The Efficiency vs. Pricing Accuracy Trade-Off in GMM Estimation of Multifactor Linear Asset Pricing Models⋆

Massimo Guidolin[a], Martin Lozano[b,∗], Juan Arismendi-Zambrano[c,d]

[a]*Bocconi University and IGIER, Bocconi University, Milan, Italy.*
[b]*School of Business, University of Monterrey - UDEM, Monterrey, Mexico.*
[c]*Department of Economics, Finance and Accounting, National University of Ireland, Maynooth, Ireland.*
[d]*ICMA Centre, Henley Business School, University of Reading, Whiteknights, Reading, United Kingdom.*

Abstract

Even though a multi-factor linear asset pricing model can be equivalently represented in a Beta or in a stochastic discount factor (SDF) form, its inferential efficiency and pricing accuracy features may differ when estimated by the generalized method of moments (GMM), both in small and in large samples. We derive the analytical asymptotic variance of the Beta and the SDF estimators under GMM and find that the SDF approach is likely to be less efficient but to yield more accurate pricing than the Beta method. We show that the main drivers of this trade-off are the higher-order moments of the factors that play an important role in the estimation process.

Keywords: Empirical asset pricing, Factor models, Higher order moments, Generalized Method of Moments, Stochastic discount factor, Beta pricing, Estimation efficiency.
JEL Classification: C51, C52, G12.

⋆An earlier version was presented at University Carlos III de Madrid, Universidad de Castilla-La Mancha, XVI AEFIN Finance Forum (ESADE Business School), the Southwestern Finance Association 48th Annual Meeting (Oklahoma City, February 2009), the Royal Economic Society Conference (London, April 2009), the Symposium in Statistics and Econometrics (Lausanne, April 2009), the Eastern Finance Association Annual Meeting (Washington D.C., May 2009), the Asian Finance Association Conference (Brisbane, July 2009), the Symposium in Economics and Finance (Geneva, July 2009), the University of Liverpool Management School; University College Dublin, the 2nd Frontiers of Factor Investing Conference (Lancaster, April 2020), and the World Congress of the Econometric Society (Milan, August 2020). We would like to thank Gregory Connor, Brendan McCabe, Genaro Sucarrat, Chien-Ting Lin, and Sheng Guo for helpful comments as discussants, and to Raymond Kan for kindly sharing complementary econometric notes on Kan and Zhou (1999).

∗Corresponding author

Email addresses: massimo.guidolin@unibocconi.it (Massimo Guidolin), martin.lozano@udem.edu (Martin Lozano), JuanCarlos.ArismendiZambrano@mu.ie;j.arismendi@icmacentre.ac.uk (Juan Arismendi-Zambrano)

I. Introduction

Any asset pricing model can be formally characterized either under a Beta or a stochastic discount factor (SDF) representation. The SDF representation states that the value of any asset equals the expected value of the product of the (stream of) payoffs on the asset and the SDF. In a Beta representation, the expected return on an asset is instead a linear function of its factor exposures (betas). The Beta approach is widespread in the finance literature and usually implemented through the two-stage cross-sectional regression methodology advocated by Black et al. (1972), Fama and MacBeth (1973), and Kan et al. (2013). The relatively more recent SDF characterization can be traced back to Dybvig and Ingersoll (1982), Hansen and Richard (1987), and Ingersoll (1987), who derive it for a number of theoretical asset pricing models formerly available only under the classical Beta framework.

Even though the two characterizations are theoretically equivalent, the parameters of interest are in general different under the two setups. In particular, the Beta representation is formulated to analyze the factor risk premia, δ, and as a residua, the Jensen's alphas, α (that can be interpreted as a measure of mean abnormal returns when the factors are tradeable). In contrast, the SDF representation is intended to analyze the parameters that enter into the assumed stochastic discount factor, λ, and the resulting pricing errors π.[1] As a matter of fact, only when the factors are standardized to have zero mean and a unit variance, and to be mutually uncorrelated, the parameters of interest coincide, i.e., $\delta = \lambda$, and $\alpha = \pi$; however these are rather a set of theoretical assumptions than circumstances commonly observed in practice.[2]

The fact that the two representations are equivalent, imply that there is a one-to-one mapping between δ and λ, and between α and π, which may facilitate the comparison of the estimators.[3] Though, this theoretical equivalence does not necessarily entail an empirical equivalence. Therefore, the experimental questions that naturally arise are: (1) *Is it better to produce inferences on δ or on λ?* and, analogously, (2) *Is it better to make inferences on α or on π?*; finally, given that there are precise links between δ and λ, and α and π, (3) *Is it better to perform estimation of the Beta representation (i.e., recover δ and α), or on the SDF representation (i.e., on λ and π?*

Our analytical and simulation results show that, in general, the Beta representation of a linear factor model is more **efficient** but less **accurate in pricing** than the SDF one. That is, the estimators of δ and π have lower simulated, asymptotic relative standard errors vs. the corresponding estimators for λ and α. The main objective of our paper is therefore to document how the choice between the Beta and the SDF methods to implement a multi-factor pricing model can be addressed in the light of this trade-off between estimation vs. pricing accuracy. We provide empirically motivated evidence about what drives this trade-off which is valuable to researchers and practitioners because it provides them with an a *priori* idea about the benefits and costs of adopting either representation in empirical work.[4].

[1]See Ferson and Jagannathan (1996) for a general discussion of the equivalence and differences between the two representations.

[2]Of course, factors may be built by the econometrician to satisfy these properties but these will be then derivative factors stemming from the observed ones and the necessary transformations are likely to severely impact their interpretation.

[3]We shall exploit the one-to-one mapping between Beta and SDF estimators in order to transform the Beta estimators into SDF units. By doing so, we are able to perform a fair comparison of the simulated standard errors because even though the values do not coincide numerically, they will have the same units of measurements.

[4]For instance, researchers and practitioners employ linear pricing models the estimate the cost of capital associated with investment and takeover decisions, which is a recurring task in accounting and corporate finance. Also, asset pricing models are used in comparative analyses of the success of different investors or to implement the performance evaluation of investment

To disentangle the source of our simulation findings, we extend Jagannathan and Wang (2002) analytical results to the case of multi-factor asset pricing models. On the one hand, Kan and Zhou (2001) already provided very reasonable intuition as to why an SDF representation may be ideal in a pricing perspective but not derive inferences on the risk premia: the SDF method does not place any restriction on λ, whereas a Beta representation explicitly incorporates the definition of δ as a subset of the moment restrictions. Therefore, it is reasonable to expect that the Beta method should be more efficient than the SDF method at estimating λ and, conversely, that the SDF method may be more robust than the Beta approach at pricing the available cross-section of test assets. On the other hand, in this paper we also find that the source of the estimation efficiency of the Beta method over the SDF method, is rooted in the higher-order moments that are likely to characterize the commonly used factors, as they impact more the SDF than the Beta estimation. In particular, we prove analytically and document through appropriate simulations that negative skewness in the factors – which is usually a characteristic of the momentum portfolio – pose an obstacle to the accurate estimation of λ even in the rather large samples typical of research with US data. Of course, the loss of precision in risk premia inferences caused by non-zero skewness grows as the samples become smaller. Yet, such problematic estimation of λ turns into an advantage when it comes to estimate the pricing errors π, given that the SDF method is essentially based on the idea of minimizing the pricing errors.

Out study contributes to both the asset pricing and financial econometrics literature, where it has become common to compare the performance of different econometric procedures either within the Beta framework or within the SDF method. For example, Jagannathan and Wang (1998) compare the asymptotic efficiency of the two-stage cross-sectional regression and of the Fama-MacBeth procedure. Shanken and Zhou (2007), analyze the finite sample properties and empirical performance of maximum likelihood applied to implement Fama-MacBeth's approach and of GMM for Beta pricing representations. Other important examples are Amsler and Schmidt (1985), Velu and Zhou (1999), Farnsworth et al. (2002), Chen and Kan (2004), Kan and Robotti (2008), and Kan and Robotti (2009). However, only recently there have been attempts to evaluate the inferential performance in finite samples of the Beta versus the SDF approaches. This is where our main interest lies.

In a first attempt to evaluate the finite sample efficiency of the Beta versus the SDF approaches, using a standardized single-factor model, Kan and Zhou (1999) show that the SDF method may be much less efficient than the Beta method. Jagannathan and Wang (2002), Cochrane (2001), and Cochrane (2005) have debated this conclusion in a non-standardized single-factor model but assuming joint normality for both the asset returns and the factors; they conclude that the SDF method is as efficient as the Beta method for estimating the risk premia. In addition, they find that standard specification tests are equally powerful in either of the two frameworks. Yet,Kan and Zhou (2001) have shown that, under more general distributional assumptions and considering non-standardized factors, the inference based on λ may be less reliable than those based on δ, especially in realistic situations where the factors are leptokurtic. Ferson (2005) have reported that when the two representations correctly exploit the same moments, they will deliver nearly

funds. They are also applied in judicial inquiries related to court decisions regarding compensation to expropriated firms whose shares are not listed on the stock market. Therefore the relative performance of the Beta vs. the SDF methods become relevant because if an investigator's choice were to fall on the technique that delivers the most precise estimators, her calculations and hypothesis tests results are more reliable and this very fact ought to be reported. For related discussion, see Kan and Zhou (1999), Kan and Zhou (2001), Jagannathan and Wang (2002), and Lozano and Rubio (2011) identical results. The interest in the topic has recently attracted additional research. For example, Lozano and Rubio (2011) show evidence suggesting that inference on δ and π is more reliable than inference on their corresponding estimators of λ and α but fail to recognize the existence of a precise mapping between λ and δ as well as π and α so that the comparisons do not occur on comparable scales. On the other hand, Peñaranda and Sentana (2015) show that a particular GMM procedure leads to numerically identical Beta and SDF estimates.

Our paper also provide an application in which inference on risk premia and pricing errors under the Beta pricing and SDF representations is performed with reference to the single-factor CAPM of Sharpe (1964), Lintner (1965) and Mossin (1966), the three-factor Fama and French (1993), the four-factor Carhart (1997), and the three-factor Lozano and Rubio (2011) (RUH) models in an application to US data. The application has an important role because it builds on our empirical results to offer two additional contributions that derive from differences in our finite sample approach with respect to previous, similar studies. The first is that we assume factors and returns are drawn from their marginal empirical distribution. The second is that we evaluate not only single, but multi-factor linear asset pricing models. By doing so, we examine the performance of the estimation methods in the presence of highly non-normal distributions, as it actually happens in realistic applications. Furthermore, we use a wide range of sets of test assets in order to address the tight factor structure problem described by Lewellen et al. (2009).

The outline of the rest of the paper is as follows. Section II presents the methodology, Section III reports our analytical results, Section IV provides a few illustrative empirical results, while Section V concludes.

II. A Review of the Outstanding Methodological Issues

In order to estimate and evaluate the Beta and SDF representations of a generic asset pricing model, we follow the GMM procedure by Hansen (1982). This guarantees that we can retrieve valid inferences even if the assumptions of independence,conditional homoskedasticity, and/or normality are not imposed, which seems to be rather realistic in practice.

A. The Beta representation

Denote r_t a vector of N stock returns in excess of the risk-free rate and f_t a vector of K economy-wide pervasive risk factors observed at time t. The mean and the covariance matrix of the factors are denoted by μ and $\mathbf{\Omega}$, respectively, where $\mu = \mathrm{E}[f_t]$, and $\mathbf{\Omega} = \mathrm{Cov}(f_t)$. Under the Beta representation, a standard linear pricing model can be written as

$$\mathrm{E}[r_t] = \mathbf{B}\delta, \tag{1}$$

where δ is the vector of factor risk premia, and $\mathbf{B}$ is the matrix of $N \times K$ factor loadings which measure the sensitivity of asset returns to the factors, defined as

$$\mathbf{B} \equiv \mathrm{E}[r_t(f_t - \mu)']\mathbf{\Omega}^{-1}. \tag{2}$$

Equivalently, we can identify $\mathbf{B}$ as a matrix of parameters in the time-series regression

$$r_t = \phi + \mathbf{B}f_t + \epsilon_t, \tag{3}$$

where the residual ϵ_t has zero mean and covariance $\mathbf{\Sigma}_{\epsilon_t}$, and it is uncorrelated with the factors f_t. We consider the general case where the factors may have non-zero higher-order moments. We define κ_3 as the coskewness tensor and κ_4 the cokurtosis tensor.[5] The specification of the asset pricing model under the Beta representation in equation (1) imposes a number of restrictions on the time-series intercept, $\phi = (\delta - \mu)\mathbf{B}$. By substituting this restriction in the regression equation, we obtain:

$$r_t = \mathbf{B}\left(\delta - \mu + f_t\right) + \epsilon_t \quad \text{where:} \quad \begin{cases} \mathrm{E}\left[\epsilon_t\right] = 0_N \\ \mathrm{E}\left[\epsilon_t f_t'\right] = 0_{N\times K} \end{cases} . \tag{4}$$

Hence, the Beta representation in equation (1) gives rise to the factor model in equation (4). The associated set of moment conditions g implied by the factor model are:

$$\begin{aligned} &\mathrm{E}[r_t - \mathbf{B}(\delta - \mu + f_t)] = 0_N, \\ &\mathrm{E}[[r_t - \mathbf{B}(\delta - \mu + f_t)]f_t'] = 0_{N\times K}. \end{aligned} \tag{5}$$

However, when the factor is the return on a portfolio of traded assets, as in the single and multi-factor models analyzed in this paper – the CAPM, Fama-French, and the Carhart's factor models – it can be easily shown that the estimate of μ (the sample mean vector of the factors) is also the estimate of the risk premium δ. Therefore, given $\delta = \mu$, the moment conditions given in equation (5) simplify to:

$$\begin{aligned} &\mathrm{E}[r_t - \mathbf{B}f_t] = 0_N, \\ &\mathrm{E}[(r_t - \mathbf{B}f_t)f_t'] = 0_{N\times K}, \\ &\mathrm{E}[f_t - \mu] = 0_K, \end{aligned} \tag{6}$$

where neither δ or μ appear in the first two set of moment conditions derived from equation (6) so that it becomes necessary to include the definition of μ as a third set of moment conditions to identify the vector of risk premia δ. [6]

Following the usual GMM notation, we define the vector of unknown parameters $\theta = [\mathrm{vec}(\mathbf{B})' \;\; \mu']'$, where the vec operator 'vectorizes' the $\mathbf{B}_{N\times K}$ matrix by stacking its columns, and the observable variables are $x_t = [r_t' \;\; f_t']'$. Then, the function g that captures the moment conditions required by the GMM can be written as:

$$g(x_t, \theta) = \begin{pmatrix} r_t - \mathbf{B}f_t \\ \mathrm{vec}[(r_t - \mathbf{B}f_t)f_t'] \\ f_t - \mu \end{pmatrix}_{(N+NK+K)\times 1} , \tag{7}$$

[5]Coskewness and cokurtosis have been investigated in asset pricing studies such as Harvey and Siddique (2000), Dittmar (2002) and Guidolin and Timmermann (2008). A tensor is an N-dimensional array: coskewness is then a 3-dimensional array while cokurtosis is a 4-dimensional array.

[6]Nevertheless, it is also possible to estimate the last moment restriction of equation (6) outside the GMM framework by computing $\mu = E[f_t]$. This is because the number of added moment restrictions in equation (6) compared with equation (5) is the same as the number of added unknown parameters. Hence, the efficiency of equation (5) and equation (6) is not affected by imposing the additional N moment restrictions in $\mathrm{E}[f_t - \mu] = 0_K$. Following this logic, we can drop the factor-mean moment conditions without ignoring that it has to be estimated. An additional moment condition to estimate the variance $\mathbf{\Omega}$ could also be added to equation (6). However the variance can also be estimated outside the GMM framework without affecting the efficiency of the conditional mean estimators.

in which, for any θ, the sample analogue of $\mathrm{E}[g(x_t, \theta)]$ is

$$g_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta). \tag{8}$$

Therefore a natural estimation strategy for θ is to choose the values that make $g_T(\theta)$ as close to the zero vector as possible. For that reason, we choose θ to solve

$$\min_{\theta}\; g_T(\theta)'\mathbf{W}^{-1}g_T(\theta). \tag{9}$$

To compute the first-stage GMM estimator θ_1 we consider $\mathbf{W} = \mathbf{I}$ in the minimization (9). The second-stage GMM estimator θ_2 is then the solution to the problem (9) when the weighting matrix is the spectral density matrix of $g(x_t, \theta_1)$:

$$\mathbf{S} = \sum_{j=-\infty}^{\infty} \mathrm{E}[g(x_t, \theta_1)g(x_t, \theta_1)'], \tag{10}$$

where $\mathbf{S}$ is of size $N \times N$. Moreover, to examine the validity of the pricing model derived from the moment restrictions in equation (6), we can test whether the vector of N Jensen's alphas, given by $\alpha = \mathrm{E}[r_t] - \delta\mathbf{B}$ is jointly equal to zero.[7] This can be done using the J-statistic which turns out to have an asymptotic χ^2 distribution. The covariance matrix of the pricing errors, $\mathrm{Cov}(g_T)$, is given by

$$\mathrm{Cov}(g_T) = \frac{1}{T}\left[(I - \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}')\mathbf{S}(I - \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}')\right], \tag{11}$$

and the test is a quadratic form in the vector of pricing errors. In particular, the Hansen (1982) J-statistic is computed as

$$\begin{array}{ll} \text{First-stage:} & g_T(\theta_1)'\mathrm{Cov}(g_T)^{-1}g_T(\theta_1) \sim \chi^2_N, \\ \text{Second-stage:} & T g_T(\theta_2)'S^{-1}g_T(\theta_2) \sim \chi^2_N. \end{array} \tag{12}$$

Both the first and second-stage statistics in equation (12) lead to the same numerical value.[8]

B. The SDF method

To derive the SDF representation from the Beta representation we follow Ferson and Jagannathan (1996), and Jagannathan and Wang (2002). First, we substitute the expression for $\mathbf{B}$ in equation (2) into equation (1) and rearrange the terms, to obtain

$$\mathrm{E}[r_t] - \mathrm{E}[r_t\delta'\mathbf{\Omega}^{-1}f_t - r_t\delta'\mathbf{\Omega}^{-1}\mu'] = \mathrm{E}[r_t(1 + \delta'\mathbf{\Omega}^{-1}\mu - \delta'\mathbf{\Omega}^{-1}f_t)] = 0_N.$$

Again, if we were considering traded factors, then $\delta = \mu$ so $1 + \delta'\mathbf{\Omega}^{-1}\mu = 1 + \mu'\mathbf{\Omega}^{-1}\mu \geq 1$, then divide each side by $1 + \delta'\mathbf{\Omega}^{-1}\mu$,[9]

$$\mathrm{E}\left[r_t\left(1 - \frac{\delta'\mathbf{\Omega}^{-1}}{1 + \delta'\mathbf{\Omega}^{-1}\mu}f_t\right)\right] = 0_N.$$

[7]This approach is known as the restricted test, see MacKinlay and Richardson (1991).

[8]However, if we weight equations (11) and (12) by any other matrix different from $\mathbf{S}$, such as $\mathrm{E}[r_t r_t']$ or $\mathrm{Cov}[r_t]$, this result will no longer hold. Given that there are $N + NK + K$ equations and $NK + K$ unknown parameters in the vector equation (7), the degrees of freedom are equal to N.

[9]Even when the factors are not traded, it is common to assume $1 + \delta'\mathbf{\Omega}^{-1}\mu \neq 0$.

If we transform the vector of risk premia, δ, into a vector of new parameters λ as follows,

$$\lambda = \frac{\mathbf{\Omega}^{-1}\delta}{1+\delta'\mathbf{\Omega}^{-1}\mu}, \tag{13}$$

then we obtain the following SDF representation, which serves at the same time the as set of moment restrictions h used to estimate the linear asset-pricing model,

$$\mathrm{E}[r_t(1-\lambda' f_t)] = 0_N, \tag{14}$$

where the random variable $m_t \equiv 1 - \lambda' f_t$ is the SDF defined as usual as $\mathrm{E}[r_t m_t] = 0_N$. Alternatively, we could derive the Beta representation from the SDF representation by expanding m and rearranging the terms, thus going in reverse compared to steps that have led us from to .

From the moment restrictions and equation (14), we obtain the vector of N pricing errors defined as $\pi \equiv \mathrm{E}[r_t] - \mathrm{E}[r_t f_t']\lambda$. The numerical estimation of the parameters implied by equation (14) can once more be obtained by GMM. Let's start by writing the sample pricing errors as

$$h_T(\lambda) = \frac{1}{T}\sum_{t=1}^{T}(-r_t + r_t f_t')\lambda, \tag{15}$$

and by defining $\mathbf{D}^U = -\frac{\partial h_T(\lambda)}{\partial \lambda'} = \frac{1}{T}\sum_{t=1}^{T} r_t f_t'$, the second-moment matrix of returns and factors. The first-order condition to minimize the quadratic form of the sample pricing errors, equation (9), is $-\left(\mathbf{D}^U\right)'\mathbf{W}[\frac{1}{T}\sum_{t=1}^{T} r_t - \mathbf{D}^U\lambda'] = 0$, where $\mathbf{W}$ is the GMM weighting matrix of size $N \times N$, equal to the identity matrix in the first-stage estimator and equal to the spectral density matrix $\mathbf{S}$, equation (10), in the second-stage estimator. Therefore, the GMM estimates of λ are:

$$\begin{aligned} \widehat{\lambda}_1^U &= \left(\left(\mathbf{D}^U\right)'\mathbf{D}^U\right)^{-1}\left(\mathbf{D}^U\right)'\tfrac{1}{T}\textstyle\sum_{t=1}^{T} r_t, \\ \widehat{\lambda}_2^U &= \left(\left(\mathbf{D}^U\right)'\mathbf{S}^{-1}\mathbf{D}^U\right)^{-1}\left(\mathbf{D}^U\right)'\mathbf{S}^{-1}\tfrac{1}{T}\textstyle\sum_{t=1}^{T} r_t. \end{aligned} \tag{16}$$

For illustrative purposes, we add an apex an U to $\widehat{\lambda}$ to indicate when the estimator is obtained from the un-centred specification and with a C to indicate when it comes from the centred specification.

Specifying the SDF as a linear function of the factors as in equation (14) is very popular in the empirical literature. However, Kan and Robotti (2008) point out that (14) is problematic because the specification test statistic will not invariant to an affine transformation of the factors; Burnside (2007) reaches a similar conclusions. Therefore, we also consider an alternative specification that defines the SDF as a linear function of the de-meaned factors. Examples of this representation can be found in Julliard and Parker (2005) and Yogo (2006). The alternative centred version of equation (14) is therefore defined as:

$$\mathrm{E}[r_t[1-\lambda(f_t-\mu)]] = 0_N. \tag{17}$$

According to Jagannathan and Wang (2002) and Jagannathan et al. (2008), it is also possible to estimate μ in equation (17) outside of the GMM estimation by the mean, $\mu = \mathrm{E}[f_t]$. This is because the number of added moment restrictions is the same as the number of added unknown parameters. Hence, the efficiency of the estimator remains the same. By following this logic, we can drop the factor-mean moment condition without ignoring that it has to be estimated, to obtain analytical expressions for $\widehat{\lambda}_1^C$ and $\widehat{\lambda}_2^C$. In fact, the procedure to enforce the moment restrictions in equation (17) and to solve the GMM minimization is similar to that for the uncentred SDF^U method. In particular, we substitute $\mathrm{E}[r_t f_t]$ for $\mathrm{Cov}[r_t f_t]$ in equation (15) and define $\mathbf{D}^C = -\frac{\partial h_T(\lambda)}{\partial \lambda'}$ as the covariance matrix of returns and factors. As a result, under SDF^C, the first and second stage GMM estimates are given by:

$$\begin{aligned} \widehat{\lambda}_1^C &= \left(\left(\mathbf{D}^C\right)' \mathbf{D}^C \right)^{-1} \left(\mathbf{D}^C\right)' \tfrac{1}{T} \textstyle\sum_{t=1}^T r_t, \\ \widehat{\lambda}_2^C &= \left(\left(\mathbf{D}^C\right)' \mathbf{S}^{-1} \mathbf{D}^C \right)^{-1} \left(\mathbf{D}^C\right)' \mathbf{S}^{-1} \tfrac{1}{T} \textstyle\sum_{t=1}^T r_t. \end{aligned} \tag{18}$$

Valid specification tests can be conducted by using (8) and (12), the only difference being that we substitute $\mathbf{B}$ by $\mathbf{D}^U = \mathrm{E}[r_t f_t]$ (the second moment matrix of returns and factors) in the SDF^U case, and by $\mathbf{D}^C = \mathrm{Cov}[r_t f_t]$ (the covariance matrix of returns and factors) in the SDF^C case. The degrees of freedom in equation (12) are specific to the Beta method, as under the SDF method the degrees of freedom is equal to $N - K$, because there are N equations and K unknown parameters in both equations (14) and (17).[10]

C. *Comparison of the methods*

There is a one-to-one mapping between the factor risk premia collected in δ and the SDF parameter vector λ, which facilitates the comparison of the two methods and that exploits the possibility to derive an indirect estimator of λ by the Beta method.[11] By the same token, we can derive an estimate of δ not only by the Beta method but also indirectly, by the SDF method. From the previous definition of λ in (13), we have:

$$\lambda = \delta' \left(\mathbf{\Omega} + \delta \mu' \right)^{-1}, \qquad \text{or} \qquad \delta = \frac{\mathbf{\Omega} \lambda}{1 - \mu' \lambda}. \tag{19}$$

In a similar way, by substituting (19) into π, we can find a one-to-one mapping between π and α, estimated from the Beta method.

$$\pi = \frac{\mathbf{\Omega}}{\mathbf{\Omega} + \delta \mu'} \alpha, \qquad \text{or} \qquad \alpha = \frac{\mathbf{\Omega} + \delta \mu'}{\mathbf{\Omega}} \pi. \tag{20}$$

Yet, we cannot directly compare λ and δ, neither π and α because they are measured in different units. An alternative to allow direct comparisons is to transform δ into λ units, and α into π units following equations (19) and (20). For convenience, we will decorate all Beta estimators with '$*$' to easily emphasize that they are Beta estimators.

III. Asymptotic, Analytical Results

In this Section, we generalize the results in Jagannathan and Wang (2002), in the sense that the vector of factors f_t is multivariate and is allowed to have a (joint) non-Gaussian distribution. Our results will be

[10] Equations (11) and (12) are weighted by (10), and this is known to be optimal. This approach was first suggested by Hansen (1982) because it maximizes the asymptotic elicitation of information in the sample about a model, given the choice of moments. However, there are also alternatives for the weighting matrix which are suitable for model comparisons because they are invariant to the nature of the model and their parameters. For instance, Hansen and Jagannathan (1997) suggest the use of the second moment matrix of excess returns $\mathbf{W} = \mathrm{E}[r_t r_t']$ instead of $\mathbf{W} = \mathbf{S}$. Also, Burnside (2007), Balduzzi and Yao (2007), and Kan and Robotti (2008) suggest that the SDF^C method should use the covariance matrix of excess returns $\mathbf{W} = \mathrm{Cov}[r_t]$. We shall investigate the implications of using these alternative weighting matrices later.

[11] We thank to Raymond Kan for kindly sharing complementary notes on Kan and Zhou (2001) that are at the roots of what follows.

useful to understand the small-sample simulation findings on the trade-off between estimation vs. pricing accuracy when using the Beta vs. the SDF representation.

PROPOSITION 1: *Let f_t represent the multivariate, systematic risk factors with mean μ, covariance $\mathbf{\Omega}$, third-order cumulant κ_3, and fourth-order cumulant κ_4, and consider the Beta representation in (1), and the SDF representation in (14). Then, the positive definite asymptotic covariance matrix of the $\hat{\lambda}$ estimators obtained under the GMM for the SDF case is*

$$ACov(\hat{\lambda}) = \left(\left(\mathbf{\Omega} + \mu\mu'\right)'\mathbf{B}'\right)^{-1}\left(\frac{1}{a_{\epsilon_t}}\mathbf{\Sigma}_{\epsilon_t}^{-1} - \frac{1}{a_{\epsilon_t}^2}\mathbf{\Sigma}_{\epsilon_t}^{-1}\mathbf{B}\left(\mathbf{A}_B^{-1} + \frac{1}{a_{\epsilon_t}}\mathbf{B}'\mathbf{\Sigma}_{\epsilon_t}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{\Sigma}_{\epsilon_t}^{-1}\right) \times \left(\mathbf{B}\left(\mathbf{\Omega} + \mu\mu'\right)\right)^{-1}, \quad (21)$$

where a_{ϵ_t} and $\mathbf{A}_B$ are

$$\begin{aligned}\mathbf{A}_B \;=\; & \otimes_R\left(\otimes_R\left(\kappa_4, \lambda\right), \lambda\right) + 2\delta\left(\otimes_R\left(\kappa_3, \lambda\right)\lambda\right)' + 2\otimes_R\left(\kappa_3, \lambda\lambda'\mu - \lambda\right) + \\ & \mathbf{\Omega}\left(\mathbf{I}_N + 4(\lambda\lambda'\mu - \lambda)\delta' + (\lambda'\mu\mu'\lambda)\mathbf{I}_N - 2\lambda'\mu\mathbf{I}_N\right) + \delta\delta'\lambda'\mathbf{\Omega}\lambda + \\ & \left(\delta\delta' + \lambda'\mu\mu'\lambda\delta\delta' - 2\lambda'\mu\delta\delta'\right),\end{aligned}$$

and

$$a_{\epsilon_t} = 1 - 2\lambda'\mu + \lambda'\mu\mu'\lambda + \lambda'\mathbf{\Omega}\lambda,$$

that turns out to depend on the (co-)skewness and (co-)kurtosis coefficients of factors. The asymptotic covariance matrix of the GMM estimator λ^* obtained from the Beta representation is

$$ACov(\hat{\lambda}^*) \;=\; \left(\mathbf{\Omega} + \mu\mu'\right)^{-1}\boldsymbol{S}_b\left(\left(\mathbf{\Omega} + \mu\mu'\right)^{-1}\right)', \quad (22)$$

where $\boldsymbol{S}_b$ is the covariance matrix of $g_b(f_t, \lambda_b) = (f_t(1 - \lambda_b' f_t))$, and $\lambda_b = \mu'\left(\mathbf{\Omega} + \mu\mu'\right)^{-1}$, which does not depend on either skewness of kurtosis tensors. In the case of a single-factor, the estimators of the GMM estimators of the asymptotic variance of the risk premium estimate from the SDF and Beta representations are,

$$Avar(\hat{\lambda}) \;=\; \frac{\sigma^2(\sigma^2 + \mu^2)}{(\sigma^2 + \mu^2)^4}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1} + \frac{\sigma^2(\sigma^4 + \mu^4)}{(\sigma^2 + \mu^2)^4} + \frac{\kappa_4\mu^2 + 2\kappa_3(\mu^3 - \mu\sigma^2) - 3\mu^2\sigma^2}{(\sigma^2 + \mu^2)^4}, \quad (23)$$

and,

$$Avar(\hat{\lambda}^*) \;=\; \frac{\left((\sigma^2 + \mu^2) - 2\lambda_b E[f_t^3] + \lambda^2 E[f_t^4]\right)}{(\sigma^2 + \mu^2)}, \quad (24)$$

respectively, with

$$\begin{aligned}E[f_t^3] &= \kappa_3 + 3\sigma^2\mu + \mu^3, \\ E[f_t^4] &= \kappa_4 + 4\kappa_3\mu + 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4.\end{aligned}$$

The GMM estimator of the asymptotic variance of the risk premium in the Beta representation (24) can be approximated using the delta method by:

$$Avar(\hat{\lambda}^*) = \frac{\sigma^2(\sigma^2 + \mu^2)}{(\sigma^2 + \mu^2)^4}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1} + \frac{\sigma^6}{(\sigma^2 + \mu^2)^4}. \quad (25)$$

The corresponding SDF and Beta pricing errors asymptotic variance estimators are

$$Avar\left(\hat{\pi}\right) \;=\; \left(\left(\mathbf{\Sigma}_{\epsilon_t}+\delta\mu'\right)\mathbf{\Sigma}_{\epsilon_t}^{-1}(\mathbf{\Sigma}_{\epsilon_t}^{-1})'\left(\mathbf{\Sigma}_{\epsilon_t}+\delta\mu'\right)\right)Q\left(S_b-D_b(D_b'S^{-1}D_b)D_b'\right)Q', \tag{26}$$

and,

$$Avar\left(\hat{\pi}^*\right)=S_s-D_s\left(D_s'S_s^{-1}D_s\right)D_s'. \tag{27}$$

where

$$Q \;=\; \left[I_n,\mathbf{0}_{n\times n},-\mathbf{B},\boldsymbol{0}_{n\times 1}\right], \tag{28}$$

$$S_s \;=\; \mathbf{B}\mathbf{A}_B\mathbf{B}'+a_{\epsilon_t}\mathbf{\Sigma}_{\epsilon_t}, \tag{29}$$

$$D_s \;=\; -\mathbf{B}\left(\mathbf{\Omega}+\mu\mu'\right). \tag{30}$$

Proof. Before calculating the asymptotic variance of the SDF and the Beta methods, we define tensor operations. Let $T1$ be a tensor of dimension $N_1\times N_2\times\cdots\times N_p$, and $T2$ a tensor of dimension $M_1\times M_2\times\cdots\times M_o$, with all the elements $N_1,\ldots,N_p,M_1,\ldots,M_o$ greater than one and $p>o$ without loss of generality, we define the **expansion tensor product**,

$$\otimes_E(T1,T2)_{i_1,\ldots,i_m,i_m+1,\ldots,i_{m+p}}=T1_{i_1,\ldots,i_p}\times T2_{i_1,\ldots,i_o}$$

as the result of the expansion of tensors $X1$ and $X2$ in a tensor of dimension $N_1\times N_2\times\cdots\times N_p\times M_1\times M_2\times\cdots\times M_o$. Consider the case where $N_1=M_1,\ldots,N_o=M_o$. The **reduction tensor product**, is defined as,

$$\otimes_R(T1,T2)_{i_{p+1},\ldots,i_{p+o}}=T1_{i_1,\ldots,i_o}\times T2_{i_1,\ldots,i_o},$$

the tensor of reduced dimension $N_{p+1}\times N_{p+2}\times\cdots\times N_o$ that results from the dot-product of tensors $T1$ and $T2$.

A.1 Asymptotic variance SDF method

In the case of the SDF method, we calculate the GMM estimator asymptotic variance. First, we consider the general case where δ, and μ in equation (13) can be different. Define $g_s(r_t,f_t,\lambda)=r_t(1-\lambda f_t)$, the covariance matrix of $g_s(r_t,f_t,\lambda)$, in the case the factor f is non-Gaussian and has higher-order moments, is:

$$\begin{aligned} S_s &= E[g_s(r_t,f_t,\lambda)g_s(r_t,f_t,\lambda)'] \\ &= \mathbf{B}\Big(\otimes_R\left(\otimes_R\left(\kappa_4,\lambda\right),\lambda\right)+ \\ &\quad 2\delta\left(\otimes_R\left(\kappa_3,\lambda\right)\lambda\right)'+2\otimes_R\left(\kappa_3,\lambda\lambda'\mu-\lambda\right)+ \\ &\quad \mathbf{\Omega}\left(\mathbf{I}_N+4(\lambda\lambda'\mu-\lambda)\delta'+(\lambda'\mu\mu'\lambda)\mathbf{I}_N-2\lambda'\mu\mathbf{I}_N\right)+\delta\delta'\lambda'\mathbf{\Omega}\lambda+ \\ &\quad \left(\delta\delta'+\lambda'\mu\mu'\lambda\delta\delta'-2\lambda'\mu\delta\delta'\right)\Big)\mathbf{B}'+ \\ &\quad (1-2\lambda'\mu+\lambda'\mu\mu'\lambda+\lambda'\mathbf{\Omega}\lambda)\,\mathbf{\Sigma}_{\epsilon_t}. \end{aligned} \tag{31}$$

The elements in (31) are sorted from the more complex (a tensor of fourth-order, to most simple (a tensor of second order – a matrix). Higher-order moments inside (31) are the result of higher-order expected values of the multivariate factor f_t. These elements will not appear in a single factor analysis such as Kan and Zhou (1999) or Jagannathan and Wang (2002). We split the elements of (31). Define:

$$\begin{aligned} \mathbf{A}_B \;=\; & \otimes_R\left(\otimes_R\left(\kappa_4,\lambda\right),\lambda\right)+2\delta\left(\otimes_R\left(\kappa_3,\lambda\right)\lambda\right)'+2\otimes_R\left(\kappa_3,\lambda\lambda'\mu-\lambda\right)+ \\ & \mathbf{\Omega}\left(\mathbf{I}_N+4(\lambda\lambda'\mu-\lambda)\delta'+(\lambda'\mu\mu'\lambda)\mathbf{I}_N-2\lambda'\mu\mathbf{I}_N\right)+\delta\delta'\lambda'\mathbf{\Omega}\lambda+ \\ & \left(\delta\delta'+\lambda'\mu\mu'\lambda\delta\delta'-2\lambda'\mu\delta\delta'\right), \end{aligned}$$

and

$$a_{\epsilon_t} = 1 - 2\lambda'\mu + \lambda'\mu\mu'\lambda + \lambda'\mathbf{\Omega}\lambda,$$

then the covariance of $g_s(r_t, f_t, \lambda)$ can be written as:

$$S_s = \mathbf{B}\mathbf{A}_B\mathbf{B}' + a_{\epsilon_t}\mathbf{\Sigma}_{\epsilon_t} \tag{32}$$

The inverse of (32) is:

$$S_s^{-1} = \frac{1}{a_{\epsilon_t}}\mathbf{\Sigma}_{\epsilon_t}^{-1} - \frac{1}{a_{\epsilon_t}^2}\mathbf{\Sigma}_{\epsilon_t}^{-1}\mathbf{B}\left(\mathbf{A}_B^{-1} + \frac{1}{a_{\epsilon_t}}\mathbf{B}'\mathbf{\Sigma}_{\epsilon_t}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{\Sigma}_{\epsilon_t}^{-1}. \tag{33}$$

The partial derivatives of g_s respect to λ will produce a matrix:

$$D_s = E\left[\frac{\partial g_s}{\partial \lambda}\right] = -\mathbf{B}\left(\mathbf{\Omega} + \delta\mu'\right), \tag{34}$$

Then, using (33) and (63), we have the asymptotic variance of the SDF model is:

$$Avar(\hat{\lambda}) = \left(D_s' S_s^{-1} D_s\right)^{-1}. \tag{35}$$

In the case of single factor models, and defining $\sigma^2 = \Omega$, the variance of the single-factor, equations (32), (33), (63), and (35), have their equivalents in:

$$\begin{aligned}
S_s &= \frac{\sigma^2(\sigma^4+\delta^4) + \kappa_4\delta^2 + 2\kappa_3(\delta^3 - \delta\sigma^2) - 3\delta^2\sigma^4}{(\sigma^2+\mu\delta)^2}\beta\beta' + \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^2}\mathbf{\Sigma}_{\epsilon_t}, && (36)\\
S_s^{-1} &= \frac{(\sigma^2+\mu\delta)^2}{\sigma^2(\sigma^2+\delta^2)}\mathbf{\Sigma}_{\epsilon_t}^{-1} - \frac{(\sigma^2+\mu\delta)^2}{\sigma^2(\sigma^2+\delta^2)} \times \\
&\quad \left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta + \frac{\sigma^2(\sigma^4+\delta^4) + \kappa_4\delta^2 + 2\kappa_3(\delta^3 - \delta\sigma^2) - 3\delta^2\sigma^2}{\sigma^2(\sigma^2+\delta^2)}\right)^{-1}\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}, && (37)\\
D_s &= E\left[\frac{\partial g_s}{\partial \lambda}\right] = -(\sigma^2+\mu\delta)\beta, && (38)\\
Avar(\hat{\lambda}) &= \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1} + \frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^4} + \frac{\kappa_4\delta^2 + 2\kappa_3(\delta^3 - \delta\sigma^2) - 3\delta^2\sigma^2}{(\sigma^2+\mu\delta)^4} && (39)
\end{aligned}$$

The equivalent asymptotic variance (39) in the single factor Gaussian case is (Jagannathan et al., 2002):

$$Avar(\hat{\lambda}) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1} + \frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^4}. \tag{40}$$

The difference in asymptotic variance of the SDF method when modelling a Gaussian factor, and a non-Gaussian vector comes from the higher-order moments terms:

$$\frac{\kappa_4\delta^2 + 2\kappa_3\delta(\delta^2 - \sigma^2) - 3\delta^2\sigma^2}{\sigma^2(\sigma^2+\delta^2)}. \tag{41}$$

The additional term (41) will increase the asymptotic variance for heavy tailed distributions (greater κ_4), and will decrease by higher negative skewness (lower values of κ_3).

A.2 Asymptotic variance Beta method - joint estimation

We calculate the GMM asymptotic variance of the risk premiums for the case of multifactor Beta models. First, we solve for the general case where the parameters, $\theta = \left(\delta, \mathbf{B}, \mu, \sigma^2\right)$, are estimated jointly as in Jagannathan and Wang (2002). Define

$$g_b(r_t, f_t, \theta) = \begin{pmatrix} g_b(1) \\ g_b(2) \\ g_b(3) \\ g_b(4) \end{pmatrix} = \begin{pmatrix} r_t - \mathbf{B}(\delta + f_t - \mu) \\ (r_t - \mathbf{B}(\delta + f_t - \mu))f_t' \\ f_t - \mu \\ (f_t - \mu)(f_t - \mu)' - \mathbf{\Omega} \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t' \\ f_t - \mu \\ (f_t - \mu)(f_t - \mu)' - \mathbf{\Omega} \end{pmatrix}, \tag{42}$$

where $\theta = (\delta, \mathbf{B}, \mu, \mathbf{\Omega})$. The covariance of g_b (42) is

$$S_b = \begin{pmatrix} \mathbf{\Sigma}_{\epsilon_t} & \otimes_E\left(\mathbf{\Sigma}_{\epsilon_t}, \mu\right) & 0 & 0 \\ \otimes_E\left(\mathbf{\Sigma}_{\epsilon_t}, \mu\right) & \otimes_E\left(\mathbf{\Sigma}_{\epsilon_t}, \mathbf{\Omega}\right) & 0 & 0 \\ 0 & 0 & \mathbf{\Omega} & \kappa_3 \\ 0 & 0 & \kappa_3 & \kappa_4 - \otimes_E\left(\mathbf{\Omega}, \mathbf{\Omega}\right) \end{pmatrix}. \tag{43}$$

We need to calculate the partial derivatives of g_b respect to the parameters θ. The first partial derivative, $\dfrac{\partial g_b(1)}{\partial \delta} = \mathbf{B}$. The partial derivative $\dfrac{\partial g_b(2)}{\partial \delta}$ will produce the third-order tensor,

$$\frac{\partial g_b(2)}{\partial \delta} = -\otimes_E\left(\mathbf{B}, \mu\right).$$

The following partial derivatives are null: $\dfrac{\partial g_b(3)}{\partial \delta} = \dfrac{\partial g_b(4)}{\partial \delta} = \dfrac{\partial g_b(3)}{\partial \mathbf{B}} = \dfrac{\partial g_b(4)}{\partial \mathbf{B}} = \dfrac{\partial g_b(1)}{\partial \mu} = \dfrac{\partial g_b(2)}{\partial \mu} = \dfrac{\partial g_b(1)}{\partial \mathbf{\Omega}} = \dfrac{\partial g_b(2)}{\partial \mathbf{\Omega}} = 0$.

In the case of $\dfrac{\partial g_b(1)}{\partial \mathbf{B}}$, it will produce the following $N \times N \times K$ third-order tensor:

$$\frac{\partial g_b(1)}{\partial \mathbf{B}} = -\left\{ \begin{pmatrix} & \delta' & & \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{pmatrix}_{N\times K}, \dots, \begin{pmatrix} 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \\ & \delta' & & \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \\ & \delta' & & \end{pmatrix} \right\}_{N\times N\times K}.$$

We need to define some tensor notations. Let us define the canonical basis:

$$\mathbf{e}' = \left[\mathbf{e}_{1,:}, \mathbf{e}_{2,:}, \dots, \mathbf{e}_{N,:}\right]_{1\times N}.$$

where every element of this vector is a matrix:

$$\mathbf{e}_{i,:} = \underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & 0 \end{pmatrix}_{N\times K}}_{i-th \text{ row equal to one.}},$$

The identity tensor can be denoted using tensor notation as $\mathbf{I}_{N\times N\times K} = \otimes_E (\mathbf{e}, \mathbf{1}_{N\times 1})$. Then, we can denote $\frac{\partial g_b(1)}{\partial \mathbf{B}}$ as

$$\delta \mathbf{I}_{N\times N\times K} \equiv \otimes_E (\delta\mathbf{e}, \mathbf{1}_{N\times 1}), \tag{44}$$

where $\delta\mathbf{e} = \left[[\delta,\dots,\delta]' \odot \mathbf{e}_{1,:},\dots,[\delta,\dots,\delta]' \odot \mathbf{e}_{N,:}\right]$ and $\odot$ is the element wise multiplication. The partial derivative $\frac{\partial g_b(2)}{\partial \mathbf{B}}$ will produce a fourth-order tensor:

$$\frac{\partial g_b(2)}{\partial \mathbf{B}} =$$

$$-\left\{\begin{matrix} \begin{pmatrix} \delta_1\mu_1 + \mathbf{\Omega}_{1,1} & \dots & \delta_1\mu_k + \mathbf{\Omega}_{1,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix}_{N\times K} & \dots & \begin{pmatrix} \delta_k\mu_1 + \mathbf{\Omega}_{k,1} & \dots & \delta_k\mu_k + \mathbf{\Omega}_{k,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & \dots & 0 \\ \delta_1\mu_1 + \mathbf{\Omega}_{1,1} & \dots & \delta_1\mu_k + \mathbf{\Omega}_{1,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix}_{N\times K} & \dots & \begin{pmatrix} 0 & \dots & 0 \\ \delta_k\mu_1 + \mathbf{\Omega}_{k,1} & \dots & \delta_k\mu_k + \mathbf{\Omega}_{k,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix} \\ & \vdots & \\ \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \\ \delta_1\mu_1 + \mathbf{\Omega}_{1,1} & \dots & \delta_1\mu_k + \mathbf{\Omega}_{1,k} \end{pmatrix}_{N\times K} & \dots & \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \\ \delta_k\mu_1 + \mathbf{\Omega}_{k,1} & \dots & \delta_k\mu_k + \mathbf{\Omega}_{k,k} \end{pmatrix} \end{matrix}\right\}_{N\times K\times N\times K} \tag{45}$$

Using a similar notation as in (44), we denote the fourth-order identity tensor and the corresponding partial derivative $\frac{\partial g_b(2)}{\partial \mathbf{B}}$ as

$$\mathbf{I}_{N\times N\times K\times K} = \otimes_E (\mathbf{e}, \mathbf{1}_{N\times K}), \tag{46}$$

$$\frac{\partial g_b(2)}{\partial \mathbf{B}} = (\delta\mu' + \mathbf{\Omega})\, \mathbf{I}_{N\times N\times K\times K} \equiv \otimes_E ((\delta\mu' + \mathbf{\Omega})\,\mathbf{e}, \mathbf{1}_{N\times N}). \tag{47}$$

The partial derivatives of the respect to the mean and the variance of the factor are $\frac{\partial g_b(3)}{\partial \mu} = -\mathbf{I}_{K\times 1}$ and

$\frac{\partial g_b(4)}{\partial \mu} = -\mathbf{I}_{K\times K}$. Then, the expected value of the partial derivatives $D_b = E\left[\frac{\partial g_b}{\partial \theta}\right]$ is the following matrix

$$D_b = E\left[\frac{\partial g_b}{\partial \theta'}\right] = \begin{pmatrix} -\mathbf{B} & -\delta \mathbf{I}_{N\times N\times K} & \mathbf{B} & 0 \\ -\otimes_E(\mathbf{B},\mu) & -(\delta\mu' + \mathbf{\Omega})\,\mathbf{I}_{N\times N\times K\times K} & \otimes_E(\mathbf{B},\mu) & 0 \\ 0 & 0 & -\mathbf{I}_{K\times 1} & 0 \\ 0 & 0 & 0 & -\mathbf{I}_{K\times K} \end{pmatrix}. \tag{48}$$

Define

$$S_b^{-1} = \text{inv}(S_b), \tag{49}$$

then, from the resulting matrix $V = \left(D_b' S_b^{-1} D_b\right)^{-1}$, the asymptotic variance of the δ^* parameter is equal to the top left corner element of the matrix:

$$Avar(\delta^*) = V_{1,1}.$$

Using the Delta method, and the definition of λ in (19), the asymptotic variance for the risk premium of the Beta method with multiple non-Gaussian factors is

$$\begin{aligned} Avar(\lambda^*) &= \left(\frac{\partial \lambda}{\partial \delta}\right)\left(\frac{\partial \lambda}{\partial \delta}\right)' Avar(\delta^*) \\ &= \left(-\delta\left(\left((\mathbf{\Omega} + \delta\mu')(\mathbf{\Omega} + \delta\mu')'\right)^{-1}\mu\right)' + \mathbf{1}_{K\times K}(\mathbf{\Omega} + \delta\mu')^{-1}\right) V_{1,1}. \end{aligned} \tag{50}$$

The calculation of the asymptotic variance of the risk premium by using the Beta method, for the case a single non-Gaussian factor has the corresponding equations to the multifactor equivalents (42), (43), (48), (49), in

$$g_b(r_t, f_t, \theta) = \begin{pmatrix} r_t - (\delta + f_t - \mu)\beta \\ (r_t - (\delta + f_t - \mu)\beta) f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix}, \tag{51}$$

$$S_b = \begin{pmatrix} \Omega & \mu\Omega & 0 & 0 \\ \mu\Omega & (\mu^2 + \sigma^2)\Omega & 0 & 0 \\ 0 & 0 & \sigma^2 & \kappa_3 \\ 0 & 0 & \kappa_3 & \kappa_4 - \sigma^4 \end{pmatrix}, \tag{52}$$

$$S_b^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} (\mu^2+\sigma^2)\Omega^{-1} & -\mu\Omega^{-1} & 0 & 0 \\ -\mu\Omega^{-1} & \Omega^{-1} & 0 & 0 \\ 0 & 0 & -\frac{\sigma^2(\kappa_4 - \sigma^4)}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} & \frac{\kappa_3\sigma^2}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} \\ 0 & 0 & \frac{\kappa_3\sigma^2}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} & \frac{\sigma^4}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} \end{pmatrix}, \tag{53}$$

$$D_b = E\left[\frac{\partial g_b}{\partial \theta'}\right] = \begin{pmatrix} -\beta & -\delta I_n & \beta & 0 \\ -\mu\beta & -(\sigma^2 + \mu\delta)I_n & \mu\beta & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{54}$$

We calculate $\left(D_b' S_b^{-1} D_b\right)^{-1}$

$$\left(D_b' S_b^{-1} D_b\right)^{-1} = \begin{pmatrix} \frac{\sigma^2+\delta^2}{\sigma^2}\left(\beta'\Omega^{-1}\beta\right)^{-1}+\sigma^2 & -\frac{\delta}{\sigma^2}\left(\beta'\Omega^{-1}\beta\right)^{-1}\beta' & \sigma^2 & \kappa_3 \\ -\frac{\delta}{\sigma^2}\left(\beta'\Omega^{-1}\beta\right)^{-1}\beta & \frac{1}{\sigma^2+\delta^2}\Omega+\frac{\delta^2\left(\beta'\Omega^{-1}\beta\right)^{-1}\beta\beta'}{\sigma^2(\sigma^2+\delta^2)} & 0 & 0 \\ \sigma^2 & 0 & \sigma^2 & \kappa_3 \\ \kappa_3 & 0 & \kappa_3 & \kappa_4-\sigma^4 \end{pmatrix}. \tag{55}$$

The asymptotic variance of the GMM estimation of δ^* for the single factor non-Gaussian case is

$$Avar(\delta^*) = \frac{\sigma^2+\delta^2}{\sigma^2}\left(\beta'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1}+\sigma^2. \tag{56}$$

Applying the Delta method,[12] the corresponding asymptotic variance of the λ^* parameter –equivalent to (50)– is[13,14]

$$Avar(\lambda^*) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1}+\frac{\sigma^6}{(\sigma^2+\mu\delta)^4}. \tag{57}$$

Considering (35), (50), and by applying some algebra with the support of equations (40) and (57) the asymptotic variance of the risk-premium estimator result is yield.

A.3 Asymptotic variance Beta method - separate estimation

In our paper, $\mu = \delta$, and the estimation of the parameter $\mathbf{B}$ is separate from the estimation of the parameters $\theta^* = \left(\delta,\sigma^2\right) = \left(\mu,\sigma^2\right)$. Then, we have that to estimate the asymptotic variance of the parameter $\delta^* = \mu^*$, we define

$$g_b(r_t,f_t,\theta) = \begin{pmatrix} f_t-\mu \\ (f_t-\mu)(f_t-\mu)'-\boldsymbol{\Omega} \end{pmatrix}, \tag{58}$$

$$S_b = \begin{pmatrix} \boldsymbol{\Omega} & \kappa_3 \\ \kappa_3 & \kappa_4-\otimes_E(\boldsymbol{\Omega},\boldsymbol{\Omega}) \end{pmatrix}, \tag{59}$$

and

$$D_b = E\left[\frac{\partial g_b}{\partial\theta^{*\prime}}\right] = \begin{pmatrix} -\mathbf{I}_{K\times 1} & 0 \\ 0 & -\mathbf{I}_{K\times K} \end{pmatrix}, \tag{60}$$

then

[12] The use of the delta method requires that the parameter estimation–given the sequence X_t–converges to a normal distribution, $\sqrt{T}\,|X_T-\theta| \xrightarrow{D} N(0,\sigma^2)$. In the case the distribution of the factor deviates from the normal distribution, the estimated parameters might deviate from the normal, and the Delta approximation might underestimate the asymptotic variance. In our case, as we estimate δ separately from $\mathbf{B}$ in the next subsection, we use the GMM asymptotic results to provide an exact estimate of the asymptotic variance of δ without using the Delta method.

[13] The equation (57) corrects the Jagannathan et al. (2002) approximation of the asymptotic variance by using the Beta method, that has a difference of $\frac{\sigma^2\delta^4}{(\sigma^2+\mu\delta)^4}$ between the Beta and the SDF methods.

[14] We can observe that in the non-Gaussian case, when using the Beta method, the higher-order moments do not affect the asymptotic variance of the risk premium estimation. This is consistent with the Beta method being a first- and second-order only asset pricing model. In the case of the SDF model, higher-order moments will discount risk premiums, therefore they will affect the asymptotic variance.

$$Avar(\lambda^*) = \left(-\delta\left(\left(\left(\mathbf{\Omega} + \delta\mu'\right)\left(\mathbf{\Omega} + \delta\mu'\right)'\right)^{-1}\mu\right)' + \mathbf{1}_{K\times K}\left(\mathbf{\Omega} + \delta\mu'\right)^{-1}\right)\mathbf{\Omega}. \tag{61}$$

This is a Delta (first-order) approximation. A better approximation is made when considering the definition of λ into the GMM,

$$g_b(f_t, \lambda_b) = \mathrm{E}[r_t m_t] = \mathrm{E}[r_t(1 - \lambda_b f_t)] = (r_t(1 - \lambda_b f_t)) = 0,$$

That can be reduced to

$$g_b(f_t, \lambda_b) = (f_t(1 - \lambda_b f_t)) = 0. \tag{62}$$

The parameter λ_b is not estimated by the SDF method, but by the Beta method and using (19), $\lambda_b = \mu'\left(\mathbf{\Omega} + \mu\mu'\right)^{-1}$. The covariance matrix of $g_b(f_t, \theta)$ in (62) is:

$$\begin{aligned} S_b &= \mathrm{E}[g_b(f_t, \lambda_b) g_b(f_t, \lambda_b)'] \\ &= \mathrm{E}[f_t f_t' - 2 \otimes_R (\lambda_b, \otimes_E(f_t f_t', f_t)) + \otimes_R(\lambda_b \lambda_b', \otimes_E(f_t f_t', f_t f_t'))]. \end{aligned}$$

The partial derivatives of g_b respect to λ_b will produce a matrix

$$D_b = \mathrm{E}\left[\frac{\partial g_b}{\partial \lambda}\right] = -\left(\mathbf{\Omega} + \mu\mu'\right), \tag{63}$$

Then, the asymptotic variance of λ^* is equal to

$$Avar(\lambda^*) = Avar(\lambda_b^*) = \left(\mathbf{\Omega} + \mu\mu'\right)^{-1} S_b \left(\left(\mathbf{\Omega} + \mu\mu'\right)^{-1}\right)'. \tag{64}$$

In the case of single factors we have

$$Avar(\lambda^*) = \frac{\left((\sigma^2 + \mu^2) - 2\lambda_b \mathrm{E}[f_t^3] + \lambda^2 \mathrm{E}[f_t^4]\right)}{(\sigma^2 + \mu^2)}, \tag{65}$$

with

$$\begin{aligned} \mathrm{E}[f_t^3] &= \kappa_3 + 3\sigma^2\mu + \mu^3, \\ \mathrm{E}[f_t^4] &= \kappa_4 + 4\kappa_3\mu + 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4. \end{aligned}$$

A.4 Asymptotic variance pricing errors

We provide the asymptotic variances of the pricing errors. In both cases, SDF and Beta methods, the asymptotic variance of the pricing error is found by defining a sample mean of the estimator:

$$e_s(\hat{\lambda}) = \frac{1}{T}\left(\sum_{i=1}^{T} g_s(r_t, f_t, \lambda)\right), \tag{66}$$

$$e_b(\theta) = \frac{1}{T}\left(\sum_{i=1}^{T} g_b(r_t, f_t, \theta)\right). \tag{67}$$

The SDF pricing error $\hat{\pi}$ will be equal to (66) (Jagannathan and Wang, 2002), then by Hansen (1982):

$$Avar\left(\hat{\pi}\right) = Avar\left(e_b(\hat{\lambda})\right) = \frac{1}{T}\left(\sum_{i=1}^{T} g_s(r_t, f_t, \lambda)\right) = S_s - D_s\left(D_s' S_s^{-1} D_s\right) D_s'. \tag{68}$$

Consider that in the Beta method, the equivalent Jensen's α is:

$$\alpha^* = Q^* e(\theta^*) = [I_n, \mathbf{0}_{n\times n}, -\beta^*, \mathbf{0}_{n\times 1}]\, e(\theta^*). \tag{69}$$

Using equations (67), (68) and (20) the asymptotic variance of the pricing error is yield:

$$Avar\left(\pi^*\right) = \left(\left(\mathbf{\Sigma}_{\epsilon_t} + \delta\mu'\right)\mathbf{\Sigma}_{\epsilon_t}^{-1}(\mathbf{\Sigma}_{\epsilon_t}^{-1})'\left(\mathbf{\Sigma}_{\epsilon_t} + \delta\mu'\right)\right) Q\left(S_b - D_b(D_b' S^{-1} D_b) D_b'\right) Q'. \tag{70}$$

□

COROLLARY 1: *When the Delta method provides an accurate approximation to the asymptotic variances of the estimators, i.e. when the third-, fourth- and higher-order moments in the Taylor expansions for the Delta approximation are negligible,*[15] *considering the typical stylized facts concerning market returns (i.e., a heavy tailed and negative skewed distribution), the asymptotic variance of the single-factor risk premium obtained from the Beta method, $Avar(\hat{\lambda}^*)$, will be lower than the asymptotic variance estimator under the SDF method, $Avar(\hat{\lambda})$. In the case of multiple risk factors, f_t, when they assumed to follow a multivariate Gaussian distribution, the Beta method will imply a lower asymptotic variance of the risk premium estimators compared to the SDF-implied estimator, but the difference will be negligible.*

Proof. By inspecting the SDF- and the Beta- (Delta approximated) implied GMM asymptotic variances for the risk premium, i.e., equations (23) and (25), and after some algebra, we observe that their difference is:

$$Avar(\hat{\lambda}) - Avar(\hat{\lambda}^*) = \frac{\sigma^2(\mu^4)}{(\sigma^2+\mu^2)^4} + \frac{\mu^2\left(\kappa_4 - 2\kappa_3(\frac{\sigma^2-\mu^2}{\mu}) - 3\sigma^2\right)}{(\sigma^2+\mu^2)^4}. \tag{71}$$

Consider the following conditions based on typical stylized facts for the returns on the market portfolio (see, e.g., Christoffersen, 2012) : $\mu < \sigma$ (volatility is higher than expected returns), $3\sigma^2 < \kappa_4$ (heavy tailed returns), and $\kappa_3 < 0$ (negatively skewed returns), then the the right-hand side of (71) is positive and this yields the result. □

By analyzing equation (3), and considering that the error ϵ_t will be assumed– at least in general – to display a standard normal distribution, we note that the source of non-normality of the returns will find its origins in the non-normality of the factors. However, even if we were in the more general case in which ϵ_t may be itself non-Gaussian and hence display non zero skewness and excess kurtosis, by Proposition 1, only the higher-order moments of f_t would be important for the efficiency properties of the Beta- vs. the SDF-based risk premia estimates.

It must be acknowledged that the results from the Corollary depend on the assumption that the Delta approximation is "very accurate". Nevertheless, in actual applications the systematic risk factors may possess distributions that are complex enough to interfere with the quality of such approximation. Even though the Corollary may not apply in an exact sense, the simulation experiments that follow show that in a qualitative sense, the implications of the Corollary hold in all (at least, most) interesting cases.

[15] For instance, higher-order moments are much lower than the second-order moments: $|\kappa_i - \hat{\kappa_i}| < 1e^{-10}\sigma^2$ for $i \geq 3$ with κ_i the i-th cumulant of the factor distribution, and $\hat{\kappa_i}$ the i-th cumulant in the Gaussian case.

IV. Simulation Experiments

Applied asset pricing research that routinely deals with estimation issues is often confronted with data sets of finite, and occasionally rather modest, sample size.[16] Even though the qualitative flavor of our results concerning asymptotic, relative efficiency is hopefully clear, it is imperative to investigate the relative small sample performance of Beta vs. SDF-based inference. Because finite-sample analytical results can be obtained only under specific, hence always fragile, distributional assumptions on returns, factors, and errors, it is customary in literature to resort to large-scale simulation designs , which allow to alter the simulation inputs and therefore develop a deep understanding of how sensitive the results may with respect to the various features of the data generating process that an empirical researcher is likely to encounter in applied work.

We use bootstrap simulations to study whether the GMM estimators and test statistics carry any biases and their relative efficiency. [17] In particular, we are interested in evaluating the standard deviations of $\widehat{\lambda}^*$, $\widehat{\lambda}$, $\widehat{\pi}^*$, $\widehat{\pi}$, denoted as $\sigma(\cdot)$ and also the thickness of the tails of the distribution of the J-statistic used to conduct specification tests. The simulations represent a (nonparametric) bootstrap in the sense that we assume that the factors f_t are drawn from their empirical distribution which allows for non-normalities, autocorrelation, heteroskedasticity and dependence of factors and residuals.

To artificially generate the excess returns we use the factor model, equation (4) where $t = 1, ..., T$. We consider two experimental set ups: (i) a Monte Carlo simulation, where the returns are generated by adding a multivariate normal error to the actual, observed empirical factors, and (ii) an empirical simulation, where the returns are generated by bootstrapping the observed historical returns, and factors are generated by bootstrapping the observed historical factors. The first experiment is designed to test the convergence of the analytical asymptotic variance results reported in Section III, and to study the relationship between: (i) the factors higher-order moments variations and (ii) the efficiency and precision in the GMM estimation of the Beta and SDF parameters ($\hat{\lambda}$ and $\lambda^*, \hat{\pi}$, and π^*) in a controlled environment; while the second set of empirical simulations serves as an empirical exercise with historical market data observations that might be useful for empirical researchers.

As far the overall sample size, T, is concerned, we consider four alternative time spans: 60, 600, 2000, and 3000 months in the first experiment (Monte Carlo simulation) and 60, 360, 600, and 1000 months in the second (bootstrapped empirical simulation). As Shanken and Zhou (2007) argue, varying T is useful in order to understand the small-sample properties of the tests and the validity of any asymptotic approximations invoked. For instance, we examine a 5 year, 60-observation window because this may show how distorted any inferences may potentially be from taking a really short sample, whilst this is a commonly adopted horizon when using rolling window recursive estimation schemes. Instead, a 30-year window corresponds approximately to the sample sizes in Fama and French (1992, 1993) and Jagannathan and Wang (1996) while the 600-month long sample matches the largest sample sized examined by Jagannathan and Wang (2002). We also examine $T = 1000$ months since this approximates the current size of the largest sample available in Kenneth French's public data library [January 1927 to December 2018 – 1104 months as of

[16] In the case of US data, the time series dimension may attain approximately 1,100 monthly historical observations; yet, already in the case of UK data, one may easily end up with a sample size capped at about 450 observations.

[17] The simulations were executed using the THOR Grid computational cluster provided by the Department of Economics,Finance and Accounting at the Maynooth University.

the writing of this paper], and could be considered as an approximation to the true asymptotic variance. The estimators and specification tests are calculated based on the T-long samples of factors and returns generated from the model. We repeat such simulation experiments independently to obtain 10,000 draws of the estimators of λ, π (the pricing errors) and J (the over-identifying restriction statistic).

While earlier studies (e.g., Kan and Zhou (1999, 2001), Jagannathan and Wang (2002) and Cochrane (2005)) simply focussed on the CAPM model to compare the efficiency of the Beta and SDF methods, we evaluate them with reference to standard multi-factor models and in particular the Carhart four-factor models, which means that $K = 1$, and $K = 4$. The factors as the excess market return (RMRF), size (SMB), value (HML) and momentum (UMD).[18] To generate the excess returns from equation (4) we first need the $N \times K$ matrix $\mathbf{B}$, capturing the sensitivity of returns to the factor(s). The $\mathbf{B}$ matrix previously defined in equation (2), represents the slope coefficients in the OLS regressions of each N-test portfolio and K-factor model. We use three values of N to generate $\mathbf{B}$, i.e., the value weighted returns of the 10 size-sorted portfolios, the 25 Fama-French portfolios (the intersections of the 5 size and 5 book-to-market portfolios) and the 30 industry portfolios.[19] As Lewellen et al. (2009) suggest, the traditional tests portfolio used in empirical work such as the size and 25 size/value sorted portfolios frequently present a strong factor structure, hence it seems reasonable to adopt other criteria (industry) for sorting. The combination of three different values for K and three values for N give rise to nine $\mathbf{B}$ matrices, allowing us to add another criteria for evaluating the method's performance, in this case measured by efficiency.[20]

In Table I we report the descriptive statistics for the time series of factors and test portfolios. These values are used to calibrate the simulations of the two sets of simulations experiments, As can be noted in Table I for US data, the additional factors characterizing the multi-factor models display rather different statistical properties vs. the classical, excess market return factor. In particular, with a sample kurtosis in excess of 30, the momentum factor is almost three times more leptokurtic than the excess market return (10.8). Therefore, in this paper we entertain the empirical distribution as the most realistic alternative to the classical, multivariate normal and resort to a bootrstap design.

[Place Table I about here]

We find that the choice of following either the Beta or the SDF method to empirically estimate and evaluate an asset pricing model can be framed in terms of a choice between efficiency and robustness. In particular, we show that frequently the Beta method dominates in terms of efficiency whereas the SDF method dominates in terms of robustness. While in a celebrated paper Jagannathan and Wang (2002) argue that the Beta and SDF approaches lead to parameter estimates with similar precision even in finite samples, in what follows we illustrate that their conclusions are only valid under rather specific conditions that cannot be uncritically generalized.

[18] See Fama and French (1993), for a complete description of these factors.

[19] 25 Fama-French and 30 industry portfolios results are provided in the online appendix.

[20] The covariance matrix $\mathrm{E}[\epsilon_t \epsilon_t' \mid f_t]$ in equation (4), is set equal to the sample covariance matrix of the residuals obtained in the N OLS regressions.

A. Monte Carlo Simulation: Convergence

Figures 1a, 1b, 2a, and 2b present the asymptotic variance results from estimating λ^* by GMM under the Beta method and λ_2^U by GMM under the SDF method. The Monte Carlo simulation parameters used are those in Tables I of the Online Appendix (16 Tables), respectively. We observe that the asymptotic variance obtained from writing the asset pricing model in a Beta framework ($Avar(\lambda^*)$) is always lower than the asymptotic variance from GMM applied under the second-stage non-centered SDF method ($Avar(\lambda_2^U)$).[21]

[Place Figure 1 about here]

[Place Figure 2 about here]

The results in Figure 1 show that, in the case of a single-factor model based on the market risk factor, for which sample skewness is closer to zero and the kurtosis is the lower vs. all single-factor models considered in our simulations, the Beta and the SDF Monte Carlo simulations and the Beta and SDF analytic estimated asymptotic variances converge towards the same value, consistent with Jagannathan and Wang (2002). Nevertheless, in the case of the size, value, and momentum single-factor models, the SDF estimated asymptotic variance is always higher than the one estimated in a Beta framework: the higher third-order cumulant (κ_3) produces an increase in the asymptotic variance, consistent with analytical results in the Section III.

B. Bootstrap simulations: Comparison of λ estimators

Tables II and III compare the performance of Beta and SDF methods at estimating λ by the single-factor CAPM model using US data, and by computing ratios of relative standard errors such as $\sigma_r(\lambda) = \sigma(\lambda)/\mathrm{E}(\lambda)$. As one would expect, our results are qualitatively and quantitatively similar to those in Jagannathan and Wang (2002). In particular, Table II shows that the expected value and the standard error of $\widehat{\lambda}^*$ and $\widehat{\lambda}$ are indeed similar. In fact, we cannot reject the null hypothesis that the standard errors of the Beta estimators are equal to the standard errors of the SDF estimators in most of the cases. Therefore, under a specific, single-factor framework there are virtually no differences in terms of efficiency between the Beta and SDF methods. One of the key implications of this result is that there are no significant advantages to applying the Beta method to nonlinear asset pricing models formerly expressed in SDF representation through linear approximations.

[Place Table II about here]

[Place Table III about here]

[21] Results for the first-stage SDF ($Avar(\lambda_1^U)$) and first-, and second-stage centered methods ($Avar(\lambda_1^C), Avar(\lambda_2^C)$) are not reported, but the resulting variances are all uniformly higher than the second-stage non-centered SDF method ($Avar(\lambda_2^U)$).

In order to conduct a clearer comparison of the standard errors of λ we present Table III. Table II allows us to compare $\sigma(\widehat{\lambda}^*)$ versus $\sigma(\widehat{\lambda})$ instead of $\sigma(\widehat{\delta})$ versus $\sigma(\widehat{\lambda})$ in order to avoid misleading conclusions driven by a scaling issue. However, if we consider the possibility of intrinsic differences among the methods, the expected values of $\widehat{\lambda}^*$ and $\widehat{\lambda}$ do not necessarily would be similar in general. For this reason it is convenient to compute ratios of relative standard errors such as $\sigma_r(\lambda) = \sigma(\lambda)/\mathrm{E}(\lambda)$. By doing so, we would have an accurate measure of the relative efficiency of the methods.

Ratios close to one represent a high degree of similarity of the efficiency of both methodologies at estimating λ. Ratios in excess of one suggest that applying the Beta method provides more accurate estimators of risk premia (even in finite samples and net of scaling effects) vs. the SDF method.

Even though our set of simulation experiments offer a clear-cut perspective on efficiency matters, our original contribution regarding the comparison of $\sigma_r(\lambda)$ across representations of the asset pricing models focuses on of multi-factor asset pricing models.

Tables IV, V and VI report the ratios of relative standard errors for the risk premia estimators derived from the Beta and SDF methods in the case of the Fama-French, RUH and Carhart models respectively.

[Place Table IV about here]

[Place Table V about here]

[Place Table VI about here]

The correspondent expected values and standard errors for the multi-factor asset pricing models are in Tables II, III and IV of the Online Appendix.

The first (upper) panel of Table IV is comparable to the Table III because in both cases the estimated λ correspond to the market factor. Thus, it is not surprisingly to find a similar pattern which reinforce the conclusion that there are no significative differences when estimating the parameter associated with the market factor. The un-centred SDF method is again more efficient than the centred SDF method at estimating λ. And this time the second-stage un-centred method is marginally more efficient than the Beta method for samples smaller than one thousand.

Contrary to the market factor case, the standard error of Beta estimators linked to the size and value factors are statistically smaller than the correspondent standard errors of SDF estimators. This become evident in the higher ratios of second and third panels of Table IV. These results suggest that the empirical equivalence of both methods is subject to the loaded factor in the asset pricing model. In particular, the market factor do not represent a challenge to the SDF method whilst the value factor can lead to significative differences according to the $\sigma_r(\lambda)$ ratios. For instance, the relative standard error of the un-centred first-stage SDF method can be more than twice as big than the relative standard error of the Beta method. Beta estimators are even more efficient than second-stage SDF estimators, which by construction are intended to increase the estimation efficiency of $\widehat{\lambda}$.

The second multi-factor asset pricing model is the RUH, which factors are market, momentum and value. The estimation of the RUH model allow us to compare the efficiency of the estimator associated with the market and value factors in previous tables and introduces the result for momentum.

The ratios of relative standard errors $\sigma_r(\lambda)$ linked to the market factor are slightly lower than one; however the standard errors $\sigma(\lambda)$ are statistically equal for most of the cases. On the other hand, the magnitudes for the value factor are similar to the ones of the Fama-French model in Table IV. The second panel of Table V show the ratios for the momentum factor, which are somewhat greater than the ratios for value factor. Here, we also find that SDF method may have difficulties in small samples which is reflected in values of 20.7 and 10.5.

The third and last multi-factor asset pricing model estimated is the Carhart model, that features as factors the market, size, value, and momentum. The estimation of Fama-French, RUH and Carhart models represent the core contribution to the field, which lead to the main original results. However, we also present results for other test portfolios and to reduced sample sizes such as the ones researchers may face at evaluating models using UK data samples. For now, we show the relative standard errors of Carhart model λ estimators in Table VI.

The results for Carhart's risk premia estimated on US data support the argument that the efficiency of the different methods/representations depends on the specification of the factors. The lower ratios of relative standard errors of the λ estimators are those related to the size factor, followed by market, value, and are the highest for the momentum factor. If the Beta and SDF representations were equally efficient, these ratios should instead be similar and close to one across all factors. Clearly, in Table VI, this fails to be the case, and this suggests that the inferences based on the SDF method may marginally be superior to Beta-driven ones in first and second panels which correspond to the results associated with the market and size factors, whilst Beta estimators outperform the SDF ones in the third and fourth panels, displaying efficiency results for value and momentum.

The evidence in Table VII shows that the Beta representation leads to more accurate inferences than the SDF one, at least in terms of inferences on the risk premia. The key implication of Table VII is that, because it is built by setting $K = 1$ in all the single.-factor models we experiment with, we artificially fix every element in the assessment of the pricing models except for the assumed risk factor. Given that ratios across panels are different, the statistical characteristics of each factor as well as their relation to the test portfolios are presumably the main drivers behind the differences of the methods at estimating λ.

[Place Table VII about here]

C. Inference on Pricing Errors

We now turn our attention to the pricing errors estimates $\widehat{\pi}$, $\widehat{\pi}^*$, to their corresponding standard errors $\sigma(\widehat{\pi})$, $\sigma(\widehat{\pi}^*)$, and the associated $\widehat{J}$, and $\widehat{J}^*$ statistics. A given representation of an asset pricing model would turn out to display a superior robustness vs. the other if the simulated standard error of the pricing errors were lower.

To better understand where the advocated trade-off between efficiency and pricing robustness comes from, we need to briefly describe the set of moments used in estimation under each representation. On the one hand, the traditional Beta GMM restrictions incorporate three sets of moments: (1) the N asset pricing restriction which define the α vector; (2) the $N \times K$ zero covariances between the errors and the factors; and (3) the K definitions of δ, which equals the mean of the traded factors. Therefore, by imposing the definition of δ, the Beta approach increases its relative estimation efficiency. On the other hand, the SDF representation is simply based on the N asset pricing restrictions for the un-centred specification which defines π; and on the N asset pricing restriction plus the K mean definitions for each of the factor risk premia in the case of the centred specification. However, the inferences based on the SDF representation fail to impose the definition of λ. As a result, they allow for freely varying risk premia estimates in order to achieve lower mean pricing errors, favoring pricing robustness over efficiency. By the same token, the specification tests derived under the Beta representation tend to under-reject in finite samples while the SDF-based tests approximately display the correct size.

Table VIII shows the relative standard errors of the pricing errors for the single and multi-factor models. Clearly, most of the ratios of the normalized standard errors of the pricing errors are now below one and this tends to be stronger in the case of smaller vs. larger sample sizes. Table VI of the Online Appendix presents the values used in the construction of Table VIII.

[Place Table VIII about here]

D. Higher-order Moments Variation

In this section, we propose to perform an alternative simulation exercise different from the one explained in subsection IV.B, to empirically reveal how the statistical characteristics of factors and their relation to the test portfolios relates to the variance of $\widehat{\lambda}^*$ and $\widehat{\lambda}$. In particular, we simulate 200 series of size $T = 996$ for the market, size, value and momentum factors. The particularity of this new factor simulation is that the mean, skewness and kurtosis are set to be less than 5% different than the original historical series (see Table I for the descriptive statistics of the original samples of returns and factors). The new resulting simulated series have different variances but virtually identical mean, skewness and kurtosis, whilst the previous factors generated from the empirical distribution allow for greater variation in the mean, variance, skewness and kurtosis since we do not impose any restriction to the data-generating process.

Figures 3 and 4 show the results for US data. For each of these Figures we plot the values of $\widehat{\lambda}^*$ and $\widehat{\lambda}_1^U$ with respect to the factor variance and the second moment of returns and factors. We do not present the comparison of the centred SDF method since results are quite similar; however, they are available upon request.

[Place Figure 3 about here]

The mean of the factor is equal to the δ estimator in the Beta method, so any fluctuation on $\widehat{\delta}$ is independent of the higher order factor moments. Therefore, the negative relation between $\widehat{\lambda}^*$ and the variance of the factor is straightforward explained by the definition of $\widehat{\lambda}^*$ in equation 19. In other words, the negative relation on the upper-left panel of Figure 3 is mainly driven by the transformation from $\widehat{\delta}$ to $\widehat{\lambda}^*$. On the other hand, the lower-left panel show that the values of $\widehat{\lambda}_1^U$ are extremely sensible to variations in the factor variance $\mathbf{\Omega}$. This high sensitivity of $\sigma(\widehat{\lambda}_1^U)$ is present for low or high values of $\mathbf{\Omega}$. In sum, the variance of the factor does not seem to have a significant role at explaining the fluctuations of Beta estimators since $\widehat{\delta}$ is independent of $\mathbf{\Omega}$. However, the fluctuations of SDF estimators seem to be quite sensible to $\mathbf{\Omega}$.

Right panels of Figure 3 relates the values of Beta and SDF λ estimators to the second moment of returns and factor. As expected, Beta estimators are independent on the covariance of returns and factors.

Contrary, SDF estimators reveal a clear relation with respect of the second moment of returns and factors. To better understand this relation we can refer to equation (15) in which we illustrate the moment conditions of the un-centred SDF method which is minimized by GMM. The equation (15) imply that $\mathrm{E}[r_t] \approx \lambda \mathrm{E}[r_t f_t]$ where $\mathrm{E}[r_t f_t]$ represent the second moment matrix of returns and factor. Thus, relative high values of $\mathrm{E}[r_t f_t]$ represent an advantage to the SDF method at estimating $\widehat{\lambda}_1^U$ with greater precision. By the same token, values of $\mathrm{E}[r_t f_t]$ close to zero would represent a drawback of the SDF method at estimating $\widehat{\lambda}_1^U$ because the method would allow all the necessary variation on the estimator in order to validate $\mathrm{E}[r_t] \approx \lambda \mathrm{E}[r_t f_t]$.

This evidence reveal that results such as the ones presented on Jagannathan and Wang (2002) which argue that there are no differences in the efficiency of the Beta and SDF methods, are not driven by intrinsic similarities of the methods. In fact, the similarity in terms of efficiency is driven by the relative high value of $\mathrm{E}[r_t f_t]$. In other words, if the Beta and SDF methods lead to similar levels of efficiency at estimating λ can be mostly explained by the statistical properties of both the returns and factors rather than any empirical equivalence of the methods.

It is worthwhile to clarify that $\mathrm{E}[r_t f_t]$ is actually a $N \times K$ matrix, which for our case is a 10×1 vector. The values of $\mathrm{E}[r_t f_t]$ in Figure 3 are in fact the average of each vector values.

Figure 4 deserves a special attention since the momentum factor has a skewness equal to -3.03, and this high and negative value has important implications in the comparison of both methods.

[Place Figure 4 about here]

As far as the factor exhibits a large and negative skewness, the chances that the covariance of returns and factor become negative increases. A negative value of $\mathrm{E}[r_t f_t]$ would force the SDF method to deliver a negative λ in order to validate $\mathrm{E}[r_t] \approx \lambda \mathrm{E}[r_t f_t]$ because in general $\mathrm{E}[r_t] > 0$. This is less likely to happen in the Beta method since the value of the estimator basically depend on the first moment of the factor. Figure 4 show that the value of $\mathrm{E}[r_t f_t]$ for momentum factor is 10.83, which is higher than those on size and value factor, but less than the correspondent of the market factor. For one side, the higher covariance with respect to the size and value factors tend to decreases the fluctuations of $\widehat{\lambda}$, but the marked negative skewness is the responsible of the negative value of the estimator, which is equal to -10.88, and equal to 2.932 for the Beta method.

Our results show that the efficiency of SDF method represented by the value of $\sigma(\widehat{\lambda})$ is highly sensible to the low covariance of the returns and factors. Besides, the higher order moments of factors such a negative skewness may cause negative values of λ which are at least not easily interpreted in economic terms. The second-stage estimators naturally increase the efficiency of the SDF method, however this gains are not enough to outperform the efficiency of Beta method. We could force the SDF method to deliver λ estimators with greater accuracy by including its definition – see equation (19) – into the moment restrictions of the un-centred SDF method in equation (15), or into the centred moment restrictions of the SDF method in equation (17). Nevertheless, this alternative is not the way the SDF method is usually implemented, and our main interest is evaluating the Beta and SDF methods as existing approaches.

Results such as in Jagannathan and Wang (2002) are conceivable because the covariance of the returns and factors is especially large for the market factor and because this particular factor is almost centred. However, we show that factors with a negative skewness, and a low covariance with respect to the returns such as momentum, entail significant differences in the estimation efficiency of λ and δ estimators.

E. Further Robustness Checks

In a first additional inquiry, we explore the specific role of the sign of the skewness of the factors. We find that such skewness is likely to determine the sign of λ rather than the sign of δ. To clarify this finding, it is practical to describe the un-centred SDF method as a cross-sectional regression of mean excess returns on the second moment matrix of returns and factors. Thus, the N moment restrictions which define π are equal to the product of λ and the second moment matrix of returns and factors minus the expected returns. It turns out that if the (single, for simplicity) factor is left-skewed, it is more likely that the second moment covariance matrix between returns and factors would be negative. When this occurs, λ should be negative in order to minimize the pricing errors. Naturally, a negative λ is not what we normally expect, and it usually remains difficult to give it an economically meaningful interpretation. This is unlikely to occur under the Beta method because a subset of the moment restrictions define the value of δ. However, what does decrease the efficiency in the estimation of the λ risk premia, actually increases the robustness of the SDF representation-based inferences.

Perhaps more important is the effect of the magnitude of the second moment matrix of returns and factors over λ. Our (unreported, but available upon request) experiments confirm that there is a strikingly close relationship between a low covariance between factors and returns (in pairs), and highly volatile estimators of λ. This is easy to grasp when such covariances are slightly in excess of zero: then λ should be considerably large in order for the product between risk premia and the risk exposures implied by such covariances to equate the expected returns. In the same way, if the value of the covariance between the factors and returns is slightly negative, then λ should be considerably negative to allow the product described above to satisfy the pricing restrictions in the sample. This of course represents a valid reason to favor the inclusion of factors which display large covariation with asset returns in SDF models.

Because we have reported results for two different implementation, centered and un-centered, of the SDF representation, it is of interest to also analyze their relative performance, even though this is not the main core of our paper. The standard error of λ is consistently lower for the un-centred representation across model representations and sample sizes. This may reflect the additional K moment restrictions that appear in the centred characterization that evidently decreases estimation efficiency of λ relative to the un-centred specification. Regarding the standard deviation of the pricing errors estimates of π, the first-stage un-centred representation also delivers lower standard errors relative to the first-stage centred representation; however this is less evident for the second-stage.

V. Conclusion

The interest in learning about the asymptotic and finite sample properties of parameter estimators (and their functions) in asset pricing models, like risk premia and pricing errors, have attracted the attention of researchers for decades. This attention is motivated by an extensive list of theoretical and empirical applications mainly – but not exclusively – in economics and finance. It is not uncommon to find examples in which different econometric approaches involve a trade-off between efficiency and robustness since the most efficient estimators may possess this property at the cost of higher pricing errors and vice versa. However, to best of our knowledge, this is the first time in which such a formal dichotomy is explicitly used to better understand the difference between the statistical properties of inferences derived from the

Beta and SDF representations. The simulation evidence that we have presented in this paper is useful to researchers and practitioners because they could choose a proper procedure given the goals of their application, i.e., whether accurate pricing and model over-identification testing may be more relevant vs. the goal of producing accurate estimates of the risk premia. For instance, in cost of capital estimation or when multiple asset pricing models need to be compared, a choice for robust pricing and testing guaranteed in many occasions by a SDF framework may be sensible, while in portfolio choice and asset management applications where to have a clear idea of what risks are compensated and in what amount, appears of primary importance and probably best obtained from a Beta representation of the model.

As always, there are a number of possible extensions that it may be worthwhile to pursue. Chiefly, it may be of high relevance to the literature to explore what happens to GMM estimators when one considers also the non-traded factors. In principle, there is no reason to expect a similar pattern to hold. For example, Kan and Robotti (2008) show that the standard errors under correctly specified vs. potentially misspecified models are similar for traded factors, while they can differ substantially for non-traded factors such as a scaled market return factor and the lagged state variable CAY. One additional extension would consist of providing examples of the economic value that can be generated, especially in the presence of many factors deviating from joint normality, in financial applications (for instance, capital budgeting vs. portfolio selection) in which accurate pricing vs. efficient estimation of the risk premia may carry differential importance.

References

Amsler, C. E. and P. Schmidt (1985). A monte carlo investigation of the accuracy of multivariate capm tests. *Journal of Financial Economics 14(3)*, 359–375.

Arismendi, J. C. and H. Kimura (2016, jul). Monte carlo approximate tensor moment simulations. *Numerical Linear Algebra with Applications 23*(5), 825–847.

Balduzzi, P. and T. Yao (2007). Testing heterogeneous-agent models: An alternative aggregation approach. *Journal of Monetary Economics 54(2)*, 369–412.

Black, F., M. Jensen, and M. S. Scholes (1972). The capital asset pricing model: Some empirical tests. *Studies in Theory of Capital Markets. Jensen, Ed. New York: Preager Publishers. 17*, 79–121.

Burnside, A. C. (2007, aug). Empirical asset pricing and statistical power in the presence of weak risk factors. Technical report, NBER.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance 52(1)*, 57–82.

Chen, R. and R. Kan (2004). Finite sample analysis of two-pass cross-sectional regressions. *The Review of Financial Studies (forthcoming)*.

Christoffersen, P. F. (2012). *Elements of Financial Risk Management.* Elsevier.

Cochrane, J. (2001, oct). A rehabilitation of stochastic discount factor methodology. Technical report, NBER.

Cochrane, J. H. (2005). *Asset Pricing - Revised Edition.* Princeton University Press.

Dittmar, R. F. (2002, feb). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance 57*(1), 369–403.

Dybvig, P. H. and J. E. Ingersoll (1982). Mean variance theory in complete markets. *Journal of Business 55(2)*, 233–251.

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

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

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

Farnsworth, H., W. Ferson, D. Jackson, and S. Todd (2002). Performance evaluation with stochastic discount factors. *Journal of Business 75(3)*, 473–503.

Ferson, W. E. (2005). *Tests of Multifactor Pricing Models, Volatility Bounds and Portfolio Performance.* Handbook of the Economics of Finance. Elsevier Science Publishers, North Holland.

Ferson, W. E. and R. Jagannathan (1996). *Econometric Evaluation of Asset Pricing Models.* G. S. Maddala and C. R. Rao, Eds. Handbook of Statistics, Vol. 14: Statistical Methods in Finance. Elsevier.

Guidolin, M. and A. Timmermann (2008, April). International asset allocation under regime switching, skew, and kurtosis preferences. *Review of Financial Studies 21*(2), 889–935.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50(4)*, 1029–1054.

Hansen, L. P. and R. Jagannathan (1997). Assessing specification errors in stochastic discount factor models. *Journal of Finance 55(2)*, 557–590.

Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable. *Econometrica 55(3)*, 587–613.

Harvey, C. R. and A. Siddique (2000, 06). Conditional skewness in asset pricing tests. *The Journal of Finance 55*(3), 1263–1295.

Ingersoll, J. E. (1987). *Theory of Financial Decision Making.* Rowman and Littlefield Publishing, Inc.

Jagannathan, R., G. Skoulakis, and Z. Wang (2002). Generalized method of moments: applications in finance. *Journal of Business and Economics Statistics 20(4)*, 470–481.

Jagannathan, R., G. Skoulakis, and Z. Wang (2008). The analysis of the cross section of security returns. *Handbook of Financial Econometrics. Edited By Yacine Ait-Sahalia and Lars Hansen 1*, ch. 14.

Jagannathan, R. and Z. Wang (1996). The conditional capm and the cross-section of expected returns. *Journal of Finance 51(1)*, 3–53.

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

Jagannathan, R. and Z. Wang (2002). Empirical evaluation of asset-pricing models: A comparison of the sdf and beta methods. *Journal of Finance 57(5)*, 2337–2367.

Julliard, C. and J. A. Parker (2005). Consumption risk and the cross-section of expected returns. *Journal of Political Economy 113(1)*, 185–222.

Kan, R. and C. Robotti (2008). Specification tests of asset pricing models using excess returns. *Journal of Empirical Finance 15(5)*, 816–838.

Kan, R. and C. Robotti (2009). A note on the estimation of asset pricing models using simple regression betas. *SSRN Electronic Journal*.

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

Kan, R. and G. Zhou (1999). A critique of the stochastic discount factor methodology. *Journal of Finance 54(4)*, 1221–1248.

Kan, R. and G. Zhou (2001). Empirical asset pricing: the beta method versus the stochastic discount factor method. *Unpublished working paper. Rotman School of Management, University of Toronto*.

Lewellen, J., S. Nagel, and J. Shanken (2009). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics Forthcoming (2009)*, 175–194.

Lintner, J. (1965). Security prices, risk and maximal gains from diversification. *Journal of Finance 20(4)*, 587–615.

Lozano, M. and G. Rubio (2011, jan). Evaluating alternative methods for testing asset pricing models with historical data. *Journal of Empirical Finance 18*(1), 136–146.

MacKinlay, A. C. and M. P. Richardson (1991). Using generalized method of moments to test mean-variance efficiency. *Journal of Finance 46(2)*, 511–527.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica 34(4)*, 768–783.

Peñaranda, F. and E. Sentana (2015, may). A unifying approach to the empirical evaluation of asset pricing models. *Review of Economics and Statistics 97*(2), 412–435.

Shanken, J. and G. Zhou (2007). Estimating and testing beta pricing models: Alternative method and their performance in simulations. *Journal of Financial Economics 84(1)*, 40–86.

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

Velu, R. and G. Zhou (1999). Testing multi-beta asset pricing models. *Journal of Empirical Finance 6(3)*, 219–241.

Yogo, M. (2006). A consumption-based explanation of expected stock returns. *The Journal of Finance 61(2)*, 539–580.

Appendix A. Tables

Table I. Descriptive statistics of factors and test portfolios.

The data on US factors and portfolios are taken from Kenneth French's library. The sample spans the period January 1927 – December 2018 ($T = 1104$). The data on UK factors and portfolios are taken from Alan Gregory's library. The sample spans the period October 1980 – December 2017 ($T = 447$).

	Factors				Portfolios		
	Market	Size	Value	Momentum	10 size	25 size-value	30 industry
				US			
Mean	0.6471	0.2096	0.3682	0.6617	0.8642	0.8786	0.7239
Std. Dev.	5.34	3.19	3.48	4.69	6.8478	6.9242	5.5991
Skewness	0.19	1.93	2.18	-3.06	1.3791	1.3957	0.3583
Kurtosis	10.81	22.28	22.18	30.90	18.3248	18.2319	11.8010
				US - Recession Periods			
Mean	-0.4320	-0.1292	0.2696	0.6361	-0.4787	-0.4481	-0.3041
Std. Dev.	8.20	3.36	5.05	7.43	10.1792	10.4137	8.4534
Skewness	0.40	0.48	3.05	-3.24	1.4465	1.3667	0.5638
Kurtosis	6.83	5.02	22.84	21.98	12.3939	11.4992	7.5031
				UK			
Mean	0.5482	0.1318	0.2889	0.9794	0.8190	-	-
Std. Dev.	4.39	3.09	3.14	4.24	4.6180	-	-
Skewness	-1.01	-0.13	-0.51	-0.93	-0.7562	-	-
Kurtosis	6.86	5.82	9.41	8.50	6.6923	-	-

Table II. Expected value and standard errors for CAPM model: US data, 10 size-sorted portfolios.

The table presents the expected value and the standard error of the GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda})/\mathrm{E}(\widehat{\lambda})$. The returns are generated by Equation (3) under the null hypothesis with the market risk factor moments being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\mathrm{E}[\widehat{\lambda}^*]$	$\mathrm{E}[\widehat{\lambda}_1^U]$	$\mathrm{E}[\widehat{\lambda}_2^U]$	$\mathrm{E}[\widehat{\lambda}_1^C]$	$\mathrm{E}[\widehat{\lambda}_2^C]$	$\sigma(\widehat{\lambda}^*)$	$\sigma(\widehat{\lambda}_1^U)$	$\sigma(\widehat{\lambda}_2^U)$	$\sigma(\widehat{\lambda}_1^C)$	$\sigma(\widehat{\lambda}_2^C)$
60	2.38	2.40	3.19	2.59	3.00	2.67	2.92	3.60	3.26	3.57
360	2.17	2.17	2.29	2.22	2.28	1.03	1.10	1.09	1.16	1.12
600	2.18	2.18	2.25	2.22	2.25	0.78	0.83	0.81	0.87	0.83
1000	2.14	2.14	2.18	2.18	2.20	0.62	0.66	0.63	0.68	0.65

Table III. Relative standard errors of risk premia estimated from the CAPM model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda})/\mathrm{E}(\widehat{\lambda})$. The returns are generated by Equation (3) under the null hypothesis with the market risk factor moments being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\widehat{\lambda}^*)$
60	1.0833	1.0045	1.1219	1.0581
360	1.0654	1.0046	1.0935	1.0324
600	1.0656	1.0030	1.0921	1.0288
1000	1.0658	1.0039	1.0918	1.0295

Table IV. Relative standard errors of of risk premia estimated from the Fama-French model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda})/\mathrm{E}(\widehat{\lambda})$. The returns are generated by Equation (3) under the null hypothesis with the factors (Market, Size, and Value) moments being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\widehat{\lambda}^*)$
		Market		
60	0.9853	0.9649	1.0102	1.0042
360	1.0158	1.0024	1.0379	1.0247
600	1.0187	1.0043	1.0412	1.0279
1000	1.0229	1.0074	1.0447	1.0304
		Size		
60	0.7075	0.6684	0.5861	0.5613
360	1.1403	1.0304	1.1123	0.9963
600	1.1606	1.0453	1.1425	1.0230
1000	1.1717	1.0562	1.1600	1.0430
		Value		
60	2.3382	1.8642	2.7692	2.8484
360	3.0934	2.3801	3.1525	2.5604
600	3.1171	2.4012	3.1628	2.5191
1000	3.2271	2.4994	3.2641	2.5791

Table V. Relative standard errors of risk premia estimated from the RUH model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda})/\mathrm{E}(\widehat{\lambda})$. The returns are generated by Equation (3) under the null hypothesis with the factors (Market, Momentum, and Value) moments being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\widehat{\lambda}^*)$
		Market		
60	0.9533	0.9134	0.9434	0.9232
360	1.0004	0.9868	1.0101	0.9992
600	0.9940	0.9842	1.0090	0.9991
1000	0.9924	0.9817	1.0109	0.9999
		Momentum		
60	2.4843	1.8939	20.7738	10.5442
360	3.3341	2.6710	5.2603	4.1297
600	3.4321	2.8270	4.5009	3.6950
1000	3.5491	2.8415	4.1864	3.3457
		Value		
60	1.6279	1.2432	2.5987	3.3452
360	2.6050	2.1012	3.0128	2.6525
600	2.7729	2.2583	3.0513	2.6226
1000	2.8326	2.2917	3.0234	2.5272

Table VI. Relative standard errors of risk premia estimated from the Carhart model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda})/\mathrm{E}(\widehat{\lambda})$. The returns are generated by Equation (3) under the null hypothesis with the factors (Market, Size, Value, and Momentum) moments being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\widehat{\lambda}^*)$
		Market		
60	0.9247	0.9078	0.9183	0.9261
360	0.9844	0.9787	0.9911	0.9861
600	0.9720	0.9658	0.9824	0.9775
1000	0.9737	0.9674	0.9877	0.9820
		Size		
60	0.0710	0.0682	0.0525	0.0525
360	0.7827	0.7292	0.6856	0.6543
600	0.8637	0.7912	0.7847	0.7320
1000	0.8916	0.8195	0.8352	0.7775
		Value		
60	1.5553	1.3347	4.1452	3.8806
360	3.0528	2.5920	4.1765	3.4226
600	3.1199	2.6862	3.7705	3.1938
1000	3.3112	2.8430	3.7456	3.1735
		Momentum		
60	2.4387	2.0441	17.4520	12.3867
360	3.1899	2.8505	5.1730	4.3170
600	3.1007	2.8569	4.1378	3.6745
1000	3.1815	2.9759	3.7893	3.4642

Table VII. Relative standard errors of risk premia estimated from four alternative single-factor models: US data, 10 size-sorted portfolios.

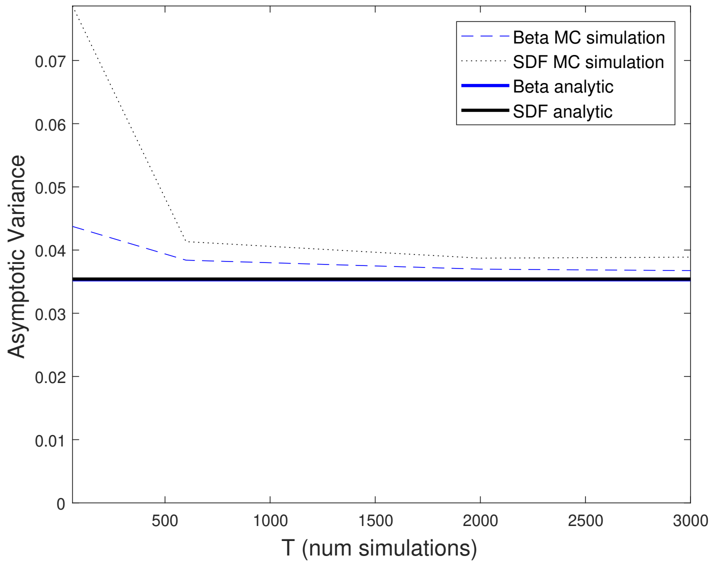
The table presents the relative standard errors of GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda})/\mathrm{E}(\widehat{\lambda})$. The returns are generated by Equation (3) under the null hypothesis with the factors (Market, Size, Value, and Momentum) moments being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\widehat{\lambda}^*)$
Single-factor model loaded with market factor				
60	1.0833	1.0045	1.1219	1.0581
360	1.0654	1.0046	1.0935	1.0324
600	1.0656	1.0030	1.0921	1.0288
1000	1.0658	1.0039	1.0918	1.0295
Single-factor model loaded with size factor				
60	1.8006	1.1771	1.7656	1.2153
360	1.7301	1.0869	1.7295	1.0965
600	1.6187	1.0576	1.6223	1.0660
1000	1.6199	1.0625	1.6249	1.0704
Single-factor model loaded with value factor				
60	3.6666	1.9399	3.4126	2.9384
360	2.9342	1.9542	2.9205	2.0909
600	2.9140	1.9689	2.9173	2.0602
1000	2.8233	1.9448	2.8397	2.0065
Single-factor model loaded with momentum factor				
60	3.0995	1.5323	3.0467	2.6295
360	1.8419	1.5265	1.8913	1.6745
600	1.8082	1.5198	1.8560	1.6095
1000	1.7837	1.5159	1.8224	1.5766

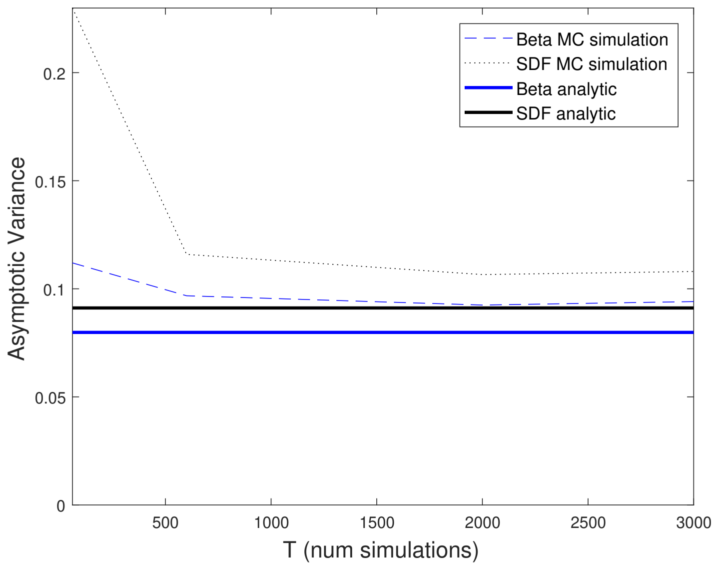
Table VIII. Relative standard errors of pricing errors for four alternative asset pricing models: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\pi}) = \sigma(\widehat{\pi})/\mathrm{E}(\widehat{\pi})$. The returns are generated by Equation (3) under the null hypothesis with the factors (Market, Momentum, and Value) being sampled from the empirical distribution applying Arismendi and Kimura's (2016) tensor moment simulation. Estimators decorated with a * are obtained by GMM from a Beta represention of the single-factor model The U and C represent the un-centred and centred SDF methods; and with 1 and 2 to the first and second-stage respectively. The results are presented for different sample sizes (T), and they are based on 10,000 bootstrapped correlated simulations.

T	$\sigma_r(\widehat{\pi}_1^U)/\sigma_r(\widehat{\pi}^*)$	$\sigma_r(\widehat{\pi}_2^U)/\sigma_r(\widehat{\pi}^*)$	$\sigma_r(\widehat{\pi}_1^C)/\sigma_r(\widehat{\pi}^*)$	$\sigma_r(\widehat{\pi}_2^C)/\sigma_r(\widehat{\pi}^*)$
		CAPM		
60	0.8246	1.1259	0.8250	1.0426
360	0.8338	1.0015	0.8340	0.9947
600	0.8237	1.0004	0.8238	0.9958
1000	0.8284	0.9996	0.8286	0.9987
		Fama-French		
60	0.6249	1.3654	0.6342	1.2793
360	0.6738	0.9997	0.6741	0.9041
600	0.6891	0.9475	0.6898	0.8846
1000	0.7220	0.9473	0.7226	0.9124
		RUH		
60	0.6245	1.2626	0.6274	1.2145
360	0.6003	1.0224	0.6014	0.9947
600	0.5896	0.9544	0.5897	0.9368
1000	0.6158	0.9572	0.6160	0.9352
		Carhart		
60	0.5807	1.2146	0.5823	1.1697
360	0.6667	1.0224	0.6650	0.9442
600	0.7317	0.9992	0.7303	0.9458
1000	0.8545	1.1034	0.8540	1.0734

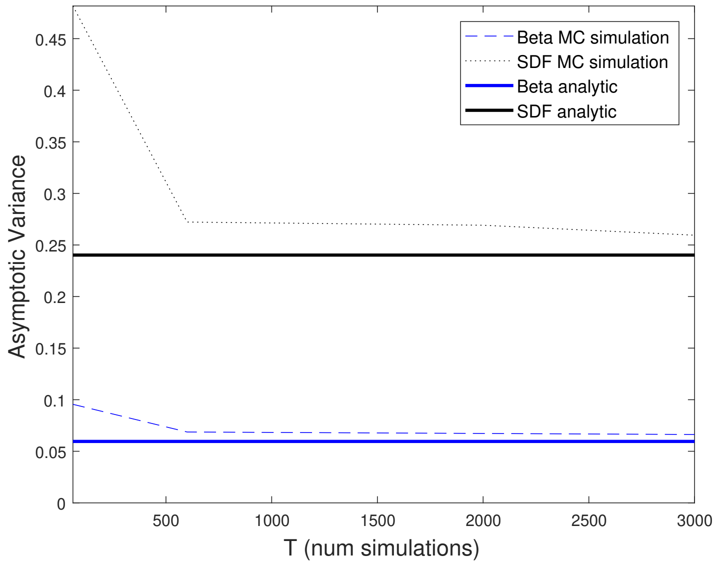


(a) Market risk factor

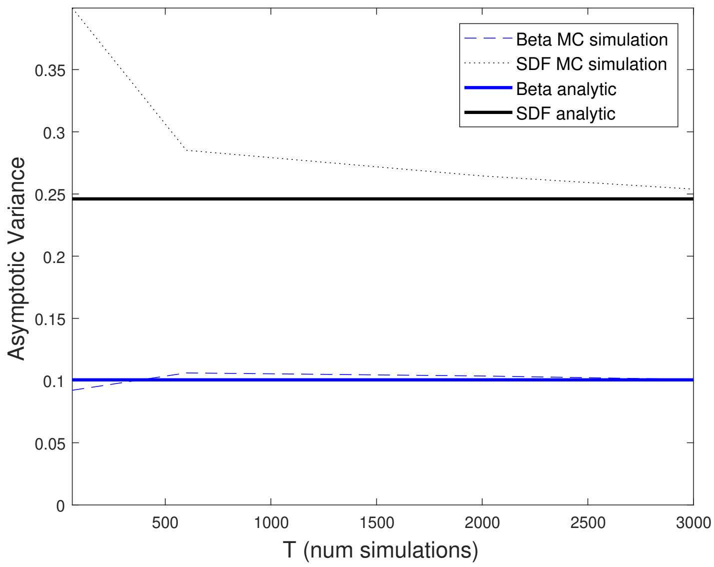


(b) Size factor

Figure 1. Asymptotic variance of the analytically and empirically estimated GMM under the Beta and SDF methods, from a set of 10,000 Monte Carlo simulation based on parameters calibrated to the observed market risk, and size factors on a sample January 1927 – December 2018. Data are downloaded from Kenneth French's library.



(a) Value factor



(b) Momentum factor

Figure 2. Asymptotic variance of the analytically and empirically estimated GMM under the Beta and SDF methods, from a set of 10,000 Monte Carlo simulation based on parameters calibrated to the observed value, and momentum factors on a sample January 1927 – December 2018. Data are downloaded from Kenneth French's library..

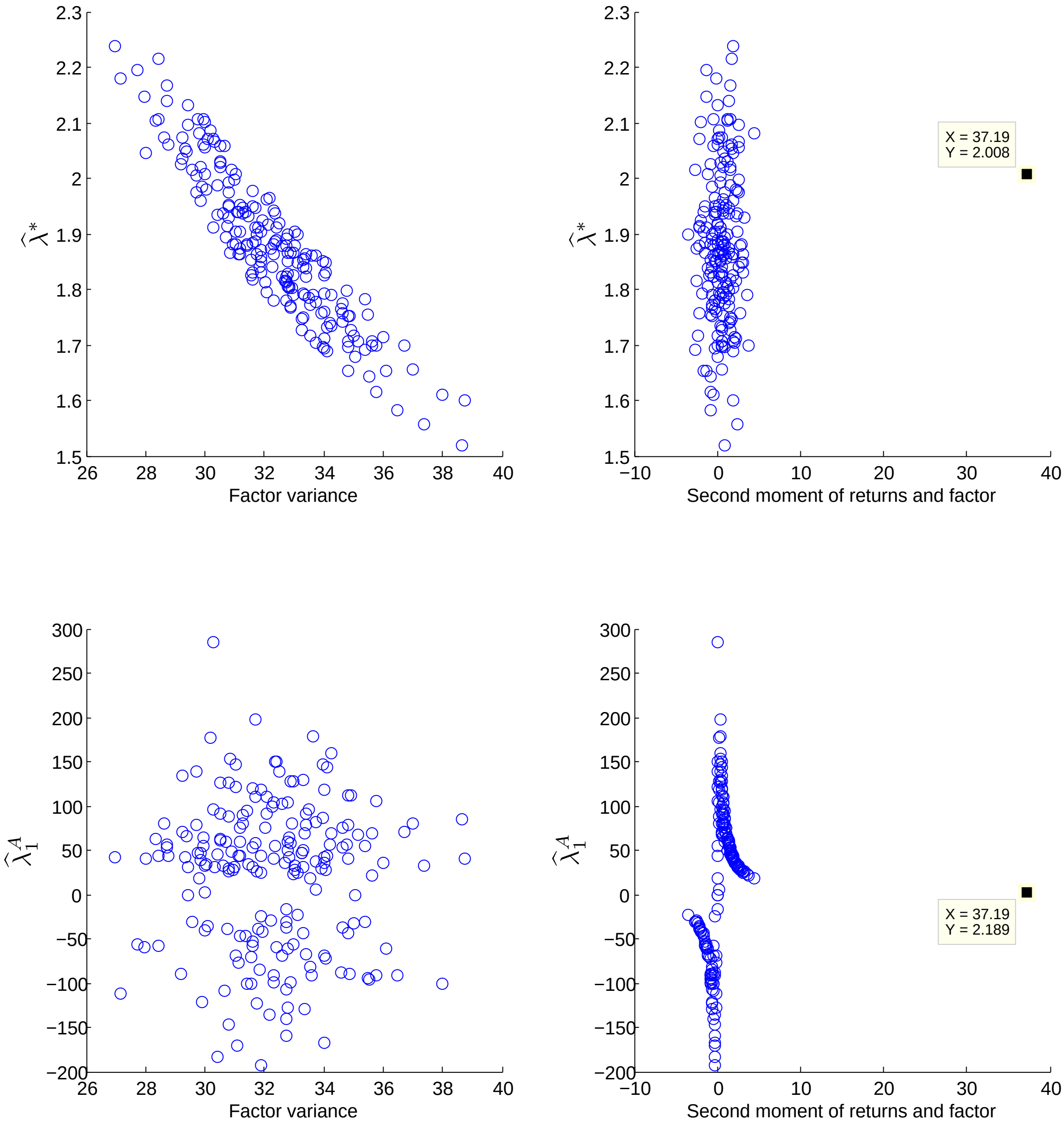


Figure 3. US: Estimators resulting from market factor draws with alternative variances and nearly independent of the portfolio returns. Actual values represented by the squared data cursor.

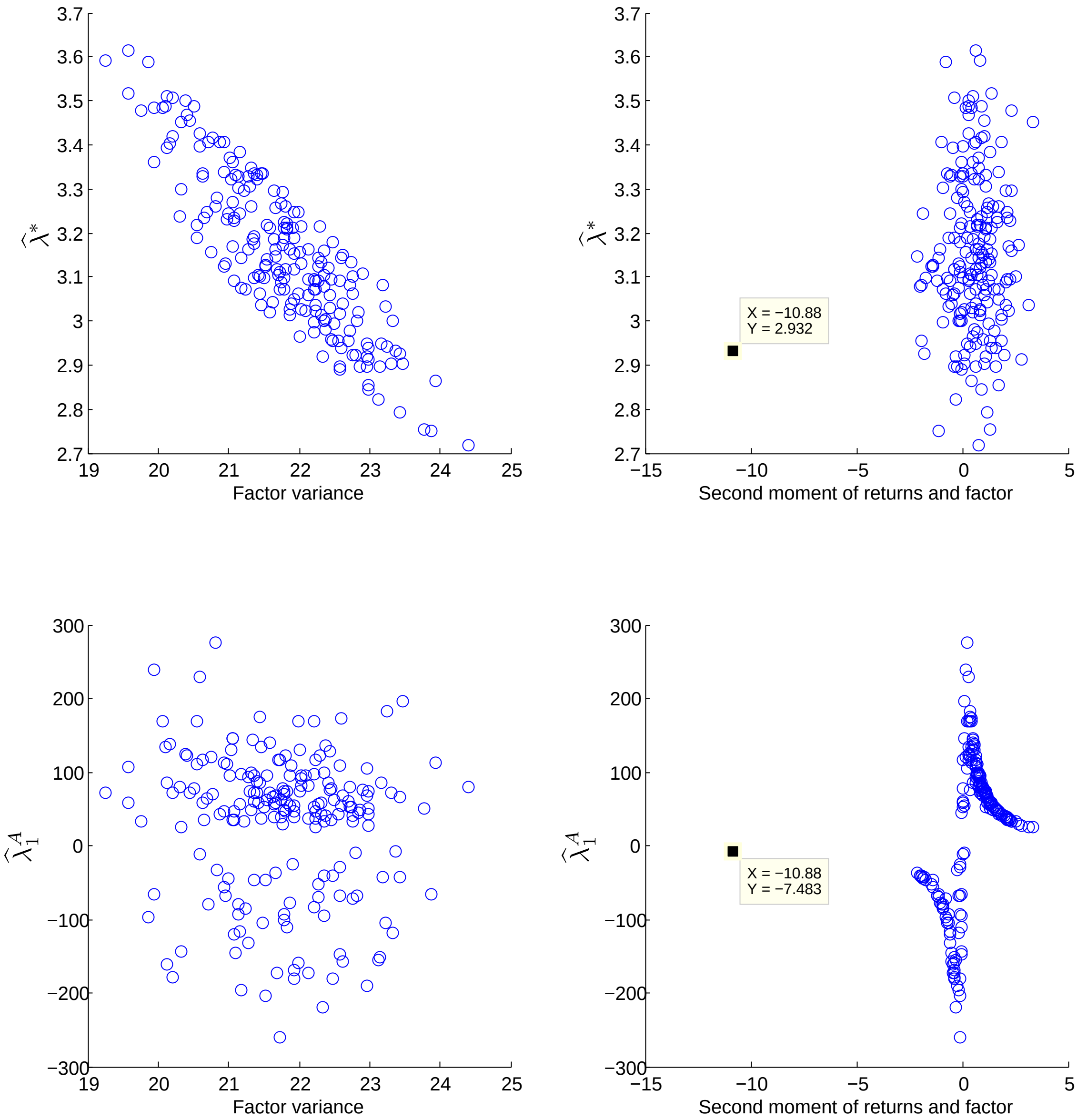


Figure 4. US: Estimators resulting from momentum factor draws with alternative variances and nearly independent of the portfolio returns. Actual values represented by the squared data cursor.