Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Linear Factor Models and the Estimation of Expected Returns

Cisil Sarisoy, Peter de Goeij and Bas Werker

DP 03/2016-020

Working paper

Linear Factor Models and the Estimation of Expected Returns

Cisil Sarisoy[a,*], Peter de Goeij[b], Bas J.M. Werker[c]

[a] *Department of Finance, Kellogg School of Management, Northwestern University*
[b] *Department of Finance, Tilburg University*
[c] *Department of Econometrics and Department of Finance, CentER, Tilburg University*

Abstract

Linear factor models of asset pricing imply a linear relationship between expected returns of assets and exposures to one or more sources of risk. We show that exploiting this linear relationship leads to statistical gains of up to 31% in variances when estimating expected returns on individual assets over historical averages. When the factors are weakly correlated with assets, i.e. β's are small, and the interest is in estimating expected excess returns, that is risk premiums, on individual assets rather than the prices of risk, the Generalized Method of Moment estimators of risk premiums does lead to reliable inference, i.e. limiting variances suffer from neither lack of identification nor unboundedness. 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 documents that using the more precise estimates of expected returns based on factor–models rather than the historical averages translates into significant improvements in the out–of–sample performances of the optimal portfolios.

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

JEL: C13, C38, G11

We thank Torben G. Andersen, Jules van Binsbergen, Joost Driessen, Thijs van der Heijden, Frank de Jong, Frank Kleibergen, Tristan Linke, Eric Renault, Enrique Sentana and participants at the SoFiE conference 2015 and seminar participants at Tilburg University for their helpful comments.

March 3, 2016

1. Introduction

Estimating expected returns on individual assets or portfolios is perhaps one of the longest standing challenges in finance. The 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. One needs very long samples for accurate estimates, which are often 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 analyze the estimation of the expected (excess) returns on individual assets or on portfolios, i.e., the product of exposures (β) and risk prices (λ). The tremendous literature on asset pricing provides potential estimators of the expected returns on individual assets and, as mentioned by Black (1993), these theoretical restrictions can help to improve the estimates of expected returns.

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 anal-

*Corresponding author.
Email addresses: cisil.sarisoy@kellogg.northwestern.edu (Cisil Sarisoy), p.c.degoeij@Tilburguniversity.nl (Peter de Goeij), b.j.m.werker@tilburguniversity.edu (Bas J.M. Werker)

ysis for both estimators, evaluating the implications of weakly correlated and omitted factors in the estimation of expected (excess) returns.

In this paper, we first provide the asymptotic statistical properties of expected excess return, that is risk–premium, estimators based on factor models of asset pricing. These limiting distributions are useful for obtaining the standard errors and, accordingly, the confidence bounds of the risk–premium estimators of individual assets or portfolios. Secondly, we 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 analyzing the precision gains over historical averages. We show when exploiting the linear relationship implied by linear factor models indeed leads to more precise estimates of expected returns over historical averages, see Corollaries (4.1-4.2). In an empirical analysis, for instance when estimating risk–premiums on 25 Fama and French (1992) size and book–to–market sorted portfolios, we document large improvements in variances of up to 31% for individual portfolios.

Second, we analyze the inference issue in estimating risk premiums in the presence of weakly correlated factors. When the factors are weakly correlated with assets, i.e. β's are small, the confidence bounds of the price of risk estimates are erroneous (see Kleibergen (2009)), which makes it difficult to make statistical inference about a specific hypothesis. The effects may be severe in empirical research, as these confidence bounds may be unbounded as documented for the case of consumption CAPM of Lettau and Ludvigson (2001)[1]. We find that such issue does not exist if the interest is in making inferences about the risk premiums on individual assets rather than the prices of risk attached to factors. In particular, the limiting variances of the risk premium estimators are not affected if the β's are small whereas the limiting variances of the risk price estimators may be unbounded.

Third, 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

[1]See also Bryzgalova (2014), Burnside (2015) on the role of weakly identified factors for making inferences about the prices of risk.

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 6.1). 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 faces: allow for misspecification and loose efficiency or run the risk of misspecification and gain efficiency.

Expected returns are not only interesting in the sense of single quantities for individual assets but they are also crucial inputs for theoretical formulations in various subfields of finance, i.e. calculations of cost of capital or valuation of cash flows. From an asset pricing perspective, the most prominent presence of expected returns is in portfolio allocation problems. We analyze the economic implications of the efficiency gains from using factor model–based estimates of expected returns in Markowitz (1952)'s setting.

Implementation of mean–variance framework of Markowitz (1952) in practice require the estimation of first two moments of asset returns. Constructing optimal portfolios with imprecise estimates of expected returns, via historical averages, and the sample covariance matrix leads 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 global minimum portfolio as they are not subject to estimation risk of expected returns (DeMiguel et al. (2009b)).[3] The mean—variance

[2]See, for example, Frost and Savarino (1988), Michaud (1989), Jobson and Korkie (1980, 1981), Best and Grauer (1991), and Litterman (2003).

[3]Several studies provide solutions on improving the covariance matrix estimates (see, e.g., Ledoit and

optimal portfolio weights can also be constructed with more precise factor–based risk–premium estimates instead of the "naive" estimates (historical averages). Accordingly, we investigate the out–of–sample performances of optimal portfolios based on factor–model based risk premium estimates in a simulation analysis. We document that the average out–of–sample Sharpe ratios of the optimal portfolios improves strikingly, with an improvement of up to 64%, if factor–model based estimates are used as estimators of risk premiums instead of historical averages. Moreover, optimal portfolios based on factor–model based estimates perform better than both the global minimum variance portfolio and 1/N strategy portfolio. Lastly, the average out–of–sample Sharpe ratios of the factor model based optimal portfolios are much more precise and significant compared to the ones based on historical averages.

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 7 documents the simulation analysis for portfolio optimization, while Section 8 concludes. All proofs are gathered in the appendix.

Wolf (2003), DeMiguel et al. (2009a) 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 has a much larger impact on the optimal portfolio weights compared to the imprecision in covariance estimates (DeMiguel et al. (2009b)).

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_i^e

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

Linear factor models additionally specify $M = a + b'F$, where $F = (F_1, ..., F_K)'$ is a vector of K factors (see, e.g., Cochrane (2001), p.69). Note that (2.1) can be written in matrix notation using the vector of excess returns $R^e = (R_1^e, ..., R_N^e)'$. 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:*

(a) *The covariance matrix of excess returns $\Sigma_{R^e R^e}$ has full rank N,*
(b) *The covariance matrix of factors Σ_{FF} has full rank K,*
(c) *The covariance matrix between excess returns and factors $Cov\left[R^e, F'\right]$ has full rank K.*

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

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

where

$$\beta = \mathrm{Cov}\left[R^e, F'\right] \Sigma_{FF}^{-1}, \tag{2.3}$$

$$\lambda = -\frac{1}{\mathrm{E}\left[M\right]} \Sigma_{FF} b. \tag{2.4}$$

Thus, (2.2) specifies a linear relationship between risk premiums, $\mathrm{E}\left[R^e\right]$, and the exposures β of the assets to the risk factors, F, with prices λ.

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

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

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*

$$E[\varepsilon_t|F_t] = 0, \tag{2.6}$$

$$Var[\varepsilon_t|F_t] = \Sigma_{\varepsilon\varepsilon}, \tag{2.7}$$

where $\Sigma_{\varepsilon\varepsilon}$ 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 $K \times N$ matrix Φ of constants and a K–vector of random variables u_t satisfying

$$F_t = \Phi_0 + \Phi R^e_t + u_t, \tag{2.8}$$

$$\mathrm{E}\left[u_t\right] = 0_{K\times 1}, \tag{2.9}$$

$$\mathrm{E}\left[u_t R^{e\prime}_t\right] = 0_{K\times N}, \tag{2.10}$$

and we define the factor–mimicking portfolios by

$$F^m_t = \Phi R^e_t. \tag{2.11}$$

We then obtain an alternative formulation of the linear factor model by replacing the original factors with factor–mimicking portfolios[4]

$$R_t^e = \alpha^m + \beta^m F_t^m + \varepsilon_t^m, \quad t = 1, 2, \ldots, T. \tag{2.12}$$

Recall that using the projection results, Φ and β are related by

$$\Phi = \Sigma_{FF} \beta' \Sigma_{R^e R^e}^{-1}, \tag{2.13}$$

while β^m and β satisfy

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

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, λ^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

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.

[4]Huberman et al. (1987), Kandel and Stambaugh (1995), Balduzzi and Robotti (2008), Hou and Kimmel(2010)

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 β and factor prices λ are

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

The first set of moment conditions identifies α and β as the regression coefficients, while the last set of 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 $A\mathrm{E}\left[h_t(\alpha,\beta,\lambda)\right] = 0$, where

$$A = \left[\begin{array}{cc} I_{N(1+K)} & 0_{N(1+K)\times N} \\ 0_{K\times(KN+N)} & \Theta_{K\times N} \end{array}\right].$$

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 $\Theta = \beta^T \Sigma_{\varepsilon\varepsilon}^{-1}$. 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, 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 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 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

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

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., 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.

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

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

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 $\Theta = \beta^T_{F,R}\Sigma^{-1}_{R^F R^F}$ with $R^F_t = \begin{bmatrix} R^e_t \\ F_t \end{bmatrix}$. 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

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

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 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 λ being zero or not and, accordingly, the properties of risk price estimates for λ have been studied and compared. Examples include Shanken (1992), Jagannathan (1998),603–621 Kleibergen (2009), Lewellen, Nagel and Shanken (2010), Kan and Robotti (2011), Kan et al. (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 β and λ 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) 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).

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 $R_1^e, R_2^e, ..., R_t^e$ is a sequence of independent and identically distributed random vectors of excess returns, 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 the 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. 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)*

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

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

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

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 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 λ (per unit variance of the factor), the exposures β, and $\Sigma_{\varepsilon\varepsilon}$. Note that the difference between the asymptotic covariance matrix of the naive estimator and the factor–based risk–premium estimator is $\left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\left(\Sigma_{\varepsilon\varepsilon} - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\right)$. 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 $\lambda'\Sigma_{FF}^{-1}\lambda < 1$, then the limiting variance of the expected (excess) return estimator $\hat{\beta}\hat{\lambda}$ is at most $\Sigma_{R^eR^e}$.*

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. Note that although $\lambda'\Sigma_{FF}^{-1}\lambda$ 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, $\lambda'\Sigma_{FF}^{-1}\lambda$ 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 $\lambda'\Sigma_{FF}^{-1}\lambda = 0.058$. Note that the smaller the value for $\lambda'\Sigma_{FF}^{-1}\lambda$, 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). 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 striking gains in variances of up to 25%.

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)*

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

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

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

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 the expected return of those factors.

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 $\left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\Sigma_{\varepsilon\varepsilon}$. 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 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* $\Sigma_{R^eR^e}$.
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 con-*

ditions (3.1).

Plugging in the parameter estimates from the analysis of Fama–French model gives $\lambda'\Sigma_{FF}^{-1}\lambda < 1 = 0.052$. Note that $\lambda'\Sigma_{FF}^{-1}\lambda < 1$ is equal to 0.058 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 λ estimates which are different than λ estimates obtained with GMM with (3.2). Comparing the variances of the risk–premium estimates based on GMM with (3.2) to those of the naive estimators (see Table 1, we see that the risk–premium estimates based on GMM with (3.2) typically have smaller asymptotic variances than the naive estimators. In particular, the size of efficiency gains is considerably large for all individual portfolios, and goes up to 31% . Moreover, consistent with Theorem 4.2, the asymptotic variances of risk–premium estimates based on GMM with (3.1) typically exceed those of the risk premium estimators based on GMM with (3.2). Specifically, the risk–premium estimates based on GMM with (3.1) have up to 7.6% larger variances than the risk–premium estimates based on GMM with (3.2). Overall, the sizeable 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 factors equal expected returns of the factors increases the 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 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 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)*

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

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

$$\begin{aligned} \Sigma_{R^eR^e} - &\left(\mu'_{R^e} \left\{ \Sigma^{-1}_{R^eR^e} - \Sigma^{-1}_{R^eR^e}\beta \left(\beta'\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \beta'\Sigma^{-1}_{R^eR^e} \right\} \mu_{R^e}\right) \\ &\times \left(\Sigma_{R^eR^e} - \beta \left(\beta'\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \Sigma^{-1}_{FF} \left(\beta\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \beta'\right) \\ &- \left(1 - \mu'_{R^e}\Sigma^{-1}_{R^eR^e}\mu_{R^e}\right) \left(\Sigma_{R^eR^e} - \beta \left(\beta'\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \beta'\right) \end{aligned} \tag{4.3}$$

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 matrices of the naive estimator and the factor–model based GMM risk–premium estimator with (3.4) is given by

$$\begin{aligned} \mu'_{R^e} &\left\{ \Sigma^{-1}_{R^eR^e} - \Sigma^{-1}_{R^eR^e}\beta \left(\beta'\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \beta'\Sigma^{-1}_{R^eR^e} \right\} \mu_{R^e} \\ &\times \left(\Sigma_{R^eR^e} - \beta \left(\beta'\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \Sigma^{-1}_{FF} \left(\beta\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \beta'\right) \\ &+ \left(1 - \mu'_{R^e}\Sigma^{-1}_{R^eR^e}\mu_{R^e}\right) \left(\Sigma_{R^eR^e} - \beta \left(\beta'\Sigma^{-1}_{R^eR^e}\beta\right)^{-1} \beta'\right). \end{aligned} \tag{4.4}$$

Efficiency gains with respect to the historical averages estimator are dependent on Eqn. (4.4) being positive semi–definite or not. Although, we were not able to prove this formally yet, the results from our empirical analysis with FF-3 factor model illustrates that there is 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 31%, smaller variances than estimating them with the naive estimator. Moreover, we find that estimating risk premiums by making use of the mimicking portfolios lead to efficiency losses over the estimation based on the general case, i.e, GMM (3.1) for all assets, ranging between 0.1% and 1.5%.

Note that one important difference between Theorem 4.4 and Theorem 4.2 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. Shanken (1992) correction in asymptotic variances for generated regressors. We should recall here that GMM standard errors automatically accounts for such effects as it solves the system of moment conditions 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 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 factors, 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). These losses range between 1% and 8.4% across portfolios.

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

A number of papers in the literature documents inference issues regarding the prices of risk when the factors are weakly correlated with the asset returns (see, e.g. Kleibergen (2009), Burnside (2015), Bryzgalova (2014), Kleibergen and Zhan (2015)). . When β's are close to zero and/or when β matrix is almost of reduced rank, the confidence bounds of the prices of risk estimates are erroneous, which leads to unreliable statistical inference in favor or against any hypothesis. The effects may be severe in empirical research, as the confidence bounds of the risk price estimates may be unbounded as documented for the case of conditional consumption CAPM of Lettau and Ludvigson (2001), see Kleibergen (2009). [5] Accordingly, Kleibergen (2009) provides identification-robust statistics and confidence sets for the risk price estimates when the β's are small.

The above-cited papers document the misleading statistical inference about the risk price estimates when the β's are small and the results are very useful to understand the significance of the pricing impact of involved factors, in particular to analyze how strong the relationship between expected returns and the candidate risk factors. However, once our interest is in estimating risk premiums on individual assets or portfolios, a natural question would be if similar effects exist for making inferences about risk premiums in the presence of weakly correlated factors. We should remember here that the focus for inference is on the multiplication of β and λ rather than λ only. In this section, we provide some results to shed light on this issue.

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

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

Stock (1997)), and Kleibergen (2009), we will assume a sequence of β's getting smaller as the sample size increases.

Remark 5.1. Suppose Assumption 1 (a), (b) and Assumption 2 hold and suppose $\beta = \frac{1}{\sqrt{T}}B$, where B is a fixed full rank $N \times K$ matrix.

- Consider the GMM estimator based on the moment conditions (3.1). The limiting variance of the expected (excess) return estimator, (4.1), depends on the space spanned by the columns of B and is not affected if the basis considered, β, has a weak value in the form of $\beta = \frac{1}{\sqrt{T}}B$.
- Suppose that all factors are traded and consider the GMM estimator based on the moment conditions (3.2). The limiting variance of the expected (excess) return estimator, (4.2), is not affected by the value of β.
- Consider the GMM estimator based on the moment conditions (3.4). The limiting variance of the expected (excess) return estimator, (4.3), depends on the space spanned by the columns of B and is not affected if the basis considered, β, has a weak value in the form of $\beta = \frac{1}{\sqrt{T}}B$.

Remark 5.2. Suppose Assumption 1 (a), (b) and Assumption 2 hold and suppose $\beta = \frac{1}{\sqrt{T}}B$, where B is a fixed full rank $N \times K$ matrix.

- Consider the GMM estimator based on the moment conditions (3.1). The limiting variance of the risk price estimator, $(1+\lambda'\Sigma_{FF}^{-1}\lambda)(\beta'\Sigma_{\varepsilon\varepsilon}\beta^{-1})^{-1}+\Sigma_{FF}$ is unbounded.
- Suppose that all factors are traded and consider the GMM estimator based on the moment conditions (3.2). The limiting variance of the risk price estimator, Σ_{FF}, is not affected by the value of β.

Remark 5.1 documents two important findings of our analysis regarding the issue of small but non-zero β. First, if the parameters of the linear factor model in focus are estimated with GMM based on the moment conditions (3.1), the limiting variances of the risk premium estimators does not suffer from either lack of identification or the unboundedness when the β has a weak value, i.e. $\beta = \frac{1}{\sqrt{T}}B$. In particular, $\frac{1}{\sqrt{T}}$ term cancels out in the limiting variance (4.1), and hence the limiting variance only depends on the space spanned by the columns of B. Since B is a full column $N \times K$ matrix, the limiting variance (4.1) is identified and does not blow up or shrink. However, the first bullet point of Remark 5.2 documents that this is not the case if one is interested in making inference about the prices of risk, λ. Specifically, it highlights that the limiting variances of the risk price estimators blow up when $\beta = \frac{1}{\sqrt{T}}B$. 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 in empirical studies, see Kleibergen (2009).

The second finding of Remark 5.1 considers 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 weak value or not. This is a straightforward result in the sense that the asymptotic variance (4.2) is independent of the value of β. Moreover, the limiting variances of the price of risk estimators, λ, in this case remains unaltered if β has a weak value.

The third finding of Remark 5.1 considers the small β issue for the estimation based on factor–mimicking portfolios. The finding is consistent with the previous two cases and the limiting variances of the risk premium estimators does not suffer from either lack of identification or the unboundedness when the β has a weak value, i.e. $\beta = \frac{1}{\sqrt{T}}B$.

The bottomline of this section is the following: if one is interested in making statistical inferences about the prices of risk, the small but non-zero β's may have detrimental effects on the inference regarding the FM, GLS two pass estimators (see, e.g.,Kleibergen (2009)) and the GMM estimators considered in this paper. However, once the interest is in estimating risk premiums, expected (excess) returns, based on the linear factor model, the limiting variances of the risk premium estimators based on GMM with (3.1), (3.2) or (3.4) suffers from neither the identification nor the unboundedness when the β has a weak value.

6. 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

$$R^e = \alpha^* + \beta^* F + \delta^* G + \varepsilon^* \tag{6.1}$$

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 be ignorant about the presence of the factors G and thus estimates the model only with the set of factors, F,

$$R^e = \alpha + \beta F + \varepsilon \tag{6.2}$$

with ε has mean–zero and $\mathrm{E}\left[F\varepsilon'\right] = 0$ and estimates the exposures, β and the prices of risk λ by incorrectly specifying $\mathrm{E}\left[R^e\right] = \beta\lambda$. Although the researcher might not know the underlying factor model exactly, she allows for misspecification by adding an N-vector of constant terms in estimation, α as in Fama and French (1993).

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

Theorem 6.1. *Assume that returns are generated by (6.1) but α, β and λ are estimated from (6.2) with GMM (3.1). Then,*

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

Lemma 6.1 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

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

It is important to note that the inconsistency of the estimates of risk prices may be caused not only by the risk prices 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.

Next, we analyse the asymptotic bias in the parameter estimates for again, α, β and λ 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 6.2. *Assume that returns are generated by (6.1) but α, β and λ are estimated from (6.2) with GMM (3.2). Then,*

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

in probability.

Theorem 6.2 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 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 6.1. *If the returns are generated by (6.1) and*

- *the model (6.2) 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* $[I_N - \beta(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}]E[R^e] = 0$.
- *all factors are traded. If the model (6.2) 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 (6.2) 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 an N–vector of constant terms, α, to the model as in (6.2). 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 6.3. *Assume that all factors in F are traded. If the returns are generated by (6.1) but the model (6.2) 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 $\Sigma_{R^e R^e}$.*

Theorem 6.3 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 an estimator for the N–vector of constant terms, $\hat{\alpha}$, to $\hat{\beta}\hat{\lambda}$. However, this estimator is not asymptotically more efficient than the naive estimator of risk premiums.

Some asset pricing studies add a one dimensional constant, henceforth λ_0, to the asset pricing specification of expected returns as in $\mathrm{E}[R^e] = 1_N \lambda_0 + \beta\lambda$, where 1_N is an N–vector of ones and make inferences about it. At this stage, we do not analyze the role of such objects. Recall that here α is an N–vector of constants; it does not represent a one dimensional object as λ_0.

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 not necessarily zero.

7. Application: Portfolio Choice with Parameter Uncertainty

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

The implementation of the mean–variance framework of Markowitz (1952) requires the estimation of first two moments of the asset returns. Although in the setting of Markowitz (1952), optimal portfolios are supposed to achieve the best performance, in practice, the estimation error in expected returns via the historical averages leads to large deterioration of the out–of–sample performance of the optimal portfolios (see, e.g.,DeMiguel et al. (2009b)). In the far end, this has led to simply abandoning the application of theoretically optimal decisions and using the naive techniques such as the $1/N$ portfolio or the global minimum variance portfolios these are not subject to estimation risk of expected returns. 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.

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,

$$\mathrm{E}\left[U\right] = w'\mu^e - \frac{\gamma}{2} w'\Sigma_{RR} w, \tag{7.3}$$

where γ measures the investor's risk aversion level, μ^e and Σ_{RR}[6] denote the expected excess returns on the assets and covariance matrix of returns. The solution to the maximization problem above is given by

$$w_{opt} = \frac{1}{\gamma}\Sigma_{RR}\mu^e. \tag{7.4}$$

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. Following the classical "plug in" approach, the moments of the excess return distribution, μ^e and Σ_{RR}, are replaced by their estimates.

Portfolios Considered: We consider four portfolios constructed with different risk–premium estimators: the optimal portfolio constructed with historical averages, the optimal portfolios constructed with the three factor model–based GMM risk premium estimates with moment conditions (3.1), (3.2) and (3.4). Note that the covariance matrix

[6]Note that Σ_{RR}=$\Sigma_{R^eR^e}$.

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 (GMV thereafter) portfolio[7] 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. Morover, we analyze the performance of the $1/N$ portfolio.

Performance Evaluation Criterion and Methodology: We compare performances of the portfolios by using their out-of-sample Sharpe Ratios[8]. We provide results both for "enlarging windows" and "rolling windows".

- *Enlarging Windows*: 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-period ahead returns, $\hat{w}_t R^e_{t+1}$, of the estimated portfolios. Next, we re–estimate the portfolio weights by including the next period's return and use this to calculate the portfolio return for the subsequent period. 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.

- *Rolling Windows*: We start with an initial window length of 120 observations over which we estimate the mean vector of excess returns, and obtain the various portfolio weights. We then calculate the one-period ahead returns, $\hat{w}_t R^e_{t+1}$, of the estimated portfolios. Next, we re–estimate the portfolio weights by including the next period's return and dropping the first period's return, and use this to

[7]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 $w_{gmv} = \Sigma_{RR}\iota_N/\iota_N\Sigma_{RR}\iota_N$

[8]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.

calculate the subsequent period's portfolio return. We obtain a time series out–of–sample excess returns for all the portfolios considered, and obtain the out–of–sample Sharpe ratios.

Simulation Setting: We use the following return–generating process:

$$R_t^e = \alpha + \beta F_t + \varepsilon_t, \quad t = 1, 2, \ldots, T, \tag{7.5}$$

with F_t and ε_t drawn from multivariate normal distributions under the null of $\mathrm{E}\left[R_t^e\right] = \beta\lambda$. To make our simulations realistic, we calibrate the parameters for the return–generating process by using the monthly data from January 1963 until October 2012 on twenty–five Fama–French (1992) portfolios 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, and the 3 Fama-French (1992) portfolios (market, book–to–market and size factors) as the risk factors. 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. We simulate independent sets of $Z = 10000$ return samples and for each set of simulated sample, we calculate the out–of–sample Sharpe ratios for the various portfolios.

Table 2 provides the simulation results for the out–of–sample Sharpe ratios of different portfolios. In particular, we provide results on the optimal portfolios based on different risk–premium estimates, GMV and 1/N portfolios. Moreover, we provide the true Sharpe ratio of the optimal portfolio, which we refer to as theoretical. 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) in parantheses, the square root of $\sum_{s=1}^{Z}(\hat{SR}_s - SR)/Z$, (third line) , where $Z = 10000$.

In order to isolate the effect of the error in risk–premium estimates, we present our results with true and estimated Σ_{RR}. Firstly, note that the true Sharpe ratio of the optimal portfolio is superior to the portfolios based on estimated risk–premiums or co-variance 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 (rolling samples), we see that the bias is striking and negative with -41.6% (-58.5%) and -44.4% (-62.3%), depending on whether 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 -12.5% $(-28.2\%)$to -11.2% (-26.1%) and ranging from -9.1% (-22.2%) to -7.7% (-19.7%) depending on true or estimated covariance matrices. In particular, with GMM–Gen estimates, average Sharpe ratio of the optimal portfolio is 0.188 (0.154) in case of true covariance matrix (with an improvement of 50% over the average Sharpe ratios with the historical averages) and 0.195 (0.167) in case of an estimated covariance matrix (with an improvement of 64% over the average Sharpe ratios with the historical averages). Among the optimal portfolios constructed with factor–model based risk–premium estimates, the one based on GMM–Tr estimates perform the best with 0.198 (0.172). However, the differences in biases 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 0.108 (given the average estimate 0.119) if the covariance matrix is estimated. The situation gets worse if the optimization is based on rolling samples, with a RMSE of 0.142 (given the average estimate 0.081). However, using factor-model based risk–premium estimators decreases the RMSEs substantially. Among the optimal portfolios based on factor–model based risk–premium estimators, GMM–Tr performs the best with a RMSE of 0.052 (given the average estimate of 0.198), as expected from the 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.

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 1/N strategy and slightly better than the GMV portfolio when the optimization is based on the enlarging samples, and performs considerably worse than both the GMV portfolios and 1/N strategy in case of rolling samples. 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 (See, e.g., Jagannathan and Ma (2003), De Miguel et al (2009) and Jorion (1985, 1986, 1991)). However, the average Sharpe ratios for all optimal portfolios based on factor model–based risk premium estimates are considerably larger than both the GMV and 1/N porfolios, with an improvement ranging from 13% to 46% for the case of enlarging samples. 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.

8. 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. These estimators are very noisy. This translates into the need for very large, in practice, mostly infeasible, samples of data in order to gain 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 (MKT, SMB, HML) setting with 25 FF portfolios, the efficiency gains are sizable and go up to 31% for indivudual portfolios. For real life applications, this translates into the benefit of using only 69% 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 estiamators are not affected, whereas the confidence bounds of the risk price estimators may be unbounded. We also 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 optimal portfolios significantly improves if factor–model 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 $\Sigma_{F^m F^m}$.

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

$$\begin{aligned} \mathrm{E}\left[M\right] &= \mathrm{E}\left[M^m\right], & \text{(A.2)} \\ \mathrm{Cov}\left[M, R^e\right] &= \mathrm{Cov}\left[M^m, R^e\right]. & \text{(A.3)} \end{aligned}$$

Thus, we have

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

which completes the proof. □

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

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

where,

$$
\begin{aligned}
J &= E\left[\frac{\partial h_t(\theta)}{\partial \theta'}\right], \qquad (A.6)\\
S &= E\left[h_t(\theta)h_t(\theta)'\right]. \qquad (A.7)
\end{aligned}
$$

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* $\theta = (\alpha', vec(\beta)', \lambda')'$*, 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 $[AJ]^{-1}ASA'[J'A']^{-1}$. $\square$

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

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

with

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

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,

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

and

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

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

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

with

$$V = \begin{bmatrix} \Sigma_{F^mF^m}^{-1}\Phi\Sigma_{R^eR^e}\Phi'\Sigma_{F^mF^m}^{-1} \otimes \beta^m\Sigma_{uu}\beta^{m\prime} + \Sigma_{F^mF^m}^{-1} \otimes \Sigma_{\varepsilon^m\varepsilon^m} & -\Sigma_{F^mF^m}^{-1}\mu_{F^m} \otimes \beta^m\Sigma_{uu} \\ -\mu_{F^m}'\Sigma_{F^mF^m}^{-1} \otimes \Sigma_{uu}\beta^{m\prime} & \mu_{R^e}'\Sigma_{R^eR^e}^{-1}\mu_{R^e}\Sigma_{uu} + \Sigma_{F^mF^m} \end{bmatrix}.$$

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

$$S = \begin{bmatrix} \begin{bmatrix} 1 & \mu_{R^e}' \\ \mu_{R^e} & \Sigma_{R^eR^e} + \mu_{R^e}\mu_{R^e}' \end{bmatrix} \otimes \Sigma_{uu} & 0_{K(1+N)\times N(K+1)} & 0_{K(1+N)\times K} \\ 0_{N(K+1)\times K(1+N)} & \begin{bmatrix} 1 & \mu_{F^m}' \\ \mu_{F^m} & \Sigma_{F^mF^m} + \mu_{F^m}\mu_{F^m}' \end{bmatrix} \otimes \Sigma_{\varepsilon^m\varepsilon^m} & 0_{N(K+1)\times K} \\ 0_{K\times K(1+N)} & 0_{K\times N(K+1)} & \Sigma_{F^mF^m} \end{bmatrix},$$

and from (A.6), we have

$$J(\theta) = \mathrm{E}\begin{bmatrix} -\begin{bmatrix} 1 & R_t^{e\prime} \\ R_t^e & R_t^e R_t^{e\prime} \end{bmatrix} \otimes I_K & 0_{K(1+N)\times N(K+1)} & 0_{K(1+N)\times K} \\ -\begin{bmatrix} 0 & R_t^{e\prime} \\ 0_{K\times 1} & \Phi(R_t^e R_t^{e\prime}) \end{bmatrix} \otimes \beta^m & -\begin{bmatrix} 1 & F_t^{m\prime} \\ F_t^m & F_t^m F_t^{m\prime} \end{bmatrix} \otimes I_N & 0_{N(K+1)\times K} \\ 0_K \quad R_t^{e\prime} \otimes I_K & 0_{K\times N(K+1)} & -I_K \end{bmatrix},$$

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 $\theta = (\mathrm{vec}\,(\beta^m)', \lambda^{m\prime})'$ 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 $(\mathrm{vec}\,(\beta)', \lambda')'$. 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

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

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

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

Remember that Lemma A.2 and A.3 give the asymptotic distributions of $\sqrt{T}(\hat{\theta}-\theta)$ where $\theta=(\alpha',\mathrm{vec}\,(\beta)',\lambda')'$ 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 $(\mathrm{vec}\,(\beta)',\lambda')'$ and calculating $\dot{g}'V_{\beta,\lambda}\dot{g}$. □

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

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

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

$$\dot{g}=\left[\begin{array}{cc}\lambda^{m\prime}\otimes I_N & \beta^m\end{array}\right]$$

Then, we have

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

The result follows from plugging the β^m and Φ respectively into the above equation via Eqn. 2.14 and Eqn. 2.13. □

The following lemma follows from Schur complement condition.

Lemma A.5. *Let*

$$\mathbf{K}=\left[\begin{array}{cc}K_{11} & K_{12}\\ K_{21} & K_{22}\end{array}\right]$$

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

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

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,

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

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

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

□

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

$$\begin{aligned}
&\Sigma_{R^eR^e} - \left(\Sigma_{R^eR^e} - \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\Sigma_{\varepsilon\varepsilon}\right) \\
&= \left(1 - \lambda'\Sigma_{FF}^{-1}\lambda\right)\Sigma_{\varepsilon\varepsilon}
\end{aligned}$$

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 that

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

is positive semi–definite. The result follows from $\Sigma_{\varepsilon\varepsilon}$ being positive semi–definite. □

Proof of Theorem 6.1. Note that $\hat{\beta}$ converges to β and $\hat{\alpha}$ converges to α in probability.

$$\begin{aligned}
\beta &= \mathrm{Cov}\left[R^e, F^T\right]\Sigma_{FF}{}^{-1}, \\
&= \mathrm{Cov}\left[R^e = \alpha^* + \beta^*F + \delta^*G + \varepsilon^*, F^T\right]\Sigma_{FF}{}^{-1}, \\
&= \beta^* + \delta^*\mathrm{Cov}\left[G, F^T\right]\Sigma_{FF}^{-1}.
\end{aligned} \tag{A.12}$$

Now, note that

$$\begin{aligned}
\alpha &= \mathrm{E}\left[R^e\right] - \beta\mathrm{E}\left[F\right], \\
&= \alpha^* + \beta^*\mathrm{E}\left[F\right] + \delta^*\mathrm{E}\left[G\right] - \beta\mathrm{E}\left[F\right], \\
&= \alpha^* + (\beta^* - \beta)\mathrm{E}\left[F\right] + \delta^*\mathrm{E}\left[G\right].
\end{aligned} \tag{A.13}$$

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

$$\hat{\lambda} = \left(\hat{\beta}'\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\hat{\beta}\right)^{-1}\hat{\beta}'\hat{\Sigma}_{\varepsilon\varepsilon}^{-1}\bar{R}^{e} \tag{A.14}$$

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

$$\begin{aligned} \lambda &= \left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\left[\beta^{*}\lambda_F + \delta^{*}\lambda_G\right] && \text{(A.15)} \\ &= \lambda_F + \left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\left[(\beta^{*}-\beta)\lambda_F + \delta^{*}\lambda_G\right] && \text{(A.16)} \end{aligned}$$

□

Proof of Theorem 6.2. Note that $\hat{\beta}$ converges to β and $\hat{\alpha}$ converges to α in probability.

$$\begin{aligned} \beta &= \text{Cov}\left[R^{e}, F^{T}\right]\Sigma_{FF}{}^{-1}, && \text{(A.17)} \\ &= \text{Cov}\left[R^{e} = \alpha^{*} + \beta^{*}F + \delta^{*}G + \varepsilon^{*}, F^{T}\right]\Sigma_{FF}{}^{-1}, \\ &= \beta^{*} + \delta^{*}\text{Cov}\left[G, F^{T}\right]\Sigma_{FF}^{-1}. \end{aligned}$$

Now, note that

$$\begin{aligned} \alpha &= \text{E}\left[R^{e}\right] - \beta\text{E}\left[F\right], && \text{(A.18)} \\ &= \alpha^{*} + \beta^{*}\lambda_F + \delta^{*}\lambda_G - \beta\lambda_F, \\ &= \alpha^{*} + (\beta^{*}-\beta)\lambda_F + \delta^{*}\lambda_G. \end{aligned}$$

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

□

Proof of Corollary 6.1. Proof of the first part of the corollary: Note that

$$\beta\lambda_F = \beta\left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\text{E}\left[R^{e}\right] \tag{A.19}$$

Note that $\hat{\beta}\hat{\lambda}$ is consistent for $\text{E}\left[R^{e}\right]$ if and only if $\text{E}\left[R^{e}\right] = \beta\left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\text{E}\left[R^{e}\right]$ which is equivalent to

$$\left[I_N - \beta\left(\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\beta\right)^{-1}\beta'\Sigma_{\varepsilon\varepsilon}^{-1}\right]\text{E}\left[R^{e}\right] = 0 \tag{A.20}$$

To prove the second part of the corollary, note that $\hat{\beta}\hat{\lambda}$ converges to $\beta\lambda$. Using A.17 and $\lambda_F = \mathrm{E}[F]$ (based on 6.2), we have

$$\begin{aligned} \beta\lambda_F &= (\beta^* + \delta^* \mathrm{Cov}\left[G, F^T\right] \Sigma_{FF}^{-1})\lambda_F, \qquad \text{(A.21)} \\ &= \mathrm{E}[R^e] - ((\beta^* - \beta)\lambda_F + \delta^* \lambda_G). \end{aligned}$$

□

Proof of Theorem 6.3. 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 α, β and γ is given in Lemma A.3 (denoted by V).

we have, by applying the delta method, that

$$\sqrt{T}\left(g(\hat{\alpha}, \hat{\beta}, \hat{\lambda}) - g(\alpha, \beta, \lambda)\right) \xrightarrow{d} \mathcal{N}(0, \dot{g}' V_{\alpha,\beta,\lambda} \dot{g}), \qquad \text{(A.22)}$$

with

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

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

□

Balduzzi, Pierluigi and Robotti, Cesare (2008). Mimicking portfolios, economic risk premia, and tests of multi-beta models. *Journal of Business & Economic Statistics*, 26, 354–368.
11

Best, Michael J and Grauer, Robert 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.
4

Black, Fischer (1993). Estimating expected return. *Financial Analysts Journal*, 36–38.
2

Douglas T. Breeden (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7, 265–296.
2, 7

Broadie, Mark (1993). Computing efficient frontiers using estimated parameters. *Annals of Operations Research: Special Issue on Financial Engineering*, 45, 21–58.

Bryzgalova, Svetlana (2014). Spurious factors in linear asset pricing models. *Working Paper.*
3, 20

Burnside, Craig (2015). Identification and inference in linear stochastic discount factor models with excess returns. *Journal of Financial Econometrics.*
3, 20

Chopra, Vijay K and Ziemba, William T (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19, 6–11.
5

Cochrane, John H (2001). *Asset pricing.* New Jersey: Princeton University Press.
6, 9

DeMiguel, Victor, Garlappi, Lorenzo, Nogales, Francisco J and Uppal, Raman (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55, 5, 798–812.
5

DeMiguel, Victor, Garlappi, Lorenzo, and Uppal, Raman (2009). Optimal versus naïve diversification: How inefficient is the 1/N portfolio strategy? *The Review of Financial Studies*, 22, 5, 1915–1953.
4, 5, 27

Fama, Eugene F (1998). Determining the number of priced state variables in the ICAPM. *Journal of Financial and Quantitative Analysis*, 33.
7

Fama, Eugene F and French, Kenneth R (1993). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47, 427–465.
3

Fama, Eugene F and French, Kenneth R (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3–56.
4, 23

Fama, Eugene F and MacBeth, James D (1973). Risk, return, and equilibrium: Empirical Tests. *The Journal of Political Economy*, 81, 3, 607–636.

Ferson, Wayne E and Harvey, Campbell R (1993). The risk and predictability of international equity returns. *The Review of Financial Studies*, 6, 527–566.

Frost, Peter A and Savarino, James E (1986). An empirical Bayes approach to efficient portfolio selection. *Journal of Financial Quantitative Analysis*, 21, 293–305.

Frost, Peter A and Savarino, James E (1988). For better performance constrain portfolio weights. *Journal of Portfolio Management*, 15, 29–34.
4

Hall, Alastair R (2005). *Generalized method of moments, Advanced Texts in Econometrics.* Oxford: Oxford University Press.
33

Hansen, Lars Peter (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 1029–1054.

8
Hou, Kewei and Kimmel, Robert L (2010). On estimation of risk premia in linear factor models. *Working Paper*
10
Huberman, Gur and Kandel, Shmuel and Stambaugh, Robert F (1987). Mimicking portfolios and exact arbitrage pricing. *The Journal of Finance*, 42, 1–9.
7
Jagannathan, Ravi and Wang, Zhenyu (1998). An asymptotic theory for estimating beta pricing models using cross–sectional regression. *The Journal of Finance*, 53, 4, 1285–1309.
12
Jorion, Philippe (1991). Bayesian and CAPM estimators of the means: Implications for portfolio selection. *Journal of Banking & Finance*, 717–727.
2
Jobson, J. D and Korkie, Bob (1980). Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, 544–554.

Jobson, J. D and Korkie, Bob (1981). Putting Markowitz theory to work. *Journal of Portfolio Management*, 7, 70–74.

Kan, Raymond and Robotti, Cesare (2011). On the estimation of asset pricing models using univariate betas. *Economics Letters*, 110, 117–121.
12
Kan, Raymond and Robotti, Cesare and Shanken, Jay (2013). Pricing model performance and the two–pass cross–sectional regression methodology. *The Journal of Finance*, 68, 6, 2617–2649.
12
Kleibergen, Frank (2009). Tests of risk premia in linear factor models. *Journal of Econometrics*, 149, 2, 149–173.
3, 12, 20, 21, 22
Lamont, Owen A (2001). Economic tracking portfolios. *Journal of Econometrics*, 105, 161–184.
7
Ledoit, Olivier and Wolf, Michael (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10, 603–621.
4
Lettau, Martin and Ludvigson, Sydney (2001). Resurrecting the (C)CAPM: A cross–sectional test when risk premia are time–varying. *Journal of Political Economy*, 109, 1238–1287.
2, 3, 4
Litterman, R. (2003). *Modern Investment Management: An Equilibrium Approach.* New Jersey: John Wiley & Sons.
4
Lewellen, Jonathan and Nagel, Stefan and Shanken, Jay (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, 96, 175–194.
12
Li, Qing and Vassalou, Maria and Xing, Yuhang (2006). Sector Investment Growth Rates and the Cross Section of Equity Returns. *The Journal of Business*, 79, 1637–1665.
4
Lustig, Hanno N and Van Nieuwerburgh, Stijn G (2005). Housing collateral, consumption insurance, and risk premia: An empirical perspective. *The Journal of Finance*, 60, 1167–1219.
4
Lustig, Hanno N and Van Nieuwerburgh, Stijn G (2005). Housing collateral, consumption insurance, and risk premia: An empirical perspective. *The Journal of Finance*, 60, 1167–1219.
4
Markowitz, Harry M. (1952). Mean-variance analysis in portfolio choice and capital markets. *the Journal of Finance*, 7, 77-91.
4, 26
Merton, Robert C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, 41, 867-887.
2, 7
Merton, Robert C. (1980). On estimating the expected return on the market. *Journal of Financial Economics*, 8, 323—361.
5

MICHAUD, RICHARD O (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45, 1, 31—42.
4

PEÑARANDA, FRANCISCO AND SENTANA, ENRIQUE (2011). Inferences about Portfolio and Stochastic Discount Factor Mean Variance Frontiers. *Working paper*.
28

ROSS, STEPHEN A. (1976a). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13, 341 - 360.

ROSS, STEPHEN A. (1976b). Return, Risk and Arbitrage. *Wharton Discussion Paperr, 1973, published in I. Friend and J. Bicksler, eds., Risk and Return in Finance. (Cambridge: Ballinger)*, 1976, 189-217.

SANTOS, TANO AND VERONESI, PIETRO (2006). Labor income and predictable stock returns. *The Review of Financial Studies*, 19, 1–44.
4

SHANKEN, JAY (1992). On the estimation of beta-pricing models. *The Review of Financial Studies*, 5, 1, 1–55.
12, 19

SHARPE, WILLIAM F. (1964). Capital Asset Prices: A Theory of market equilibrium under conditions of Risk. *The Journal of Finance*, 19, 425–442.
2

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–French (1992,1993). The factors are the three factors from Fama–French (1992), market, size and book–to–market. The results are based on monthly data from January 1963 until October 2012, i.e. 597 observations for each portfolio. The first column (RP_{GMM}*–Gen over N*) 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 N*) and the third (RP_{GMM}*–Mim over N*) 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. Fourth column (RP_{GMM}*–Tr over* RP_{GMM}*–Gen*) corresponds to the precision gains from estimating the risk premiums based on GMM with (3.2) over the case based on GMM with (3.2). The last column (RP_{GMM}*–Mim over* RP_{GMM}*–Gen*) presents the gains from making use of mimicking portfolios and estimate the system with GMM (3.4) over estimation with GMM (3.1)

Assets	RP_{GMM}–Gen over N	RP_{GMM}–Tr over N	RP_{GMM}–Mim over N	RP_{GMM}–Tr over RP_{GMM}–Gen	RP_{GMM}–Mim over RP_{GMM}–Gen	RP_{GMM}–Mim over RP_{GMM}–Tr
1	19	28	17	7.6	-1.5	-8.4
2	18	25	17	5.6	-1.1	-6.3
3	12	16	11	4.0	-0.8	-4.5
4	10	13	10	3.0	-0.6	-3.5
5	6	9	5	2.9	-0.6	-3.4
6	8	15	6	6.8	-1.4	-7.7
7	6	11	6	4.3	-0.9	-4.9
8	5	8	4	3.0	-0.6	-3.5
9	6	7	5	1.5	-0.2	-1.8
10	5	6	5	1.3	-0.2	-1.5
11	8	13	7	5.2	-1.1	-5.9
12	7	10	7	2.3	-0.4	-2.6
13	9	10	9	1.1	-0.1	-1.2
14	9	10	9	1.0	-0.1	-1.0
15	10	11	10	1.0	-0.1	-1.1
16	10	13	10	2.2	-0.4	-2.6
17	10	11	10	0.9	-0.1	-1.0
18	12	13	12	1.0	-0.1	-1.1
19	11	12	10	1.6	-0.2	-1.8
20	14	15	14	1.5	-0.2	-1.6
21	5	6	5	0.9	-0.1	-1.0
22	11	13	11	1.6	-0.2	-1.7
23	18	22	18	2.9	-0.5	-3.2
24	15	21	14	5.3	-1.0	-6.0
25	25	31	24	4.7	-0.8	-5.3

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

This table provides the average out–of–sample Sharpe ratio (first line), its percentage error compared to the true Sharpe ratio (second line) and the root–mean–squared errors (third line) over 10000 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 $E[R_t] = \beta\lambda$. 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 October 2012. 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 597 and risk aversion is 5. The upper panel presents the results for enlarging samples and the lower panel presents the results for the rolling windows based on window sizes of 120 observations.

$\gamma = 5$			*Enlarging Samples*					
	True μ^e	***Naive***	RP_{GMM}***–Gen***	RP_{GMM}***–Tr***	RP_{GMM}***–Mim***	***GMV***	***1/N***	***Theoretical***
	0.218	0.125	0.188	0.190	0.188	0.111	0.136	0.214
True $\boldsymbol{\Sigma_{RR}}$	0.019	-0.416	-0.125	-0.112	-0.125			
	0.004	0.103	0.056	0.055	0.056	0.045	0.046	
	0.208	0.119	0.195	0.198	0.195	0.111	0.136	0.214
$\boldsymbol{\hat{\Sigma}_{RR}}$	-0.030	-0.444	-0.091	-0.077	-0.091			
	0.006	0.108	0.053	0.052	0.053	0.046	0.046	
				Rolling-sample Window Length=120				
	0.218	0.089	0.154	0.158	0.154	0.111	0.136	0.214
True $\boldsymbol{\Sigma_{RR}}$	0.019	-0.585	-0.282	-0.261	-0.282			
	0.004	0.135	0.080	0.076	0.080	0.045	0.046	
	0.193	0.081	0.167	0.172	0.167	0.103	0.136	0.214
$\boldsymbol{\hat{\Sigma}_{RR}}$	-0.101	-0.623	-0.222	-0.197	-0.222			
	0.022	0.142	0.071	0.067	0.071	0.046	0.046	