Abstract

Estimating expected returns on individual assets or portfolios is one of the most fundamental problems of finance research. The standard approach, using historical averages, produces noisy estimates. Linear factor models of asset pricing imply a linear relationship between expected returns and exposures to one or more sources of risk. We show that exploiting this linear relationship leads to statistical gains of 36% in standard deviations when estimating expected returns over historical averages. If the factor model is misspecified in the sense of an omitted factor, we show that factor model–based estimates may be inconsistent. However, we show that adding an alpha to the model capturing mispricing only leads to consistent estimators in case of traded factors. Moreover, our simulation experiment shows that using factor–model based estimates of expected returns significantly improves the out–of–sample performance of the optimal portfolios.

Keywords: Factor Pricing Models, Risk–Return Models, Omitted Factors, Misspecified models

JEL: C21, G12

1. Introduction

One of the key problems of finance studies is the estimation of risk premiums, that is expected excess returns, on individual securities or portfolios. The standard approach,
which has been favoured by researchers, investors and analysts, is to use historical averages. However, it is also known that these estimates are generally very noisy. Even using daily data does not help much, if at all. One needs very long samples for accurate estimates, which often are unavailable.

The asset pricing literature provides a wide variety of linear factor models motivating certain risks that explain the cross section of expected returns on assets. Examples include Sharpe (1964)’s CAPM, Merton (1973)’s ICAPM, Breeden (1979)’s CCAPM, Ross (1976a,b)’ APT, Lettau and Ludvigson (2001)’s conditional CCAPM among many others. These models all imply that expected returns of assets are linear in their exposures to the risk factors. The coefficients in this linear relationship are the prices of the risk factors. The literature on factor models mainly concentrates on determining these prices of risk and evaluating the ability of the models in explaining the cross section of expected returns on assets.

In this study, the focus is different: we assess the precision gains in the estimation of the expected (excess) returns on an individual asset and on portfolios, i.e., the product of exposures ($\beta$) and risk prices ($\lambda$), vis-à-vis the historical averages approach. As mentioned by Black (1993), theory can help to improve the estimates of expected returns. We show when exploiting the linear relationship implied by linear factor models indeed leads to more precise estimates of expected returns over historical averages.

Estimating expected returns using factor models is not a new idea and was, to our knowledge, first suggested by Jorion (1991). In his empirical analysis, he compares CAPM—based estimators with classical sample averages of past returns finding the former outperforming the latter in estimating expected stock returns for his data. Our paper complements his work by providing the first detailed asymptotic efficiency analysis for both estimators, and evaluating the implications of omitted factors on the estimation of expected (excess) returns.

First, we investigate the issue of accurate estimation of risk-premiums, i.e. expected excess returns on individual assets and portfolios by providing a detailed (asymptotic)
analysis of risk–premium estimators based on factor models. Comparing the limiting covariance matrices of factor–based risk–premium estimators with those of the historical averages estimator, we find sizeable efficiency gains from imposing the factor structure, see Corollaries (4.1-4.2). In an empirical analysis, for instance when estimating risk–premiums on 25 size and book–to–market Fama–French portfolios, we document large gains in standard deviations of 36% on average.

Secondly, we consider the issue of estimating risk premiums in the ubiquitous situation where one may face omitted factors in the specification of the linear factor model. After the Capital Asset Pricing Model (CAPM) had been substantially criticized, researchers have come up with new risk factors to help explaining the cross section of expected returns. See, e.g., Fama and French (1993), Lettau and Ludvigson (2001), Lustig and Van Nieuwerburgh (2005), Li, Vassalau and Xing (2006), Santos and Veronesi (2006). While it is doubtful that “the correct” factors have been found, the literature points to the existence of missing factors. We show that when a model is misspecified, in the sense that a relevant pricing factor is omitted, standard methods will generally not even provide consistent estimates of risk premiums on the individual assets or portfolios (see Theorem 5.1). However, we show that adding an alpha capturing the misspecification leads to a consistent estimator only in case of traded factors, but there is no efficiency gain over historical averages. Thus, our paper documents precisely the trade-off any empirical researcher always faces: allow for misspecification and loose efficiency or run the risk of misspecification and gain efficiency.

The mean—variance framework of Markowitz (1952) is still a very popular model for portfolio allocation used in practice. However, it is also well known that the practical applications suffer from uncertainty in the parameter estimates. In particular, portfolios constructed with sample counterparts of first two moments in general have poor out of sample performance.\(^1\) Already Merton (1980), followed by Chopra and Ziemba (1993), pointed out that estimation error in asset return means is more severe than er-

rors in covariance estimates. Moreover, imprecision in estimates of the mean has a much larger impact on portfolio weights compared to the imprecision in covariance estimates (DeMiguel et al. (2009)). The mean—variance portfolio weights could also be constructed with factor–based risk–premium estimates instead of the “naive” estimates (historical averages). Accordingly, we investigate if it is possible to achieve performance gains based on the higher precision of factor–based risk–premium estimates. In particular, we analyze the out–of–sample performances of tangency portfolios based on various risk–premium estimators in a simulation study. Our results document that the average out–of–sample Sharpe ratio of the tangency portfolio increases strikingly if the portfolio weights are constructed with factor–based risk–premium estimates rather than the naive estimates. Moreover, out–of–sample Sharpe ratios of the factor–based tangency portfolios is more precise than the tangency portfolios based on historical averages. Our simulation results also document that these portfolios, in contrast to the tangency portfolios based on historical averages, perform considerably better than the global minimum variance portfolio.

The rest 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 expected return obtained with non–traded factors and with factor–mimicking portfolios. Section 3 discusses in detail the estimators we consider. In particular, we recall the different sets of moment conditions for various cases such as all factors being traded and factor–mimicking portfolios. Section 4 derives the asymptotic properties of these induced GMM estimators. In particular, we derive the efficiency gains over and above the risk–premium estimator based on historical averages. Section 5 adresses the question of using misspecified factor pricing models. Section 6 documents the simulation analysis for portfolio optimization, while Section 7 concludes. All proofs are gathered in the appendix.
2. Model

It is well known that in the absence of arbitrage, there exists a stochastic discount factor $M$ such that for any traded asset $i = 1, 2, \ldots, N$ with excess return $R^e_i$

$$E[MR^e_i] = 0.$$

(2.1)

Linear factor models additionally specify $M = a + b'F$, where $F = (F_1, ..., F_K)'$ is a vector of K factors (see, e.g., Cochrane (2001), p.69). Note that (2.1) can be written in matrix notation using the vector of excess returns $R^e = (R^e_1, ..., R^e_N)'$. 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 $\Sigma_{FF}$ has full rank $K$,
3. The covariance matrix between excess returns and factors $\text{Cov}[R^e, F']$ has full rank $K$.

Given the linear factor model and Assumption 1, it is classical to show

$$E[R^e] = \beta \lambda,$$

(2.2)

where

$$\beta = \text{Cov}[R^e, F'] \Sigma_{FF}^{-1},$$

(2.3)

$$\lambda = -\frac{1}{E[M]} \Sigma_{FF} b.$$  \hspace{1cm} (2.4)

Thus, (2.2) specifies a linear relationship between risk premiums, $E[R^e]$, and the exposures $\beta$ of the assets to the risk factors, $F$, with prices $\lambda$.

In empirical work, we need to make assumptions about the time–series behavior of consecutive returns and factors. In this paper, we focus on the simplest, and most used, setting where returns are i.i.d. over time. Express the excess asset returns

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

(2.5)
Working paper

where α is an N–vector of constants, ε_t is an N–vector of idiosyncratic errors and T is the number of time–series observations. We then, additionally, impose the following.

**Assumption 2.** The disturbance ε_t and the factors F_t, are independently and identically distributed over time with

\[
E[ε_t|F_t] = 0, \quad (2.6)
\]

\[
\text{Var}[ε_t|F_t] = Σ_{εε}, \quad (2.7)
\]

where Σ_{εε} has full rank.

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 suggested to refer to alternative formulations of such factor models replacing the factors by their projections on the linear span of the returns. This is commonly referred to as factor mimicking portfolios and early references go back to Huberman (1987) (see also, e.g., Fama (1998) and Lamont (2001)). We analyze, in this paper, the role 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 in estimating risk premiums.

We project the factors F_t onto the space of excess asset returns, augmented with a constant. In particular, given Assumption 1, there exists a K–vector Φ_0 and a K × N matrix Φ of constants and a K–vector of random variables u_t satisfying

\[
F_t = Φ_0 + Φ R_t^ε + u_t, \quad (2.8)
\]

\[
E[u_t] = 0_{K×1}, \quad (2.9)
\]

\[
E[u_t R_t^ε]' = 0_{K×N}, \quad (2.10)
\]

and we define the factor–mimicking portfolios by

\[
F_t^m = Φ R_t^ε. \quad (2.11)
\]
We, then, obtain an alternative formulation of the linear factor model by replacing the original factors with factor–mimicking portfolios:

$$R_t^* = \alpha^m + \beta^m F_t^m + \varepsilon_t^m, \quad t = 1, 2, \ldots, T.$$  (2.12)

Recall that using the projection results, $\Phi$ and $\beta$ are related by

$$\Phi = \Sigma^{FF} \beta' \Sigma_{RR}^{-1} R,$$  (2.13)

while $\beta^m$ and $\beta$ satisfy

$$\beta^m = \beta \left( \beta' \Sigma_{RR}^{-1} \beta \right)^{-1} \Sigma_{FF}^{-1}.$$  (2.14)

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

**Theorem 2.1.** Under Assumptions 1 and 2, we have $\beta \lambda = \beta^m \lambda^m$, where $\lambda^m = E[F_t^m]$.

Note that since the factor–mimicking portfolio is an excess return, asset pricing theory implies that the price of risk attached to it, $\lambda^m$, equals its expectation. This can be imposed in the estimation of expected (excess) returns and thus one may hope that the expected (excess) return estimators obtained with factor–mimicking portfolios are more efficient than the expected (excess) return estimators obtained with the non-traded factors themselves.

3. Estimation

As indicated in the introduction, 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 $\beta$ and $\lambda$ 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 resulting moment conditions to estimate both factor loadings $\beta$ and factor prices $\lambda$ are

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

The first moment conditions identifies $\alpha$ and $\beta$ as the regression coefficients, while the last conditions represent the pricing restrictions. Note that there are $N \times (1 + K + 1)$ moment conditions although there are $N \times (1 + K) + K$ parameters, which implies that the system is overidentified. Again following Cochrane (2001), we set a linear combination of the given moment conditions to zero, that is, we set $AE[h_t(\alpha, \beta, \lambda)] = 0$, where

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

Note that the matrix $A$ specified above combines the last $N$ moment conditions into $K$ moment conditions so that the system becomes exactly identified. Following Cochrane (2001), we take $\Theta = \beta^T \Sigma_{\epsilon \epsilon}^{-1}$. The advantage of this particular choice is that the resulting $\lambda$ 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, its price equals its expectation. For example, the price of market risk is equal to the expected market return over the risk-free rate, 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 (2010) 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 portfolios 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 may use the additional moment condition that their expectation equals $\lambda$. Then, the relevant moment conditions are given by

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

(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., 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 $\lambda$, then the
results are just those of the non-traded case in Section 3.1.

Note that 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. Under this setting, the moment conditions would be given by

$$E[h_t(\alpha, \beta)] = E \left[ \frac{1}{F_t} \otimes [R_t^e - \alpha - \beta F_t] \right] = 0,$$

(3.3)

where $\beta_{F,R} = \begin{bmatrix} \beta \\ I_K \end{bmatrix}$. Following the same procedure as in the general case, we specify an $A$ matrix and set $\Theta = \beta_{F,R}^T \Sigma^{-1}_{R^F} R^F$ with $R^F = \begin{bmatrix} R_t^e \\ F_t \end{bmatrix}$. Because we find that the GMM based on (3.3) leads to the same asymptotic variance covariance matrices for risk premiums as the GMM based on (3.2), we omit the GMM based on (3.3) in the rest of the paper and present results for the GMM based on (3.2).

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 to be used are

$$E[h_t(\alpha, \beta, \Phi)] = E \left[ \frac{1}{F_t^m} \otimes [R_t^e - \alpha - \beta m F_t^m] \right] = 0,$$

(3.4)

$$E[h_t(\alpha, \beta, \Phi, \lambda)] = E \left[ \frac{1}{F_t} \otimes [R_t^e - \alpha - \beta m F_t^m] \right] = 0,$$

$$\Phi R_t^e - \lambda^m = 0.$$
with \( F^m_t = \Phi R^e_t \). In this case, there are \( K(1 + N) + N(1 + K) + K \) moment conditions and parameters, which makes the system again 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. However, 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) of \( \lambda \) being zero or not and, accordingly, the properties of risk price estimates for \( \lambda \) have been studied and compared. Examples include Shanken (1992), Jagannathan and Wang (1998), Shanken and Zhou (2007), Kleibergen (2009), Lewellen, Nagel and Shanken (2010), Kan and Robotti (2011), Kan, Robotti and Shanken (2013).

In the current 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 \( \beta \) and \( \lambda \) for the three GMM estimators introduced in Sections 3.1 to 3.3. 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 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) standard deviations 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).

**Data for Empirical Results:** The asset data used in this paper consists of 25 portfolios formed by Fama-French (1992,1993), downloaded from Kenneth French’s website. These portfolios are value–weighted and formed from the intersections of five size and five book–to–market (B/M) portfolios and they include the stocks of the New York Stock Exchange, the American Stock Exchange, and NASDAQ. For details, we refer the reader to the Fama–French articles (1992,1993). The factors are the 3 factors of Fama-French (1992) (market, book–to–market and size). Our analysis is based on monthly data from
January 1963 until October 2012, i.e., we have 597 observations for each Fama–French portfolio.

The following theorem provides the limiting distribution of the historical averages estimator. It’s classical and provided for reference only.

**Theorem 4.1.** Given that $R^e_1, R^e_2, \ldots, R^e_T$ is a sequence of independent and identically distributed random vectors of excess returns, we have $\sqrt{T} (\bar{R}^e - E[R^e]) \xrightarrow{d} N(0, \Sigma_{R^eR^e})$.

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 the Asset Pricing models. Note that the joint distributions of $\lambda$ and $\beta$ are different for each set of moment conditions, which leads to different asymptotic distributions. 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 as in Section 3.1. Note that this result is valid for both traded and non-traded factors.

**Theorem 4.2.** Impose Assumptions 1 and 2, and consider the moment conditions (3.1)

\[
E[h_t(\alpha, \beta, \lambda)] = E \left[ \begin{array}{c}
1 \\
F_t \\
R^e_t - \beta \lambda \\
\end{array} \right] \otimes [R^e_t - \alpha - \beta F_t] = 0.
\]

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

\[
\Sigma_{R^eR^e} = (1 - \lambda' \Sigma_E^{-1} \lambda) \left( \Sigma_{\epsilon\epsilon} - \beta (\beta' \Sigma_{\epsilon\epsilon}^{-1} \beta)^{-1} \beta' \right).
\]
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 can be obtained. Second, it allows us to study the precision gains for estimating the risk premiums from incorporating the information about the factor model.

In case of a one–factor model and there is one–test asset, the (asymptotic) variances of both the naive risk–premium estimator and the factor–model based risk–premium estimator with (3.1) are the same. When more assets/portfolios are available, $N > 1$, observe that size of the asymptotic variances of risk–premium estimators depends on the magnitude of the prices of risk associated with the factor $\lambda$ (per unit variance of the factor), the exposures $\beta$, and $\Sigma_{\epsilon\epsilon}$. Note that the difference between the asymptotic covariance matrix of the naive estimator and the factor–based risk–premium estimator is $(1 - \lambda'\Sigma^{-1}_{FF}\lambda) \left( \Sigma_{\epsilon\epsilon} - \beta'(\Sigma^{-1}_{\epsilon\epsilon}\beta)^{-1}\beta' \right)$. In order to understand the efficiency gains from adding the information on the factor model, we will next analyse this formula. The following corollary formalizes the relation between the asymptotic covariance matrices of the naive estimator and the factor–model based risk–premium estimator.

**Corollary 4.1.** Impose Assumptions 1 and 2, and consider the moment conditions (3.1). Then, we have the following.

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

Corollary 4.1 shows that there may be precision gains for estimating risk premiums from the added information about the factor model if $\lambda'\Sigma^{-1}_{FF}\lambda$ is smaller than one. Note that although $\lambda'\Sigma^{-1}_{FF}\lambda$ can be larger than one mathematically, it is typically smaller than one given the parameters found in empirical research. Observe that in the one–factor case with a traded factor, $\lambda'\Sigma^{-1}_{FF}\lambda$ 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 \( \lambda' \Sigma_{FP}^{-1} \lambda = 0.06 \). Note that the smaller the value for \( \lambda' \Sigma_{FP}^{-1} \lambda \), the larger the efficiency gains from imposing a factor model.

As mentioned earlier, we study the empirical relevance of our results by using the parameter values from the FF 3-factor model estimated with FF 25 size–B/M portfolios. In particular, we estimate the parameters by using GMM with the moment conditions (3.1). We, then, 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). Comparing the standard deviation 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. In particular, using the 3–factor model in estimating risk premiums of 25 FF portfolios leads to striking gains in standard deviations with 32\% on average over assets.

4.2. Precision with Moment Conditions for Traded Factors

When the risk factors are traded, meaning that the factor 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 expect efficiency gains in estimating both the prices of risk and the expected (excess) returns. In this section, we consider such case and derive the asymptotic variances of the expected (excess) return estimators with the moment conditions for the case all factors are traded.

**Theorem 4.3.** Suppose that all factors are traded. Under Assumptions 1 and 2, consider the moment conditions (3.2)

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

14
Then, the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is given by

$$\Sigma_{R^eR^e} - (1 - \lambda'(\Sigma^{-1}_{F_F}\lambda))\Sigma_{ee}. \quad (4.2)$$

The theorem above shows that when the factors are traded, the asymptotic covariance matrices of the factor–based risk–premium estimators may change. This is because we incorporate, in the estimation, the restriction that prices of risk associated with factors equal to the expected return of that factor.

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 (3.2) to the one of the naive estimator, we observe that the difference is given by $(1 - \lambda'(\Sigma^{-1}_{F_F}\lambda))\Sigma_{ee}$. Moreover, observe that asymptotic covariance matrix of 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), which indicates that there may be efficiency gains from the information about the factors being traded. The following corollary formalizes these issues.

**Corollary 4.2.** Suppose that all factors are traded. Under Assumption 1 and 2, consider the GMM estimator based on the moment conditions (3.2). Then, we have the following.

1. If $\lambda'(\Sigma^{-1}_{F_F}\lambda) < 1$, then the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is at most $\Sigma_{R^eR^e}$.
2. The limiting variance of this expected (excess) return estimator is at most the limiting variance of the expected (excess) return estimator based on the moment conditions (3.1).

Plugging in the parameter estimates from the analysis of Fama–French model gives $\lambda'(\Sigma^{-1}_{F_F}\lambda) < 1 = 0.05$. Note that $\lambda'(\Sigma^{-1}_{F_F}\lambda) < 1$ is equal to 0.06 in the general case based on GMM 3.1. This happens because estimation based on GMM with the set of moment conditions 3.1 leads to $\lambda$ estimates which are different than $\lambda$ estimates obtained with GMM.
with 3.2. Comparing the standard deviations of the risk–premium estimates based on GMM with (3.2) to those of the naive estimators, we see that the risk–premium estimates based on GMM with (3.2) typically have smaller asymptotic standard deviations than the naive estimators. In particular, the size of efficiency gains in standard deviations is striking with 36% on average (over assets). Moreover, consistent with Theorem 4.2, the standard deviations of risk–premium estimates based on GMM with (3.1) typically exceed those of the naive estimator. Specifically, the risk–premium estimates based on GMM with (3.1) have, on average, 16% larger standard deviations than the risk–premium estimates based on GMM with (3.2). Overall, there are indeed sizeable precision gains from estimating risk premiums based on factor models based on 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 factors equal expected returns of the factors increases the preciseness 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 bring efficiency gains since the additional restriction on the price of the factor risk can be incorporated into the estimation. In this subsection, we derive the asymptotic variances of expected (excess) return estimators obtained with factor–mimicking portfolios.

**Theorem 4.4.** Under Assumption 1 and 2, consider the GMM estimator based on the moment conditions (3.4)

\[
E[h_t(\alpha^m, \beta^m, \Phi_0, \Phi, \lambda^m)] = E \left[ \begin{array}{c} \frac{1}{R_t^c} \\
\Phi R_t^c - \lambda^m \\
\end{array} \right] \otimes \left[ \begin{array}{c} [F_t - \Phi_0 - \Phi R_t^c] \\
[R_t^c - \alpha^m - \beta^m F_t^m] \\
\end{array} \right] = 0.
\]
Then, the limiting variance of the expected (excess) return estimator, $\hat{\beta}^m \hat{\lambda}^m$, is given by

$$\Sigma_{\text{Re-Re}} - (1 - \lambda' \left( \beta' \Sigma_{\text{Re-Re}}^{-1} \beta \right) \lambda) \left( \Sigma_{\text{Re-Re}} - \beta \left( \beta' \Sigma_{\text{Re-Re}}^{-1} \beta \right) \beta' \right).$$  \hspace{1cm} (4.3)

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 matrix of the risk-premium estimator based on GMM with (3.4) and the asymptotic covariance matrix of the naive estimator is given by

$$\left( 1 - \lambda' \left( \beta' \Sigma_{\text{Re-Re}}^{-1} \beta \right) \lambda \right) \left( \Sigma_{\text{Re-Re}} - \beta \left( \beta' \Sigma_{\text{Re-Re}}^{-1} \beta \right) \beta' \right).$$

Note that efficiency gains are dependent on this quantity being positive semi-definite or not. Although we haven’t found an answer to this yet, the results from our empirical analysis with FF model illustrates that there is considerable efficiency gains over naive estimation. In particular, estimating risk premiums with GMM (3.4) leads to, on average, 32% smaller standard deviations than estimating them with naive estimator. The gains in standard deviation is about 2% compared to the case where risk premiums are estimated with GMM (3.1). Note that small gains are expected in our particular empirical example because the three factors of Fama French are zero investment portfolios themselves.

5. Risk Premium Estimation with Omitted Factors

The asymptotic results in the previous section 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 independently and identically distributed returns. However, the pricing model may be misspecified and this might induce inconsistent risk-premium estimates. We investigate this issue and its solution in the present section.

We consider model misspecification due to ommitted 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

$$R^e = \alpha^* + \beta^* F + \delta^* G + \varepsilon^*$$  \hfill (5.1)

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

However, a researcher may be ignorant about the presence of the factors $G$ and thus estimates the model only with the set of factors, $F$,

$$R^e = \alpha + \beta F + \varepsilon$$  \hfill (5.2)

with $\varepsilon$ has mean–zero and $E[F\varepsilon'] = 0$ and estimates the exposures, $\beta$ and the prices of risk $\lambda$ by incorrectly specifying $E[R^e] = \beta\lambda$. Although the researcher might not know the underlying factor model exactly, she allows for misspecification by adding a constant term in estimation, $\alpha$ (see, e.g., Fama French (1993)).

The asymptotic bias in the parameter estimates for, $\alpha$, $\beta$ and $\lambda$ are presented in the following theorem:

**Theorem 5.1.** Assume that returns are generated by (5.1) but $\alpha$, $\beta$ and $\lambda$ are estimated from (5.2) with GMM (3.1). Then,

1. $\hat{\alpha}$ converges to $\alpha^* + (\beta^* - \beta)E[F] + \delta^* E[G]$,
2. $\hat{\beta}$ converges to $\beta^* + \delta^* \text{Cov}[G,F]^T \Sigma_F^{-1}$,
3. $\hat{\lambda}$ converges to $\lambda_F + (\beta^* \Sigma_{ee}^{-1}\beta)^{-1}\beta^* \Sigma_{ee}^{-1} [(\beta^* - \beta)\lambda_F + \delta^* \lambda_G]$

in probability.

The Lemma 5.1 shows that, if a researcher ignores some risk factors $G$, then the risk price estimators associated with factors $F$ are inconsistent if and only if

$$\beta^* \Sigma_{ee}^{-1} [(\beta^* - \beta)\lambda_F + \delta^* \lambda_G] \neq 0.$$
It is important to note that the inconsistency of the estimates of risk prices may be caused not only by the risk prices of omitted factors but also the bias in betas of the factors $F$. This result has an important implication: even if the ignored factors are associated with risk prices of zero, 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, which is often the case.

Next, we analyse the asymptotic bias in the parameter estimates for again, $\alpha$, $\beta$ and $\lambda$ but this time, in case the factors are traded and the estimation is based on GMM with moment conditions (3.2) of Section 3.2:

**Theorem 5.2.** Assume that returns are generated by (5.1) but $\alpha$, $\beta$ and $\lambda$ are estimated from (5.2) with GMM (3.2). Then,

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

in probability.

Theorem 5.2 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 $\beta$ disappears. Moreover, if the ignored factors are associated with zero prices of risk and uncorrelated with $F$, then the $\hat{\alpha}$ will converge to zero.

What happens to the risk–premium estimators on individual assets or portfolios if some true factors are ignored? The following corollary provides the consistency condition for risk–premium estimators of individual assets or portfolios.

**Corollary 5.1.** If the returns are generated by (5.1) and
The model (5.2) is estimated with GMM (3.1), then the vector of resulting risk-premium estimators \( \hat{\beta} \hat{\lambda} \) converges to \( E[R^e] \) if and only if
\[
I_N - \beta (\beta' \Sigma_{ee}^{-1} \beta)^{-1} \beta' \Sigma_{ee}^{-1} E[R^e] = 0.
\]

All factors are traded. If the model (5.2) 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 model (5.2) is estimated with GMM (3.1), the consistency of the risk-premium estimators is dependent on a specific condition that may not be satisfied. Moreover, if the factors are traded and the estimation is via GMM with moment conditions (3.2), then the risk-premium estimator obtained may be biased.

In order to capture misspecification, it is a common approach to add \( \hat{\alpha} \) to the model as in (5.2). In the following theorem, we will show that in case of traded factors, it is possible to achieve the consistency for estimating risk premiums.

**Theorem 5.3.** Assume that all factors in \( F \) are traded. If the returns are generated by (5.1) but the model (5.2) is estimated with GMM (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 \( E[R^e] \). However, the asymptotic variance of such estimator equals \( \Sigma_{R^e} \).

Theorem 5.3 shows that when all the factors in the estimation (\( F \)) are traded and if the estimation is based on GMM with moment conditions (3.2), then we obtain a consistent estimator for risk premiums by adding \( \hat{\alpha} \) to \( \hat{\beta} \hat{\lambda} \). However, this estimator is not asymptotically more efficient than the naive estimator of risk premiums.

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 \( E[R^e] \). In particular, \( \hat{\alpha} + \hat{\beta} \hat{\lambda} \) converges to \( E[R^e] - \beta(\lambda - E[F]) \) and \( \lambda - E[F] \) is not necessarily zero.
6. Application: Portfolio Choice with Parameter Uncertainty

This section analyzes the performances of portfolios based on different risk–premium estimates in the optimization problem of Markowitz (1952). The implementation of the mean–variance framework of Markowitz (1952) requires the estimation of first two moments of the asset returns. Mean–variance portfolios could be constructed by plugging in both factor–based risk–premium estimates or historical averages. Because we showed in previous sections that factor–model based risk–premium estimators are more precise than the naive estimator, the following questions arise: how is the performance of the mean–variance portfolio affected by the improvement in the precision of risk–premium estimates? To answer this, we analyze, in this section, the out of sample performances of the tangency portfolios based on the various risk–premium estimators in a simulation analysis.

**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 fully characterized by the expected return and variance of his selected portfolio, \( w \). The investor maximizes his expected utility, by choosing the vector of portfolio weights \( w \),

\[
E[U] = w'\mu^e - \frac{\gamma}{2} w'\Sigma_{RR}w, \quad (6.3)
\]

where \( \gamma \) measures the investor’s risk aversion level, \( \mu^e \) and \( \Sigma_{RR} \) denote the expected excess returns on the assets and covariance matrix of returns. The solution to the maximization problem above is given by \( w_{opt} = \frac{1}{2} \Sigma_{RR} \mu^e \). From this expression, the vector of tangency portfolio weights can be derived by incorporating the constraint that portfolio weights of risky assets sum to one and is given by

\[
w_{tg} = \frac{\Sigma_{RR} \mu^e}{\sqrt{N} \Sigma_{RR} \mu^e}, \quad (6.4)
\]

---

3Note that \( \Sigma_{RR} = \Sigma_{R^e - R^e} \).

4Because it lies on the mean variance frontier.
where $\iota_N$ is an $N$–vector of ones.

In the optimization problem above, since the true risk premium vector, $\mu^e$, and the true covariance matrix of asset returns, $\Sigma_{RR}$, are unknown, in empirical work, one needs to estimate them. Following the classical “plug in” approach, the moments of the excess return distribution, $\mu^e$ and $\Sigma_{RR}$, are replaced by their estimates.

**Portfolios Considered:** We consider four portfolios constructed with different risk–premium estimators: the tangency portfolio constructed with historical averages, the tangency portfolio constructed with the factor–model based estimates (GMM–Gen, GMM–Tr, GMM-Mim). Note that the covariance matrix is estimated using the traditional sample counterpart, $1/(T-1) \sum_1^T (R_t - \bar{R}_t)(R_t - \bar{R}_t)'$, where $\bar{R}_t$ is the sample average of returns. We also consider the global minimum variance portfolio\(^5\) 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.

**Performance Evaluation Criterion and Methodology:** We compare performances of the portfolios considered by using out-of-sample Sharpe Ratios. We set an initial window length over which we estimate the mean vector of excess returns and covariance matrix, and obtain the various portfolio weights. For our analysis, the initial window length is of 120 data points, corresponding to 10 years of data. We then calculate the one-month ahead returns, $\hat{w}_t R_{t+1}$, of the estimated portfolios. Next, we reestimate the portfolio weights by including the next month’s return and use this to calculate the return for the subsequent month. We continue doing this and obtain the time series of out–of–sample excess returns for each portfolio considered, from which we calculate the out–of–sample Sharpe ratios.

**Simulation Experiment:** We consider twenty–five Fama–French (1992) portfolios

\[^5\text{This portfolio is obtained by minimizing the portfolio variance with respect to the weights with the only constraint that weights sum to 1 and the N–vector of portfolio weights is given by } w_{\text{gmv}} = \Sigma_{RR\iota_N}/\iota_N\Sigma_{RR\iota_N}^2\]
sorted by size and book-to-market as risky assets and the nominal 1-month Treasury bill rate as a proxy for risk-free rate (both available on French’s website). We use the 3 Fama-French (1992) portfolios (market, book-to-market and size factors) as our factors. To make our simulations realistic, we calibrate the parameters by using the monthly data of the aforementioned portfolios, from January 1963 until December 2012. Specifically, we estimate $\alpha, \beta, \mu_F, \Sigma_{FF}, \Sigma_{\varepsilon \varepsilon}, \lambda$ and take them to be the truth in the simulation exercise to generate samples of 597 observations. To be precise, we use the following return-generating process:

$$R_t^e = \alpha + \beta F_t + \varepsilon_t, \quad t = 1, 2, \ldots, T, \quad (6.5)$$

with $F_t$ and $\varepsilon_t$ drawn from multivariate normal distributions with the true moments. Note that we set $\alpha$ equal to zero for all simulations. We simulate independent sets $Z = 5000$ return samples with the full sample size of 597. For each set of simulated sample, we calculate the out-of-sample Sharpe ratios for the various portfolios.

Table 3 provides the simulation results for the out-of-sample Sharpe ratios of different portfolios. In particular, we provide results on the tangency portfolios based on different risk-premium estimates and global minimum variance portfolios. Moreover, we provide the true Sharpe ratio of the tangency portfolio, which we refer as theoretical. For each portfolio, we present the average estimate over simulations, $\overline{SR}$, the bias as the percentage of the population Sharpe ratios, $(\overline{SR} - SR)/SR$ and the root-mean-square error (RMSE) in parentheses, the square root of $\sum_{s=1}^{Z} (\hat{SR}_s - SR)^2 / Z$, where $Z = 5000$.

In order to isolate the effect of the error in risk-premium estimates, we present our results with true and estimated $\Sigma_{RR}$. Firstly, note that the true Sharpe ratio of the tangency portfolio is superior to the portfolios based on estimated risk-premiums or covariance matrix of asset returns. Comparing the average Sharpe ratio of the tangency portfolio based on historical averages to the true Sharpe ratio of tangency portfolio, we see that the bias is striking and negative with $-56.26\%$ and $-56.88\%$, depending on the covariance matrix of asset returns is the true one or the estimated one. However, using
factor–models to estimate risk–premiums reduces the bias in Sharpe ratios substantially to a level ranging from $-18.02\%$ to $-26.25\%$. In particular, with GMM–Gen estimates, average Sharpe ratio of the tangency portfolio is 0.1541 in case of true covariance matrix (with an improvement of 69% over the average Sharpe ratios with the historical averages) and 0.1670 in case of an estimated covariance matrix (with an improvement of 90% over the average Sharpe ratios with the historical averages). Among the tangency portfolios constructed with factor–model based risk–premium estimates, GMM–Tr estimates perform, in terms of bias, the best given that the covariance matrix is known and GMM–Gen estimates perform, in terms of bias, the best given that the covariance matrix is estimated. However, the differences in biases are minimal for all tangency 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 tangency portfolios based on historical averages is extremely volatile across simulations. That is, it has a RMSE of 0.1353 (given the average estimate 0.0879) if the covariance matrix is estimated. However, using factor–based risk–premium estimators decreases the RMSEs substantially. Among the tangency portfolios based on factor–model based risk–premium estimators, GMM–Tr performs the best with a RMSE of 0.0881 (given the average estimate of 0.1668), as expected from our asymptotic analyses of risk–premium estimators in previous sections. However, the differences in RMSEs are minor among the portfolios with factor–based risk–premium estimates.

We also compare the performance of the tangency portfolio estimates to the global minimum variance portfolios. Jagannathan and Ma (2003) and De Miguel et al (2009) note that the estimation error in expected returns is so large that focusing on the minimum variance portfolios, which ignore the expected returns completely, is less sensitive to the estimation error than the mean–variance portfolios. In particular, it has been shown in empirical studies that minimum–variance portfolios usually has better out–of–sample performance than any other mean–variance portfolios.\footnote{DeMiguel et al.(2009), Jagannathan and Ma(2003), Jorion(1985, 1986, 1991)} Consistent with them, we find
Working paper

that global minimum variance portfolio with an estimated covariance matrix has a higher average Sharpe ratio, 0.0984 and substantially lower RMSE, 0.0469, compared to the tangency portfolios constructed with historical averages. However, the average Sharpe ratios of the tangency portfolios are considerably larger than the average Sharpe ratio of the global minimum variance portfolio when the factor–based risk–premium estimates are used. Specifically, average Sharpe ratios of the GMM–Gen, GMM–Tr, GMM–Mim are 0.1670, 0.1668, 0.1670 respectively. Overall, using the factor–model based risk–premium estimators improves the performance of tangency portfolios substantially over the plug in estimates of historical averages, in terms of both bias and RMSEs. Moreover, in contrast to the tangency portfolios with historical averages, these portfolios perform considerably better than the global minimum variance portfolio.

7. Conclusions

It has been the standard technique in the literature to use average historical returns as estimates of expected excess returns, that is risk premiums, on individual assets or portfolios. 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. Moreover, we show that using a misspecified asset pricing model in the sense that some factors are forgotten generally leads to inconsistent estimates. However, in case the factors are traded, then adding an alpha to the model capturing mispricing leads to consistent estimators.

Out of sample performance of tangency portfolios significantly improve if factor–based estimates of risk premium are used in portfolio weights instead of the classical historical averages.
A. Proofs

In the rest of the paper, the covariance matrix of the factor–mimicking portfolios is denoted by \( \Sigma_{FmFm} \).

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

\[
E[M] = E[M^m], \tag{A.2}
\]

\[
\text{Cov}[M, R^e] = \text{Cov}[M^m, R^e]. \tag{A.3}
\]

Thus, we have

\[
\beta \lambda = \text{Cov}[R^e, F' \Sigma^{-1}_{FF} \left( -\frac{1}{E[M]} \Sigma_{FF} b \right) \tag{A.4}
\]

\[
= -\frac{1}{E[M]} \text{Cov}[R^e, F'] b
\]

\[
= -\frac{1}{E[M^m]} \text{Cov}[R^e, F^m'] b
\]

\[
= -\frac{1}{E[M^m]} \text{Cov}[R^e, F^m'] \Sigma^{-1}_{FmFm} \Sigma_{FmFm} b
\]

\[
= \beta^m \lambda^m,
\]

which completes the proof. \( \square \)

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 weights 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 $h_t(\theta) \in \mathbb{R}^q$, independently and identically distributed over time. Given a prespecified matrix $A \in \mathbb{R}^{p \times q}$, its consistent estimator $\hat{A}$ and $\hat{A}_T \sum_{t=1}^{T} h_t(\hat{\theta}) = 0$,

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

where,

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

$$S = E[h_t(\theta)h_t(\theta)']. \quad (A.7)$$

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

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

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

with

$$V = \begin{bmatrix} 1 + \mu_F'\Sigma^{-1}_{FF}\mu_F & -\mu_F'\Sigma^{-1}_{FF} \sigma_E & \otimes \Sigma_{\xi\xi} & V_c \\ -\Sigma^{-1}_{FF}\mu_F & \Sigma^{-1}_{FF} & V_c' \end{bmatrix}$$

$$V_c' = \begin{bmatrix} (1 + \lambda'\Sigma^{-1}_{FF}\lambda)(\beta'\Sigma^{-1}_{\xi\xi}\beta)^{-1} + \Sigma_{FF} \end{bmatrix}$$

where $\mu_F = E[F_t]$ and $V_c = \begin{bmatrix} 1 + \mu_F'\Sigma^{-1}_{FF}\lambda & \otimes \beta(\beta'\Sigma^{-1}_{\xi\xi}\beta)^{-1} \end{bmatrix}$.

Proof. The proof follows from plugging the appropriate matrices for the moment condi-
tions 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,

\[
S = \begin{bmatrix}
\Sigma_{\varepsilon\varepsilon} & \mu_F' \otimes \Sigma_{\varepsilon\varepsilon} & \Sigma_{\varepsilon\varepsilon} \\
\mu_F' \otimes \Sigma_{\varepsilon\varepsilon} & \left[ \Sigma_{\varepsilon F} + \mu_F \mu_F' \right] \otimes \Sigma_{\varepsilon\varepsilon} & \mu_F' \otimes \Sigma_{\varepsilon\varepsilon} \\
\Sigma_{\varepsilon\varepsilon} & \mu_F' \otimes \Sigma_{\varepsilon\varepsilon} & \beta \Sigma_{\varepsilon F} \beta' + \Sigma_{\varepsilon\varepsilon}
\end{bmatrix}.
\]

\[
J(\theta) = \mathbb{E} \left[ \frac{\partial h_t(\theta)}{\partial \theta} \right] = \begin{bmatrix}
-1 & \mu_F' \\
\mu_F & \Sigma_{\varepsilon F} + \mu_F \mu_F' \\
0_{N \times N} & -\lambda' \otimes I_N
\end{bmatrix} \otimes I_N \quad 0_{N(K+1) \times K}.
\]

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 \( \theta \) is obtained by performing the matrix multiplications \([AJ]^{-1}ASA'[J'A']^{-1}\).

Lemma A.3. Suppose that all factors are traded. Then, under Assumptions 1, 2 and the moment conditions (3.2) with parameter vector \( \theta = (\alpha', \text{vec}(\beta)' , \lambda)' \), we have

\[
\sqrt{T}(\hat{\theta} - \theta) \overset{d}{\to} \mathcal{N}(0, V),
\]

(A.9)

with

\[
V = \begin{bmatrix}
1 + \mu_F' \Sigma_{\varepsilon F}^{-1} \mu_F & -\mu_F' \Sigma_{\varepsilon F}^{-1} \\
-\Sigma_{\varepsilon F}^{-1} \mu_F & \Sigma_{\varepsilon F}^{-1}
\end{bmatrix} \otimes \Sigma_{\varepsilon\varepsilon} 0_{N(K+1) \times K}
\]

\[
0_{K \times N(K+1)} \quad \Sigma_{\varepsilon F} \quad \Sigma_{\varepsilon F}
\]

Proof. The proof follows from plugging the appropriate matrices for the moment con-
ditions (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,

\[
S = \begin{bmatrix}
\Sigma_{\varepsilon \varepsilon} & \mu_F' \otimes \Sigma_{\varepsilon \varepsilon} & 0_{N \times K} \\
\mu_F \otimes \Sigma_{\varepsilon \varepsilon} & [\Sigma_{F \varepsilon} + \mu_F \mu_F'] \otimes \Sigma_{\varepsilon \varepsilon} & 0_{NK \times K} \\
0_{K \times N} & 0_{K \times NK} & \Sigma_{F F}
\end{bmatrix},
\]

and

\[
J(\theta) = \begin{bmatrix}
-1 & \mu_F' \\
-\mu_F & \Sigma_{F F} + \mu_F \mu_F' \\
0_{K \times N(K+1)} & I_K
\end{bmatrix} \otimes I_{N(K+1) \times K}.
\]

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

The next lemma provides the asymptotic properties of the GMM estimator with factor-mimicking portfolios.

**Lemma A.4.** Given that Assumption 1, 2 are satisfied and that (2.8)–(2.10) hold, then under the moment conditions (3.4), for \( \theta = (\text{vec}(\beta^m)', \lambda^m)' \), we have

\[
\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} \mathcal{N}(0, V), \quad \text{(A.10)}
\]

with

\[
V = \begin{bmatrix}
\Sigma_{F F}^{-1} \otimes \beta^m & -\Sigma_{F F}^{-1} \mu_F' \otimes \beta^m \\
-\mu_F' \Sigma_{F F}^{-1} \otimes \Sigma_{\varepsilon \varepsilon} \beta^m & \mu_F' \Sigma_{F F}^{-1} \mu_F' \Sigma_{\varepsilon \varepsilon} + \Sigma_{F F}^{-1} F F
\end{bmatrix}.
\]

**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

\[
S = \begin{bmatrix}
1 & \mu'_R e \\
\mu_R e & \Sigma_{RR} + \mu_R e \mu'_R e \\
0_{N(K+1) \times K(1+N)} & 0_{K \times K(1+N)}
\end{bmatrix} \otimes \Sigma_{ee} \begin{bmatrix}
1 & \mu'_F m \\
\mu_F m & \Sigma_{Fm} F_m + \mu_F m \mu_F m' \\
0_{N(1+N) \times N(1+N)} & 0_{K \times N(1+N)}
\end{bmatrix} \otimes \Sigma_{m m'} \begin{bmatrix}
0_{K(1+N) \times N(1+N)} & 0_{K(1+N) \times K}
\end{bmatrix},
\]

and from (A.6), we have

\[
J(\theta) = E \begin{bmatrix}
-1 & R_t' e' \\
R_t e' & R_t e' R_t e' \\
0_{K \times 1} \Phi(R_t e' R_t e') & \beta^m \\
0_K & R_t e' \otimes I_K
\end{bmatrix} \otimes I_K \begin{bmatrix}
0_{K(1+N) \times N(1+N)} & 0_{K(1+N) \times K}
\end{bmatrix} - I_K
\]

with \( A = I_K(1+N) + N(K+1) + K \). Thus, the limiting variance of the GMM estimator for \( \theta = (\text{vec}(\beta^m)', \lambda^m)' \) 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

\[
(\Phi_0', \text{vec}(\Phi)', \alpha^m, \text{vec}(\beta^m)', \lambda^m')
\]  

(A.11)

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

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.

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

$$ \text{vec} \left( \hat{\beta}' \right)' \lambda' - \text{vec} \left( \beta' \right)' \lambda' \rightarrow \mathcal{N}(0, V_{\beta, \lambda}), \quad (A.12) $$

we have, by applying the delta method, that

$$ \sqrt{T} \left( g(\hat{\beta}, \hat{\lambda}) - g(\beta, \lambda) \right) \rightarrow \mathcal{N}(0, \hat{g}' V_{\beta, \lambda} \hat{g}), \quad (A.13) $$

with

$$ \hat{g} = \left[ \lambda' \otimes I_N \quad \beta \right]. $$

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 $\hat{g}' V_{\beta, \lambda} \hat{g}$.  

The following theorem follows from the formula of partitioned inverses.

**Lemma A.5.** Let

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

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

31
Proof of Theorem 4.4. We are interested in \( g(\beta^m, \lambda^m) = \beta^m \lambda^m \). Given
\[
\text{vec} \left( \hat{\beta}^m \right)' \hat{\lambda}^m - (\text{vec} (\beta^m)' \lambda^m)' \overset{d}{\rightarrow} \mathcal{N}(0, V_{\beta^m,\lambda^m}), \tag{A.14}
\]
Then, by applying the delta method, we have
\[
\sqrt{T} (g(\hat{\beta}^m, \hat{\lambda}^m) - g(\beta^m, \lambda^m)) \overset{d}{\rightarrow} \mathcal{N}(0, \dot{g}' V_{\beta^m,\lambda^m} \dot{g}) \tag{A.15}
\]
and note that here \( \dot{g} = \left[ \lambda^m \otimes I_N \beta^m \right] \)
Then, we have
\[
\dot{g}' V_{\hat{\beta}^m, \hat{\lambda}^m} \dot{g} = \lambda^m \Sigma_F \Sigma_{\pi F m}^{-1} \lambda^m \Sigma_{\pi \epsilon m} + \beta^m \Sigma_F \Sigma_{\pi F m} \beta^m \\
+ (\mu_\pi^{-1} \Phi \Sigma_{\pi R m}^{-1} \Phi \mu_\pi - \lambda^m \Sigma_F \Sigma_{\pi F m}^{-1} \lambda^m) \beta^m \Sigma_{\mu \pi} \beta^m \tag{A.16}
\]
\[
= (1 - \lambda^m \Sigma_{\pi F m}^{-1}) \left[ \Sigma_{\pi \epsilon} - \beta (\beta' \Sigma_{\pi \epsilon}^{-1} \beta)^{-1} \beta' \right] \tag{A.17}
\]
The result follows from plugging the \( \beta^m \) and \( \Phi \) respectively into the above equation. \( \square \)

Proofs of Corollaries 4.1 and 4.2. We need to study the difference between

1. The limiting variance of the historical averages and the limiting variance of the expected (excess) return estimator based on (3.1), referring to Corollary 4.1
2. The limiting variance of the historical averages and the limiting variance of the expected (excess) return estimator based on (3.2), referring to the Corollary 4.2
3. 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), referring to the Corollary 4.2

Suppose \( 1 - \lambda' \Sigma_{\pi F m}^{-1} \lambda < 1 \) and \( 1 - \lambda' \beta' \Sigma_{\pi R m}^{-1} \beta \lambda < 1 \). In the following, we will show that the differences between the limiting variance above are positive semi–definite. Lemma A.5 will be used to establish the positive semi–definiteness of the differences.

To prove 1, we need to study the difference
\[
\Sigma_{R^+ R^+} - (\Sigma_{R^+ R^+} - (1 - \lambda' \Sigma_{\pi F m}^{-1} \lambda) \left[ \Sigma_{\epsilon \epsilon} - \beta (\beta' \Sigma_{\epsilon \epsilon}^{-1} \beta)^{-1} \beta' \right]) \tag{A.18}
\]
\[
= (1 - \lambda' \Sigma_{\pi F m}^{-1} \lambda) \left[ \Sigma_{\epsilon \epsilon} - \beta (\beta' \Sigma_{\epsilon \epsilon}^{-1} \beta)^{-1} \beta' \right]
\]
32
In order to show that $\Sigma_{ee} - \beta(\beta'\Sigma_{ee}^{-1}\beta)^{-1}\beta'$ is positive semi–definite, we will use Lemma A.5. Now, define $K_1 = \Sigma_{ee}^{1/2}$ and $K_2 = \beta'\Sigma_{ee}^{-1/2}$. Then,

$$K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \begin{bmatrix} K_1' & K_2' \\ K_2' & K_2'' \end{bmatrix} = \begin{bmatrix} K_1K_1' & K_1K_2' \\ K_2K_1' & K_2K_2' \end{bmatrix}$$

so that

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

Then, Lemma A.5 yields that

$$\Sigma_{ee} - \beta(\beta'\Sigma_{ee}^{-1}\beta)^{-1}\beta' \geq 0 \quad (A.19)$$

To prove 2 referring to Corollary 4.2, we need to study the difference

$$\Sigma_{R^cR^c} - (\Sigma_{R^cR^c} - (1 - \lambda'(\Sigma_{F^c}^{-1}\lambda)) \Sigma_{ee})$$

$$= (1 - \lambda'(\Sigma_{F^c}^{-1}\lambda)) \Sigma_{ee} \quad (A.20)$$

Since $\Sigma_{ee}$ is positive semi-definite, the first part of the Corollary 4.2 follows. In order to prove the second part of the Corollary 4.2, we need to show that $\beta(\beta'\Sigma_{ee}^{-1}\beta)^{-1}\beta'$ is positive semi–definite. Since $\Sigma_{ee}$ is positive semi–definite, $\Sigma_{ee}^{-1}$ is also positive semi–definite. There exists a positive semi–definite matrix $Z$ such that $Z^2 = \Sigma_{ee}^{-1}$. Then, $((Z\beta)'(Z\beta)) = \beta'\Sigma_{ee}^{-1}\beta$ and it is positive semi–definite. The result follow from applying the same property of positive semi–definite matrices once more for $\beta'\Sigma_{ee}^{-1}\beta$.

Finally, concerning 4, we need to study the difference

$$\Sigma_{R^cR^c} - (\Sigma_{R^cR^c} - (1 - \lambda'(\beta'\Sigma_{R^cR^c}^{-1}\beta')\lambda) \left[\Sigma_{R^cR^c} - \beta(\beta'\Sigma_{R^cR^c}^{-1}\beta')^{-1}\beta'\right])$$

$$= (1 - \lambda'(\beta'\Sigma_{R^cR^c}^{-1}\beta')\lambda) \left[\Sigma_{R^cR^c} - \beta(\beta'\Sigma_{R^cR^c}^{-1}\beta')^{-1}\beta'\right] \quad (A.21)$$

Positive definiteness of $[\Sigma_{R^cR^c} - \beta(\beta'\Sigma_{R^cR^c}^{-1}\beta')^{-1}\beta']$ follows from Lemma A.5.  

\[\square\]
Proof of Theorem 5.1. Note that \( \hat{\beta} \) converges to \( \beta \) and \( \hat{\alpha} \) converges to \( \alpha \) in probability.

\[
\beta = \text{Cov} \left[ R^*, F^T \right] \Sigma_{FF}^{-1},
\]
(A.22)

\[
= \text{Cov} \left[ R^* = \alpha^* + \beta^* F + \delta^* G + \epsilon^*, F^T \right] \Sigma_{FF}^{-1},
\]

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

Now, note that

\[
\alpha = \mathbb{E} \left[ R^* \right] - \beta \mathbb{E} \left[ F \right],
\]
(A.23)

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

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

Furthermore, for \( \hat{\lambda} \), first notice that

\[
\hat{\lambda} = \left( \hat{\beta}^T \hat{\Sigma}^{-1} \hat{\beta} \right)^{-1} \hat{\beta}^T \hat{\Sigma}^{-1} \mathbb{E} [R^*]
\]
(A.24)

Then, observe following equality

\[
\hat{R}^* = \hat{\beta} \lambda_F + (\hat{R}^* - \mathbb{E} [R^*]) - (\hat{\beta} - \beta) \lambda_F + (\beta^* - \beta) \lambda_F + \delta^* \lambda_G
\]
(A.25)

Multiplying both sides by \( \left( \hat{\beta}^T \hat{\Sigma}^{-1} \hat{\beta} \right)^{-1} \hat{\beta}^T \hat{\Sigma}^{-1} \) gives

\[
\hat{\lambda} = \lambda_F + \left( \hat{\beta}^T \hat{\Sigma}^{-1} \hat{\beta} \right)^{-1} \hat{\beta}^T \hat{\Sigma}^{-1} \left[ (\hat{R}^* - \mathbb{E} [R^*]) - (\hat{\beta} - \beta) \lambda_F \right] + \left( \hat{\beta}^T \hat{\Sigma}^{-1} \hat{\beta} \right)^{-1} \hat{\beta}^T \hat{\Sigma}^{-1} [(\beta^* - \beta) \lambda_F + \delta^* \lambda_G]
\]
(A.26)

Hence, the probability limit of \( \hat{\lambda} \) from GMM (3.1) is given by
Proof of Theorem 5.2. Note that \( \hat{\beta} \) converges to \( \beta \) and \( \hat{\alpha} \) converges to \( \alpha \) in probability.

\[
\beta = \text{Cov} \left[ R^c, F^T \right] \Sigma_{FF}^{-1},
\]

(A.28)

Now, note that

\[
\alpha = E \left[ R^c \right] - \beta E \left[ F \right],
\]

(A.29)

\[
= \beta^* \lambda_F + \delta^* \lambda_G - \beta \lambda_F,
\]

\[
= (\beta^* - \beta) \lambda_F + \delta^* \lambda_G.
\]

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

Proof of Corollary 5.1. Proof of the first part of the corollary: Note that

\[
\beta \lambda_F = \beta \left( \beta^* \Sigma_{ee}^{-1} \beta \right)^{-1} \beta^* \Sigma_{ee}^{-1} \text{E} \left[ R^c \right]
\]

(A.30)

Note that \( \hat{\beta} \hat{\lambda} \) is consistent for \( E \left[ R^c \right] \) if and only if

\[
E \left[ R^c \right] = \beta \left( \beta^* \Sigma_{ee}^{-1} \beta \right)^{-1} \beta^* \Sigma_{ee}^{-1} \text{E} \left[ R^c \right]
\]

which is equivalent to

\[
\left[ I_N - \beta \left( \beta^* \Sigma_{ee}^{-1} \beta \right)^{-1} \beta^* \Sigma_{ee}^{-1} \right] E \left[ R^c \right] = 0
\]

(A.31)

To prove the second part of the corollary, note that \( \hat{\beta} \hat{\lambda} \) converges to \( \beta \lambda \). Using A.28 and \( \lambda_F = E \left[ F \right] \) (based on 5.2), we have

\[
\beta \lambda_F = (\beta^* + \delta^* \text{Cov} \left[ G, F^T \right] \Sigma_{FF}^{-1}) \lambda_F,
\]

(A.32)

\[
= E \left[ R^c \right] - (\beta^* - \beta) \lambda_F + \delta^* \lambda_G.
\]
Proof of Theorem 5.3. Consistency of $\hat{\alpha} + \beta \lambda_F$ is straightforward. The asymptotic variance is given by the delta method for the function $g$. Assume $g(\alpha, \beta, \lambda_F) = \alpha + \beta \lambda_F$. The asymptotic covariance matrix of $\alpha, \beta \text{ and } \gamma$ is given in Lemma A.3 (denoted by $V$).

we have, by applying the delta method, that

$$\sqrt{T} \left( g(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) - g(\alpha, \beta, \lambda) \right) \xrightarrow{d} N(0, \check{g}'V_{\alpha, \beta, \lambda}\check{g}), \quad (A.33)$$

with

$$\check{g} = \left[ \begin{array}{c} 1 \\ \lambda' \otimes I_N \\ \beta \end{array} \right].$$

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

□


Jagannathan, Ravi and Wang, Zhenyu (1998). An asymptotic theory for estimating beta pricing...


Table 1: Efficiency Gains for Factor–Model Based Risk Premium Estimators

<table>
<thead>
<tr>
<th>Adjusted</th>
<th>RP with GMM–Gen</th>
<th>RP with GMM–Tr</th>
<th>RP with GMM–Mim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive</td>
<td>32</td>
<td>36</td>
<td>32</td>
</tr>
<tr>
<td>RP with GMM–Gen</td>
<td>-</td>
<td>16</td>
<td>2</td>
</tr>
</tbody>
</table>

Notes: This table presents the average gains in standard deviations for the various risk premium estimates. The test assets are the 25 Fama–French size and book–to–market portfolios and the factors are the three factors of Fama French (1992). The first row illustrates the gains for three different factor–model based risk–premium estimates (GMM–Gen, GMM–Tr and GMM–Mim) over the historical averages. The table presents the average gains over 25 assets and all numbers are in percentages.
## Table 2: Efficiency Gains for Factor Based Risk Premium Estimates for 25 FF Assets

<table>
<thead>
<tr>
<th>Assets</th>
<th>Gen–N</th>
<th>Tr–N</th>
<th>Mim–N</th>
<th>Tr–Gen</th>
<th>Mim–Gen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4374</td>
<td>0.5305</td>
<td>0.4390</td>
<td>0.2750</td>
<td>0.0339</td>
</tr>
<tr>
<td>2</td>
<td>0.4269</td>
<td>0.4980</td>
<td>0.4284</td>
<td>0.2359</td>
<td>0.0331</td>
</tr>
<tr>
<td>3</td>
<td>0.3404</td>
<td>0.4001</td>
<td>0.3416</td>
<td>0.1990</td>
<td>0.0264</td>
</tr>
<tr>
<td>4</td>
<td>0.3179</td>
<td>0.3665</td>
<td>0.3190</td>
<td>0.1738</td>
<td>0.0246</td>
</tr>
<tr>
<td>5</td>
<td>0.2348</td>
<td>0.2933</td>
<td>0.2356</td>
<td>0.1711</td>
<td>0.0182</td>
</tr>
<tr>
<td>6</td>
<td>0.2751</td>
<td>0.3856</td>
<td>0.2759</td>
<td>0.2066</td>
<td>0.0213</td>
</tr>
<tr>
<td>7</td>
<td>0.2536</td>
<td>0.3313</td>
<td>0.2544</td>
<td>0.2067</td>
<td>0.0196</td>
</tr>
<tr>
<td>8</td>
<td>0.2169</td>
<td>0.2797</td>
<td>0.2176</td>
<td>0.1725</td>
<td>0.0168</td>
</tr>
<tr>
<td>9</td>
<td>0.2394</td>
<td>0.2714</td>
<td>0.2402</td>
<td>0.1243</td>
<td>0.0185</td>
</tr>
<tr>
<td>10</td>
<td>0.2250</td>
<td>0.2534</td>
<td>0.2257</td>
<td>0.1138</td>
<td>0.0174</td>
</tr>
<tr>
<td>11</td>
<td>0.2808</td>
<td>0.3668</td>
<td>0.2817</td>
<td>0.2271</td>
<td>0.0218</td>
</tr>
<tr>
<td>12</td>
<td>0.2705</td>
<td>0.3125</td>
<td>0.2714</td>
<td>0.1510</td>
<td>0.0209</td>
</tr>
<tr>
<td>13</td>
<td>0.3044</td>
<td>0.3239</td>
<td>0.3054</td>
<td>0.1059</td>
<td>0.0236</td>
</tr>
<tr>
<td>14</td>
<td>0.3005</td>
<td>0.3174</td>
<td>0.3015</td>
<td>0.0978</td>
<td>0.0233</td>
</tr>
<tr>
<td>15</td>
<td>0.3127</td>
<td>0.3299</td>
<td>0.3137</td>
<td>0.1006</td>
<td>0.0242</td>
</tr>
<tr>
<td>16</td>
<td>0.3174</td>
<td>0.3542</td>
<td>0.3184</td>
<td>0.1498</td>
<td>0.0246</td>
</tr>
<tr>
<td>17</td>
<td>0.3234</td>
<td>0.3385</td>
<td>0.3244</td>
<td>0.0951</td>
<td>0.0250</td>
</tr>
<tr>
<td>18</td>
<td>0.3458</td>
<td>0.3619</td>
<td>0.3470</td>
<td>0.1009</td>
<td>0.0268</td>
</tr>
<tr>
<td>19</td>
<td>0.3244</td>
<td>0.3507</td>
<td>0.3254</td>
<td>0.1269</td>
<td>0.0251</td>
</tr>
<tr>
<td>20</td>
<td>0.3713</td>
<td>0.3935</td>
<td>0.3725</td>
<td>0.1221</td>
<td>0.0288</td>
</tr>
<tr>
<td>21</td>
<td>0.2257</td>
<td>0.2460</td>
<td>0.2264</td>
<td>0.0953</td>
<td>0.0175</td>
</tr>
<tr>
<td>22</td>
<td>0.3368</td>
<td>0.3615</td>
<td>0.3379</td>
<td>0.1245</td>
<td>0.0261</td>
</tr>
<tr>
<td>23</td>
<td>0.4259</td>
<td>0.4640</td>
<td>0.4274</td>
<td>0.1694</td>
<td>0.0330</td>
</tr>
<tr>
<td>24</td>
<td>0.3868</td>
<td>0.4591</td>
<td>0.3881</td>
<td>0.2308</td>
<td>0.0300</td>
</tr>
<tr>
<td>25</td>
<td>0.5039</td>
<td>0.5596</td>
<td>0.5058</td>
<td>0.2173</td>
<td>0.0390</td>
</tr>
</tbody>
</table>

Notes: This table illustrates the gains in standard deviations for the various risk premium estimates for 25 portfolios formed by Fama–French (1992,1993). The factors are the three factors (market, size and book–to–market) of Fama–French (1992). The results are based on monthly data from January 1963 until October 2012, i.e. 597 observations for each portfolio. The first column (Gen–N) presents the gains of the factor–model based estimates of risk premiums based on GMM with 3.1 over the naive estimate of historical averages. The second and third columns present the gains of factor model based estimates of risk premiums based on GMM with 3.2 and with 3.4 over naive estimates respectively. Fourth column corresponds to the gains from estimating the system based on GMM with 3.2 over the case of estimating the system based on GMM with 3.2. The last column presents the gains from making use of mimicking portfolios and estimate the system with GMM (3.4) over estimation with GMM (3.1)
Table 3: Tangency out of Sample Sharpe Ratio Estimates with different risk-premium estimates

<table>
<thead>
<tr>
<th></th>
<th>True $\mu^*$</th>
<th>Estimated $\Sigma_{RR}$</th>
<th>Theoret</th>
</tr>
</thead>
<tbody>
<tr>
<td>True $\Sigma$</td>
<td>0.2129</td>
<td>0.2038</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td>0.0188</td>
<td>-0.0249</td>
<td></td>
</tr>
<tr>
<td>Naive</td>
<td>0.0914</td>
<td>0.0879</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td>-0.5626</td>
<td>-0.5688</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1375)</td>
<td>(0.1353)</td>
<td></td>
</tr>
<tr>
<td>GMM–Gen</td>
<td>0.1541</td>
<td>0.1670</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td>-0.2625</td>
<td>-0.1802</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0977)</td>
<td>(0.0867)</td>
<td></td>
</tr>
<tr>
<td>GMM–Tr</td>
<td>0.1552</td>
<td>0.1668</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td>-0.2574</td>
<td>-0.1814</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0966)</td>
<td>(0.0881)</td>
<td></td>
</tr>
<tr>
<td>GMM–Mim</td>
<td>0.1541</td>
<td>0.1670</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td>-0.2625</td>
<td>-0.1802</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0966)</td>
<td>(0.0867)</td>
<td></td>
</tr>
<tr>
<td>GMV</td>
<td>0.0974</td>
<td>0.0984</td>
<td>0.2000</td>
</tr>
<tr>
<td></td>
<td>(0.0461)</td>
<td>(0.0469)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table provides, average sharpe ratio estimate over simulations, its percentage error, compared to the true sharpe ratio and the RMSE (in parenthesis) of the sharpe ratios constructed with various mean estimates. Note that the variance covariance matrix is estimated by the sample variance covariance matrix.