^m F^m}^{-1}\lambda^m\right)\beta^m \Sigma_{uu}\beta^{m\prime}$$

The result follows from plugging the β^m and Φ respectively into the above equation via (2.6).

$\square$

The following lemma follows from the Schur complement condition, see Boyd and Vandenberghe (2004).

Lemma A.5. *Let*

$$\mathbf{K} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$$

be a symmetric matrix and assume that K_{22}^{-1} exists. Then $K \geq 0$ is equivalent to $K_{22} \geq 0$ and $K_{11} - K_{12}K_{22}^{-1}K_{21} \geq 0$.

Proof of Corollary 4.1. Suppose $\lambda'\Sigma_{FF}^{-1}\lambda < 1$. We need to study the difference between the limiting variance of the historical averages and the limiting variance of the expected (excess) return estimator based on (3.1). In particular, we need to study

$$\Sigma_{R^e R^e} - \left(\Sigma_{R^e R^e} - \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\left[\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\right]\right)$$
$$= \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\left[\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\right].$$

In order to show that $\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'$ is positive semi–definite, we will use Lemma A.5.

Now, let $K_1 = \Sigma_{\varepsilon\varepsilon}^{1/2}$ and $K_2 = \beta'\Sigma_{\varepsilon\varepsilon}^{-1/2}$. Then,

$$
\mathbf{K} = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \begin{bmatrix} K_1' & K_2' \end{bmatrix} = \begin{bmatrix} K_1 K_1' & K_1 K_2' \\ K_2 K_1' & K_2 K_2' \end{bmatrix}
$$

so that

$$
\mathbf{K} = \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & \beta \\ \beta' & \beta'\Sigma\varepsilon\varepsilon^{-1}\beta \end{bmatrix}.
$$

Then, Lemma A.5 yields that

$$
\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta' \geq 0
$$

$\square$

Proof of Corollary 4.2. Suppose $\lambda'\Sigma_{FF}^{-1}\lambda < 1$.

In order to prove Corollary 4.2–1, we need to study the difference between the limiting variance of the historical averages and the limiting variance of the expected (excess) return estimator based on (3.2). In particular, we need to show that

$$
\Sigma_{R^e R^e} - \left(\Sigma_{R^e R^e} - \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\Sigma_{\varepsilon\varepsilon}\right)
$$
$$
= \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\Sigma_{\varepsilon\varepsilon}
$$

is positive semi–definite. Since $\Sigma_{\varepsilon\varepsilon}$ is positive semi-definite, Corollary 4.2–1 follows.

In order to prove Corollary 4.2–2, we need to study the difference between the limiting variance of the expected (excess) return estimator based on (3.1) and the limiting variance of the expected (excess) return estimator based on (3.2). In particular, we need to show

39

that

$$\left(\Sigma_{R^eR^e} - \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\left[\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\right]\right) - \left(\Sigma_{R^eR^e} - \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\Sigma_{\varepsilon\varepsilon}\right)$$

$$= \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'$$

is positive semi–definite. This follows immediately from $\Sigma_{\varepsilon\varepsilon}$ being positive semi–definite.

$\square$

Following the literature on weak instruments (see, e.g., Staiger and Stock (1997)), and Kleibergen (2009), we consider a sequence of β's getting smaller as the sample size increases.

Proof of Corollary 5.1. Suppose Assumption 2 and Assumption 3 hold and consider the small β case where $\beta = \frac{1}{\sqrt{T}}B$ for a fixed full rank $N \times K$ matrix B. Then

1. Consider the GMM estimator based on the moment conditions (3.1). We want to study

$$\left(\hat{\beta}\hat{\lambda} - \beta\lambda\right) = \hat{\Sigma}_{R^eF}\hat{\Sigma}_{FF}^{-1}\left(\hat{\beta}'\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\hat{\beta}\right)^{-1}\left(\hat{\beta}'\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\bar{R}^e\right) - \beta\lambda. \tag{A.12}$$

Note, using $R_t^e = \alpha + \beta F_t + \varepsilon_t$,

$$\hat{\Sigma}_{R^eF} = \frac{1}{T}\sum_{t=1}^{T}\left(R_t^e - \bar{R}^e\right)\left(F_t - \bar{F}\right)' \tag{A.13}$$

$$= \frac{1}{T}\sum_{t=1}^{T}\left[\beta\left(F_t - \bar{F}\right) + (\varepsilon_t - \hat{\varepsilon})\right]\left(F_t - \bar{F}\right)'$$

$$= \beta\hat{\Sigma}_{FF} + \hat{\Sigma}_{\varepsilon F},$$

so that

$$\hat{\beta} = \beta + \hat{\Sigma}_{\varepsilon F}\hat{\Sigma}_{FF}^{-1}. \tag{A.14}$$

Moreover, note

$$\hat{\beta}' \hat{\Sigma}_{\varepsilon\varepsilon}^{-1} \bar{R}^e = \left(\beta + \hat{\Sigma}_{\varepsilon F} \hat{\Sigma}_{FF}^{-1}\right)' \hat{\Sigma}_{\varepsilon\varepsilon} \left(\alpha + \beta \bar{F} + \varepsilon\right),$$ (A.15)

and

$$\hat{\beta}' \hat{\Sigma}_{\varepsilon\varepsilon}^{-1} \hat{\beta} = \left(\beta + \hat{\Sigma}_{\varepsilon F} \hat{\Sigma}_{FF}^{-1}\right)' \hat{\Sigma}_{\varepsilon\varepsilon} \left(\beta + \hat{\Sigma}_{\varepsilon F} \hat{\Sigma}_{FF}^{-1}\right).$$ (A.16)

Considering $\beta = B/\sqrt{T}$ and assuming that $\sqrt{(T)}\hat{\Sigma}_{\varepsilon F} \xrightarrow{d} Y$, and $\sqrt{T}\bar{\varepsilon} \xrightarrow{d} U$ for some random U and Y, from (A.16), we get

$$\hat{\beta}' \hat{\Sigma}_{\varepsilon\varepsilon}^{-1} \hat{\beta} \xrightarrow{d} \left(B + Y\Sigma_{FF}^{-1}\right)' \Sigma_{\varepsilon\varepsilon} \left(B + Y\Sigma_{FF}^{-1}\right).$$ (A.17)

Furthermore, note

$$ER_t^e = \beta\lambda = \alpha + \beta EF,$$ (A.18)

so that

$$\alpha = \beta\left(\lambda - EF\right) = B/\sqrt{T}\left(\lambda - EF\right).$$ (A.19)

As a result, (A.15) leads to

$$T\hat{\beta}' \hat{\Sigma}_{\varepsilon\varepsilon}^{-1} \bar{R}^e \xrightarrow{d} \left(B + Y\Sigma_{FF}^{-1}\right)' \Sigma_{\varepsilon\varepsilon} \left(B\lambda + U\right).$$ (A.20)

Hence, (A.12) implies

$$\left(\hat{\beta}\hat{\lambda} - \beta\lambda\right) \xrightarrow{d} \left(B + Y\Sigma_{FF}^{-1}\right) \left[\left(B + Y\Sigma_{FF}^{-1}\right)' \Sigma_{\epsilon\epsilon}^{-1} \left(B + Y\Sigma_{FF}^{-1}\right)\right]^{-1}$$
$$\times \left(B + Y\Sigma_{FF}^{-1}\right)' \Sigma_{\epsilon\epsilon}^{-1}[B\lambda + U] - \beta\lambda.$$

Under the assumptions in the paper, U is independent of Y, $U \sim N(0, \Sigma_{\epsilon\epsilon})$ and $vec(Y) \sim N(0, \Sigma_{FF} \otimes \Sigma_{\epsilon\epsilon})$.

2. Suppose that factors are traded, and consider the GMM estimator based on the moment conditions (3.2). Then,

$$\hat{\beta}\hat{\lambda} = \hat{\Sigma}_{R^e F}\hat{\Sigma}_{FF}^{-1}\bar{F} \tag{A.21}$$

$$= \left(\beta + \hat{\Sigma}_{\varepsilon F}\hat{\Sigma}_{FF}^{-1}\right)\bar{F},$$

so that

$$\sqrt{T}\left(\hat{\beta}\hat{\lambda} - \beta\lambda\right) = \beta\sqrt{T}\left(\bar{F} - EF\right) + \sqrt{T}\hat{\Sigma}_{\varepsilon F}\hat{\Sigma}_{FF}^{-1}\bar{F}. \tag{A.22}$$

Now, taking $\beta = B/\sqrt{T}$,

$$\sqrt{T}\left(\hat{\beta}\hat{\lambda} - \beta\lambda\right) \xrightarrow{d} Y\Sigma_{FF}^{-1}\lambda. \tag{A.23}$$

Using $vec(Y) \sim N(0, \Sigma_{FF} \otimes \Sigma_{\epsilon\epsilon})$,

$$V\Sigma_{FF}^{-1}\lambda = vec\left(Y\Sigma_{FF}^{-1}\lambda\right)$$

$$= \left(\lambda^T\Sigma_{FF}^{-1} \otimes I_N\right)vec(Y)$$

$$\sim N\left(0, \left(\lambda^T\Sigma_{FF}^{-1} \otimes I_n\right)(\Sigma_{FF} \otimes \Sigma_{\varepsilon\varepsilon})(\Sigma_{FF}^{-1}\lambda \otimes I_N)\right)$$

$$= N(0, \lambda^T\Sigma_{FF}^{-1}\Sigma_{FF}\Sigma_{FF}^{-1}\lambda \otimes \Sigma_{\varepsilon\varepsilon})$$

$$= N(0, \lambda^T\Sigma_{FF}^{-1}\lambda\Sigma_{\varepsilon\varepsilon}) \tag{A.24}$$

$\square$

B Inference about Risk–Premiums with Omitted Factors

The asymptotic results in the main paper are based on the assumption that the pricing model is correctly specified. The researcher is assumed to know the true factor model that explains expected excess returns on the assets. In that case, the risk–premium estimators are consistent certainly under our maintained assumption of stationary and ergodic returns. However, the pricing model may be misspecified and this might induce inconsistent risk–premium estimates. We investigate this issue and its solution in this section.

We consider model misspecification due to omitted factors. An example of such type of misspecification would be to use Fama–French three factor model if the true pricing model is the four factor Fama–French–Carhart Model. Formally, assume that excess returns are generated by a factor model with two different sets of distinct factors, F and G such that

$$R^e = \alpha^* + \beta^* F + \delta^* G + \varepsilon^*, \tag{B.1}$$

where ε^* is a vector of residuals with mean zero and $\mathrm{E}\left[F\varepsilon^{*\prime}\right] = 0$ and $\mathrm{E}\left[G\varepsilon^{*\prime}\right] = 0$. Note that the sets of factors F and G perfectly explain the expected excess returns of the test assets, i.e., $\mathrm{E}\left[R^e\right] = \beta^*\lambda_F + \delta^*\lambda_G$.

However, a researcher may forget about the presence of the factors G and thus estimates the model only with factors F. Then, the estimated model is

$$R^e = \alpha + \beta F + \varepsilon, \tag{B.2}$$

with zero–mean ε, and $\mathrm{E}\left[F\varepsilon^\prime\right] = 0$. As the researcher might not know the underlying factor model exactly, she allows for misspecification by adding an N-vector of constant terms, α, in the estimation as in Fama and French (1993).

The bias in the parameter estimates for, α, β and λ are presented in the following

theorem:

Theorem B.1. *Assume that returns are generated by (B.1) but α, β, and λ are estimated from (B.2) with GMM based on (3.1). Then,*

1. *$\hat{\alpha}$ converges to $\alpha^* + (\beta^* - \beta)E[F] + \delta^*E[G]$,*
2. *$\hat{\beta}$ converges to $\beta^* + \delta^* Cov\left[G, F^T\right]\Sigma_{FF}^{-1}$,*
3. *$\hat{\lambda}$ converges to $\lambda_F + (\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\left[(\beta^* - \beta)\lambda_F + \delta^*\lambda_G\right]$,*

in probability.

Theorem B.1 shows that, if a researcher ignores some risk factors G, then the risk price estimators associated with the factors F are inconsistent if and only if

$$\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\left[(\beta^* - \beta)\lambda_F + \delta^*\lambda_G\right] \neq 0.$$

It is important to note that the inconsistency of the estimates of risk prices may be caused not only by the risk prices λ of the omitted factors but also by the bias in betas of the factors F. This result has an important implication: even if the ignored factors have zero price of risk, the cross–sectional estimates of the prices of risk on the true factors included in the estimation (F) can still be asymptotically biased. This happens in case F and G are correlated.

Next, we analyse the asymptotic bias in the parameter estimates for α, β and λ in case the factors are traded and the estimation is based on GMM with the moment conditions (3.2) of Section 3.2.

Theorem B.2. *Assume that returns are generated by (B.1) but α, β, and λ are estimated from (B.2) with GMM based on (3.2). Then,*

1. *$\hat{\alpha}$ converges to $\alpha^* + (\beta^* - \beta)\lambda_F + \delta^*\lambda_G$,*
2. *$\hat{\beta}$ converges to $\beta^* + \delta^* Cov\left[G, F^T\right]\Sigma_F^{-1}$,*

3. $\hat{\lambda}$ converges to λ_F,

in probability.

Theorem B.2 illustrates that, even if the researcher forgets some risk factors, risk price estimators will still be asymptotically unbiased. Notice that this is in contrast with the estimator based on GMM with moment conditions (3.1) of Section 3.1. It is important to note that, if the forgotten factors G, are uncorrelated with the factors, then the bias in β disappears. Moreover, if the ignored factors are associated with zero prices of risk and are uncorrelated with F, then $\hat{\alpha}$ will converge to α.

This raises the question what happens to the risk–premium estimators on individual assets or portfolios if some true factors are ignored in the estimation? The following corollary provides consistency conditions for risk–premium estimators of individual assets or portfolios.

Corollary B.1. *If the returns are generated by (B.1) and*

1. *the model (B.2) is estimated with GMM based on (3.1), then the vector of resulting risk–premium estimators $\hat{\beta}\hat{\lambda}$ converges to $E[R^e]$ if and only if*
$$[I_N - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}]E[R^e] = 0.$$

2. *all factors are traded. If the model (B.2) is estimated with GMM (3.2), then the vector of resulting risk–premium estimators $\hat{\beta}\hat{\lambda}$ converges to $E[R^e]$ if and only if $(\beta^* - \beta)\lambda_F + \delta^*\lambda_G = 0$.*

In the view of the theorem above, if the factors are traded, and the estimation is via GMM with moment conditions (3.2), then the risk–premium estimator is unbiased when the omitted factors are uncorrelated with the factors, F, and the the omitted factors are associted with zero prices of risk. In order to capture misspecification, it is a common approach to add an N–vector of constant terms, α, to the model as in (B.2). In the following theorem, we will show that in case of traded factors, it is possible to achieve the

consistency for estimating risk premiums, however, this comes at the cost of loosing all efficiency gains.

Theorem B.3. *Assume that all factors in F are traded. If the returns are generated by (B.1) but the model (B.2) is estimated with GMM based on (3.2) where the risk price estimates are given by the factor averages, then the estimator $\hat{\alpha} + \hat{\beta}\hat{\lambda}$ is consistent for $E[R^e]$. The asymptotic variance of such estimator equals $\Sigma_{R^e R^e}$.*

It is important to note that adding the $\hat{\alpha}$ to $\hat{\beta}\hat{\lambda}$ does not solve the inconsistency problem if the system is estimated via GMM with (3.1). If some factors are non–traded and the parameters are estimated via GMM with (3.1), adding the $\hat{\alpha}$ capturing the misspecification to $\hat{\beta}\hat{\lambda}$ doesn't lead to consistent estimates of $\mathrm{E}[R^e]$. In particular, $\hat{\alpha} + \hat{\beta}\hat{\lambda}$ converges to $\mathrm{E}[R^e] - \beta(\lambda - \mathrm{E}[F])$ and $\lambda - \mathrm{E}[F]$ is generally nonzero. Table A1 documents empirical values of the bias caused by estimating the CAPM model on 25 Fama French portfolios. There is a slight bias in ignoring the other factors, SMB and HML, about 1% on average across 25 Fama French portfolios for the general case and mimicking factor case, and about 1.5% for the traded factor case. Observe that adding the $\hat{\alpha}$ to $\hat{\beta}\hat{\lambda}$ corrects for the bias for traded factors whereas it does not correct for the bias for the general case.

C   Proofs for Section "Inference about Risk–Premiums with Omitted Factors"

Proof of Theorem B.1. Note that $\hat{\beta}$ converges to β and $\hat{\alpha}$ converges to α in probability, where

$$
\begin{aligned}
\beta &= \mathrm{Cov}\left[R^e, F^T\right]\Sigma_{FF}{}^{-1}, &\text{(C.3)}\\
&= \mathrm{Cov}\left[R^e = \alpha^* + \beta^* F + \delta^* G + \varepsilon^*, F^T\right]\Sigma_{FF}{}^{-1},\\
&= \beta^* + \delta^*\mathrm{Cov}\left[G, F^T\right]\Sigma_{FF}^{-1}.
\end{aligned}
$$

and

$$\alpha \quad = \quad \mathrm{E}\left[R^e\right] - \beta \mathrm{E}\left[F\right], \tag{C.4}$$

$$= \quad \alpha^* + \beta^* \mathrm{E}\left[F\right] + \delta^* \mathrm{E}\left[G\right] - \beta \mathrm{E}\left[F\right],$$

$$= \quad \alpha^* + (\beta^* - \beta)\mathrm{E}\left[F\right] + \delta^* \mathrm{E}\left[G\right].$$

Furthermore, for $\hat{\lambda}$, first notice that

$$\hat{\lambda} = \left(\hat{\beta}'\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\hat{\beta}\right)^{-1}\hat{\beta}'\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\bar{R}^e. \tag{C.5}$$

The probability limit of $\hat{\lambda}$ from GMM (3.1) is thus given by

$$\lambda \quad = \quad \left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\left[\beta^*\lambda_F + \delta^*\lambda_G\right]$$

$$= \quad \lambda_F + \left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\left[(\beta^* - \beta)\lambda_F + \delta^*\lambda_G\right]. \tag{C.6}$$

$\square$

Proof of Theorem B.2. Note that $\hat{\beta}$ converges to β and $\hat{\alpha}$ converges to α in probability, with

$$\beta \quad = \quad \mathrm{Cov}\left[R^e, F^T\right]\Sigma_{FF}^{-1} \tag{C.7}$$

$$= \quad \mathrm{Cov}\left[R^e - \alpha^* + \beta^* F + \delta^* G + \varepsilon^*, F^T\right]\Sigma_{FF}^{-1}$$

$$= \quad \beta^* + \delta^* \mathrm{Cov}\left[G, F^T\right]\Sigma_{FF}^{-1}.$$

and

$$\alpha \quad = \quad \mathrm{E}\left[R^e\right] - \beta \mathrm{E}\left[F\right] \tag{C.8}$$

$$= \quad \alpha^* + \beta^*\lambda_F + \delta^*\lambda_G - \beta\lambda_F$$

$$= \quad \alpha^* + (\beta^* - \beta)\lambda_F + \delta^*\lambda_G.$$

Furthermore, for $\hat{\lambda}_F$, notice that $\hat{\lambda}_F = \bar{F}$, which converges to $\lambda_F = \mathrm{E}\left[F\right]$ in probability.

$\square$

Proof of Corollary B.1. For the first part of the corollary, note that

$$\beta\lambda_F = \beta \left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\mathrm{E}\left[R^e\right]. \tag{C.9}$$

Hence, $\hat{\beta}\hat{\lambda}$ is consistent for $\mathrm{E}\left[R^e\right]$ if and only if $\mathrm{E}\left[R^e\right] = \beta\left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\mathrm{E}\left[R^e\right]$. This, in turn, equivalent to

$$\left[I_N - \beta\left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\right]\mathrm{E}\left[R^e\right] = 0. \tag{C.10}$$

To prove the second part of the corollary, note that $\hat{\beta}\hat{\lambda}$ converges to $\beta\lambda$. Using (C.7) and $\lambda_F = \mathrm{E}\left[F\right]$, we have

$$\begin{aligned}
\beta\lambda_F &= (\beta^* + \delta^*\mathrm{Cov}\left[G, F^T\right]\Sigma_{FF}^{-1})\lambda_F, \tag{C.11}\\
&= \mathrm{E}\left[R^e\right] - ((\beta^* - \beta)\lambda_F + \delta^*\lambda_G).
\end{aligned}$$

$\square$

Proof of Theorem B.3. Consistency of $\hat{\alpha}+\beta\lambda_F$ is straightforward. The asymptotic variance is given by the delta method using $g(\alpha, \beta, \lambda_F) = \alpha+\beta\lambda_F$. The asymptotic covariance matrix of α, β, and γ is given in Lemma A.3 (denoted by V).

Thus,

$$\sqrt{T}\left(g(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) - g(\alpha, \beta, \lambda)\right) \overset{d}{\to} \mathcal{N}(0, \dot{g}'V_{\alpha,\beta,\lambda}\dot{g}), \tag{C.12}$$

with

$$\dot{g} = \left[\begin{array}{cc} [1 \quad \lambda'] \otimes I_N & \beta \end{array}\right].$$

Matrix multiplication of calculating $\dot{g}'V_{\alpha,\beta,\lambda}\dot{g}$ gives $\Sigma_{R^e R^e}$.

$\square$

Table A1: **Omitted Factors**

This table provides the average risk premium estimates on average across assets, $\hat{\beta}\hat{\lambda}$ (first line), its bias, the average risk premium estimates augmented with an alpha on average across assets, $\hat{\alpha} + \hat{\beta}\hat{\lambda}$,and its bias. The test assets are the 25 portfolios formed by Fama–French (1992,1993). The factor is the market factor in a standard CAPM. The results are based on monthly data from January 1963 until August 2020, i.e., 692 observations for each portfolio. The first column presents the improvements for the factor–model based risk–premium estimates based on GMM with (3.1) over the naive estimate of historical averages. The second and the third columns present the gains of factor–model based risk–premium estimates based on GMM with (3.2) and with (3.4) over naive estimates, respectively.

|	General Case	Traded Factors	Mimicking Factors
$\hat{\beta}\hat{\lambda}$	7.73	7.38	7.73
Bias	-1.05	-1.39	-1.05
$\hat{\alpha} + \hat{\beta}\hat{\lambda}$	9.12	8.77	8.77
Bias	0.35	0.00	0.00

Figure 1: **Factor-Model-Based Standard Error Estimates of Expected Returns**
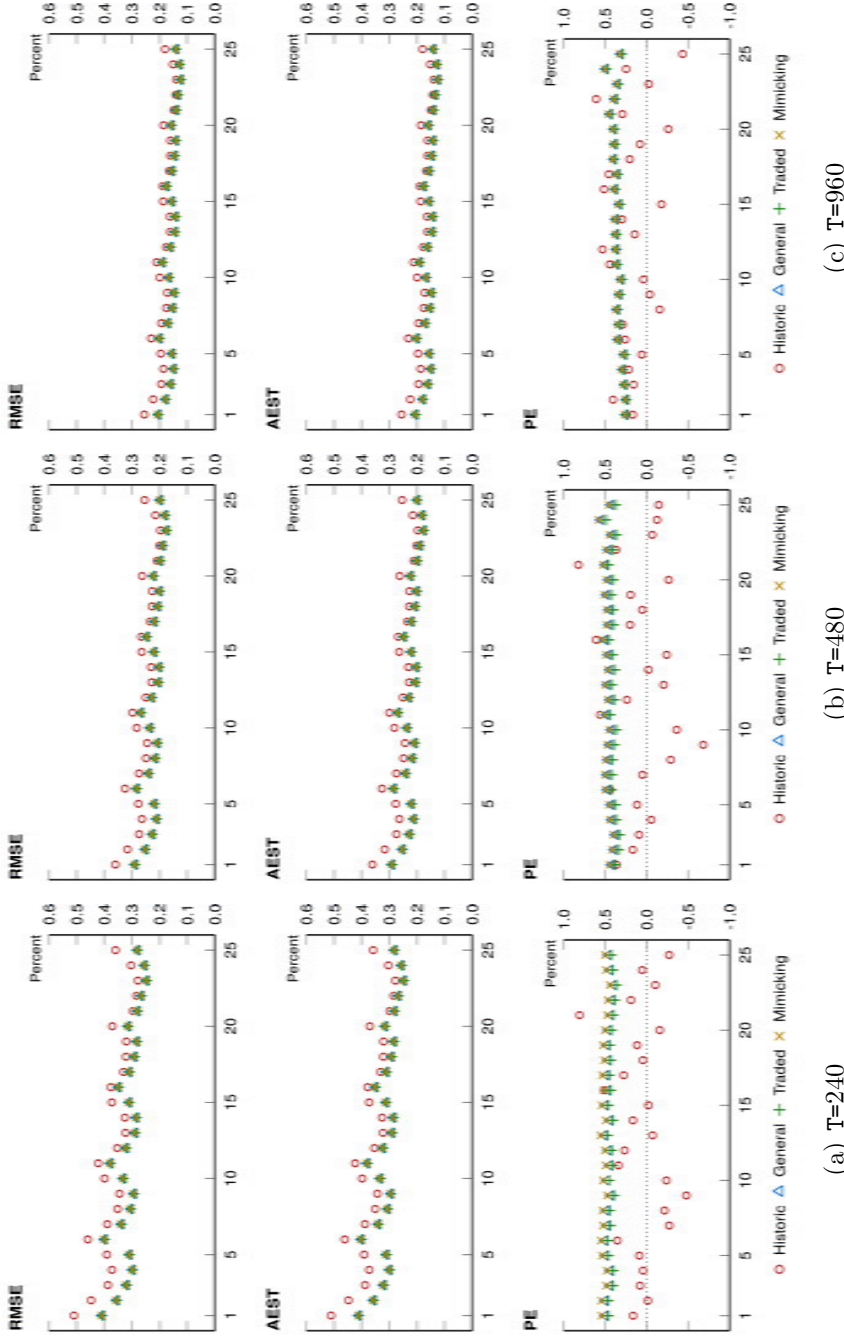
(a) T=240

(b) T=480

(c) T=960

This figure reports the standard errors of the various expected return estimators for the 25 Fama-French portfolios from 20000 simulations with sample size $T = 240$ (panel a), $T = 480$ (panel b), and $T = 960$ (panel c). The first exhibit shows the root-mean-square errors (RMSE) of the expected return estimates based on naive, GMM-Gen, GMM-Tr and GMM-Mim. The second exhibit illustrates the simulation average of the estimated asymptotic standard errors (AEST), and the third exhibit reports the percentage error (PE) of the AEST as compared to the RMSE. Returns are generated from a one-factor CAPM with a normally distributed market factor and the error terms. The moments of the market factor and the error terms and the parameters of the data generating process are calibrated to CAPM model with 25 FF portfolios with data from 1963:1-2020:8.

Figure 2: **Factor-Model-Based Standard Error Estimates of Expected Returns: Weak Identification**
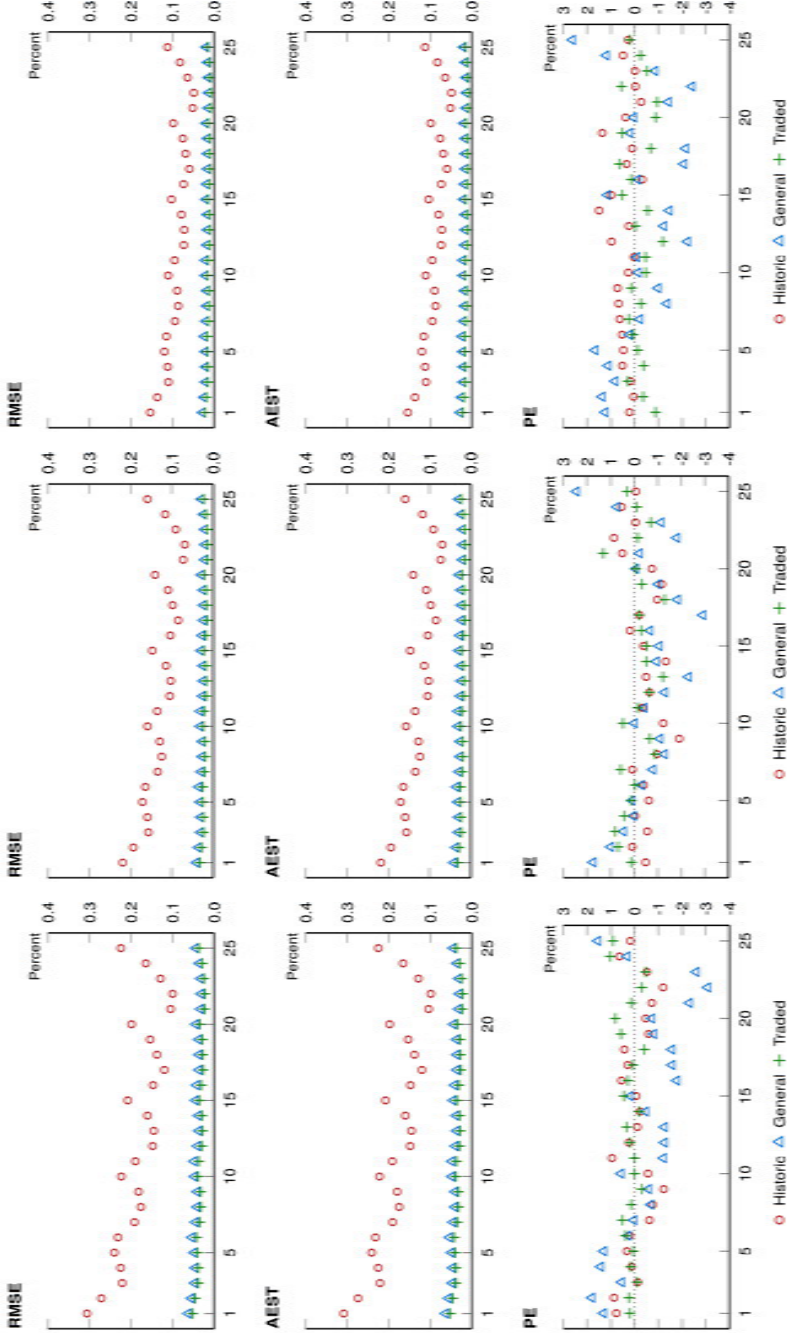
(a) T=240

(b) T=480

(c) T=960

This figure reports the standard errors of the various expected return estimators for the 25 Fama-French portfolios in a one-factor model with a weak factor. The results are from 20000 simulations with sample sizes $T = 240$ (panel a), $T = 480$ (panel b), and $T = 960$ (panel c). The first exhibit shows the root-mean-square errors (RMSE) of the expected return estimates based on naive (historic), GMM-Gen (General), and GMM-Tr (Traded). The second exhibit illustrates the simulation average of the estimated asymptotic standard errors (AEST), and the third exhibit reports the percentage error (PE) of the AEST as compared to the RMSE. Returns are generated from a 1-factor CAPM with

$$R_t^e = \alpha + \frac{B}{\sqrt{T}} R_{m,t}^e + \varepsilon_t, \quad t = 1, 2, \ldots, T,$$

with normally distributed F_t and ε_t and the asset pricing restriction $E[R_t^e] = \frac{B}{\sqrt{T}} \lambda$, with data from 1963:1-2020:8.

Table 1: **Improvements in Efficiency for the 25 Fama–French Portfolios (in percentage)**

This table illustrates the gains in variances (in percentage) for the various risk–premium estimates for the 25 portfolios formed by Fama and French (1992,1993). The factors are the three factors from Fama and French (1992): market, size and book–to–market. The results are based on monthly data from January 1963 until August 2020, i.e., 692 observations for each portfolio. The first column (RP_{GMM}–Gen over $Naive$) presents the improvements for the factor–model based risk–premium estimates based on GMM with (3.1) over the naive estimate of historical averages. The second (RP_{GMM}–Tr over $Naive$) and the third (RP_{GMM}–Mim over $Naive$) columns present the gains of factor–model based risk–premium estimates based on GMM with (3.2) or (3.4) over naive estimates, respectively. The fourth column (RP_{GMM}–Tr over RP_{GMM}–Gen) corresponds to the precision gains from estimating the risk premiums based on GMM using the moment conditions (3.2) over the case based on GMM with (3.1). The last column (RP_{GMM}–Mim over RP_{GMM}–Gen) presents the gains from making use of mimicking portfolios using (3.4) over estimation based on (3.1)

Assets	RP_{GMM}–Gen over Naive	RP_{GMM}–Tr over Naive	RP_{GMM}–Mim over Naive	RP_{GMM}–Tr over RP_{GMM}–Gen	RP_{GMM}–Mim over RP_{GMM}–Gen	RP_{GMM}–Mim over RP_{GMM}–Tr
1	8.9	9.6	8.8	0.6	-0.1	-0.7
2	6.9	7.5	6.9	0.5	-0.1	-0.6
3	4.3	4.8	4.3	0.5	-0.1	-0.5
4	4.7	5.2	4.6	0.5	-0.1	-0.5
5	4.8	5.4	4.8	0.5	-0.1	-0.6
6	4.8	5.5	4.7	0.7	-0.1	-0.8
7	5.0	5.5	4.9	0.5	-0.1	-0.6
8	7.0	7.6	6.9	0.5	-0.1	-0.6
9	4.9	5.5	4.8	0.6	-0.1	-0.7
10	4.2	4.8	4.1	0.6	-0.1	-0.7
11	4.6	5.4	4.5	0.7	-0.1	-0.8
12	7.6	8.2	7.5	0.6	-0.1	-0.6
13	9.9	10.6	9.8	0.7	-0.1	-0.7
14	8.2	9.0	8.1	0.7	-0.1	-0.8
15	10.1	11.0	10.0	0.8	-0.1	-0.9
16	6.1	6.9	6.0	0.8	-0.1	-0.8
17	10.8	11.6	10.7	0.7	-0.1	-0.8
18	12.3	13.3	12.2	0.8	-0.1	-0.9
19	11.3	12.3	11.2	0.9	-0.1	-1.0
20	12.2	13.3	12.1	0.9	-0.1	-1.1
21	4.8	5.8	4.7	0.9	-0.1	-1.0
22	9.5	10.5	9.3	0.9	-0.1	-1.0
23	15.1	16.3	15.0	1.1	-0.1	-1.2
24	10.3	11.8	10.1	1.3	-0.2	-1.5
25	22.7	24.3	22.5	1.3	-0.2	-1.5

Table 2: **Out–of–Sample Sharpe Ratios based on various risk–premium estimates**

This table provides the average out–of–sample Sharpe ratio over simulations, $\overline{SR}$ (first line), its percentage error compared to the true Sharpe ratio, $(\overline{SR} - SR)/SR$ (second line) and the root–mean–squared errors, the square root of $\sum_{s=1}^{Z}(\hat{SR}_s - SR)/Z$ (third line) over 20,000 simulated data sets for optimal portfolios constructed with various risk premium estimates. The data generating process is

$$R_t^e = \alpha + \beta F_t + \varepsilon_t, \quad t = 1, 2, \ldots, T,$$

with normally distributed F_t and ε_t and the asset pricing restriction $E[R_t^e] = \beta\lambda$. R^e is the N–vector of asset returns at period t, F_t is K–vector of factors and T is the number of periods. The moments of factors and residuals and the parameters of data generating process are obtained from a calibration of Fama–French 3 factor model from January 1963 to August 2020. The risk premium estimates are based on naive, GMM with moment conditions (3.1)-RP_{GMM}–Gen , (3.2)- RP_{GMM}–Tr and (3.4)-RP_{GMM}–Mim. The variance–covariance matrix is estimated by the sample variance covariance matrix. T is assumed to be 692 and risk aversion is 5.

$\gamma = 5$								
				Enlarging Samples				
	True μ^e	Naive	RP_{GMM}–Gen	RP_{GMM}–Tr	RP_{GMM}–Mim	GMV	1/N	Theoretical
True Σ_{RR}	0.181	0.104	0.157	0.157	0.157	0.11	0.147	0.178
	0.015	-0.419	-0.118	-0.117	-0.118			
	0.003	0.091	0.054	0.054	0.054	0.047	0.047	
$\hat{\Sigma}_{RR}$	0.175	0.101	0.161	0.162	0.161	0.109	0.147	0.178
	-0.016	-0.433	-0.095	-0.093	-0.095			
	0.003	0.093	0.053	0.053	0.053	0.048	0.047	