Department of Economics, Finance & Accounting

Working Paper N304-20

Multifactor Empirical Asset Pricing Under Higher-Order Moment Variations

Massimo Guidolin*, Martin Lozano†, and
Juan Arismendi Zambrano‡ ,§

June 2020

ABSTRACT

Even though an asset pricing model can be equivalently expressed in a classical Beta or in the relatively new stochastic discount factor (SDF) representation, their key empirical features - efficiency and robustness - may differ when estimated by the generalized method of moments. Using US and UK data we find that the SDF approach is likely to be less efficient but also more robust than the Beta method. We derive the analytical asymptotic variance and show that the main drivers of this trade-off are the higher-order moments of the factors, by which skewness and covariance between returns and factors play an important role.

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.

---

*Bocconi University and IGIER, Bocconi University. Address: V1a Roetgen 1, 20136 Milan, Italy. email: massimo.guidolin@unibocconi.it.

†University of Monterrey - UDEM, Monterrey, Mexico. email: martin.lozano@udem.edu

‡Department of Economics, Finance, and Accounting, National University of Ireland, Maynooth. ICMA Centre, Henley Business School, University of Reading, Reading, UK. email: juancarlos.arismendizambrano@mu.ie; j.arismendi@icmacentre.ac.uk

§An earlier version was presented (or accepted for presentation) at the 2008 Annual Doctoral Conference at Manchester Business School; Internal Seminar (University Carlos III de Madrid, April 2008); Jornada de Investigación en Finanzas (Universidad de Castilla-La Mancha, October 2008); XVI AEFIN Finance Forum (ESADE Business School, Barcelona, November 2008); International Conference on Finance (National Taiwan University, December 2008); Southwestern Finance Association 48th Annual Meeting; (Oklahoma City, February 2009); Royal Economic Society Conference (London, April 2009); Symposium in Statistics and Econometrics (Lausanne, April 2009); Eastern Finance Association Annual Meeting (Washington D.C., May 2009); Asian Finance Association Conference (Brisbane, July 2009); and at the Symposium in Economics and Finance (Geneva, July 2009); University of Liverpool Management School Seminar Series (Liverpool, November 2019); UCD Seminar Series (Dublin, December 2019); 2nd Frontiers of Factor Investing Conference (Lancaster, April 2020); and 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).

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 widely implemented in the finance literature using the two-stage cross-sectional regression methodology advocated by Black et al. (1972), Fama and MacBeth (1973), and Kan et al. (2013). The relatively new 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 both 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 residual, Jensen's alpha, α. 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, a variance equal to one, 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. Though, this theoretical equivalence does not necessarily entail an empirical equivalence. Therefore, the experimental questions that 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?*

[1] See Ferson and Jagannathan (1996) for a general approach to any equivalence and differences between the two approaches.

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

Our empirical and analytical results show that, in general, the Beta method is more **efficient** but less **robust** than the SDF method. That is, the estimators of δ and π have lower simulated (relative) standard errors than their corresponding estimators for λ and α. The main objective of this paper is then to document that the choice of the Beta or the SDF method can be condensed into a trade-off between efficiency versus robustness. In this sense, we provide empirically motivated evidence about what drives this trade-off. This "easy way" to validate the information in asset prices is valuable for researchers and practitioners because they might have a *priori* rough idea about the benefits and the pitfalls of following either method in order to conduct their empirical analysis.

To disentangle the source of the empirical results, we extend Jagannathan and Wang (2002) analytical results to the case of multifactor asset pricing models. We find that the source of the efficiency of the Beta method over the SDF method, is rooted in the higher-order moments effects on factors, that impact more the SDF than the Beta estimation.[3] We show that the factor negative skewness – which is usually a characteristic of the momentum portfolios, constitute a drawback to the SDF method at estimating λ even if the sample is as large as one thousand. Additionally, the application to the UK data illustrates that this relative disadvantage increases as we consider smaller samples. In any case, this drawback at estimating λ represents an advantage at estimating π given that the method is basically devoted to minimize the pricing errors. In this paper, we take advantage of 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 measure units.

Out study contributes to the asset pricing and financial econometrics literature, where it has become common to compare the performance of different econometric procedures

[3] Graham and Harvey (2001) find in a survey that almost 74% of the respondents (Chief Financial Officers - CFOs) used "always or most of the times" the CAPM to calculate the cost of capital for their projects versus 40% of the CFOs that used the average historical returns of a stock. This second method – average historical returns of a stock – can be considered a particular version of the SDF method. Our results have an economic importance due to the estimation of the cost of capital, as every increase/decrease of 1% in the cost of capital due to the estimation variance, can be translated into an increased/decreased $1MM additional cost per every $100MM project valuation.

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 the Fama-MacBeth maximum likelihood approach and of the GMM for Beta pricing models. Other key 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 initial attempts to evaluate the performance in finite samples of the Beta versus the SDF approaches, and this is where our main interest lies, because this has become a growing research field in the econometrics of asset pricing models.

In a first attempt to evaluate the performance in finite samples of the Beta versus the SDF approaches, using a standardized single-factor model, Kan and Zhou (1999) show that the SDF method is 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 and 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. Furthermore,Kan and Zhou (2001) have showm that, under more general distributional assumptions and considering non-standardized factors, inference based on λ can still be less reliable than inferences based on δ, especially in realistic situations where the factors are leptokurtic. Contributing to this debate, Ferson (2005) have reported that when the two methods correctly exploit the same moments they deliver nearly identical results. The interest about 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. On the other hand, Peñaranda and Sentana (2015) show that a particular GMM procedure leads to numerically identical Beta and SDF estimates.

In this paper, we advocate that the core difference between the SDF and the Beta methods is a matter of trade-off between efficiency versus robustness. Commonly in econometrics, a robust approach places fewer restrictions on an estimator but pays a precise in terms of resulting precision. In this sense, the SDF method does not place any restriction to λ, whereas the Beta method explicitly incorporates the definition of δ as a subset of the moment restrictions. Therefore, it may be reasonable to expect that the Beta method should be more efficient than the SDF method and, conversely, that the SDF method may be more robust than the Beta method. As a result, the choice of estimate and evaluate an asset pricing model by either the SDF or the Beta methods, boils down to a choice between efficiency versus robustness and the decision ought to depend on the goals of the researcher. In particular, if the main interest lies with the inference of the risk premia, the Beta method is likely to provide more accurate estimates vs. the SDF method; yet, if the main concern were to achieve lower pricing errors, the SDF method should be implemented. Hence, there is no method that unconditionally dominates the other–rather they are complementary, and they should not be considered as empirically equivalent.

Regarding the Beta method, there is no disagreement in the literature about which set of moments have to be included in GMM estimation. Yet, there are two dominant and subtly different SDF representations in the literature that leads to exploit alternative sample moments. The first strategy defines the SDF as a linear function of the factors. However, Kan and Robotti (2008) point out that this is problematic because the specification test statistic is not invariant to an affine transformation of the factors. Therefore, in our study we also consider an alternative specification that defines the SDF as a linear function of the de-meaned factors. These two SDF variants are also known as the un-centred and centred forms, respectively. In our paper, these alternative SDF definitions and methods are then applied to estimate and evaluate 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 of Lozano and Rubio (2011) (RUH) models in an application to US data. We also provide some illustrative results using a reduced UK sample.

There are two additional contributions that we offer that derive from differences in our finite sample approach with respect to previous, similar studies that allow us to obtain a number of insights. The first is that we assume factors and returns are drawn from their 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 debate on the differential performance of alternative estimation methods in asset pricing research is far from being exhausted. Furthermore, the available results in the extant literature usually rely on the evaluation of a single-factor model which, while still an important benchmark, fails to span the breadth of factor-based models used in the literature as well in practice. In fact, multi-factor pricing models are by far the most commonly used in empirical applications and they incorporate factors with more extreme non-normalities than the market portfolio factor itself. For instance, in our sample, the momentum factor has similar mean and variance as the market factor; nonetheless, it is almost three times more leptokurtic. Likewise, the market factor is close to displaying a symmetric distribution, although skewness is positive in the case of the size and value factors, but markedly negative in the case of the momentum factor. Table I contains descriptive statistics for the factors and test portfolios used in this paper and supports our remarks above.

[Place Table I about here]

The outline of the rest of the paper is as follows. Section II presents the methodology, Section IV 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 serial independence, conditional homoskedasticity, and/or normality are not fully realistic in practice. [4]

A. Beta method

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 during period 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 asset-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 parameter 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 spec-

[4] For a in-depth discussion of GMM applied to the estimation and specification testing of asset pricing models see Campbell et al. (1997), Jagannathan et al. (2002), and Cochrane (2005).

[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

ification 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 of 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, RUH, 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. 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

array: coskewness is then a 3-dimensional array while cokurtosis is a 4-dimensional array.

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.

Following the usual GMM notation, we define the vector of unknown parameters $\theta = [\text{vec}(\mathbf{B})' \;\; \mu']'$, 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:

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

in which, for any θ, the sample analogue of $\text{E}[g(x_t,\theta)]$ 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 $g(x_t,\theta_1)$:

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

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 $\alpha =$E$[r_t] - \delta\mathbf{B}$ is jointly equal to zero.[6] 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, Cov(g_T), is given by

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

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

$$\begin{aligned} &\text{First-stage:} && g_T(\theta_1)'\text{Cov}(g_T)^{-1}g_T(\theta_1) \sim \chi^2_N, \\ &\text{Second-stage:} && Tg_T(\theta_2)'S^{-1}g_T(\theta_2) \sim \chi^2_N. \end{aligned} \tag{12}$$

Both the first and second-stage statistics in equation (12) lead to the same numerical value. However, if we weight equations (11) and (12) by any other matrix different to $\mathbf{S}$, such as E$[r_t r_t']$ or Cov$[r_t]$, this result no longer holds. 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.

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

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

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 $1 + \delta'\mathbf{\Omega}^{-1}\mu$,[7]

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

[6] This approach is known as the restricted test, see MacKinlay and Richardson (1991).
[7] Even when the factors are not traded, it is common to suppose $1 + \delta'\mathbf{\Omega}^{-1}\mu \neq 0$.

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 set of moment restrictions g used to estimate the linear asset-pricing model,

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

where the random variable $m_t \equiv 1 - \lambda' f_t$ is the SDF defined as usual as $\text{E}[r_t m_t] = 0_N$. Alternatively, we could derive the Beta representation from the SDF representation by expanding m and rearranging the terms.

From the moment restrictions and equation (14), we obtain the vector of N pricing errors defined as $\pi = \text{E}[r_t] - \text{E}[r_t f_t']\lambda$. The numerical solution to equation (14) can once more be obtained by GMM.[8] Let's start by writing the sample pricing errors as

$$g_T(\lambda) = -\text{E}[r_t] + \text{E}[r_t f_t']\lambda, \tag{15}$$

and by defining $\mathbf{D}^U = -\frac{\partial g_T(\lambda)}{\partial \lambda'} = \text{E}[r_t f_t']$, 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 $-\left(\mathbf{D}^U\right)' \mathbf{W}[\text{E}[r_t] - \mathbf{D}^U \lambda'] = 0$, 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:

$$\begin{aligned} \widehat{\lambda}_1^U &= \left(\left(\mathbf{D}^U\right)' \mathbf{D}^U\right)^{-1} \left(\mathbf{D}^U\right)' \text{E}[r_t], \\ \widehat{\lambda}_2^U &= \left(\left(\mathbf{D}^U\right)' \mathbf{S}^{-1}\mathbf{D}^U\right)^{-1} \left(\mathbf{D}^U\right)' \mathbf{S}^{-1}\text{E}[r_t]. \end{aligned} \tag{16}$$

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

[8] This is useful given the need to undertake vast numbers of simulations. Similar simplifications of multi-dimensional optimization problems for Beta models can be found in Shanken and Zhou (2007).

the centred specification.

Specifying the SDF as a linear function of the factors as in equation (14) has been very popular in the empirical literature. However, Kan and Robotti (2008) point out that this is problematic because the specification test statistic is not invariant to an affine transformation of the factors; Burnside (2007) also reaches 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 computing μ =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 solve the GMM minimization is similar to that for the uncentred SDF^U method. In particular, we substitute E$[r_t f_t]$ for Cov$[r_t f_t]$ in equation (15) and define $\mathbf{D}^C = -\frac{\partial g_T(\lambda)}{\partial \lambda'}$ as the covariance matrix of returns and factors. As a result, under SDF^C, the first and second stage GMM estimates are given by:

$$\begin{aligned}\widehat{\lambda}_1^C &= \left((\mathbf{D}^C)'\mathbf{D}^C\right)^{-1}(\mathbf{D}^C)'\,\mathrm{E}[r_t],\\ \widehat{\lambda}_2^C &= \left((\mathbf{D}^C)'\mathbf{S}^{-1}\mathbf{D}^C\right)^{-1}(\mathbf{D}^C)'\,\mathbf{S}^{-1}\mathrm{E}[r_t].\end{aligned} \tag{18}$$

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).

Equations (11) and (12) are weighted by (10), as this is statistically 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 investigate the implications of using these alternative weighting matrices later on.

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.[9] 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.

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

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

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.

In a first formal attempt to compare both methods, Kan and Zhou (1999) assumed that the only factor had zero mean and unit variance, that is $\mu = 0$ and $\boldsymbol{\Omega} = 1$. In this standardized, single factor model, equations (19) and (20) simplify to $\lambda = \delta$ and $\pi = \alpha$. By assuming that the mean and the variance of the factor are predetermined without a need for estimating them, Kan and Zhou (1999) ignored the sampling error associated with the estimation of μ and $\boldsymbol{\Omega}$ and concludes that the estimators obtained through the Beta method were more efficient. Jagannathan and Wang (2002) and Cochrane (2001) further investigated the effects of standardizing the factors, showing that in general, predetermining the factor moments reduces the sampling error in the estimates from the Beta method but not from the SDF method.[10] Predetermining the values of μ and $\boldsymbol{\Omega}$ to be known constants, not necessarily $\mu = 0$ and $\boldsymbol{\Omega} = 1$, gives an informational advantage to the Beta method in terms of efficiency. Predetermining without estimation implies ignoring the sampling errors associated with μ^* and $\boldsymbol{\Omega}^*$: as a result, λ^* becomes considerably more efficient than when we solve the GMM problem implied by (6). In our simulation analysis to follow, we therefore consider the case where μ and $\boldsymbol{\Omega}$ must be estimated.

To summarize, the Beta method gives the GMM estimate δ^* while the SDF method gives the GMM estimate λ. In this paper, we plan to we transform the estimate δ^* into an estimate of λ and then compare the variances of the sampling distribution of $\widehat{\lambda}^*$ and $\widehat{\lambda}$. In the same way, we shall transform α^* into an estimate of π and then compare the efficiency of $\widehat{\pi}^*$ and $\widehat{\pi}$. We also compare the distributions of the Hansen (1982) classical test of overidentification using the J-statistic of the transformed Beta $\widehat{J}^*$ and $\widehat{J}$ from the

[10] However, under the moment restrictions derived from the Beta representation, (6), we only can make inference on δ, not on λ. Yet to compare the methods using equation (19) requires an estimator of $\boldsymbol{\Omega}$. One solution is to add an additional moment condition to (6) to estimate $\boldsymbol{\Omega}$. An alternative is to estimate μ and $\boldsymbol{\Omega}$ outside the GMM. In simulation results not reported in this paper, we find that the efficiency of both alternatives is the same. Hence, in what follows, we elect to estimate $\boldsymbol{\Omega}$ outside the GMM.

SDF method. The null hypothesis is that all pricing errors are zero. But before attacking the issue through a set of simulation experiments, in the next Section we provide necessary analytical background to our key Monte carlo results to follow.

III. Analytical results

In this Section we derive a general extension to the results in Jagannathan and Wang (2002), in the sense that the factor f_t is multivariate and is allowed to have a (joint) non-Gaussian distribution. Our results will be useful to understand the simulation findings on the trade-off between efficiency and robustness when using the Beta vs. the SDF methods.

PROPOSITION 1: *Let f_t represent the multivariate, systematic risk factors, and consider the Beta representation in (1), and the SDF representation in (14). Then, the positive definite asymptotic covariance matrix of the $\hat{\lambda}$ parameters in the SDF model 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 a scalar whose expression is provided in Appendix A). The asymptotic covariance matrix of the Beta representation λ^* parameter vector is

$$ACov(\lambda^*) \quad = \quad \left(\mathbf{\Omega} + \mu\mu'\right)^{-1} S_b \left(\left(\mathbf{\Omega} + \mu\mu'\right)^{-1}\right)', \quad (22)$$

where S_b is the covariance matrix of $g_b(f_t, \lambda_b) = (f_t(1 - \lambda_b f_t))$, and $\lambda_b = \mu'\left(\mathbf{\Omega} + \mu\mu'\right)^{-1}$. 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,

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

and,

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

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:

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

The corresponding SDF and Beta pricing errors asymptotic variance estimators are

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

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. See appendix A. □

The impact of higher-order moments is the result of second order terms of f_t in equation (14): the GMM estimation requires the estimation of the variance in (14) and therefore the third- and fourth-order terms of f_t appear in the risk premia estimator. In the case of the Beta method, higher-order moments do not impact the estimation of the asymptotic variance, $Avar(\lambda^*)$, because although the moments restrictions in (6) do imply second-order terms for f_t –and as a consequence there will be fourth-order terms involved in the estimation– these do not affect the estimation of δ^* nor the equivalent risk premium λ^*, but rather simply affect the estimation of the other parameters (for example, $\mathbf{\Omega}$). We now extract the implications of A for applied work in the form of a

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 have 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 interesting cases.

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.[11] 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. Even though the qualitative flavor of

[11]See Ahn and Gadarowski (2004) for an examination of finite-sample properties of several methods to test model specification.

our results concerning relative efficiency is hopefully clear, it remains imperative to grasp the relative small sample performance of the two methods and therefore re-introduce robustness consideration in our assessment of Beta vs. SDF-based methods. 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. [12] 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 analytical asymptotic variance result reported in Section , while the second set of simulations serves to measure the asymptotic variance of the different multivariate factor models analyzed.

As fas 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-

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

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 inFama 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 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 Fama-French, RUH, and the Carhart four-factor models, which means that $K = 1, 3$, and 4. The factors as the excess market return (RMRF), size (SMB), value (HML) and momentum (UMD).[13] 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. 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

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

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.[14]

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, Tables II and III provide details about the parameters used in the simulations. 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). Moreover, while the market factor implies an unconditional distribution which is essentially symmetric (sample skewness is 0.2), the size and value portfolio returns are strongly right-skewed (1.9 and 2.2, respectively), while momentum returns exhibit massive left-skewness (-3.1). Thus, it is important to consider inferential and testing methods for multi-factor asset pricing models that are able to reflect relative efficiency and robustness in the light of the empirical properties of the data, such as extreme asymmetries and excess kurtosis, when factors over and beyond market risk are considered. Studies such as Kan and Zhou (2017) have considered asset pricing models under a Student-t distribution, even though the magnitude of kurtosis is still limited for a t-distribution when asymptotics requires a finite fourth moment.[15] For instance, in unreported simulations, we have experimented with a Student-t distribution with five degrees of freedom that implies a kurtosis of 6 for the RMRF factor, which is still much lower than the empirical value of 11 and found 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 II about here]

[Place Table III about here]

[14]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.

[15]The asymptotic distribution theory for the GMM requires that returns and factors have finite fourth moments. Hence, any marginal t-Student distribution for errors and factors must be characterized by more than four degrees of freedom, which limits considerably tail thickness

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.

A. Monte Carlo Simulation: Convergence

Figures 1a, 1b, 1c, and 1d 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 II and III, 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 ($Avar(\lambda_2^U)$).[16]

[Place Figure 1 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:

[16]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 ($Avar(\lambda_2^U)$).

the higher third-order cumulant (κ_3) produces an increase in the asymptotic variance, consistent with analytical results in the Appendix A.

B. Bootstrap simulations: Comparison of λ *estimators*

Tables IV and V compare the performance of Beta and SDF methods at estimating λ by the single-factor CAPM model using US data. As one would expect, our results are qualitatively and quantitatively similar to those in Jagannathan and Wang (2002). In particular, Table IV 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 IV about here]

Table IV 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 scaling issues. However, if we consider the possibility of intrinsic differences among the methods, the expected values of $\widehat{\lambda}^*$ and $\widehat{\lambda}$ are not necessarily similar, at least 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 obtain an accurate measure of the relative efficiency of the methods. These results are presented in Table V.

[Place Table V about here]

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. Given that one is generally interested in testing the null hypothesis that the estimator is equal to zero, ratios such as those reported in Table V offer a good indication of which method may lead to the most accurate inferences. In particular, all ratios in Table V are slightly in excess of one, which means that the Beta method leads to lower standard errors when estimating the CAPM market risk premium. In general, the values of the ratios decline as we increase the size of the sample while, as expected, second-stage SDF estimators of λ are more efficient than first-stage estimators. In addition, the un-centred SDF specification reveals a marginal advantage when compared to the centred specification, probably as a result of the additional moment restrictions that derive from centering.

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 VI, VII and VIII 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 VI about here]

[Place Table VII about here]

[Place Table VIII about here]

The corresponding expected values and standard errors for the estimators of λ derived from the multi-factor asset pricing models are in Tables IX, X and XI. We begin by describing the results obtained in the case of the Fama-French model, which is characterized by $K = 3$, i.e., the market, size, and the value factors.

[Place Table IX about here]

[Place Table X about here]

[Place Table XI about here]

The first (upper) panel of Table VI is comparable to Table V because in both cases the estimated λ corresponds to the market factor. Thus, it is not surprising to find a similar pattern which reinforces the conclusion that there are no significant differences in the inferential precision obtained by alternative asset pricing model representations when estimating the parameters associated with the market risk premium. The un-centred SDF method is again more efficient than the centred SDF method at estimating market risk λ. In this case, the second-stage un-centred method is marginally more efficient than the Beta method for samples smaller than $T = 1000$.

Contrary to the case of the market risk factor, the standard error of the Beta estimators of the risk premia associated to the size and value factors are significantly smaller than the corresponding standard errors of the SDF-based estimators. This is evident from the ratios in excess of one appearing in the second and third panels of Table VI, especially in the case of the ratios derived for the value risk premium. These results suggest that the empirical equivalence of the two ways to represent linear factor models may strongly depend on which factors are included in the specification of the asset pricing model. In particular, the market factor does not pose a challenge to the SDF method whilst the value factor can show significant variation across representations, according to the $\sigma_r(\lambda)$ ratios. For instance, the relative standard error the λ estimates from the un-centred, first-stage SDF method can be more than twice as big as the relative standard errors from the Beta representation. 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 we examine in this paper is the RUH, in which the priced factors are market, momentum, and value. The estimation of the RUH model allows us to compare the efficiency of the estimator associated with the market and value factors in previous tables and introduces novel results for momentum.

The ratios of the relative, normalized standard errors $\sigma_r(\lambda)$ associated to the market risk premium are now slightly below one; however the standard errors $\sigma(\lambda)$ are not significantly different in most cases. However, the magnitude of the efficiency ratios in the case of the value risk premium are similar to the ones of the Fama-French model in Table VI. The second panel of Table VII shows the ratios for the momentum premium, which turn out to be somewhat larger than the ratios obtained for the value factor. Here, we find that the SDF representation may face severe difficulties in small samples, a fact that is reflected in massive (in-)efficiency scores ranging between 20.7 and 10.5.

The third and last multi-factor asset pricing model we have analyzed is Carhart's model, that features as factors the market, size, value, and momentum. Although the elaboration of simulation results for the well-known multi-factor models by Fama-French, RUH, and Carhart represents our core contribution, in the case of Carhart's model we also present results for test portfolios and the typical sample size faced by researchers using shorter, UK data. We start by displaying the normalized, relative standard errors of Carhart's model λ estimators in Table VIII, with reference to US data.

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 VIII, 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.

One might counter that the results in Tables VI, VII and VIII may be partially driven by the additional factors that had failed to be investigated in earlier literature, and not by any structural differences of GMM estimators of risk premia across alternative representations of the linear factor models. To address this possibility, we estimate four alternative single-factor asset pricing models. In each case, the model just includes one of the Fama-French-Carhart risk factors at the time, i.e., the market, size, value, and momentum. The ratios of the relative, normalized standard errors of the Beta- vs. the SDF-based inferences for these single-factor models are reported in Table XII. The corresponding expected values and standard errors are shown in Table XIII.

[Place Table XII about here]

[Place Table XIII about here]

The evidence in Table XII 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 XII 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.

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 XIV 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 XV presents the values used in the construction of Table XIV.

[Place Table XIV about here]

[Place Table XV about here]

D. 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 favour 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.

The debate on the relative, finite sample performance of the GMM as applied to alternative ways to write linear factor pricing models appears to be open. Given the findings in Jagannathan and Wang (2002) to Kan and Zhou (1999), common wisdom is that alternative ways to represent standard linear pricing models does not matter much in terms of inferential efficiency, see for example, Cochrane (2001), Smith and Wickens (2002), Cochrane (2005), Nieto and Rodríguez (2005), Vassalou et al. (2006), Wang and Zhang (2006), Balvers and Huang (2007), Jagannathan et al. (2008), Cai and Hong (2009), and Brandt and Chapman (2018). On this background, our driving motivation is that in the existing literature the comparisons between Beta and SDF representations are conducted under sensible but rather specific conditions that appear to be insufficient to differentiate the relative performance of GMM inferences, in particular concerning the number of factors (typically one, as in the CAPM) and their distributional departures from normality (typically, modest). Once we relax these conditions, we show that important differences between the strength and usefulness of the inferences on the two otherwise equivalent methods of representation emerge. We find evidence suggesting that the Beta method leads to more precise risk premia GMM estimators while the SDF method leads to better pricing error estimators, in terms of efficiency. We evaluate the magnitude of the resulting efficiency losses and illustrate what are the drivers of such differential performance.

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

Ahn, S. C. and C. Gadarowski (2004). Small sample properties of the model specification test based on the hansen-jagannathan distance. *Journal of Empirical Finance 11(1)*, 109–132.

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.

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

Balvers, R. J. and D. Huang (2007). Productivity-based asset pricing: Theory and evidence. *Journal of Financial Economics 86(2)*, 405–445.

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.

Brandt, M. W. and D. A. Chapman (2018, jan). Linear approximations and tests of conditional pricing models. *Review of Finance*.

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

Cai, Z. and Y. Hong (2009). Some recent developments in nonparametric finance. *Advances in Econometrics 25*, 379–432.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton NJ: Princeton University Press.

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)*.

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.

Graham, J. R. and C. R. Harvey (2001, may). The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics 60*(2-3), 187–243.

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*.

Kan, R. and G. Zhou (2017, feb). Modeling non-normality using multivariate t: implications for asset pricing. *China Finance Review International 7*(1), 2–32.

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.

Nieto, B. and R. Rodríguez (2005). Modelos de valoración de activos condicionales: un panorama comparativo. *Investigaciones Económicas XXIX (1)*, 33–72.

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.

Smith, P. and M. Wickens (2002). Asset pricing with observable stochastic discount factors. *Journal of Economic Surveys 16(3)*, 397–446.

Vassalou, M., Q. Li, and Y. Xing (2006). Sector investment growth rates and the cross-section of equity returns. *Journal of Business 79(3)*, 1637–1636.

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

Wang, Z. and X. Zhang (2006). Empirical evaluation of asset pricing models: Arbitrage and pricing errors over contingent claims. Technical report, Staff Report, Federal Reserve Bank of New York.

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

Appendix A. Proof Proposition 1 – Higher-order Moments in Empirical Asset Pricing Estimation

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:

$$\begin{aligned} S_s &= E[g_s(r_t, f_t, \lambda)g_s(r_t, f_t, \lambda)'] \\ &= \mathbf{B}\Big(\otimes_R \left(\otimes_R \left(\kappa_4, \lambda\right), \lambda\right) + \\ &\quad 2\delta \left(\otimes_R \left(\kappa_3, \lambda\right) \lambda\right)' + 2 \otimes_R \left(\kappa_3, \lambda\lambda'\mu - \lambda\right) + \\ &\quad \mathbf{\Omega} \left(\mathbf{I}_N + 4(\lambda\lambda'\mu - \lambda)\delta' + (\lambda'\mu\mu'\lambda)\mathbf{I}_N - 2\lambda'\mu\mathbf{I}_N\right) + \delta\delta'\lambda'\mathbf{\Omega}\lambda + \\ &\quad (\delta\delta' + \lambda'\mu\mu'\lambda\delta\delta' - 2\lambda'\mu\delta\delta')\Big)\mathbf{B}' + \\ &\quad (1 - 2\lambda'\mu + \lambda'\mu\mu'\lambda + \lambda'\mathbf{\Omega}\lambda)\, \mathbf{\Sigma}_{\epsilon_t}. \end{aligned} \tag{A1}$$

The elements in (A1) 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 (A1) 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 (A1). Define:

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

and

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

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

$$S_s=\mathbf{B}\mathbf{A}_B\mathbf{B}'+a_{\epsilon_t}\boldsymbol{\Sigma}_{\epsilon_t} \tag{A2}$$

The inverse of (A2) is:

$$S_s^{-1}=\frac{1}{a_{\epsilon_t}}\boldsymbol{\Sigma}_{\epsilon_t}^{-1}-\frac{1}{a_{\epsilon_t}^2}\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\mathbf{B}\left(\mathbf{A}_B^{-1}+\frac{1}{a_{\epsilon_t}}\mathbf{B}'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}. \tag{A3}$$

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

$$D_s=E\left[\frac{\partial g_s}{\partial\lambda}\right]=-\mathbf{B}\left(\boldsymbol{\Omega}+\delta\mu'\right), \tag{A4}$$

Then, using (A3) and (A33), we have the asymptotic variance of the SDF model is:

$$Avar(\hat{\lambda})=\left(D_s'S_s^{-1}D_s\right)^{-1}. \tag{A5}$$

In the case of single factor models, and defining $\sigma^2=\Omega$, the variance of the single-factor, equations (A2), (A3), (A33), and (A5), have their equivalents in:

$$\begin{aligned}S_s &= \frac{\sigma^2(\sigma^4+\delta^4)+\kappa_4\delta^2+2\kappa_3(\delta^3-\delta\sigma^2)-3\delta^2\sigma^4}{(\sigma^2+\mu\delta)^2}\beta\beta'+\frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^2}\boldsymbol{\Sigma}_{\epsilon_t}, &\text{(A6)}\\ S_s^{-1} &= \frac{(\sigma^2+\mu\delta)^2}{\sigma^2(\sigma^2+\delta^2)}\boldsymbol{\Sigma}_{\epsilon_t}^{-1}-\frac{(\sigma^2+\mu\delta)^2}{\sigma^2(\sigma^2+\delta^2)}\times\\ &\quad\left(\beta'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\beta+\frac{\sigma^2(\sigma^4+\delta^4)+\kappa_4\delta^2+2\kappa_3(\delta^3-\delta\sigma^2)-3\delta^2\sigma^2}{\sigma^2(\sigma^2+\delta^2)}\right)^{-1}\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\beta\beta'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}, &\text{(A7)}\\ D_s &= E\left[\frac{\partial g_s}{\partial\lambda}\right]=-(\sigma^2+\mu\delta)\beta, &\text{(A8)}\\ Avar(\hat{\lambda}) &= \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\boldsymbol{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1}+\frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^4}+\frac{\kappa_4\delta^2+2\kappa_3(\delta^3-\delta\sigma^2)-3\delta^2\sigma^2}{(\sigma^2+\mu\delta)^4} &\text{(A9)}\end{aligned}$$

The equivalent asymptotic variance (A9) in the single factor Gaussian case is (Jagannathan et al., 2002):

$$Avar(\hat{\lambda}) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1} + \frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^4}. \tag{A10}$$

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:

$$\frac{\kappa_4\delta^2 + 2\kappa_3\delta(\delta^2-\sigma^2) - 3\delta^2\sigma^2}{\sigma^2(\sigma^2+\delta^2)}. \tag{A11}$$

The additional term (A11) 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, $\theta = (\delta, \mathbf{B}, \mu, \sigma^2)$, are estimated jointly as in Jagannathan and Wang (2002). Define

$$g_b(r_t, f_t, \theta) = \begin{pmatrix} g_b(1) \\ g_b(2) \\ g_b(3) \\ g_b(4) \end{pmatrix} = \begin{pmatrix} r_t - \mathbf{B}(\delta + f_t - \mu) \\ (r_t - \mathbf{B}(\delta + f_t - \mu))f_t' \\ f_t - \mu \\ (f_t-\mu)(f_t-\mu)' - \mathbf{\Omega} \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t' \\ f_t - \mu \\ (f_t-\mu)(f_t-\mu)' - \mathbf{\Omega} \end{pmatrix}, \tag{A12}$$

where $\theta = (\delta, \mathbf{B}, \mu, \mathbf{\Omega})$. The covariance of g_b (A12) is

$$S_b = \begin{pmatrix} \mathbf{\Sigma}_{\epsilon_t} & \otimes_E(\mathbf{\Sigma}_{\epsilon_t}, \mu) & 0 & 0 \\ \otimes_E(\mathbf{\Sigma}_{\epsilon_t}, \mu) & \otimes_E(\mathbf{\Sigma}_{\epsilon_t}, \mathbf{\Omega}) & 0 & 0 \\ 0 & 0 & \mathbf{\Omega} & \kappa_3 \\ 0 & 0 & \kappa_3 & \kappa_4 - \otimes_E(\mathbf{\Omega}, \mathbf{\Omega}) \end{pmatrix}. \tag{A13}$$

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

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

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

In the case of $\frac{\partial g_b(1)}{\partial \mathbf{B}}$, 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:

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

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

$$\delta \mathbf{I}_{N\times N\times K} \equiv \otimes_E (\delta \mathbf{e}, \mathbf{1}_{N\times 1}), \tag{A14}$$

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:

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

$$-\left\{\begin{matrix} \begin{pmatrix} \delta_1\mu_1+\boldsymbol{\Omega}_{1,1} & \dots & \delta_1\mu_k+\boldsymbol{\Omega}_{1,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix}_{N\times K} & \dots & \begin{pmatrix} \delta_k\mu_1+\boldsymbol{\Omega}_{k,1} & \dots & \delta_k\mu_k+\boldsymbol{\Omega}_{k,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & \dots & 0 \\ \delta_1\mu_1+\boldsymbol{\Omega}_{1,1} & \dots & \delta_1\mu_k+\boldsymbol{\Omega}_{1,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix}_{N\times K} & \dots & \begin{pmatrix} 0 & \dots & 0 \\ \delta_k\mu_1+\boldsymbol{\Omega}_{k,1} & \dots & \delta_k\mu_k+\boldsymbol{\Omega}_{k,k} \\ 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \end{pmatrix} \\ & \vdots & \\ \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \\ \delta_1\mu_1+\boldsymbol{\Omega}_{1,1} & \dots & \delta_1\mu_k+\boldsymbol{\Omega}_{1,k} \end{pmatrix}_{N\times K} & \dots & \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \\ 0 & \dots & 0 \\ \delta_k\mu_1+\boldsymbol{\Omega}_{k,1} & \dots & \delta_k\mu_k+\boldsymbol{\Omega}_{k,k} \end{pmatrix} \end{matrix}\right\}_{N\times K\times N\times K} \quad \text{(A15)}$$

Using a similar notation as in (A14), we denote the fourth-order identity tensor and the corresponding partial derivative $\frac{\partial g_b(2)}{\partial \mathbf{B}}$ as

$$\mathbf{I}_{N\times N\times K\times K} = \otimes_E\left(\mathbf{e}, \mathbf{1}_{N\times K}\right), \quad \text{(A16)}$$

$$\frac{\partial g_b(2)}{\partial \mathbf{B}} = \left(\delta\mu' + \boldsymbol{\Omega}\right)\mathbf{I}_{N\times N\times K\times K} \equiv \otimes_E\left(\left(\delta\mu' + \boldsymbol{\Omega}\right)\mathbf{e}, \mathbf{1}_{N\times N}\right). \quad \text{(A17)}$$

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

$$D_b = E\left[\frac{\partial g_b}{\partial \theta'}\right] = \begin{pmatrix} -\mathbf{B} & -\delta\mathbf{I}_{N\times N\times K} & \mathbf{B} & 0 \\ -\otimes_E\left(\mathbf{B},\mu\right) & -\left(\delta\mu'+\boldsymbol{\Omega}\right)\mathbf{I}_{N\times N\times K\times K} & \otimes_E\left(\mathbf{B},\mu\right) & 0 \\ 0 & 0 & -\mathbf{I}_{K\times 1} & 0 \\ 0 & 0 & 0 & -\mathbf{I}_{K\times K} \end{pmatrix}. \quad \text{(A18)}$$

Define

$$S_b^{-1} = \text{inv}(S_b), \quad \text{(A19)}$$

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

$$\begin{aligned} Avar(\lambda^*) &= \left(\frac{\partial \lambda}{\partial \delta}\right)\left(\frac{\partial \lambda}{\partial \delta}\right)' Avar(\delta^*) \\ &= \left(-\delta\left(\left((\mathbf{\Omega} + \delta\mu')(\mathbf{\Omega} + \delta\mu')'\right)^{-1}\mu\right)' + \mathbf{1}_{K\times K}(\mathbf{\Omega} + \delta\mu')^{-1}\right) V_{1,1}. \quad \text{(A20)} \end{aligned}$$

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 (A12), (A13), (A18), (A19), in

$$g_b(r_t, f_t, \theta) = \begin{pmatrix} r_t - (\delta + f_t - \mu)\beta \\ (r_t - (\delta + f_t - \mu)\beta) f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix} = \begin{pmatrix} \epsilon_t \\ \epsilon_t f_t \\ f_t - \mu \\ (f_t - \mu)^2 - \sigma^2 \end{pmatrix}, \quad \text{(A21)}$$

$$S_b = \begin{pmatrix} \Omega & \mu\Omega & 0 & 0 \\ \mu\Omega & (\mu^2 + \sigma^2)\Omega & 0 & 0 \\ 0 & 0 & \sigma^2 & \kappa_3 \\ 0 & 0 & \kappa_3 & \kappa_4 - \sigma^4 \end{pmatrix}, \quad \text{(A22)}$$

$$S_b^{-1} = \frac{1}{\sigma^2} \begin{pmatrix} (\mu^2 + \sigma^2)\Omega^{-1} & -\mu\Omega^{-1} & 0 & 0 \\ -\mu\Omega^{-1} & \Omega^{-1} & 0 & 0 \\ 0 & 0 & -\frac{\sigma^2(\kappa_4 - \sigma^4)}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} & \frac{\kappa_3\sigma^2}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} \\ 0 & 0 & \frac{\kappa_3\sigma^2}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} & \frac{\sigma^4}{\sigma^6 - \kappa_4\sigma^2 + \kappa_3^2} \end{pmatrix}, \quad \text{(A23)}$$

$$D_b = E\left[\frac{\partial g_b}{\partial \theta'}\right] = \begin{pmatrix} -\beta & -\delta I_n & \beta & 0 \\ -\mu\beta & -(\sigma^2 + \mu\delta) I_n & \mu\beta & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \text{(A24)}$$

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

$$\left(D_b' S_b^{-1} D_b\right)^{-1} = \begin{pmatrix} \frac{\sigma^2+\delta^2}{\sigma^2}\left(\beta'\Omega^{-1}\beta\right)^{-1}+\sigma^2 & -\frac{\delta}{\sigma^2}\left(\beta'\Omega^{-1}\beta\right)^{-1}\beta' & \sigma^2 & \kappa_3 \\ -\frac{\delta}{\sigma^2}\left(\beta'\Omega^{-1}\beta\right)^{-1}\beta & \frac{1}{\sigma^2+\delta^2}\Omega+\frac{\delta^2\left(\beta'\Omega^{-1}\beta\right)^{-1}\beta\beta'}{\sigma^2(\sigma^2+\delta^2)} & 0 & 0 \\ \sigma^2 & 0 & \sigma^2 & \kappa_3 \\ \kappa_3 & 0 & \kappa_3 & \kappa_4-\sigma^4 \end{pmatrix} \tag{A25}$$

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

$$Avar(\delta^*) = \frac{\sigma^2+\delta^2}{\sigma^2}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1}+\sigma^2. \tag{A26}$$

Applying the Delta method,[17] the corresponding asymptotic variance of the λ^* parameter –equivalent to (A20)– is[18,19]

$$Avar(\lambda^*) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Sigma}_{\epsilon_t}^{-1}\beta\right)^{-1}+\frac{\sigma^6}{(\sigma^2+\mu\delta)^4}. \tag{A27}$$

Considering (A5), (A20), and by applying some algebra with the support of equations (A10) and (A27) 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 $\theta^* = (\delta,\sigma^2) = (\mu,\sigma^2)$. Then, we have that to estimate the asymptotic variance of the parameter $\delta^* = \mu^*$, we define

$$g_b(r_t, f_t, \theta) = \begin{pmatrix} f_t-\mu \\ (f_t-\mu)(f_t-\mu)'-\mathbf{\Omega} \end{pmatrix}, \tag{A28}$$

$$S_b = \begin{pmatrix} \mathbf{\Omega} & \kappa_3 \\ \kappa_3 & \kappa_4-\otimes_E(\mathbf{\Omega},\mathbf{\Omega}) \end{pmatrix}, \tag{A29}$$

[17] 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.

[18] The equation (A27) 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.

[19] 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.

and

$$D_b = E\left[\frac{\partial g_b}{\partial \theta^{*\prime}}\right] = \begin{pmatrix} -\mathbf{I}_{K\times 1} & 0 \\ 0 & -\mathbf{I}_{K\times K} \end{pmatrix}, \tag{A30}$$

then

$$Avar(\lambda^*) \;=\; \left(-\delta\left(\left((\mathbf{\Omega}+\delta\mu')\,(\mathbf{\Omega}+\delta\mu')'\right)^{-1}\mu\right)' + \mathbf{1}_{K\times K}\,(\mathbf{\Omega}+\delta\mu')^{-1}\right)\mathbf{\Omega}. \tag{A31}$$

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

$$g_b(f_t, \lambda_b) \;=\; (f_t(1-\lambda_b f_t)) = 0. \tag{A32}$$

The parameter λ_b is not estimated by the SDF method, but by the Beta method and using (19), $\lambda_b = \mu'\,(\mathbf{\Omega}+\mu\mu')^{-1}$. The covariance matrix of $g_b(f_t,\theta)$ in (A32) 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

$$D_b = \mathrm{E}\left[\frac{\partial g_b}{\partial \lambda}\right] = -\left(\mathbf{\Omega}+\mu\mu'\right), \tag{A33}$$

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

$$Avar(\lambda^*) \;=\; Avar(\lambda_b^*) = (\mathbf{\Omega}+\mu\mu')^{-1}\,S_b\left((\mathbf{\Omega}+\mu\mu')^{-1}\right)'. \tag{A34}$$

In the case of single factors we have

$$Avar(\lambda^*) \;=\; \frac{((\sigma^2+\mu^2) - 2\lambda_b\mathrm{E}[f_t^3] + \lambda^2\mathrm{E}[f_t^4])}{(\sigma^2+\mu^2)}, \tag{A35}$$

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:

$$e_s(\hat{\lambda}) = \frac{1}{T}\left(\sum_{i=1}^{T} g_s(r_t, f_t, \lambda)\right), \tag{A36}$$

$$e_b(\theta) = \frac{1}{T}\left(\sum_{i=1}^{T} g_b(r_t, f_t, \theta)\right). \tag{A37}$$

The SDF pricing error $\hat{\pi}$ will be equal to (A36) (Jagannathan and Wang, 2002), then by Hansen (1982):

$$Avar\left(\hat{\pi}\right) = Avar\left(e_b(\hat{\lambda})\right) = \frac{1}{T}\left(\sum_{i=1}^{T} g_s(r_t, f_t, \lambda)\right) = S_s - D_s\left(D_s' S_s^{-1} D_s\right) D_s'. \tag{A38}$$

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

$$\alpha^* = Q^* e(\theta^*) = \left[I_n, \mathbf{0}_{n\times n}, -\beta^*, \mathbf{0}_{n\times 1}\right] e(\theta^*). \tag{A39}$$

Using equations (A37), (A38) and (20) the asymptotic variance of the pricing error is yield:

$$Avar\left(\pi^*\right) = \left(\left(\mathbf{\Sigma}_{\epsilon_t} + \delta\mu'\right) \mathbf{\Sigma}_{\epsilon_t}^{-1} (\mathbf{\Sigma}_{\epsilon_t}^{-1})' \left(\mathbf{\Sigma}_{\epsilon_t} + \delta\mu'\right)\right) Q \left(S_b - D_b(D_b' S^{-1} D_b) D_b'\right) Q'. \tag{A40}$$

□

Appendix B. Tables

Table I. Descriptive statistics of factors and portfolios.

US factors and portfolios taken from Kenneth R. French library, January 1927 to December 2018 ($T = 1104$). UK factors and portfolios taken from Alan Gregory library, October 1980 to 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
Variance	28.58	10.22	12.17	22.02	46.8924	47.9440	31.3499
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
Variance	67.35	11.31	25.52	55.30	103.6166	108.4459	71.4597
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	-	-
Variance	19.35	9.59	9.90	18.03	21.3262	-	-
Skewness	-1.01	-0.13	-0.51	-0.93	-0.7562	-	-
Kurtosis	6.86	5.82	9.41	8.50	6.6923	-	-

Table II. Parameter Values used in Monte Carlo Simulation: US data, 10 size-sorted portfolios.

The table presents the parameters used for the Monte Carlo simulation to estimate GMM under Beta and SDF methods. Parameters are estimated from single US factors and portfolios taken from Kenneth R. French library, January 1927 to December 2018 ($T = 1104$).

Panel A: Market Factor									
μ =0.6471, σ =5.3457, δ =0.6471, λ =0.0223, κ_3 =28.4438, κ_4 =8830.8782									
Decile Portfolios									
Small	2	3	4	5	6	7	8	9	Large
β									
1.4187	1.3854	1.3285	1.2572	1.2287	1.2037	1.1514	1.1143	1.0628	0.9311
α									
0.1951	0.0750	0.1006	0.1094	0.0772	0.1037	0.0708	0.0675	0.0288	-0.0051
$\mathbf{\Sigma}_\epsilon$									
40.9683	26.4672	19.2810	15.9265	11.5231	8.7891	6.9172	3.6288	1.7423	-2.9924
26.4672	21.0432	14.8231	12.5335	9.4669	7.1121	5.5034	3.0992	1.2590	-2.5093
19.2810	14.8231	12.7128	10.2017	7.7451	6.0527	4.5158	2.9101	1.2151	-2.1518
15.9265	12.5335	10.2017	9.9662	6.9964	5.6539	4.3061	2.7373	1.1963	-1.9928
11.5231	9.4669	7.7451	6.9964	6.3529	4.4908	3.4261	2.2858	1.0150	-1.6416
8.7891	7.1121	6.0527	5.6539	4.4908	4.5537	2.9280	2.1043	1.0706	-1.3824
6.9172	5.5034	4.5158	4.3061	3.4261	2.9280	3.2914	1.7930	0.9149	-1.1456
3.6288	3.0992	2.9101	2.7373	2.2858	2.1043	1.7930	2.1077	0.8522	-0.8831
1.7423	1.2590	1.2151	1.1963	1.0150	1.0706	0.9149	0.8522	1.2128	-0.4942
-2.9924	-2.5093	-2.1518	-1.9928	-1.6416	-1.3824	-1.1456	-0.8831	-0.4942	0.6792

Panel B: Size Factor									
μ =0.2096, σ =3.1976, δ =0.2096, λ =0.0204, κ_3 =63.0820, κ_4 =2329.3833									
Decile Portfolios									
Small	2	3	4	5	6	7	8	9	Large
β									
2.1964	1.9059	1.6321	1.4843	1.3097	1.1070	0.9908	0.8134	0.6319	0.3032
α									
0.6529	0.5722	0.6183	0.6119	0.5979	0.6507	0.6082	0.6182	0.5841	0.5339
$\mathbf{\Sigma}_\epsilon$									
49.1522	39.8210	36.4736	33.5470	31.9046	32.7070	31.3220	30.5135	30.6135	27.9171
39.8210	38.7365	35.5968	33.3644	32.5677	33.1717	31.7548	31.3391	30.9958	28.4157
36.4736	35.5968	35.8935	33.1421	32.5133	33.2535	31.6684	31.6150	30.9928	28.1090
33.5470	33.3644	33.1421	32.5886	31.2411	32.0761	30.6112	30.4023	29.7651	26.8311
31.9046	32.5677	32.5133	31.2411	31.9316	31.9058	30.5595	30.4927	29.8437	26.9634
32.7070	33.1717	33.2535	32.0761	31.9058	33.4033	31.2926	31.2015	30.4507	27.1876
31.3220	31.7548	31.6684	30.6112	30.5595	31.2926	31.1114	30.1903	29.4565	26.3927
30.5135	31.3391	31.6150	30.4023	30.4927	31.2015	30.1903	30.7998	29.4141	26.2191
30.6135	30.9958	30.9928	29.7651	29.8437	30.4507	29.4565	29.4141	29.3838	25.8014
27.9171	28.4157	28.1090	26.8311	26.9634	27.1876	26.3927	26.2191	25.8014	24.4915

Table III. Parameter Values used in Monte Carlo Simulation: US data, 10 size-sorted portfolios.

The table presents the parameters used for the Monte Carlo simulation to estimate GMM under Beta and SDF methods. Parameters are estimated from single US factors and portfolios taken from Kenneth R. French library, January 1927 to December 2018 ($T = 1104$).

Panel C: Value Factor									
μ =0.3682, σ =3.4880, δ =0.3682, λ =0.0299, κ_3 =92.6417, κ_4 =3282.2015									
Decile Portfolios									
Small	2	3	4	5	6	7	8	9	Large
β									
1.3215	1.0101	0.8781	0.7841	0.6631	0.6651	0.5611	0.5238	0.4976	0.2951
α									
0.6266	0.5996	0.6370	0.6343	0.6282	0.6377	0.6093	0.5958	0.5333	0.4888
$\mathbf{\Sigma}_\epsilon$									
77.2036	66.3564	58.9870	54.2554	50.6370	46.8608	44.5395	40.3482	36.7966	29.9806
66.3564	63.4395	56.5904	52.6359	49.9237	46.5579	44.1559	40.7427	37.1873	30.6963
58.9870	56.5904	53.7316	49.5204	47.2705	44.6110	42.1987	39.5844	36.2158	30.0150
54.2554	52.6359	49.5204	47.6225	44.7794	42.5227	40.2872	37.7428	34.6031	28.6167
50.6370	49.9237	47.2705	44.7794	44.1080	41.3551	39.2925	37.1520	34.2859	28.6418
46.8608	46.5579	44.6110	42.5227	41.3551	40.5448	37.9611	36.1645	33.5728	28.2310
44.5395	44.1559	42.1987	40.2872	39.2925	37.9611	37.3131	34.8500	32.4578	27.4493
40.3482	40.7427	39.5844	37.7428	37.1520	36.1645	34.8500	34.2225	31.4955	26.8597
36.7966	37.1873	36.2158	34.6031	34.2859	33.5728	32.4578	31.4955	30.4520	25.9737
29.9806	30.6963	30.0150	28.6167	28.6418	28.2310	27.4493	26.8597	25.9737	24.3724

Panel D: Momentum Factor									
μ =0.6617, σ =4.6930, δ =0.6617, λ =0.0295, κ_3 =-315.7859, κ_4 =14986.8391									
Decile Portfolios									
Small	2	3	4	5	6	7	8	9	Large
β									
-0.8723	-0.7138	-0.6682	-0.5838	-0.5502	-0.5373	-0.4702	-0.4723	-0.4548	-0.3434
α									
1.6904	1.4439	1.4025	1.3093	1.2364	1.2382	1.1270	1.1012	1.0175	0.8247
$\mathbf{\Sigma}_\epsilon$									
81.6857	68.8806	60.2655	55.6434	50.7276	47.2306	44.5265	39.6965	36.0591	28.1299
68.8806	64.6311	56.8770	53.0932	49.4241	46.2851	43.6598	39.7564	36.1533	28.9266
60.2655	56.8770	53.2794	49.3051	46.2587	43.8098	41.2739	38.2313	34.8395	28.1163
55.6434	53.0932	49.3051	47.5951	44.0312	41.9589	39.5941	36.6677	33.5025	27.0181
50.7276	49.4241	46.2587	44.0312	42.7920	40.2110	38.1225	35.6563	32.7902	26.8634
47.2306	46.2851	43.8098	41.9589	40.2110	39.5691	36.9378	34.8153	32.2180	26.5570
44.5265	43.6598	41.2739	39.5941	38.1225	36.9378	36.2747	33.5359	31.1455	25.9092
39.6965	39.7564	38.2313	36.6677	35.6563	34.8153	33.5359	32.6494	29.9370	25.1703
36.0591	36.1533	34.8395	33.5025	32.7902	32.2180	31.1455	29.9370	28.9096	24.3221
28.1299	28.9266	28.1163	27.0181	26.8634	26.5570	25.9092	25.1703	24.3221	22.8367

Table IV. 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 λ GMM estimates under the Beta and the SDF methods. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. The first estimator decorated with * are from the Beta method; the second and third correspond to the first and second-stage un-centred SDF method; and the fourth and fifth to the first and second-stage centred SDF method. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$E[\widehat{\lambda}^*]$	$E[\widehat{\lambda}_1^U]$	$E[\widehat{\lambda}_2^U]$	$E[\widehat{\lambda}_1^C]$	$E[\widehat{\lambda}_2^C]$	$\sigma(\widehat{\lambda}^*)$	$\sigma(\widehat{\lambda}_1^U)$	$\sigma(\widehat{\lambda}_2^U)$	$\sigma(\widehat{\lambda}_1^C)$	$\sigma(\widehat{\lambda}_2^C)$
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 V. Relative standard errors for CAPM model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of λ GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda}) \div \mathrm{E}(\widehat{\lambda})$. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. Estimators decorated with * are from the Beta method; with A and B to the un-centred and centred SDF method; and with 1 and 2 to the first and second-stage respectively. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\sigma_r(\widehat{\lambda}_1^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C) \div \sigma_r(\widehat{\lambda}^*)$
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 VI. Relative standard errors for Fama-French model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of λ GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda}) \div \mathrm{E}(\widehat{\lambda})$. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. Estimators decorated with * are from the Beta method; with A and B to the un-centred and centred SDF method; and with 1 and 2 to the first and second-stage respectively. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\sigma_r(\widehat{\lambda}_1^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C) \div \sigma_r(\widehat{\lambda}^*)$
		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 VII. Relative standard errors for RUH model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of λ GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda}) \div \mathrm{E}(\widehat{\lambda})$. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. Estimators decorated with * are from the Beta method; with A and B to the un-centred and centred SDF method; and with 1 and 2 to the first and second-stage respectively. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\sigma_r(\widehat{\lambda}_1^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C) \div \sigma_r(\widehat{\lambda}^*)$
		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 VIII. Relative standard errors for Carhart model: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of λ GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda}) \div \mathrm{E}(\widehat{\lambda})$. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. Estimators decorated with * are from the Beta method; with A and B to the un-centred and centred SDF method; and with 1 and 2 to the first and second-stage respectively. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\sigma_r(\widehat{\lambda}_1^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C) \div \sigma_r(\widehat{\lambda}^*)$
		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 IX. Expected value and standard errors for Fama-French model: US data, 10 size-sorted portfolios.

The table presents the expected value and the standard error of λ GMM estimates under the Beta and the SDF methods. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. The first estimator decorated with * are from the Beta method; the second and third correspond to the first and second-stage un-centred SDF method; and the fourth and fifth to the first and second-stage centred SDF method. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\mathrm{E}[\widehat{\lambda}^*]$	$\mathrm{E}[\widehat{\lambda}_1^U]$	$\mathrm{E}[\widehat{\lambda}_2^U]$	$\mathrm{E}[\widehat{\lambda}_1^C]$	$\mathrm{E}[\widehat{\lambda}_2^C]$	$\sigma(\widehat{\lambda}^*)$	$\sigma(\widehat{\lambda}_1^U)$	$\sigma(\widehat{\lambda}_2^U)$	$\sigma(\widehat{\lambda}_1^C)$	$\sigma(\widehat{\lambda}_2^C)$
					Market					
60	2.26	2.39	2.93	2.67	2.95	2.79	2.91	3.48	3.33	3.66
360	2.07	2.14	2.22	2.23	2.27	1.01	1.06	1.09	1.13	1.14
600	2.06	2.12	2.18	2.20	2.23	0.77	0.81	0.82	0.86	0.86
1000	2.06	2.13	2.16	2.19	2.21	0.59	0.62	0.62	0.66	0.65
					Size					
60	0.86	1.69	1.94	2.16	2.34	4.29	5.95	6.46	6.30	6.54
360	1.40	1.76	1.83	1.87	1.93	1.59	2.29	2.14	2.36	2.19
600	1.44	1.78	1.82	1.86	1.90	1.24	1.79	1.65	1.84	1.69
1000	1.48	1.81	1.84	1.88	1.91	0.93	1.34	1.23	1.38	1.26
					Value					
60	2.18	2.79	3.14	2.49	2.13	3.94	11.77	10.58	12.46	10.97
360	2.57	2.89	2.95	2.95	2.83	1.38	4.80	3.78	5.01	3.90
600	2.63	2.90	2.93	2.98	2.89	1.07	3.66	2.85	3.81	2.95
1000	2.64	2.86	2.87	2.94	2.88	0.81	2.81	2.19	2.92	2.26

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

The table presents the expected value and the standard error of λ GMM estimates under the Beta and the SDF methods. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. The first estimator decorated with * are from the Beta method; the second and third correspond to the first and second-stage un-centred SDF method; and the fourth and fifth to the first and second-stage centred SDF method. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\mathrm{E}[\widehat{\lambda}^*]$	$\mathrm{E}[\widehat{\lambda}_1^U]$	$\mathrm{E}[\widehat{\lambda}_2^U]$	$\mathrm{E}[\widehat{\lambda}_1^C]$	$\mathrm{E}[\widehat{\lambda}_2^C]$	$\sigma(\widehat{\lambda}^*)$	$\sigma(\widehat{\lambda}_1^U)$	$\sigma(\widehat{\lambda}_2^U)$	$\sigma(\widehat{\lambda}_1^C)$	$\sigma(\widehat{\lambda}_2^C)$
					Market					
60	1.94	2.33	2.74	2.67	2.85	2.71	3.10	3.50	3.52	3.69
360	1.95	2.13	2.20	2.26	2.28	0.98	1.07	1.09	1.15	1.15
600	1.94	2.10	2.14	2.21	2.22	0.76	0.82	0.83	0.88	0.87
1000	1.95	2.10	2.13	2.20	2.21	0.59	0.63	0.63	0.67	0.67
					Momentum					
60	3.73	4.32	5.23	0.55	0.98	3.86	11.12	10.26	11.77	10.70
360	2.77	2.98	3.25	1.99	2.14	1.62	5.80	5.08	6.12	5.16
600	2.59	2.80	2.92	2.26	2.31	1.23	4.56	3.92	4.84	4.05
1000	2.52	2.69	2.80	2.41	2.48	0.94	3.57	2.97	3.77	3.09
					Value					
60	1.51	2.80	3.03	1.90	1.19	4.00	12.10	10.01	13.14	10.56
360	2.26	2.83	2.89	2.60	2.41	1.40	4.56	3.75	4.85	3.95
600	2.35	2.81	2.85	2.72	2.60	1.06	3.51	2.90	3.74	3.07
1000	2.39	2.84	2.87	2.83	2.76	0.80	2.70	2.20	2.87	2.34

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

The table presents the expected value and the standard error of λ GMM estimates under the Beta and the SDF methods. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. The first estimator decorated with * are from the Beta method; the second and third correspond to the first and second-stage un-centred SDF method; and the fourth and fifth to the first and second-stage centred SDF method. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\mathrm{E}[\widehat{\lambda}^*]$	$\mathrm{E}[\widehat{\lambda}_1^U]$	$\mathrm{E}[\widehat{\lambda}_2^U]$	$\mathrm{E}[\widehat{\lambda}_1^C]$	$\mathrm{E}[\widehat{\lambda}_2^C]$	$\sigma(\widehat{\lambda}^*)$	$\sigma(\widehat{\lambda}_1^U)$	$\sigma(\widehat{\lambda}_2^U)$	$\sigma(\widehat{\lambda}_1^C)$	$\sigma(\widehat{\lambda}_2^C)$
					Market					
60	1.99	2.36	2.69	2.75	2.88	2.89	3.16	3.53	3.66	3.87
360	1.90	2.10	2.16	2.24	2.27	0.99	1.08	1.10	1.16	1.16
600	1.89	2.08	2.11	2.20	2.21	0.76	0.82	0.82	0.87	0.87
1000	1.90	2.08	2.10	2.19	2.20	0.59	0.63	0.64	0.68	0.67
					Size					
60	-0.08	1.64	1.80	2.43	2.48	4.47	6.41	6.75	7.01	7.15
360	0.90	1.76	1.79	2.08	2.06	1.61	2.46	2.33	2.55	2.41
600	0.99	1.73	1.77	1.97	1.97	1.25	1.89	1.77	1.96	1.82
1000	1.04	1.77	1.78	1.95	1.94	0.96	1.45	1.34	1.50	1.38
					Value					
60	1.40	2.72	2.96	1.12	1.10	4.13	12.47	11.64	13.65	12.64
360	2.17	2.76	2.85	2.14	2.29	1.43	5.55	4.87	5.89	5.16
600	2.26	2.89	2.90	2.53	2.59	1.09	4.33	3.75	4.59	3.98
1000	2.29	2.81	2.84	2.63	2.69	0.81	3.30	2.86	3.49	3.03
					Momentum					
60	3.83	4.25	4.98	0.65	0.89	4.07	11.01	10.81	11.98	11.69
360	2.69	2.89	3.11	1.89	2.13	1.62	5.54	5.33	5.86	5.53
600	2.57	2.79	2.90	2.21	2.35	1.24	4.17	4.01	4.41	4.17
1000	2.50	2.68	2.74	2.37	2.46	0.95	3.25	3.11	3.42	3.25

Table XII. Relative standard errors for four alternative single-factor models: US data, 10 size-sorted portfolios.

The table presents the relative standard errors of λ GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\lambda}) = \sigma(\widehat{\lambda}) \div \mathrm{E}(\widehat{\lambda})$. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. Estimators decorated with * are from the Beta method; with A and B to the un-centred and centred SDF method; and with 1 and 2 to the first and second-stage respectively. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\sigma_r(\widehat{\lambda}_1^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^U) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_1^C) \div \sigma_r(\widehat{\lambda}^*)$	$\sigma_r(\widehat{\lambda}_2^C) \div \sigma_r(\widehat{\lambda}^*)$
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 XIII. Expected value and standard errors for four alternative single-factor models: US data, 10 size-sorted portfolios.

The table presents the expected value and the standard error of λ GMM estimates under the Beta and the SDF methods. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. The first estimator decorated with * are from the Beta method; the second and third correspond to the first and second-stage un-centred SDF method; and the fourth and fifth to the first and second-stage centred SDF method. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\mathrm{E}[\widehat{\lambda}^*]$	$\mathrm{E}[\widehat{\lambda}_1^U]$	$\mathrm{E}[\widehat{\lambda}_2^U]$	$\mathrm{E}[\widehat{\lambda}_1^C]$	$\mathrm{E}[\widehat{\lambda}_2^C]$	$\sigma(\widehat{\lambda}^*)$	$\sigma(\widehat{\lambda}_1^U)$	$\sigma(\widehat{\lambda}_2^U)$	$\sigma(\widehat{\lambda}_1^C)$	$\sigma(\widehat{\lambda}_2^C)$
Single-factor model loaded with market factor										
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
Single-factor model loaded with size factor										
60	1.66	1.69	2.02	1.82	1.84	4.31	7.90	6.19	8.34	5.81
360	1.80	1.77	1.87	1.79	1.85	1.64	2.78	1.85	2.82	1.84
600	1.83	1.86	1.90	1.88	1.88	1.26	2.06	1.37	2.09	1.38
1000	1.85	1.87	1.88	1.89	1.88	0.97	1.60	1.05	1.61	1.06
Single-factor model loaded with value factor										
60	2.84	2.79	3.33	3.22	2.09	3.94	14.17	8.97	15.24	8.51
360	2.93	2.90	3.05	2.99	2.87	1.44	4.19	2.93	4.30	2.95
600	2.92	2.91	2.97	2.98	2.88	1.09	3.16	2.18	3.24	2.21
1000	2.91	2.96	2.98	3.01	2.94	0.84	2.41	1.67	2.46	1.70
Single-factor model loaded with momentum factor										
60	4.30	4.39	5.79	5.51	3.25	3.96	12.52	8.15	15.44	7.85
360	3.01	3.06	3.29	3.22	3.00	1.75	3.26	2.91	3.53	2.91
600	2.89	2.91	3.05	3.02	2.89	1.35	2.45	2.16	2.61	2.17
1000	2.77	2.76	2.87	2.84	2.80	1.02	1.82	1.61	1.91	1.63

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

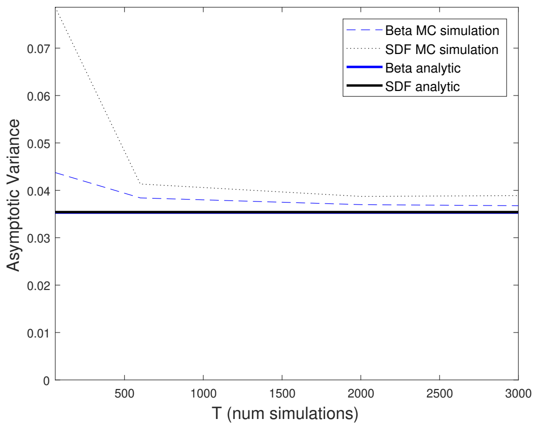
The table presents the relative standard errors of π GMM estimates under the Beta and the SDF methods, computed as $\sigma_r(\widehat{\pi}) = \sigma(\widehat{\pi}) \div \mathrm{E}(\widehat{\pi})$. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. Estimators decorated with * are from the Beta method; with A and B to the un-centred and centred SDF method; and with 1 and 2 to the first and second-stage respectively. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

T	$\sigma_r(\widehat{\pi}_1^U) \div \sigma_r(\widehat{\pi}^*)$	$\sigma_r(\widehat{\pi}_2^U) \div \sigma_r(\widehat{\pi}^*)$	$\sigma_r(\widehat{\pi}_1^C) \div \sigma_r(\widehat{\pi}^*)$	$\sigma_r(\widehat{\pi}_2^C) \div \sigma_r(\widehat{\pi}^*)$
		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

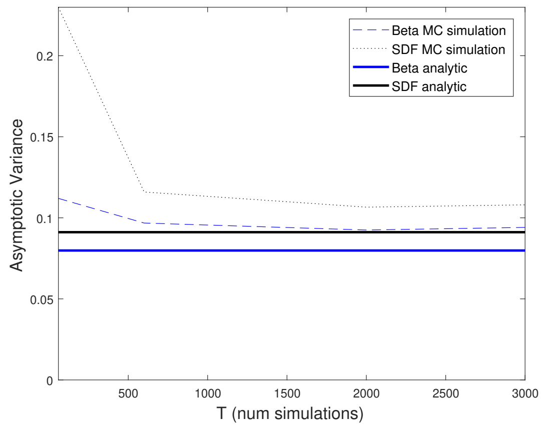
Table XV. Expected value and standard errors for four alternative asset pricing models: US data, 10 size-sorted portfolios.

The table presents the expected value and the standard error of π GMM estimates under the Beta and the SDF methods. The returns and factors are generated under the null hypothesis with the factors sampled from the empirical distribution. The first estimator decorated with * are from the Beta method; the second and third correspond to the first and second-stage un-centred SDF method; and the fourth and fifth to the first and second-stage centred SDF method. The results are presented for different lengths of time series observations, and they are based on 10,000 simulations.

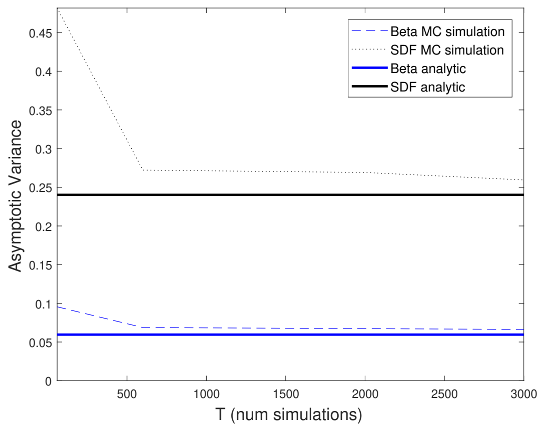
T	$\mathrm{E}[\widehat{\pi}^*]$	$\mathrm{E}[\widehat{\pi}_1^U]$	$\mathrm{E}[\widehat{\pi}_2^U]$	$\mathrm{E}[\widehat{\pi}_1^C]$	$\mathrm{E}[\widehat{\pi}_2^C]$	$\sigma(\widehat{\pi}^*)$	$\sigma(\widehat{\pi}_1^U)$	$\sigma(\widehat{\pi}_2^U)$	$\sigma(\widehat{\pi}_1^C)$	$\sigma(\widehat{\pi}_2^C)$
					CAPM					
60	35.39	23.50	47.11	24.27	41.30	20.05	10.98	30.04	11.34	24.39
360	14.63	9.74	15.13	9.90	15.01	8.23	4.57	8.52	4.64	8.40
600	11.25	7.50	11.48	7.61	11.48	6.35	3.48	6.48	3.54	6.45
1000	8.80	5.86	8.90	5.94	8.96	5.00	2.76	5.06	2.80	5.08
					Fama-French					
60	24.10	11.50	30.48	12.41	27.23	12.37	3.69	21.36	4.04	17.88
360	10.12	4.61	7.26	4.77	6.92	4.54	1.39	3.25	1.44	2.80
600	8.37	3.57	5.25	3.69	5.11	3.61	1.06	2.14	1.10	1.95
1000	7.07	2.76	3.85	2.84	3.81	2.89	0.82	1.49	0.84	1.42
					RUH					
60	36.68	12.38	36.33	13.49	35.80	20.01	4.22	25.02	4.61	23.71
360	15.46	5.02	10.59	5.27	10.66	8.28	1.62	5.80	1.70	5.68
600	12.33	3.89	7.79	4.07	7.93	6.65	1.24	4.01	1.29	4.00
1000	10.15	3.03	5.87	3.16	6.02	5.31	0.98	2.94	1.02	2.95
					Carhart					
60	32.97	9.99	27.49	11.08	26.98	19.16	3.37	19.41	3.75	18.34
360	16.28	4.17	6.87	4.39	6.78	7.98	1.36	3.44	1.43	3.14
600	14.49	3.24	4.89	3.40	4.91	6.31	1.03	2.13	1.08	2.02
1000	13.28	2.52	3.59	2.63	3.65	4.99	0.81	1.49	0.84	1.47



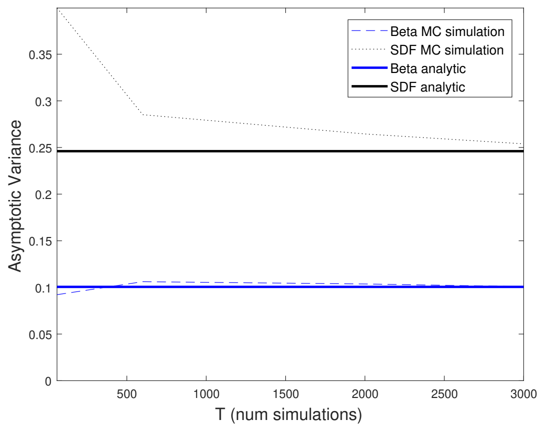
(a) Market risk factor



(b) Size factor



(c) Value factor



(d) Momentum factor

Figure 1. Asymptotic variance of the analytic and empirical estimated GMM with Beta and SDF methods, from a Monte Carlo simulation with data calibrated to the empirical observed market risk, size, value, and momentum factors from January 1927 to December 2018. Data from Kenneth R. French library.