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

EEA/ESEM, Erasmus University Rotterdam – August 2024

Juan Arismendi-Zambrano,* Massimo Guidolin, Martin Lozano

27 August 2024

School of Business, University College Dublin, Dublin, Ireland

Contact: juan.arismendi-zambrano@ucd.ie

Question (?) Hansen, 1982 (Appendix)

and we let

$$Dg_N(\beta^1, \ldots, \beta^r) = \begin{bmatrix} \frac{\partial g_N^1}{\partial \beta}(\beta^1) \\ \vdots \\ \frac{\partial g_N^r}{\partial \beta}(\beta^r) \end{bmatrix}.$$

Using Taylor's theorem and Assumptions 3.2–3.4, with probability arbitrarily close to one for sufficiently large N we can write

$$(12) \qquad g_N(b_N^*) = g_N(\beta_0) + Dg_N(\bar{b}_N^1, \ldots, \bar{b}_N^r)(b_N^* - \beta_0)$$

where b_N^i is between β_0 and b_N^* for $i = 1, \ldots, r$. Premultiplying by a_N^*, we obtain

$$a_N^* g_N(b_N^*) = a_N^* g_N(\beta_0) + a_N^* Dg_N(\bar{b}_N^1, \ldots, \bar{b}_N^r)(b_N^* - \beta_0).$$

Dovonon and Hall (2018) (JoE)? Hansen and Lee (2021) (ECMA)? I believe, but results still need it for financial economics (in particular empirical finance).

Motivation

- Even though both characterizations are theoretically equivalent, the estimators of the parameters of interest may in general be different under the two representations:

    1. The **Beta representation** is formulated to analyze the (estimates of the) factor risk premia, δ, and the Jensen alphas, α;

    2. While the **SDF representation** is intended to analyze the (estimates of the) parameters that enter into the assumed stochastic discount factor, λ, and the resulting pricing errors, π.

- The fact that both representations are equivalent, imply that there is a one-to-one mapping between δ and λ; and between α and π, which may facilitates the comparison of the estimators.
- However, this **theoretical equivalence** does not necessarily entail an **empirical equivalence**.

Question

- Therefore, the empirical/experimental questions that arise are:[1]

  1. **Is it better to make inferences on δ or on λ?** And analogously:
  2. **Is it better to make inferences on α or on π?**
  3. Or, equivalently, is it better to perform estimation of the Beta representation (i.e., recover δ and α), or on the SDF representation (i.e., on λ and π?)

[1] From now on, we will refer only to GMM estimators of either the Beta or the SDF representations, although MLE remains as possible as "risky" in terms of misspecifications.

Contribution

- Our empirical and analytical results show that, in general, the **Beta method** is more **efficient** but **less accurate in pricing** vs. the **SDF method**; that is, δ and π have lower simulated, adjusted standard errors (coefficients of variation) than the λ and α estimators.
- To disentangle the source of the results, we extend **Jagannathan and Wang (2002) univariate Gaussian and Kan and Zhou (2001) univariate non-Gaussian analytical results** to the case of multi-factor models.
- Valuable as researchers and practitioners would have an a priori idea about the benefits and costs of either representation. We provide an **out-of-sample (OOS) trading exercise** that helps on this.
- The source of the efficiency of the Beta method over the SDF method, is rooted in the **higher-order moments effects**, that impact the SDF-based estimation, but hardly the Beta.

Literature Review

- First attempt to assess the **performance in finite samples** of the **Beta** versus the **SDF** approaches, and using a standardized **single-factor model**, is Kan and Zhou (1999) who show that SDF method may be much less efficient than the Beta method.
- Jagannathan and Wang (2002) and Cochrane (2001) adverse this conclusion in a **non-standardized single-factor model** and based on **joint normality assumption** for both the asset returns and the factors.
  - They conclude that the SDF method is as efficient as the Beta method for estimating risk premiums.
  - Model specification tests are equally powerful.
- Yet Kan and Zhou (2001) show that under more general distributional assumptions, inference based on λ can still be less reliable than inference based on δ in realistic situations where the factors are leptokurtic.

The Beta Method

- The standard linear asset-pricing model under the Beta representation is given by
$$\mathsf{E}[r_t] = \mathbf{B}\delta. \tag{1}$$
- Equivalently, we can identify $\mathbf{B}$ as a parameter in the time-series regression
$$r_t = \phi + \mathbf{B} f_t + \epsilon_t, \tag{2}$$
- Then, we can derive the associated set of moment conditions $\mathbf{g}_b$ of the factor model as:
$$\begin{array}{lcl} \mathsf{E}\left[r_t - \mathbf{B}(\delta - \mu + f_t)\right] & = & 0_N, \\ \mathsf{E}\left[[r_t - \mathbf{B}(\delta - \mu + f_t)] f_t'\right] & = & 0_{N\times K}, \\ \mathsf{E}\left[f_t - \mu\right] & = & 0_{K\times 1}, \\ \mathsf{E}\left[(f_t - \mu)(f_t - \mu)' - \mathbf{\Sigma}\right] & = & 0_{K\times K}, \end{array} \tag{3}$$
- The corresponding unknown parameters are $\theta^* = \left[\delta^{*\prime}, \mathrm{vec}(\mathbf{B}^*)', \mu^{*\prime}, \mathrm{vec}(\mathbf{\Sigma}^*)\right]'$. The observable variables are $x_t = [r_t', f_t']'$.

The SDF Method

- To derive the SDF representation from the Beta one, we substitute the expression for $\mathbf{B} \equiv \mathsf{E}[r_t(f_t - \mu)']\mathbf{\Sigma}^{-1}$ into equation (22) and rearrange terms:

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

  For traded factors, $\delta = \mu$ so $1 + \delta'\mathbf{\Sigma}^{-1}\mu = 1 + \mu'\ \mathbf{\Sigma}^{-1}\mu \geq 1$ and

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

- Now transform the vector of risk premium δ into a vector λ as

  $$\lambda = \frac{\mathbf{\Sigma}^{-1}\delta}{1 + \delta'\mathbf{\Sigma}^{-1}\mu}, \qquad (4)$$

  to obtain the following SDF representation (g_s),

  $$\mathsf{E}[r_t(1 - \lambda' f_t)] = 0_N, \qquad (5)$$

  where $m_t \equiv 1 - \lambda' f_t$ as $\mathsf{E}[r_t m_t] = 0_N$.

Beta vs. SDF Method

- There is a one-to-one mapping between the factor risk premium δ and the SDF parameters λ, which facilitates the comparison of the two methods.
- Hence, we can derive an estimate of λ not only by the SDF but also by the Beta representation; by the same token we can derive an estimate of δ not only by the Beta but also by the SDF method.
- From the previous definition of λ in equation (28), we have:

$$\lambda = \delta' \left(\mathbf{\Sigma} + \delta \mu' \right)^{-1}, \qquad \text{or} \qquad \delta = \frac{\mathbf{\Sigma} \lambda}{1 - \mu' \lambda}. \tag{6}$$

- In a similar way, by substituting (30) into π, we can find a one-to-one mapping between π from the SDF and α from the Beta method:

$$\pi = \left(1 + \delta' \mathbf{\Sigma}^{-1} \mu\right)^{-1} \alpha, \qquad \text{or} \qquad \alpha = \left(1 + \delta' \mathbf{\Sigma}^{-1} \mu\right) \pi. \tag{7}$$

Asymptotic Results

- As in Jagannathan and Wang (2002), if one is ready to use asymptotics, then clear-cut analytical results can be derived.

Theorem (1 – Single-factor – From Kan and Zhou (2001) but we use it to explain the structure of our general multifactor result)

In the case of a single-factor, the asymptotic variance of the GMM risk premium estimate from the SDF representation is,

$$Avar(\hat{\lambda}) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Omega}^{-1}\beta\right)^{-1} + \frac{\sigma^2(\sigma^4+\delta^4)}{(\sigma^2+\mu\delta)^4} + \frac{2\kappa_3(\delta^3-\delta\sigma^2)+\delta^2\left(\kappa_4-3\sigma^4\right)}{(\sigma^2+\mu\delta)^4}, \quad (8)$$

where σ^2 is the variance of the single-factor f_t, and κ_3, κ_4 are the skewness and kurtosis.

Asymptotic Results (Single-factor)

Theorem (1 – Single-factor)

The equivalent asymptotic covariance for the Beta representation is,

$$Avar(\lambda^*) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Omega}^{-1}\beta\right)^{-1} + \frac{\sigma^6}{(\sigma^2+\mu\delta)^4}, \quad (9)$$

when using a first-order Delta approximation, and,

$$Avar(\lambda^*) = \frac{\sigma^2(\sigma^2+\delta^2)}{(\sigma^2+\mu\delta)^4}\left(\beta'\mathbf{\Omega}^{-1}\beta\right)^{-1} + \frac{\sigma^6}{(\sigma^2+\mu\delta)^4} - 2\frac{(\sigma^4\mu)}{(\sigma^2+\mu\delta)^5}\kappa_{3,\delta}, \quad (10)$$

when using a second-order Delta approximation, where $\kappa_{3,\delta}$ is the asymptotic third-order central moments of the estimator of δ.

Proof.: See Appendix A of the paper.

Single-factors Gaussian, Beta vs. SDF

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

$$Avar(\hat{\lambda}) = \frac{\sigma^2(\sigma^2 + \delta^2)}{(\sigma^2 + \mu\delta)^4} \left(\beta' \Omega^{-1} \beta\right)^{-1} + \frac{\sigma^2(\sigma^4 + \delta^4)}{(\sigma^2 + \mu\delta)^4}. \quad (11)$$

Corollary (1)

Consider a single-factor linear asset pricing model (as in (22) but f_t univariate), where f_t returns are from a portfolio of non-traded assets, and consider these returns to have a Gaussian distribution. Then, when GMM is used for the estimation of the parameters (obtained under the uncentered first-stage ($\mathbf{W} = \mathbf{I}$)), the Beta representation has a higher efficiency than the SDF representation when measuring the risk premia, but the difference is negligible for asset pricing tests.

Proof.: In this case $Avar(\hat{\lambda}) - Avar(\lambda^*) = \sigma^2\delta^4 / \left(\sigma^2 + \mu\delta\right)^4$, that is always positive.

GMM Structure of Covariance Matrix of Moments ($\mathbf{S_s} = \mathbf{g_s g_s'}$ and $\mathbf{S_b} = \mathbf{g_b g_b'}$)

But why we get this systematic difference in results?

- If we note the covariance matrix structure of the Beta representation ($\mathbf{S_b} = \mathbf{g_b g_b'}$), the matrix is **divided in two blocks**, and the higher-order moments have no effects on the results on the first block of the equation, that is where the δ is estimated:

$$S_b = \begin{pmatrix} \boldsymbol{\Omega} & \mu\boldsymbol{\Omega} & 0 & 0 \\ \mu\boldsymbol{\Omega} & (\mu^2+\sigma^2)\boldsymbol{\Omega} & 0 & 0 \\ 0 & 0 & \sigma^2 & \kappa_3 \\ 0 & 0 & \kappa_3 & \kappa_4 - \sigma^4 \end{pmatrix}, \qquad (12)$$

- This doesn't happen in the covariance matrix of the SDF representation ($\mathbf{S_s} = \mathbf{g_s g_s'}$).

Bootstrap – DGP

- We develop an empirical simulation, where the returns are generated by bootstrapping the observed historical returns, and factors are generated by bootstrapping the observed historical factors:

$$r_t = \mathbf{B}(\delta - \mu + f_t) + \epsilon_t, \qquad \epsilon_t | f_t \sim N(\mathbf{0}, \Omega), \qquad f_t \sim \mathbf{F}, \quad (13)$$

where $\mathbf{F}$ is the sample factor matrix observed $(T \times K)$.

  - The factors f_t are drawn from their empirical distribution which allows for non-normalities, autocorrelation, heteroskedasticity and dependence.
  - To artificially generate the excess returns we use the model in (13).
  - Repeat this independently to obtain 10,000 draws of the estimators.

Bootstrap Test

- We estimate a ratio of the relative standard errors of the method,

$$\sigma_r(\hat{\lambda}) = \sigma(\hat{\lambda})/E(\hat{\lambda}), \qquad \text{and,} \qquad \sigma_r(\lambda^*) = \sigma(\lambda^*)/E(\lambda^*), \tag{14}$$

and with them we measure the four ratios, $\sigma_r(\hat{\lambda}_1^U)/\sigma_r(\lambda^*)$, $\sigma_r(\hat{\lambda}_2^U)/\sigma_r(\lambda^*)$, $\sigma_r(\hat{\lambda}_1^C)/\sigma_r(\lambda^*)$, and $\sigma_r(\hat{\lambda}_2^C)/\sigma_r(\lambda^*)$, where the U and C indicate estimators obtained by the GMM from the uncentred and centered SDF representations, and with 1 and 2 represent estimators obtained by the first and second-stage methods.

- With the **bootstrap of the ratios**, we test the null hypothesis that the interval of confidence (measured at p-values of 10%, 5% and 1%) is equal to 1, by rejecting the test when both extremes of the interval are superior to 1, or both extremes are inferior to 1.

Market Risk Factor Modified (Arismendi and Kimura, 2016) – Increasing Third-order Moments

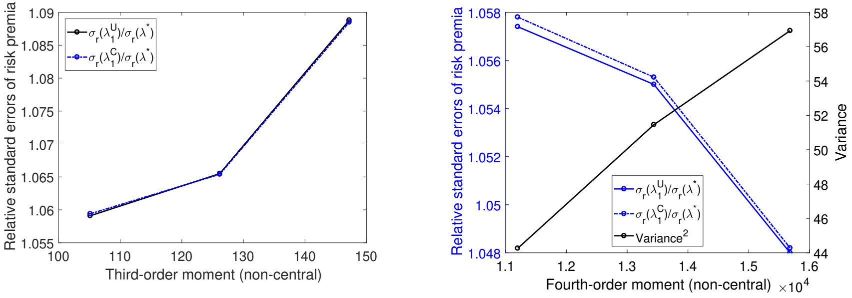

Figure 1: Effects of the increasing factror skewness and kurtosis on the ratio of the relative standard error between the Beta and the SDF methods risk premium estimation.

Gaussian Non-traded Factors Simulation

Surprise? (No, Ravi's results are great.)

T	$\sigma_r(\widehat{\lambda_1}^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda_2}^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda_1}^C)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda_2}^C)/\sigma_r(\lambda^*)$
60	0.9948***	0.9677***	1.0507***	0.9961
360	0.9888***	0.9263***	1.0368***	0.9696***
600	0.9949***	0.9349***	1.0436***	0.9810***
1000	0.9919***	0.9374***	1.0405***	0.9849***
2000	1.0000	0.9315***	1.0483***	0.9794***
3000	1.0000	0.9335***	1.0483***	0.9815***
4000	0.9999	0.9321***	1.0480***	0.9801***
5000	1.0000	0.9306***	1.0478***	0.9784***
15000	0.9999***	0.9364***	1.0481***	0.9850***

Table 1: Relative standard errors of risk premium estimated from the CAPM model: US data, 10 size-sorted (non-traded, Gaussian) portfolios.

Relative standard errors for the CAPM model

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\lambda^*)$
60	1.0735***	1.0003	1.1101***	1.0522***
360	1.0750***	1.0047***	1.1010***	1.0306***
600	1.0613***	1.0026***	1.0871***	1.0280***
1000	1.0652***	1.0020**	1.0912***	1.0279***
Asym	1.0863 (1st)			

Table 2: Relative standard errors of risk premium estimated from the CAPM model: US data, 10 size-sorted portfolios.

Relative Standard Errors for Carhart Model

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\lambda^*)$
		Market		
60	1.1823***	1.0553***	1.5959***	1.4850***
360	1.3860***	1.2834***	1.7107***	1.5805***
600	1.4097***	1.2949***	1.6881***	1.5452***
1000	1.4567***	1.3483***	1.7079***	1.5760***
Asym	1.4157			
		Size		
60	2.9360***	2.6395***	1.8828***	1.9595***
360	0.3931***	0.3665***	0.2863***	0.2762***
600	0.2440***	0.2245***	0.1891***	0.1794***
1000	0.1456***	0.1334***	0.1204***	0.1133***
Asym	1.4635			
		Value		
60	1.5132***	1.2923***	12.6684***	10.5390***
360	3.2375***	2.8995***	6.8216***	5.6345***
600	3.4535***	3.1444***	5.4927***	4.7474***
1000	3.6626***	3.3574***	4.9674***	4.3756***
Asym	4.4710			
		Momentum		
60	2.5915***	2.1321***	17.5486***	13.2040***
360	4.3878***	4.0154***	8.4040***	7.2284***
600	4.7195***	4.3607***	7.2355***	6.4094***
1000	4.8740***	4.6044***	6.5327***	5.9530***
Asym	4.9860			

Table 3: Relative standard errors of risk premium estimated from the Carhart model: US data, 10 size-sorted portfolios.

Relative Standard Errors Pricing Accuracy

T	$\sigma_r(\widehat{\pi}_1^U)/\sigma_r(\pi^*)$	$\sigma_r(\widehat{\pi}_2^U)/\sigma_r(\pi^*)$	$\sigma_r(\widehat{\pi}_1^C)/\sigma_r(\pi^*)$	$\sigma_r(\widehat{\pi}_2^C)/\sigma_r(\pi^*)$
		CAPM		
60	0.8430***	1.1292***	0.8425***	1.0438***
360	0.8337***	0.9962	0.8334***	0.9901**
600	0.8286***	1.0018	0.8286***	0.9987
1000	0.8319***	0.9992	0.8317***	0.9987
Asymptotic$^{\text{ave}}$ (1st Ord)	0.9951			
		Fama-French		
60	0.6475***	1.5690***	0.6539***	1.6775***
360	0.7197***	1.1929***	0.7198***	1.1480***
600	0.7415***	1.1035***	0.7420***	1.0298*
1000	0.7575***	1.0008	0.7576***	0.9768**
Asymptotic$^{\text{ave}}$ (1st Ord)	0.7779			
		Asness-Moskowitz-Pedersen		
60	0.6307***	1.3723***	0.6282***	1.4370***
360	0.6233***	1.1571***	0.6200***	1.1372***
600	0.6414***	1.1250***	0.6436***	1.1055***
1000	0.6670***	1.1054***	0.6684***	1.0843***
Asymptotic$^{\text{ave}}$ (1st Ord)	0.6644			
		Carhart		
60	0.6053***	1.4469***	0.6065***	1.4902***
360	0.7910***	1.6388***	0.7868***	1.4804***
600	0.9292***	1.6034***	0.9204***	1.4802***
1000	1.1151***	1.6688***	1.1073***	1.5735***
Asymptotic$^{\text{ave}}$ (1st Ord)	0.7337			

Table 4: RELATIVE STANDARD ERRORS OF PRICING ERRORS FOR FOUR ALTERNATIVE ASSET PRICING MODELS: US DATA, 10 SIZE-SORTED PORTFOLIOS.

Mean-Variance Estimation (Optimal Portfolio)

- Consider a linear asset pricing model as in equation (33), and sample asset returns observations $\mathbf{R}$ of dimension $T \times N$ from the asset returns r_t.
- From $\mathbf{R}$, using the SDF representation as in equation (29), we
  - Estimate the K-factor pricing model using the known factors f_t from which we have a sample factor matrix $\mathbf{F}$ of dimension $K \times N$, and
  - Estimate the corresponding sample error $\tilde{\mathbf{E}}$ with dimension $T \times N$.

Mean-Variance Estimation (Cont.)

- The new filtered returns are $\tilde{\mathbf{R}} = \mathbf{R} - \tilde{\mathbf{E}}$, that with the SDF representation is estimated as,

$$\tilde{\mathbf{R}} = \mathbf{R} - \underbrace{\mathbf{R}(\mathbf{1} - \tilde{\lambda}'\mathbf{F})}_{\text{estimated error } \tilde{\mathbf{E}}}, \tag{15}$$

where $\tilde{\lambda}$ is the dimension $K \times 1$ estimated risk premium, and $\mathbf{1}$ is the unit vector.

- With the resulting filtered excess returns $\tilde{\mathbf{R}}$ we compute the required inputs for the optimal mean-variance portfolio, the sample mean and sample covariance as estimates of the expected mean and variance

$$\mu_{r_t} = E[r_t] = \bar{\tilde{R}}, \tag{16}$$

$$\mathbf{\Sigma}_{r_t} = E\left[\left(r_t - \mu_{r_t}\right)\left(r_t - \mu_{r_t}\right)'\right] = COV(\tilde{R}), \tag{17}$$

and obtain optimal allocation rules $\tilde{\omega}$.

Results (OOS) – Smooth Initialization

Model difference	Initial window size 360	480	600
	Fama-French		
SR^*	7.15	9.23	10.32
$SR^* - SR_1^U$	-1.876	-0.812	-0.176
$SR^* - SR_2^U$	-0.603	0.449	1.124
$SR^* - SR_1^C$	-1.818	-0.747	-0.105
$SR^* - SR_2^C$	-0.563	0.498	1.175
	Asness-Moskowitz-Pedersen		
SR^*	19.65	20.90	25.36
$SR^* - SR_1^U$	9.590	7.568	9.672
$SR^* - SR_2^U$	7.052	6.445	8.060
$SR^* - SR_1^C$	10.497	8.228	10.566
$SR^* - SR_2^C$	7.162	6.347	8.168
	Carhart		
SR^*	19.85	20.85	25.18
$SR^* - SR_1^U$	5.981	5.250	6.733
$SR^* - SR_2^U$	6.004	5.427	7.099
$SR^* - SR_1^C$	5.919	5.224	6.822
$SR^* - SR_2^C$	6.259	5.671	7.416

Table 5: Sharpe ratio of mean-variance optimal portfolios with parameters estimated with Beta and SDF representations (smooth expected covariance initialization model).

Conclusions

- One of few paper in which a **trade-off** between **estimation efficiency** and **pricing accuracy** is explicitly used to better understand the differences of the Beta and SDF representations.
  - This evidence is useful for researchers and practitioners because they could choose a proper procedure in terms of a given application.
- Previous published studies that compared the two approaches were performed under conditions (single factor, normality) that are empirically insufficient to differentiate their performance .
- Once we relax those conditions, we find evidence suggesting that Beta representations may easily lead to **better risk premium** estimators while the SDF method leads to **better pricing error estimators** and hence pricing accuracy.

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.

Ang, A., J. Liu, and K. Schwarz (2020, may). Using stocks or portfolios in tests of factor models. *Journal of Financial and Quantitative Analysis 55*(3), 709–750.

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

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.

Dovonon, P. and A. R. Hall (2018, jul). The asymptotic properties of GMM and indirect inference under second-order identification. *Journal of Econometrics 205*(1), 76–111.

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

Hansen, B. E. and S. Lee (2021). Inference for iterated GMM under misspecification. *Econometrica 89*(3), 1419–1447.

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

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. 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.

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. 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*.

Ledoit, O. and M. Wolf (2017, jun). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies 30*(12), 4349–4388.

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.

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

Gaussian Factors vs. Non-Gaussian Factors

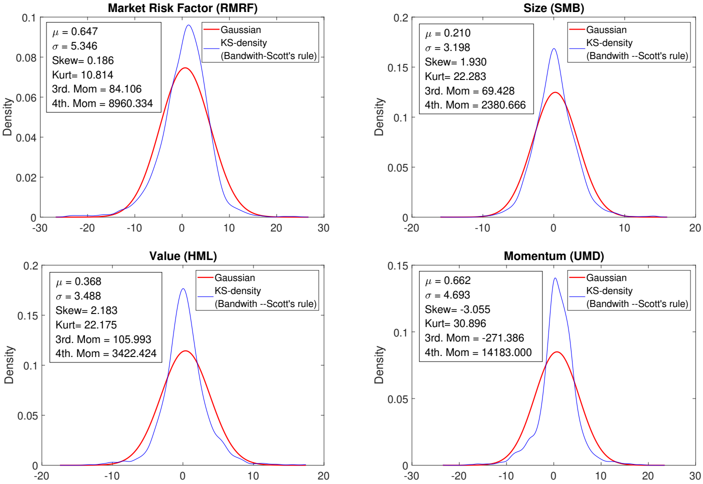

Figure 2: US: Kernel density estimators vs. benchmark Gaussian distributions – Fama-French factors + Carhart's momentum

Motivation: Small Sample Statistics (Skewness of α and δ (size factor – smb)) – (Some Answers)

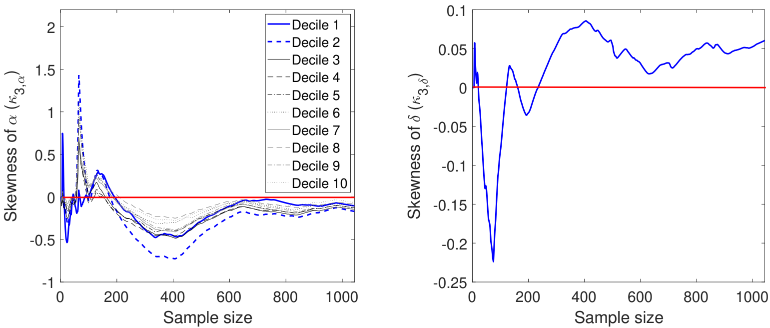

Figure 3: Skewness of α and δ for the size factor: US data, 10 size-sorted portfolios.

Motivation

- Any asset pricing model can be formally characterized either under a **Beta** or a **stochastic discount factor (SDF)** representation.
- The **SDF representation**, states that in the absence of arbitrage opportunities, the value of any asset should equal the expected value of the product of the payoff of the asset and the SDF.
- Whereas according to the **Beta representation**, the expected return on any asset is a linear function of its beta exposures, each multiplied by the corresponding factor risk premia.

Trade-off

- But why a trade-off?
- Let us see the following relationship in the SDF representation:

$$\text{Error} = \text{returns} \times (1 - \text{factor} \times \text{risk premia}), \tag{18}$$

that means the variance,

$$\sigma^2_{\text{Error}} = \sigma^2_{\text{returns}} - \sigma^2_{\text{returns}\times\text{factor}\times\text{risk premia}} + \text{cov's}. \tag{19}$$

Similarly, in Beta representation:

$$\text{returns} = \text{alpha} + \text{beta} \times (\text{risk premia} - \mu + \text{factor}) + \epsilon, \tag{20}$$

that means the variance,

$$\sigma^2_{\text{returns}} = \sigma^2_{\text{alpha}} + \text{beta} \times \sigma^2_{(\text{risk premia}-\mu+\text{factor})} + \sigma^2_{\epsilon} + \text{cov's}. \tag{21}$$

Literature Review

- It is a common in empirical finance to compare the performance of different econometric procedures **either** within the Beta framework, **or** within the SDF method.

- For example, Jagannathan and Wang (2002) compare the **asymptotic efficiency** of the two-stage cross-sectional regression and the Fama-MacBeth procedures with a Beta pricing model.

- Shanken and Zhou (2007), analyze the **finite sample properties** and **pricing performance** of the Fama-MacBeth, maximum likelihood, and GMM for Beta pricing models.

- Other examples are Farnsworth et al. (2002), Velu and Zhou (1999), Kan and Robotti (2008), Kan and Robotti (2009), Chen and Kan (2004), Amsler and Schmidt (1985), and **Ang et al. (2020)** just to mention a few.

The Beta Method

- The standard linear asset-pricing model under the Beta representation is given by

$$\mathsf{E}[r_t] = \mathbf{B}\delta, \tag{22}$$

where

r_t	vector of N **excess stock returns**
f_t	vector of K economy-wide **pervasive risk factors**
μ and $\boldsymbol{\Sigma}$	the **mean** and the **covariance** matrix of the factors
δ	is the vector of **factor risk premia**

and $\mathbf{B}$ is the matrix of $N \times K$ **factor loadings** which measure the exposure of asset returns to the factors, defined as

$$\mathbf{B} \equiv \mathsf{E}[r_t(f_t - \mu)']\boldsymbol{\Sigma}^{-1}. \tag{23}$$

The Beta Method (Cont.) – Higher-order Moments of Factors

- We consider that factors have higher-order moments defined by:
  - $E[f_t \otimes (f_t f_t')] = m_3$ as the **third-order uncentered co-moment tensor** of f_t (related to co-skewness, defined as $E[(f_t - \mu) \otimes ((f_t - \mu)(f_t - \mu)')] = \kappa_3$), and
  - $E[(f_t f_t') \otimes (f_t f_t')] = m_4$ as the **fourth-order uncentered co-moment tensor** of f_t (related to co-kurtosis, defined as $E[(f_t - \mu)(f_t - \mu)' \otimes ((f_t - \mu)(f_t - \mu)')] = \kappa_4$)

The Beta Method (Cont.)

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

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

- Hence, the Beta representation in equation (22) imposes the following restriction on the time-series intercept

$$\phi = (\delta - \mu)\mathbf{B}.$$

- By substituting this restriction in the regression equation, we obtain:

$$r_t = \mathbf{B}\left(\delta - \mu + f_t\right) + \epsilon_t \quad \text{where} \quad \left\{ \begin{array}{l} \mathsf{E}\left[\epsilon_t\right] = 0_N \\ \mathsf{E}\left[\epsilon_t f_t'\right] = 0_{N\times K} \end{array} \right. . \qquad (24)$$

The Beta Method (Cont.)

- Hence, the associated set of moment conditions $\mathbf{g}_b$ of the factor model are:

$$\begin{array}{lcl} \mathsf{E}\left[r_t - \mathbf{B}(\delta - \mu + f_t)\right] & = & 0_N, \\ \mathsf{E}\left[\left[r_t - \mathbf{B}(\delta - \mu + f_t)\right] f_t'\right] & = & 0_{N\times K}, \\ \mathsf{E}\left[f_t - \mu\right] & = & 0_{K\times 1}, \\ \mathsf{E}\left[\left(f_t - \mu\right)\left(f_t - \mu\right)' - \mathbf{\Sigma}\right] & = & 0_{K\times K}, \end{array} \tag{25}$$

- The corresponding unknown parameters are $\theta^* = [(\delta^*)', \text{vec}(\mathbf{B}^*)', \mu^{*\prime}, \text{vec}(\mathbf{\Sigma}^*)]'$.

- The $\mathrm{vec}(\cdot)$ operator 'vectorizes' the $\mathbf{B}^*_{N\times K}$ and the $\mathbf{\Sigma}^*$ matrices by stacking their columns.

- The observable variables are $x_t = [r_t',\ f_t']'$.

The Beta Method (Cont.) – GMM

- Then, the function g in the moment restriction is given by

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

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

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

- We choose θ to solve

$$\min_{\theta} \; g_T(\theta)' \mathbf{W}^{-1} g_T(\theta). \tag{26}$$

The Beta Method (Cont.) – GMM

- We choose θ to solve

$$\min_{\theta} g_T(\theta)'\mathbf{W}^{-1}g_T(\theta). \tag{27}$$

- We consider $\mathbf{W} = \mathbf{I}$ in equation (27).

- The second-stage GMM estimator θ_2 is the solution of equation (27) when the weighting matrix is the spectral density matrix of $g(x_t, \theta_1)$:

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

where $\mathbf{S}$ is of size $N \times N$.

The SDF Method

- To derive the SDF representation from the Beta one, we substitute the expression for $\mathbf{B}$ in (23) into equation (22) and rearrange terms:

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

  For traded factors, $\delta = \mu$ so $1 + \delta'\mathbf{\Sigma}^{-1}\mu = 1 + \mu'\ \mathbf{\Sigma}^{-1}\mu \geq 1$ and

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

- Now transform the vector of risk premia δ into a vector λ as

  $$\lambda = \frac{\mathbf{\Sigma}^{-1}\delta}{1+\delta'\mathbf{\Sigma}^{-1}\mu}, \tag{28}$$

  to obtain the following SDF representation (g_s),

  $$\mathsf{E}[r_t(1-\lambda' f_t)] = 0_N, \tag{29}$$

  where $m_t \equiv 1 - \lambda' f_t$ as $\mathsf{E}[r_t m_t] = 0_N$.

The SDF Method (Cont.) – GMM

- From the moment restrictions, equation (29), we obtain the vector of N pricing errors defined as

$$\pi = E[r_t] - E[r_t f_t']\lambda$$

- Writing the sample pricing errors as

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

  define $\mathbf{D}^U = -\frac{\partial g_T(\lambda)}{\partial \lambda'} = \mathsf{E}[r_t f_t']$, the second-moment matrix of returns and factors.

- The first-order condition to minimize the quadratic form of the sample pricing errors, equation (27), is $-\left(\mathbf{D}^U\right)' \mathbf{W}[\mathsf{E}[r_t] - \mathbf{D}^U \lambda'] = 0$, where $\mathbf{W}$ is the GMM weighting matrix of size $N \times N$.

Beta vs. SDF Method

- There is a one-to-one mapping between the factor risk premia δ and the SDF parameters λ, which facilitates the comparison of the two methods.
- Hence, we can derive an estimate of λ not only by the SDF but also by the Beta representation; by the same token we can derive an estimate of δ not only by the Beta but also by the SDF method.
- From the previous definition of λ in equation (28), we have:

$$\lambda = \delta' \left(\mathbf{\Sigma} + \delta \mu' \right)^{-1}, \qquad \text{or} \qquad \delta = \frac{\mathbf{\Sigma}\lambda}{1 - \mu'\lambda}. \tag{30}$$

- In a similar way, by substituting (30) into π, we can find a one-to-one mapping between π from the SDF and α from the Beta method:

$$\pi = \left(1 + \delta' \mathbf{\Sigma}^{-1} \mu\right)^{-1} \alpha, \qquad \text{or} \qquad \alpha = \left(1 + \delta' \mathbf{\Sigma}^{-1} \mu\right) \pi. \tag{31}$$

Single-factors Gaussian, Beta vs. SDF

Corollary (2)

Consider a single-factor linear asset pricing model (as in (22) *but* f_t *univariate), where* f_t *returns are from a portfolio of non-traded assets, and consider these returns to have higher-order moments that deviate from a Gaussian distribution, but where the first-order Delta approximation is precise (error of the approximation is non-detectable in statistical tests). Consider the returns of the factors to have stylized facts such as daily returns (negative skewness and heavy tailed). Then, when GMM is used for the estimation of the parameters (obtained under the uncentred first-stage* ($\mathbf{W} = \mathbf{I}$)*), the Beta representation has a higher efficiency than the SDF representation when measuring the risk premium.*

Single-factors Gaussian, Beta vs. SDF

Proof.: Consider that the first-order Delta method provides an accurate approximation to the asymptotic variances of the estimators, i.e. the third-, fourth- and higher-order moments in the Taylor expansions for the Delta approximation are negligible. In this case, the difference between the asymptotic variance of the SDF and the Beta methods can be expressed as,

$$Avar(\hat{\lambda}) - Avar(\lambda^*) = \frac{\sigma^2(\delta^4)}{(\sigma^2+\mu\delta)^4} + \frac{\mu^2\left(\kappa_4 - 2\kappa_3(\frac{\sigma^2-\delta^2}{\delta}) - 3\sigma^2\right)}{(\sigma^2+\mu\delta)^4}. \quad (32)$$

Consider the following conditions: $\delta < \sigma$ (volatility is higher than expected returns), $3\sigma^2 < \kappa_4$ (heavy tailed returns), and $\kappa_3 < 0$ (negatively skewed returns), then the second term of right-hand side of equation (32) is positive and the result is yield.

But why we get this systematic difference in results?

Single-factors Gaussian, Beta vs. SDF

Proof.: In this case,

$$Avar(\hat{\lambda}) - Avar(\lambda^*) = \sigma^2\delta^4 / \left(\sigma^2 + \mu\delta\right)^4 , \tag{33}$$

that is always positive. However, for the 'standard' moment values typical values of factors, this term is small.

- For example, consider single-factor models with **Gaussian returns**, where the mean and variance follow the values of the factors (for the period of the dataset), (market risk, size, value, and momentum factors).

- Then, the differences in equation (33), **proportional** to the SDF representation asymptotic variances, for the 'Gaussian' single-factor model loaded with the market risk, size, value, and momentum factors, are equal to 0.021, 0.0016, 0.0037, 0.0088 respectively (in percentage)

Asymptotic Results (Multifactor)

Theorem (1 – Multifactor)

Part A (factors on non-traded assets, risk premium): *Let f_t represent the multivariate, systematic risk factors with mean $\boldsymbol{\mu}$, covariance $\boldsymbol{\Sigma}$, third-order central moment κ_3, and fourth-order central moment κ_4, and consider the Beta representation in equation* (22) *with risk premium δ on the factors f_t, and the SDF representation in equation* (29)*. Then, the asymptotic covariance matrix of the $\hat{\lambda}$ estimators obtained under the first-stage ($\mathbf{W} = \mathbf{I}$) uncentred GMM for the SDF case is,*

$$
\begin{aligned}
Acov(\hat{\lambda}) \;&=\; \left(\left(\boldsymbol{\Sigma} + \mu\mu'\right)'\mathbf{B}'\right)^{-1}\left(\frac{1}{a_{\epsilon_t}}\boldsymbol{\Omega}^{-1} - \right. \\
&\qquad \left. \frac{1}{a_{\epsilon_t}^2}\boldsymbol{\Omega}^{-1}\mathbf{B}\left(\mathbf{A}_s^{-1} + \frac{1}{a_{\epsilon_t}}\mathbf{B}'\boldsymbol{\Omega}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}'\boldsymbol{\Omega}^{-1}\right) \times \\
&\quad \left(\mathbf{B}\left(\boldsymbol{\Sigma} + \mu\mu'\right)\right)^{-1},
\end{aligned}
$$

Asymptotic Results (Multifactor)

Theorem (1 – Multifactor (Cont.))

where

$$a_{\epsilon_t} = 1 - 2\lambda'\mu + (\lambda \odot \lambda)'\mathrm{diag}\left(\mathbf{\Sigma} + \mu\mu'\right) + \mathrm{triu\_vec}\left(\lambda\lambda'\right) + 2\left(\mathrm{triu\_vec}\left(\mathbf{\Sigma} + \mu\mu'\right)\right)$$

and $\mathbf{A}_s$ *is,*

$$\mathbf{A}_s = m_4^{reduced} + m_3^{reduced} + m_2^{reduced} + m_1^{reduced} + (\delta - \mu)(\delta - \mu)',$$

Asymptotic Results (Multifactor)

Theorem (1 – Multifactor (Cont.))

- *with* $m_2 = \mathbf{\Sigma} + \mu\mu'$ *is the second-order uncentred moment of* f_t *(*$E[f_t f_t']$*),*
- $m_{2(:,i)}$ *the* i*-th column of* m_2*,* $\odot$ *is the element wise multiplication,*
- $\mathrm{diag}(\cdot)$ *is the operator that returns the diagonal of a matrix as a vector,*
- $\mathrm{triu\_vec}(\cdot)$ *is the operator that return the upper triangular matrix (without the diagonal) in a vector form, and*
- $m_4^{reduced}, m_3^{reduced}, m_2^{reduced}$*, and* $m_1^{reduced}$ *are matrices resulting from tensor operations (see Appendix A of the paper for details).*

Asymptotic Results (Multifactor)

Theorem (1 – Multifactor (Cont.))

The asymptotic covariance matrix of the λ^* estimators obtained under the first-stage uncentred GMM estimator for the Beta case is

$$Acov(\lambda^*) = \left(|\mathbf{\Sigma}| \times \left|\left(\mathbf{\Sigma} + \mu\delta'\right)^{-1}\right| \times \left(\mathbf{\Sigma} + \mu\delta'\right)^{-1}\right)^2 V_{1,1}, \quad (34)$$

with $V_{1,1}$ a matrix of dimension $K \times K$ with the asymptotic covariance of δ (see Appendix A of the paper for details).

Tensor Moments

Definition (reduced tensor moments)

Let λ be as in equation (28). Let define the resulting reduced tensors,

$$m_4^{\text{reduced}} = \sum_i \sum_j (\lambda \odot \lambda)' \text{diag}\left(\otimes_R (m_4, i, j)\right) + 2\left(\text{triu\_vec}\left(\lambda\lambda'\right)' \text{triu\_vec}\left(\otimes_R (m_4, i, j)\right)\right),$$

$$m_3^{\text{reduced}} = \sum_i \sum_j -2\lambda' \otimes_R (m_3, i, j) + \left(\delta_j - \mu_j\right) \times \left((\lambda \odot \lambda)' \text{diag}\left(\otimes_R (m_3, i)\right) + 2\left(\text{triu\_vec}\left(\lambda\lambda'\right)' \text{triu\_vec}\left(\otimes_R (m_3, i)\right)\right)\right) + \left(\delta_i - \mu_i\right) \times \left((\lambda \odot \lambda)' \text{diag}\left(\otimes_R (m_3, j)\right) + 2\left(\text{triu\_vec}\left(\lambda\lambda'\right)' \text{triu\_vec}\left(\otimes_R (m_3, j)\right)\right)\right),$$

Tensor Moments

Definition (reduced tensor moments)

$$m_2^{\text{reduced}} = m_2 + \sum_i \sum_j -2\left(\delta_j - \mu_j\right)\left(\lambda'\left(m_{2(:,i)}\right)\right) + \left(\delta_i - \mu_i\right)\left(\lambda'\left(m_{2(:,j)}\right)\right) + \left(\delta - \mu\right)\left(\delta - \mu\right)' + 2(\lambda \odot \lambda)'\text{triu\_vec}\left(m_2\right),$$
$$m_1^{\text{reduced}} = \mu\left(\delta - \mu\right)' + \left(\delta - \mu\right)\mu' - 2\left(\delta - \mu\right)\left(\delta - \mu\right)'\left(\lambda'\mu\right),$$

where m_4^{reduced}, m_3^{reduced}, and m_2^{reduced} are matrices resulting from fourth-, third-, and second-order tensor reduction operations of the $g_s(r_t, f_t, \lambda)g_s(r_t, f_t, \lambda)'$ tensor.

Finite Sample Data and Simulation Design

- Since finite-sample analytical results can be obtained only under certain distributional assumptions of the returns, factors and errors, it is customary to resort to **simulation techniques**.
    - Interested in evaluating the standard deviations of λ^*, $\widehat{\lambda}$, π^*, $\widehat{\pi}$.
    - In the case of US, the size of the data sets that can be used are as large as 1,000 monthly historical observations, but in other cases (e.g., the UK), we may end up with as few as 300.

Simulations – DGP

- In our case, we focus in the case $\delta = \mu$. We provide additional results for factors on portfolios of non-traded assets in the Online Appendix, but we demonstrate that results in this case are similar to the traded case.

Relative Standard Errors for Fama-French 3-Factor Model

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\lambda^*)$
		Market		
60	1.1432***	1.0292***	1.1832***	1.0957***
360	1.1262***	1.0538***	1.1362***	1.0631***
600	1.1166***	1.0476***	1.1243***	1.0548***
1000	1.1191***	1.0377***	1.1263***	1.0456***
Asym	1.0460 (1st)			
		Size		
60	0.4084**	0.3901**	0.2338***	0.2128***
360	0.9521	0.8670***	0.9052***	0.7844***
600	1.0609**	0.9492**	1.0298	0.8921***
1000	1.0910***	0.9737*	1.0719***	0.9378***
Asym	1.3673 (1st)			
		Value		
60	2.2088***	1.8250***	2.7403***	3.0190***
360	2.9783***	2.3351***	3.0486***	2.5326***
600	3.0086***	2.3650***	3.0523***	2.4803***
1000	3.0798***	2.3943***	3.1119***	2.4688***
Asym	2.6219 (1st)			

Table 6: Relative standard errors of risk premium estimated from the Fama-French model: US data, 10 size-sorted portfolios.

Relative Standard Errors for Asness-Moskowitz-Pedersen Model

T	$\sigma_r(\widehat{\lambda}_1^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^U)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_1^C)/\sigma_r(\lambda^*)$	$\sigma_r(\widehat{\lambda}_2^C)/\sigma_r(\lambda^*)$
		Market		
60	1.1442***	0.9799*	1.4931***	1.3178***
360	1.4210***	1.2566***	1.7415***	1.5107***
600	1.4773***	1.2893***	1.7597***	1.5149***
1000	1.5400***	1.3293***	1.7963***	1.5370***
Asym	1.4004 (1st)			
		Momentum		
60	2.8205***	2.0929***	25.1919***	12.9843***
360	4.7186***	3.8545***	8.7920***	6.9631***
600	5.1621***	4.2124***	7.6393***	6.2049***
1000	5.4884***	4.4432***	7.0397***	5.7554***
Asym	4.6360 (1st)			
		Value		
60	1.4554***	1.0929***	5.8393***	6.3630***
360	2.9517***	2.4302***	4.8139***	4.2775***
600	3.2012***	2.6355***	4.3874***	3.7936***
1000	3.4854***	2.8286***	4.2691***	3.5995***
Asym	3.6640 (1st)			

Table 7: Relative standard errors of risk premium estimated from the Asness-Moskowitz-Pedersen model: US data, 10 size-sorted portfolios.

Specification Tests (J-test) – Size Test

	1 %				
	λ^*	$\widehat{\lambda}_1^U$	$\widehat{\lambda}_2^U$	$\widehat{\lambda}_1^C$	$\widehat{\lambda}_2^C$
			CAPM		
60	3.30	5.21	5.21	5.91	5.91
360	0.79	1.52	1.52	1.54	1.54
600	0.70	1.40	1.40	1.43	1.43
1000	0.51	1.05	1.05	1.03	1.03

Table 8: GMM estimation specification tests (size, $W = \mathbf{S}$): US data, 10 size-sorted portfolios.

Mean-Variance Estimation (Cont.)

- Using exactly the same procedure for estimation of the sample mean and sample covariance,[3] we apply the two different representations for the estimation of $\tilde{\lambda}$:
  - The **Beta estimated** ($\tilde{\lambda}^*$), and
  - The **SDF estimated** ($\tilde{\lambda}_1^U$, $\tilde{\lambda}_2^U$, $\tilde{\lambda}_1^C$, and $\tilde{\lambda}_2^C$ for the corresponding uncentred first- and second-stage, and centered first- and second-stage respectively).

[3] We use the Ledoit and Wolf (2017) shrinkage method for estimating the covariance matrix in both representations, Beta and SDF, to avoid problems with small sample covariance estimates.

Data

- We consider an expanding window, that resembles the learning process of an investor on the underlying properties on the distribution of the portfolio returns:
    - In our case, the US 10 size-sorted portfolios extracted from the Kenneth French website that spans the period January 1927 – December 2018 ($T = 1104$).
- The data of the factors (rmrf, smb,hml, and umd) correspond to the same library a time span.

Noisy vs. Smooth Expected Covariance Initialization Inference

- To add an additional layer of partial knowledge of the system, we divide our trading experiments in two:

  1. One where the investor is poorly knowledgeable of the initial properties of the returns, in particular, the expected covariance (uses initial windows of 60, 120, and 240 months) and we define this setup as a **'noisy expected covariance initialization'** model.
  2. In contrast to poor knowledge of the properties, the **'smooth expected covariance initialization'** has an initial window of 360, 480 and 600 months, and we are able to verify the OOS portfolio variance returns bounded values close to the values of the whole sample.

Results (OOS) – Smooth Initialization

Surprise! (No, Ravi's results are great again!)

Model difference	Initial window size		
	360	480	600
	CAPM		
SR^*	5.97	8.21	8.70
$SR^* - SR_1^U$	-6.73×10^{-10}	-9.91×10^{-10}	8.54×10^{-11}
$SR^* - SR_2^U$	-1.37×10^{-9}	-1.33×10^{-9}	-2.37×10^{-9}
$SR^* - SR_1^C$	-1.82×10^{-9}	-3.50×10^{-9}	-1.68×10^{-9}
$SR^* - SR_2^C$	-2.73×10^{-9}	-3.32×10^{-9}	-2.25×10^{-9}

Table 9: Sharpe ratio of mean-variance optimal portfolios with parameters estimated with Beta and SDF representations (smooth expected covariance initialization model).

Results (OOS) – Noisy Initialization

Model difference	Initial window size		
	60	120	240
	Fama-French		
SR^*	14.63	14.06	17.11
$SR^* - SR_1^U$	−1.542	−1.729	−2.006
$SR^* - SR_2^U$	−1.818	−0.721	−0.617
$SR^* - SR_1^C$	−1.485	−1.681	−1.953
$SR^* - SR_2^C$	−1.395	−0.672	−0.584
	Asness-Moskowitz-Pedersen		
SR^*	23.13	25.61	30.91
$SR^* - SR_1^U$	6.235	8.596	10.390
$SR^* - SR_2^U$	2.980	5.481	6.845
$SR^* - SR_1^C$	8.500	9.365	11.469
$SR^* - SR_2^C$	2.828	6.003	7.303
	Carhart		
SR^*	25.74	25.66	31.09
$SR^* - SR_1^U$	7.411	5.357	5.979
$SR^* - SR_2^U$	3.370	4.330	5.855
$SR^* - SR_1^C$	7.912	5.417	6.149
$SR^* - SR_2^C$	5.763	4.541	6.263

Table 10: Sharpe ratio of mean-variance optimal portfolios with parameters estimated with Beta and SDF representations (noisy expected covariance initialization model).

Convergence of Results – Market Factor

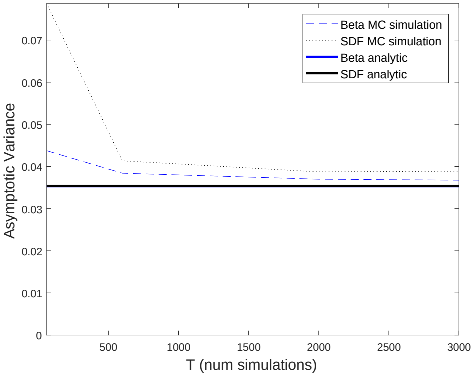

Figure 4: Asymptotic variance of the analytically and empirically estimated GMM under the Beta and SDF methods, from a set of 10,000 Monte Carlo simulation based on parameters calibrated to the observed market risk, and size factors on a sample January 1927 – December 2018. Data are downloaded from Kenneth French's library.

Convergence of Results – Value

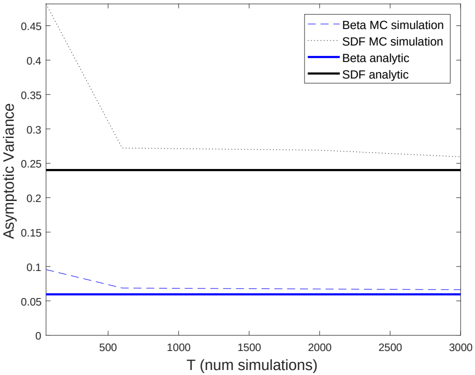

Figure 5: Asymptotic variance of the analytically and empirically estimated GMM under the Beta and SDF methods, from a set of 10,000 Monte Carlo simulation based on parameters calibrated to the observed market risk, and size factors on a sample January 1927 – December 2018. Data are downloaded from Kenneth French's library.

Convergence of Results – Size

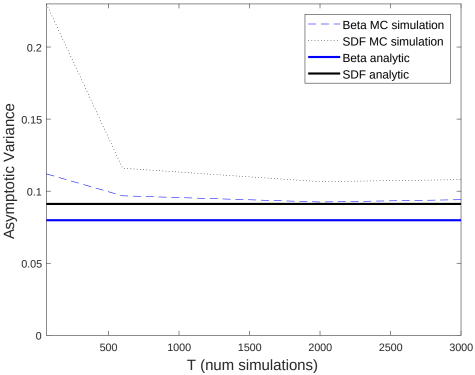

Figure 6: Asymptotic variance of the analytically and empirically estimated GMM under the Beta and SDF methods, from a set of 10,000 Monte Carlo simulation based on parameters calibrated to the observed market risk, and size factors on a sample January 1927 – December 2018. Data are downloaded from Kenneth French's library.

Convergence of Results – Momentum

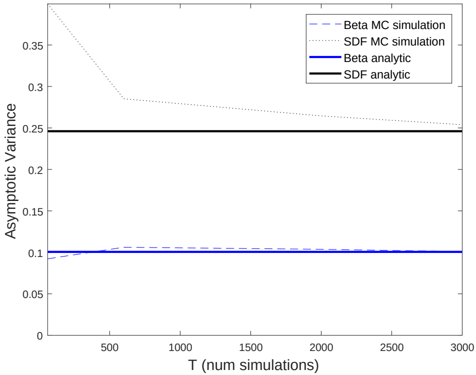

Figure 7: Asymptotic variance of the analytically and empirically estimated GMM under the Beta and SDF methods, from a set of 10,000 Monte Carlo simulation based on parameters calibrated to the observed market risk, and size factors on a sample January 1927 – December 2018. Data are downloaded from Kenneth French's library.

Momentum Factor and Second-order Moments

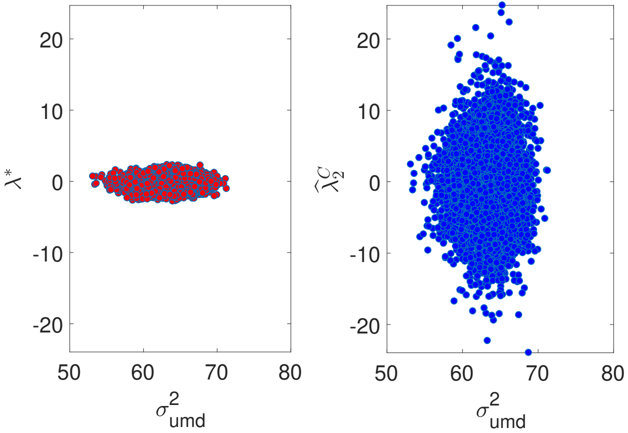

Figure 8: Variance of momentum factor vs. estimated risk premium (λ^* and $\widehat{\lambda}_2^C$).

Momentum Factor and Third- and Fourth-order Moments

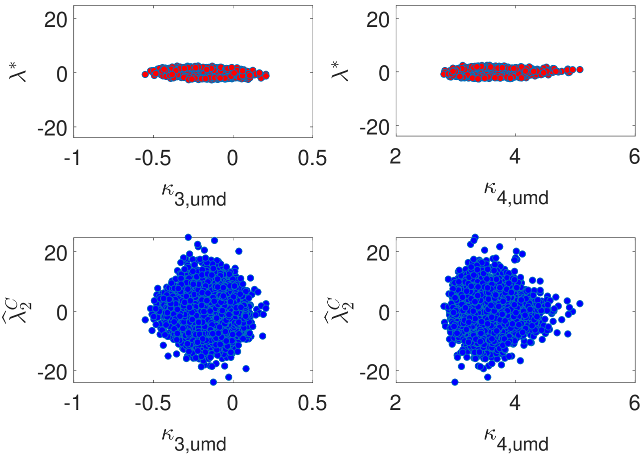

Figure 9: Skewness and kurtosis of momentum factor vs. estimated risk premium (λ^* and $\widehat{\lambda}_2^C$).