A Unifying Approach to the Empirical Evaluation of Asset Pricing Models

Francisco Peñaranda
Enrique Sentana

July 2010

Barcelona Economics Working Paper Series

Working Paper nº 488

A Unifying Approach to the Empirical Evaluation of Asset Pricing Models[*]

Francisco Peñaranda

UPF, Ramón Trias Fargas 25-27, 08005 Barcelona, Spain.

<francisco.penaranda@upf.edu>

Enrique Sentana

CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain.

<sentana@cemfi.es>

Revised: July 2010

Abstract

Two main approaches are commonly used to empirically evaluate linear factor pricing models: regression and SDF methods, with centred and uncentred versions of the latter. We show that unlike standard two-step or iterated GMM procedures, single-step estimators such as continuously updated GMM yield numerically identical values for prices of risk, pricing errors, Jensen's alphas and overidentifying restrictions tests irrespective of the model validity. Therefore, there is arguably a single approach regardless of the factors being traded or not, or the use of excess or gross returns. We illustrate our results by revisiting Lustig and Verdelhan's (2007).empirical analysis of currency returns.

Keywords: CU-GMM, Factor pricing models, Forward premium puzzle, Generalised Empirical Likelihood, Stochastic discount factor.

JEL: G11, G12, C12, C13.

1   Introduction

Asset pricing theories are concerned with determining the expected returns of assets whose payoffs are risky. Specifically, these models analyse the relationship between risk and expected returns, and address the crucial question of how to value risk. The most popular empirically oriented asset pricing models effectively assume the existence of a common stochastic discount factor (SDF) that is a linear function of some risk factors, which discounts uncertain payoffs differently across different states of the world. Those factors can be either the excess returns on some traded securities, as in the traditional CAPM of Sharpe (1964), Lintner (1965) and Mossin (1966) or the so-called Fama and French (1993) model, non-traded economy wide sources of uncertainty related to macroeconomic variables, like in the Consumption CAPM (CCAPM) of Breeden (1979), Lucas (1978) or Rubinstein (1976), or a combination of the two, as in the exact version of Ross' (1976) APT.

There are two main approaches to formally evaluate linear factor pricing models from an empirical point of view. The traditional method relies on regressions of excess returns on factors, and exploits the fact that an asset pricing model imposes certain testable constraints on the intercepts. More recent methods rely on the SDF representation of the model instead, and exploit the fact that the corresponding pricing errors should be zero. There are in fact two variants of the SDF method, one that demeans the factors and another one that does not. Although the initial asset pricing tests tended to make the assumption that asset returns and factors were independently and identically distributed as a multivariate normal vector, nowadays these approaches are often implemented by means of the generalized method of moments (GMM) of Hansen (1982), which has the advantage of yielding valid inferences even if the assumptions of serial independence, conditional homoskedasticity or normality are not totally realistic in practice (see Campbell, Lo and MacKinlay (1996) or Cochrane (2001a) for advanced textbook treatments).

Unfortunately, though, each approach typically yields different estimates of prices of risk, pricing errors and Jensen's alphas, and different values for the overidentifying restrictions test. This begs the question of which approach is best, and there is some controversy surrounding the answer. For example, Kan and Zhou (2000) advocated the use of the regression method over the uncentred SDF method because the former provides more reliable risk premia estimators and more powerful specification tests than the latter. However, Cochrane (2001b) and Jagannathan and Wang (2002) criticised their conclusions on the grounds that they did not consider the estimation of factor means and variances. Specifically, Jagannathan and Wang (2002) showed that once one takes the sampling uncertainty of those estimators into account, the regression

approach and the (uncentred) SDF approach are asymptotically equivalent in terms of standard errors and power if excess returns and factors are independently and identically distributed as a multivariate normal random vector.

More recently, Burnside (2007) and Kan and Robotti (2008) have also pointed out that in certain cases there may be dramatic differences between the results obtained by applying standard two-step or iterated GMM procedures to the centred and uncentred versions of the SDF approach. At the same time, Kan and Robotti (2008) prove the numerical equivalence of the continuously updated GMM (CU-GMM) of Hansen, Heaton and Yaron (1996) implementation of the overidentification tests associated to the centred and uncentred versions of the SDF approach (see their footnote 3). As is well known, CU-GMM is a single-step method that integrates the heteroskedasticity and autocorrelation consistent (HAC) estimator of the long-run covariance matrix as part of the objective function.

In this context, the main contribution of our paper is to show the more subtle result that the application to both the regression and SDF approaches of single-step GMM methods, including CU-GMM, gives rise to numerically identical estimates of prices of risk, pricing errors, Jensen's alphas and overidentifying restrictions tests irrespective of the validity of the asset pricing model.

Therefore, one could argue that there is effectively a single method to empirically evaluate asset-pricing models. Although the rationale for our results is the well-known functional invariance of maximum likelihood estimators, our results do not depend on any distributional assumption, and remain true regardless of whether or not the asset pricing restrictions hold, the factors are traded or non-traded, or the researcher works with excess returns or gross returns. To ease the exposition, we centre our discussion on models with a single priced factor. Nevertheless, our numerical equivalence results do not depend in any way on this simplification. In fact, the proofs of our main results explicitly consider the general multifactor case. Similarly, our empirical application includes both single and multifactor models.

Importantly, our results apply to optimal GMM inference procedures. In particular, we do not consider sequential GMM methods that fix the factor means to their sample counterparts. We do not consider either procedures that use alternative weighting matrices such as the uncentred second moment of returns. Although the choice in Hansen and Jagannathan (1997) is reasonable in SDF contexts when one wants to compare several misspecified models, it does not have a natural regression counterpart. In any case, those suboptimal GMM estimators fall outside the realm of single-step methods such as CU-GMM, and therefore they would typically give rise to numerically different statistics.

While single-step methods are not widespread in empirical finance applications, this situation

is likely to change in the future because they do not require to arbitrarily choose the number of iterations, and they often yield more reliable inferences in finite samples.[1] In fact, the recent papers by Julliard and Gosh (2008), Almeida and Garcia (2009) or Campbell, Gilgio and Polk (2010) attest their increasing popularity. However, the CU-GMM estimator and other single-step methods such as empirical likelihood or exponentially-tilted methods are often more difficult to compute than two-step estimators, particularly in linear models. For that reason, we also develop simple, intuitive consistent parameter estimators that can be used to obtain good initial values, and which will be efficient for elliptically distributed returns and factors.

We illustrate our results by revisiting the empirical analysis of currency returns in Lustig and Verdelhan (2007). Our findings confirm that the conflict among criteria for testing asset pricing models that we have previously mentioned is not only a theoretical possibility, but a hard reality. Nevertheless, such a conflict disappears when one uses single-step methods. A different issue, though, is the interpretation of the restrictions that are effectively tested. In this sense, our results also confirm Burnside's (2007) findings that the favourable empirical evidence on the asset pricing models with non-traded factors that Lustig and Verdelhan (2007) considered should be interpreted with some care because their consumption-based pricing factors seem to be uncorrelated to currency returns. We also find that if we force the models with traded factors to price those factors, and in particular the US market portfolio, then we reject the asset pricing restrictions.

The rest of the paper is organised as follows. We study the case of traded factors from the theoretical and empirical perspectives in section 2, while in section 3 we analyse the case of non-traded factors in an analogous manner. In section 4 we show that our results hold not only for excess returns but also for gross returns. Finally, we summarise our conclusions in section 5. A brief description of CU-GMM as an example of single-step methods, multifactor models and formal proofs are gathered in appendices A, B and C, respectively.

2 Traded factors

Let $\mathbf{r}$ be an $n \times 1$ vector of excess returns. Standard arguments such as lack of arbitrage opportunities or the first order conditions of a representative investor imply that

$$E\left(m\mathbf{r}\right) = \mathbf{0}$$

for some random variable m called SDF, which discounts uncertain payoffs in such a way that their expected discounted value equals their cost.

---

[1] Newey and Smith (2004) highlight the finite sample advantages of CU and other generalised empirical likelihood estimators over two-step GMM by going beyond the usual first-order asymptotic equivalence results.

The standard approach in empirical finance is to model m as an affine transformation of some risk factors, even though this ignores that m must be positive with probability 1 to avoid arbitrage opportunities (see Hansen and Jagannathan (1991)). With a single factor f, we can express the pricing equation as

$$E\left[(a + bf)\,\mathbf{r}\right] = \mathbf{0} \tag{1}$$

for some real numbers (a, b).

Although $\mathbf{r}$ only contains assets with 0 cost, which leaves the *scale* and *sign* of m undetermined, we would like our candidate SDF to price other assets with positive prices. Therefore, we require a *scale normalisation* to rule out the trivial solution $(a, b) = (0, 0)$ (see Cochrane (2001a), pp. 256-258). For example, we could choose the popular asymmetric normalisations $a = 1$ or $E(m) = a + b\mu = 1$, where $\mu = E(f)$. Alternatively, we could choose the symmetric normalisation $a^2 + b^2 = 1$, together with a sign restriction on one of the nonzero coefficients. As we shall see below, this seemingly innocuous issue may have important empirical consequences.

In this section we assume that the pricing factor f is itself the excess return on another asset, such as the market portfolio in the CAPM.[2] As forcefully argued by Farnsworth et al. (2002) and Lewellen, Nagel and Shanken (2009) among others, the pricing model applies to f too, which means that

$$E\left[(a + bf)\,f\right] = 0. \tag{2}$$

Following Chamberlain (1983b) we also know that $a + bf$ will constitute an admissible SDF if and only if f lies on the mean-variance frontier generated by f and $\mathbf{r}$. Then, the well-known properties of mean-variance frontiers imply that the least squares projection of $\mathbf{r}$ onto a constant and f should be proportional to f. As result, we can equivalently write the above pricing equations as

$$E\left[\begin{pmatrix} 1 \\ f \end{pmatrix}(\mathbf{r} - \boldsymbol{\beta}f)\right] = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \tag{3}$$

where the parameter vector $\boldsymbol{\beta}$ represents the slopes of the projection.

2.1 Existing approaches

Equations (1) and (2) are particularly amenable to GMM estimation once we choose a normalisation for (a, b). As we mentioned before, there are two widespread asymmetric choices in empirical finance: $a = 1$ and $a + b\mu = 1$, with the corresponding SDFs typically expressed as $m = 1 - f\lambda$ and $m = 1 - (f - \mu)\tau$, respectively. Kan and Zhou (2000), Cochrane (2001b) and

---

[2]It is important to mention that our assets could include managed portfolios. Similary, the factor could also be a scaled version of a primitive excess return to accommodate conditioning information; see the discussion in chapter 8 of Cochrane (2001a).

Jagannathan and Wang (2002) only study the first variant, but the second one is also widely used in the literature (see e.g. Parker and Julliard (2005) or Yogo (2006)). Burnside (2007) refers to the first approach as the a-normalisation and to the second one as the ξ-normalisation. We will refer to them instead as the uncentred and centred SDF parametrisations since they differ in their use of either $E(\mathbf{r}f)$ or $Cov(\mathbf{r}, f)$ in explaining the cross-section of risk premia.

Specifically, the uncentred SDF test relies on the overidentified $n + 1$ moment conditions

$$E \begin{bmatrix} \mathbf{r}(1 - f\lambda) \\ f(1 - f\lambda) \end{bmatrix} = \mathbf{0}, \tag{4}$$

where the only unknown parameter is λ. Given that the last moment condition implies that $E(f) = E(f^2)\lambda$, we will have that

$$\lambda = \frac{\mu}{\gamma}, \tag{5}$$

where γ is the second moment of the factor, which allows us to interpret λ as a "price of risk" for the factor. Under standard regularity conditions, the overidentifying restrictions (J) test will follow an asymptotic chi-square distribution with n degrees of freedom (χ_n^2) when (4) is correctly specified.

In contrast, the centred SDF test relies on the $n + 2$ overidentified moment conditions

$$E \begin{bmatrix} \mathbf{r}(1 - (f - \mu)\tau) \\ f(1 - (f - \mu)\tau) \\ f - \mu \end{bmatrix} = \mathbf{0}, \tag{6}$$

where the additional moment condition allows the estimation of μ. The two unknown parameters are (μ, τ). Once again, we can use the middle moment condition to show that:

$$\tau = \frac{\mu}{\sigma^2}, \tag{7}$$

where $\sigma^2 = \gamma - \mu^2$ denotes the variance of f, which means that τ also has a "price of risk" interpretation. Not surprisingly, the corresponding J test also converges in distribution to a χ_n^2 under correct specification.

The regression (or beta) representation of the pricing model is also amenable to GMM estimation. In particular, we can follow MacKinlay and Richardson (1991) in regarding (3) as $2n$ overidentified moment conditions, where the n unknown parameters are the slope coefficients $\boldsymbol{\beta}$, which under the null coincide with both $E(\mathbf{r}f)/E(f^2)$ and $Cov(\mathbf{r}, f)/V(f)$. Therefore, the J test will be asymptotically distributed as χ_n^2 under the null.

The regression method identifies μ with the expected excess return of a portfolio whose "beta" is equal to 1. Hence, this parameter may also be interpreted as an alternative "price of risk". To estimate it, we can add $f - \mu$ to (3) as in (6), and simultaneously obtain $\boldsymbol{\beta}$ and μ.

The overidentification tests are regularly complemented by three standard evaluation measures, which correspond to the value of the different moment conditions when the linear factor pricing model is incorrect. In this way, we define Jensen's alpha:

$$\boldsymbol{\alpha} = E\left(\mathbf{r}\right) - \boldsymbol{\beta}E\left(f\right) \tag{8}$$

for the regression method, the pricing errors obtained from the uncentred SDF representation:

$$\boldsymbol{\pi} = E\left(\mathbf{r}\right) - E\left(\mathbf{r}f\right)\lambda,$$

or from the centred SDF representation:

$$\boldsymbol{\psi} = E\left(\mathbf{r}\right) - E\left(\mathbf{r}\left(f - \mu\right)\right)\tau.$$

Under the null hypothesis these three measures should be simultaneously 0, but otherwise their values will be different.

2.2 Numerical equivalence results

As we mentioned in the introduction, Kan and Zhou (2000), Cochrane (2001b), Jagannathan and Wang (2002), Burnside (2007) and Kan and Robotti (2008) compare some of the aforementioned approaches when researchers rely on traditional, two-step or iterated GMM procedures.

In contrast, our main result is that all the aforementioned methods coincide if one uses instead single-step procedures such as CU-GMM. More formally

Proposition 1 *If we apply single-step procedures to the uncentred SDF method based on the moment conditions (4), the centred SDF method based on the moment conditions (6), and the regression method based on the moment conditions (3), then for a common choice of HAC weighting matrix the following numerical equivalences hold:*
1) The three overidentification restrictions (J) tests.
2) The direct estimate of the "price of risk" λ from (4), the indirect estimate

$$\lambda = \frac{\sigma^2}{\gamma}\tau,$$

from (6) extended to include γ, and the indirect estimate

$$\lambda = \frac{\mu}{\gamma}$$

from (3) extended to include (μ, γ). Analogous results apply to τ and μ.
3) The estimates of Jensen's alphas in (8) obtained by replacing $E\left(\cdot\right)$ by an unrestricted sample average and $\boldsymbol{\beta}$ by their direct estimates obtained from the regression method, or the indirect estimates obtained from SDF methods extended to include $\boldsymbol{\beta}$. Analogous results apply to the alternative pricing errors $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$.

In order to provide some intuition for the numerical equivalence between all the CU-GMM J tests that we have discussed, imagine that for estimation purposes we assumed that the joint distribution of $\mathbf{r}$ and f is $i.i.d.$ multivariate normal. In that context, we could test the mean-variance efficiency of f by means of a likelihood ratio (LR) test. We could then factorise the joint log-likelihood function of $\mathbf{r}$ and f as the marginal log-likelihood of f, whose parameters μ and σ^2 would be unrestricted, and the conditional log-likelihood of $\mathbf{r}$ given f. As a result, the LR version of the original Gibbons, Ross and Shanken (1989) test would be numerically identical to the LR test in the joint system irrespective of the chosen parametrisation. The CU-GMM overidentification test, which implicitly uses the Gaussian scores as influence functions, inherits the invariance of the LR test. The advantage, though, is that we can make it robust to departures from normality, serial independence or conditional homoskedasticity.

Alternatively, we could understand the CU-GMM procedure as being based on the original moment conditions (1) and (2), which are valid for any normalisation of the SDF scale of the form $a^2 + b^2 \neq 0$. In this light, the SDF result in footnote 3 of Kan and Robotti (2008) is a direct consequence of the fact that single-step procedures are numerical invariant to normalisation, while the additional, less immediate results relating the regression and SDF approaches in Proposition 1 follow from the fact that those GMM procedures are also invariant to parameter dependent linear transformations of the moments and reparametrisations(see Appendix A).

One drawback of CU-GMM procedures, though, is that they induce a non-linearity in the GMM objective function. In this sense, the numerical equivalence of the three approaches gives a non-trivial computational advantage to the uncentred SDF method, which only contains a single unknown parameter.[3]

Still, it is convenient to have consistent initial parameter values. For that reason, we propose a computationally simple intuitive estimator that is always consistent, but which would become efficient for i.i.d. elliptical returns, a popular assumption in finance because it guarantees the compatibility of mean-variance preferences with expected utility maximisation regardless of investors' preferences (see Chamberlain (1983a) and Owen and Rabinovitch (1983)). For the uncentred SDF model in particular:

Lemma 1 *If $(\mathbf{r}_t, f_t)$ is an i.i.d. elliptical random vector with bounded fourth moments, and the null hypothesis of linear factor pricing holds, then the most efficient estimator of λ obtained from (4) will be given by*

$$\dot{\lambda}_T = \frac{\sum_{t=1}^{T} f_t}{\sum_{t=1}^{T} f_t^2}. \tag{9}$$

---

[3]This advantage becomes more relevant as the number of factors k increases because the centred SDF method requires the additional estimation of k factor means, while the regression method the estimation of $n \times k$ factor loadings.

Intuitively, this means that under those circumstances (2), which is the moment involving f, exactly identifies the parameter λ, while (1), which are the moments corresponding to $\mathbf{r}$, provide the n overidentification restrictions to test. Although the elliptical family is rather broad (see e.g. Fang, Kotz and Ng (1990)), and includes the multivariate normal and Student t distribution as special cases, it is important to stress that $\dot{\lambda}_T$ will remain consistent under linear factor pricing even if the assumptions of serial independence and ellipticity are not totally realistic in practice.

2.3 Empirical application

Over the last thirty years most empirical studies have rejected the hypothesis of uncovered interest parity, which in its basic form implies that the expected return to speculation in the forward foreign exchange market conditioned on available information should be zero. Specifically, many of those studies find support for the so-called the "forward premium puzzle", which implies that, contrary to the theory, high domestic interest rates relative to those in the foreign country predict a future appreciation of the home currency. In fact, the so-called "carry trade", which involves borrowing low-interest-rate currencies and investing in high-interest-rate ones, constitutes a very popular currency speculation strategy developed by financial market practitioners to exploit this "anomaly" (see Burnside et al. (2006)).

One of the most popular explanations among economists is that such a seemingly anomalous pattern might reflect a reward to the exposure of foreign currency positions to certain systematic risk factors. To study this possibility, Lustig and Verdelhan (2007) constructed eight portfolios of currencies sorted at the end of the previous year by their nominal interest rate differential to the US dollar, creating in this way annual excess returns (in real terms) on foreign T-Bill investments for a US investor over the period 1953-2002. Then they tested if some popular empirical asset pricing models that rely on certain domestic US risk factors were able to explain the cross-section of risk premia.

Table 1 reports some simple descriptive statistics for those portfolios. Interestingly, the broadly monotonic relationship between the level of interest rates differentials and risk premia provides informal evidence on the failure of uncovered interest rate parity.

(TABLE 1)

Given that so far we have only considered traded factors in our theoretical analysis, we initially focus on the CAPM and the three-factor Fama-French model. Following Lustig and Verdelhan (2007), in the first case we take the pricing factor to be the US market portfolio

8

(MK), which we identify with the CRSP value-weighted excess return. In contrast, in the second model we add two additional factors labelled SMB and HML that capture the so-called size and value effects. Specifically, SMB reflects the returns on small cap firms in excess of the returns on large cap firms, while HML measures the returns on value firms in excess of the returns on growth firms.

Table 2 contains the results of applying the different inference procedures previously discussed to this data set. In all cases, we estimate the asymptotic covariance matrix of the relevant influence functions by means of its sample counterpart, as in Hansen, Heaton and Yaron (1996), except for the first-step estimators, for which we use the identity matrix as initial weighting matrix.[4] The first thing to note is that while the two-step and CU-GMM estimates of λ and μ are fairly similar, the CU-GMM estimate of τ is higher than its two-step counterpart, although the t-ratio is lower. In addition, there are marked numerical differences between the three two-step versions of the overidentification restrictions tests. In particular, the centred SDF yields a much higher J statistic. Nevertheless, all tests reject the null hypothesis of linear factor pricing, including the single CU-GMM statistic that we have introduced.

(TABLE 2)

Panel B of Table 2 presents the different pricing errors, with the α of a traded factor being computed as the difference between its sample mean and its GMM estimate. As can be seen, there are very few statistically significant values of $\boldsymbol{\alpha}$, $\boldsymbol{\pi}$ or $\boldsymbol{\psi}$ for the 8 currency portfolios in the two GMM versions. Only the CU-GMM estimates of α for the 7th portfolio are significant at the 1% level, while the corresponding two-step GMM estimates are significant at the 5%. In contrast, all three pricing errors for MK are highly significant. Therefore, in this empirical application the failure of the CAPM to price the market portfolio provides the clearest source of model rejection, thereby confirming the relevance of the recommendation in Farnsworth et al. (2002) and Lewellen, Nagel and Shanken (2009).

Finally, Table 3 extends the analysis to the multifactor model of Fama-French. Once again, there are marked numerical differences between the three two-step approaches (see Panel A). Specifically, the estimates of $\boldsymbol{\lambda}$, $\boldsymbol{\tau}$ and $\boldsymbol{\mu}$ differ between two-step and single-step methods, especially in the case of the centred SDF. As for the overidentification tests, the centred SDF also yields a much higher J statistic. Nevertheless, the three two-step GMM approaches reject the null hypothesis of linear factor pricing, and CU-GMM confirms the result.

---

[4]We have also considered another two-step GMM procedure that uses (9) as the first-step estimator, but the results are qualitatively similar.

(TABLE 3)

Panel B of Table 3 displays the pricing errors. Once again, there are very few statistically significant values of $\boldsymbol{\alpha}$, $\boldsymbol{\pi}$ or $\boldsymbol{\psi}$ for the 8 currency portfolios in the two GMM versions. The main exception is the 7th portfolio, whose two-step GMM pricing errors are significant at the 1% level. As in the case of the CAPM, the main reason behind the rejection seems to be the failure of the model to price the factors, and in the particular the market portfolio.

3 Non-traded factors

Let us now consider a situation in which f is a scalar non-traded factor, such as the growth rate of per capita consumption. The main difference with the analysis in section 2 is that the factor may not satisfy the pricing equation (2), so that the SDF is simply defined by (1).

As in the case of a traded factor, we can equivalently write this pricing condition as a restriction on the least squares projection of $\mathbf{r}$ onto a constant and f. Specifically, if we define

$$\boldsymbol{\phi} = E(\mathbf{r}) - \boldsymbol{\beta}\mu$$

and

$$\boldsymbol{\beta} = Cov\left(\mathbf{r},f\right)/\sigma^2 \tag{10}$$

as the vectors of intercepts and slopes in that projection, respectively, then (1) is equivalent to

$$E\left[(a + bf)\left(\boldsymbol{\phi} + \boldsymbol{\beta}f\right)\right] = c\boldsymbol{\phi} + d\boldsymbol{\beta} = \mathbf{0}, \tag{11}$$

where c and d are two scalars not simultaneously equal to 0, so that the projection and the SDF should be orthogonal. Intuitively, (1) implies that we can find a non-trivial linear combination of $E(\mathbf{r})$ and $E(\mathbf{r}f)$ (or $Cov(\mathbf{r}, f)$) that is zero, which in turn implies that we can find a non-trivial linear combination of $\boldsymbol{\phi}$ and $\boldsymbol{\beta}$ that is zero too.[5]

In practice, we can easily impose the implicit constraint (11) by writing the moment conditions that define the projection as

$$E\left[\begin{pmatrix} 1 \\ f \end{pmatrix}\left(\mathbf{r} - \boldsymbol{\varphi}\left(d - cf\right)\right)\right] = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \tag{12}$$

where $\boldsymbol{\varphi}$ is some $n \times 1$ vector such that $\boldsymbol{\phi} = d\boldsymbol{\varphi}$ and $\boldsymbol{\beta} = -c\boldsymbol{\varphi}$. This moment condition closely resembles (3), except for the fact that when f is traded, the additional condition (2) implies that $\boldsymbol{\phi} = \mathbf{0}$ and $d = 0$. More generally, we need to solve the scale indeterminacy of c and d by

---

[5]Note that $c = a + b\mu$ is the mean of the SDF, while $d = a\mu + b\gamma$ would be the price of f if it was traded.

choosing either the popular asymmetric normalisation $c = -1$, so that $\boldsymbol{\beta} = \boldsymbol{\varphi}$, or the symmetric normalisation $c^2 + d^2 = 1$. As we shall see below, this seemingly innocuous issue may once again have important empirical consequences.

3.1 Existing approaches

As we mentioned in the case of traded factors, some normalisation is required to identify (a, b) from (1). Specifically, the uncentred SDF approach implicitly sets $a = 1$, and relies on the n overidentified moment conditions

$$E \left[\ \mathbf{r} \left(1 - f\lambda \right) \ \right] = \mathbf{0}, \tag{13}$$

where the only parameter to estimate is λ. This parameter still has a "price of risk" interpretation, as in (5), but this time in terms of factor mimicking portfolios. In particular

$$\lambda = \frac{E(r^+)}{E(r^{+2})}, \tag{14}$$

where

$$r^+ = E(f\mathbf{r}')E^{-1}(\mathbf{rr}')\mathbf{r} \tag{15}$$

is the uncentred least squares projection of f on $\mathbf{r}$. As expected, the asymptotic distribution of the J test will be χ^2_{n-1} under the null. For that reason, in what follows we assume that the number of assets exceeds the number of factors to ensure that the linear factor pricing model imposes testable restrictions on asset returns.

In contrast, the centred SDF test implicitly sets $a + b\mu = 1$, and relies on the $n + 1$ overidentified moment conditions

$$E \begin{bmatrix} \mathbf{r} \left(1 - (f - \mu)\tau \right) \\ f - \mu \end{bmatrix} = \mathbf{0}, \tag{16}$$

where the two parameters to estimate are (μ, τ). Once again, τ has a "price of risk" interpretation as in (7), but in terms of factor mimicking portfolios. Specifically

$$\tau = \frac{E(r^{++})}{V(r^{++})}, \tag{17}$$

where

$$r^{++} = cov(f, \mathbf{r}')V^{-1}(\mathbf{r})\mathbf{r}$$

is the centred least squares projection of f on $\mathbf{r}$. The link to the uncentred "price of risk" parameter is simply

$$\tau = \frac{\lambda}{1 - \lambda\mu},$$

which effectively divides λ by the SDF mean since the centred SDF approach normalises that mean to 1. As before, the J test will asymptotically converge to a χ^2_{n-1} when the asset pricing model is correct.

On the other hand, if we normalise $c = -1$ then we can write (12) as the following $2n$ overidentified moment conditions:

$$E\left[\begin{pmatrix} 1 \\ f \end{pmatrix}(\mathbf{r} - \boldsymbol{\beta}(\varkappa + f))\right] = \mathbf{0}, \tag{18}$$

where the $n+1$ parameters to estimate are $(\varkappa, \boldsymbol{\beta})$, with $\varkappa = \delta - \mu$, so that we can interpret δ as the expected excess return of a portfolio whose "beta" is equal to 1.[6] Given that the risk premium of r^{++} is $\tau V(r^{++})$ and its beta is $V(r^{++})/\sigma^2$, there is again a simple connection to the "price or risk" of the centred SDF. Specifically,

$$\delta = \sigma^2 \tau,$$

which simply re-scales τ by the factor variance. This is the usual regression (or beta) test of the pricing model, which implicitly exploits the restrictions that the model imposes on the regression intercepts (see Campbell, Lo and MacKinlay (1996, chap. 5)).

Finally, the centred and uncentred pricing errors $\boldsymbol{\psi}$ and $\boldsymbol{\pi}$ are defined as in section 2, while Jensen's alpha is now defined as

$$\boldsymbol{\alpha} = E(\mathbf{r}) - \boldsymbol{\beta}\delta. \tag{19}$$

Unfortunately, the existing approaches may run into difficulties in those cases in which their implicit normalisations are invalid. For instance, if $E(\mathbf{r}f) = \mathbf{0}$ so that the true SDF would be proportional to f and $r^+ = 0$, then the normalisation of (1) with $a = 1$ in the uncentred SDF approach (13) is not well defined in population terms. Similarly, if $Cov(\mathbf{r}, f) = \mathbf{0}$ so that the true SDF would be proportional to $(f - \mu)$ and $r^{++} = 0$, then neither the normalisation of (1) with $a + b\mu = 1$ in the centred SDF approach (16), nor the normalisation of (12) with $c = -1$ in the centred regression approach (18) are properly defined. In contrast, the symmetric normalisations $a^2 + b^2 = 1$ and $c^2 + d^2 = 1$ continue to be well defined in those circumstances. We shall return to this issue in the empirical application.

3.2 Numerical equivalence results

As in the case of traded factors, we can show that all the approaches discussed in the previous subsection coincide if one uses single-step methods such as CU-GMM. More formally

---

[6]Jagannathan and Wang (2002) use $\delta - \mu$ instead of $\varkappa$, and add the influence functions $f - \mu$ and $(f - \mu)^2 - \sigma^2$ to estimate μ and σ^2 too. The addition of these moments is irrelevant for the estimation of $\varkappa$ and the J test because they exactly identify μ and σ^2. Consequently, we will ignore them for the time being to simplify the exposition, although we will use them in our proofs to link the regression and SDF approaches.

Proposition 2 *If we apply single-step procedures to the uncentred SDF method based on the moment conditions (13), the centred SDF method based on the moment conditions (16), and the regression method based on the moment conditions (18), then for a common choice of HAC weighting matrix the following numerical equivalences hold:*
1) The three overidentification restrictions (J) tests.
2) The direct estimate of the "price of risk" λ from (13), the indirect estimate

$$\lambda = \frac{\tau}{1 + \tau\mu}$$

from (16), and the indirect estimate of

$$\lambda = \frac{\delta}{\sigma^2 + \delta\mu}$$

from (18) parametrised in terms of δ and extended to include γ, provided that $1 + \tau\mu = \left(\sigma^2 + \delta\mu\right)/\sigma^2 \neq 0$. Analogous results apply to τ and δ.
3) The estimates of Jensen's alphas in (19) obtained by replacing $E(\cdot)$ by an unrestricted sample average and $\boldsymbol{\beta}$ and δ by their direct estimates obtained from the regression method or the indirect estimates obtained from SDF methods extended to include $\boldsymbol{\beta}$, γ and μ. Analogous results apply to the pricing errors $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$.

It is important to distinguish this proposition from the results in Jagannathan and Wang (2002). These authors showed that if $(\mathbf{r}, f)$ were i.i.d. normal, then the centred regression and uncentred SDF approaches would lead to asymptotically equivalent inferences. Although it would be fairly straightforward to show that their results would continue to be true under more general distributional assumptions, Proposition 2 shows that in fact both SDF approaches and the regression method yield numerically identical conclusions if we work with one-step GMM procedures such as CU-GMM.[7]

As we mentioned before, though, one drawback of CU-GMM is that it induces a non-linearity in the GMM objective function. In this sense, the numerical equivalence of the three approaches gives once more a non-trivial computational advantage to the uncentred SDF method, which only contains the single unknown parameter λ.[8]

Still, it is convenient to have consistent initial values. For that reason, we propose a computationally simple intuitive estimator that is always consistent, but which would become efficient when the returns and factors are i.i.d. elliptical, which nests the joint Gaussian assumption in Jagannathan and Wang (2002). For the uncentred SDF model, in particular:

---

[7]Kan and Robotti (2008) also show that CU-GMM versions of the SDF approach are numerically invariant to affine transformations of the factors with known coefficients, which is not necessarily true of two-step or iterated GMM methods. Not surprisingly, it is easy to adapt the proof of Proposition 2 to show that the regression approach is also numerically invariant to such transformations.

[8]This advantage becomes more relevant as the number of factors k increases because the centred SDF method requires the additional estimation of k factor means, while the regression method the estimation of $n \times k$ factor loadings if we use the parametrisation in terms of $\varkappa$, and even more elements if we use the parametrisation in terms of $\boldsymbol{\delta}$.

Lemma 2 *If $(\mathbf{r}_t, f_t)$ is an i.i.d. elliptical random vector with bounded fourth moments, and the null hypothesis of linear factor pricing holds, then the most efficient estimator of λ obtained from (13) will be given by*

$$\mathring{\lambda}_T = \frac{\sum_{t=1}^{T} r_t^+}{\sum_{t=1}^{T} r_t^{+2}} \tag{20}$$

where r_t^+ is the uncentred factor mimicking portfolio defined in (15), whose sample counterpart would be

$$\tilde{r}_t^+ = \left(\sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_t'\right)^{-1} \left(\sum_{t=1}^{T} \mathbf{r}_t\right).$$

Once again, it is important to stress that the feasible version of $\mathring{\lambda}_T$ will remain consistent under linear factor pricing even if the assumptions of serial independence and a multivariate elliptical distribution are not totally realistic in practice.

3.3 Empirical application

We look again at the eight currency portfolios in Lustig and Verdelhan (2007), but this time we focus on linear factor pricing models with non-traded factors. Specifically, we consider a linearised version of the CCAPM, which defines the US aggregate consumption growth of nondurable as the only pricing factor, as well as an extension (DCAPM) along the lines of Yogo (2006), which adds the consumption growth of durables as a second pricing factor.

Table 4 displays the results from the application of the different inference procedures previously discussed to this data set for the purposes of testing the CCAPM. The computation of the weighting matrix in two-step and CU-GMM is the same as we explained for the case of traded factors.

Not surprisingly, there are again numerical differences between the standard two-step GMM implementation of the three existing approaches. Unlike what happened in the case of traded factors, though, in this case the numerical differences lead to different conclusions at conventional significance levels. Specifically, while the centred SDF approach rejects the null hypothesis, the uncentred SDF and the regression approach do not. In turn, the single-step GMM version of the J test does not reject the null hypothesis implicit in (1) and (12), which is in agreement with the empirical results in Lustig and Verdelhan (2007). Similarly, the two-step and CU-GMM estimates of λ and δ are fairly close, while the CU-GMM estimate of τ is higher than its two-step counterpart, although the t-ratio is much lower.

The numerical differences between the three two-step overidentification tests are substantially reduced but not eliminated if we use these GMM estimates to compute a new weighting matrix, which we then use to obtain three-step parameter estimates, and so on and so forth.[9]   In

---

[9] As Hansen, Heaton and Yaron (1996) show, though, such iterated GMM estimators do not generally coincide with the CU ones.

particular, after just a few iterations the uncentred SDF method converges to an estimate of λ of 0.49 and a J statistic of 5.69, which are very similar to their CU-GMM counterparts. In turn, the regression method converges to an estimate of δ of 5.61 and a J statistic of 5.68, which are again rather close to their CU-GMM counterparts. However, the centred SDF converges after many iterations to an estimate of τ of 1.15 and a J statistic of 16.92, whose p-value is 2%.

(TABLE 4)

Panel B of Table 4 shows the pricing errors. Once again, there are few statistically significant values of $\boldsymbol{\alpha}$, $\boldsymbol{\pi}$ or $\boldsymbol{\psi}$ for the 8 currency portfolios in the two GMM versions. The main exception is the two-step GMM estimates of $\boldsymbol{\psi}$, with 3 significant pricing errors at the 5% level.

The evidence on the DCAPM shown in Table 5 is broadly similar to the evidence on the CCAPM in Table 4. Specifically, the numerical differences across two-step implementations reported in Panel A yield different conclusions at conventional significance levels: while the centred SDF approach rejects the null hypothesis, the uncentred SDF and the regression approach do not. In turn, the single-step GMM test does not reject the null hypothesis implicit in (1) and (12). Similarly, Panel B shows that there are not significant values at the 5% level for $\boldsymbol{\alpha}$ or $\boldsymbol{\pi}$ for the 8 currency portfolios in the two GMM versions. However, we can find 4 portfolios with two-step GMM estimates of $\boldsymbol{\psi}$ that are significant at 1%. In contrast, none of the CU-GMM estimates of those pricing errors are significant at the 5% level.

(TABLE 5)

Burnside (2007) argues that the usual two-step implementation of the uncentred SDF approach has no power against potentially misspecified SDFs when the population covariance of the pricing factors with the excess returns on the assets is 0. Similarly, one could easily modify his arguments to say that the usual two-step implementation of the centred SDF approach would have no power if the cross moment between pricing factors and excess returns were 0, and the same would apply to the centred regression approach.

Given the numerical equivalence of the single-step implementation of the three approaches in Proposition 2, in our view the focus should not be on the statistical properties of the different estimators and tests, but rather, on the interpretation of the restrictions that are effectively tested in those two special cases.

15

For the sake of clarity, let us study these issues with a single factor. Specifically, when $Cov(\mathbf{r}, f) = \mathbf{0}$ but $E(\mathbf{r}) \neq \mathbf{0}$, (12) will be satisfied with $c = 0$ and $\boldsymbol{\varphi}d = E(\mathbf{r})$, while the moment conditions (1) will be satisfied by any affine transformation of f such that $a + b\mu = 0$. Therefore, the value of λ given by (14) will trivially satisfy (13), as Burnside (2007) shows. In contrast, one cannot find any finite value of τ that will satisfy (16) because the centred mimicking representing portfolio r^{++} will be 0. Likewise, one cannot find any finite value of $\varkappa$ that will satisfy (18) because $\boldsymbol{\beta} = \mathbf{0}$.

Similarly, when $E(\mathbf{r}f) = \mathbf{0}$ but $E(\mathbf{r}) \neq \mathbf{0}$, (12) will be satisfied with $d\mu - c\gamma = 0$ and $\boldsymbol{\varphi}d \left(\sigma^2/\gamma\right) = E(\mathbf{r})$, while the moment conditions (1) will be satisfied by any SDF which is exactly proportional to f (so that $a = 0$). Therefore, the value of τ given by (17) will trivially satisfy (16), and the same applies to (18) with (10) and $\varkappa = -\gamma/\mu$. In contrast, one cannot find any finite value of λ that will satisfy (13) because the uncentred mimicking portfolio r^+ will be 0.

From an economic point of view those solutions are clearly unsatisfactory, but strictly speaking the corresponding SDF's correctly price the vector of excess returns at hand. In our view, the best way to solve these problems would be to add a risky or riskless asset whose cost is not 0, which would implicitly fix the scale of the SDF by fixing its mean; see e.g. Hodrick and Zhang (2001) and Farnsworth et al. (2002). For that reason, we devote the next section to this case. Unfortunately, most empirical studies do not include such an asset.

From an econometric point of view, though, the truly problematic case arises when (1) and (12) hold but $E(\mathbf{r}) = \mathbf{0}$, in which case both $Cov(\mathbf{r}, f)$ and $E(\mathbf{r}f)$ must be $\mathbf{0}$ too. In this situation, the SDF parameters a and b and the projection parameters c and d are underidentified even after normalisation, which renders standard GMM inferences invalid.[10]

To investigate whether these theoretical situations are empirically relevant, we perform single-step GMM overidentification tests of the following null hypotheses:

(a) the mean excess return is 0, i.e.

$$E(\mathbf{r}) = \mathbf{0},$$

(b) the cross moment between excess returns and factors are 0, i.e.

$$E(\mathbf{r}f) = \mathbf{0},$$

and (c) the covariance between excess returns and factors are 0, i.e.

$$E\left[(\mathbf{r} - \boldsymbol{\varphi}) f\right] = \mathbf{0},$$

---

[10]See Khan and Zhang (1999) or Burnside (2007) for the implications that identification failures have for two-step GMM procedures.

where $\boldsymbol{\varphi}$ is a vector of parameters to estimate. We do so for each of the two consumption factors separately, and also perform a joint test.

Hypothesis (a) is clearly rejected with a J statistic of 39.97, whose p-value is essentially 0. Therefore, there are statistically significant risk premia in search of pricing factors to explain them. The next step is to investigate if the consumption factors considered by Lustig and Verdelhan (2007) can play such a role.

Table 6 shows that hypothesis (b) is clearly rejected, but there is not much evidence against hypothesis (c) for non-durable consumption, durable consumption or both factors together. Thus, we cannot reject that those factors are jointly uncorrelated with the currency portfolios, which indicates that the seemingly positive evaluation of the consumption based asset pricing models in Tables 4 and 5 must be interpreted with some care.

(TABLE 6)

The last column of Table 6 shows that the joint hypothesis (a)+(b), or equivalently (a)+(c), is rejected for each factor, which confirms that the parameters appearing in (1) and (12) are point identified in a single factor model.[11]   Interestingly, this test coincides with a simple version of the underidentification test of Arellano, Hansen and Sentana (2009) adapted to linear factor pricing models by Manresa (2008). In this sense, it is worth noting that once again CU-GMM proves useful to unify the empirical results because we obtain the same J test whether we regress $\mathbf{r}$ on f or f on $\mathbf{r}$.

In summary, our empirical results with non-traded factors indicate that although we cannot reject the overidentification restrictions implicit in (1) and (12), the main reason seems to be that the aggregate consumption growth of nondurable and the consumption growth of durables are uncorrelated to the eight portfolios of currencies in Lustig and Verdelhan (2007).

In fact, we find similar results with the traded factors MK, HML and SMB if we treat them as non-traded by discarding the moment condition (2). In particular, the CAPM would give rise to a CU-GMM J test with a p-value of 42%, while the p-value of the test of the Fama-French model would become 74%.

4   Evaluation with gross returns

Let us define $\mathbf{R}$ as an $n \times 1$ vector of gross returns, which are such that their cost is given by the $n \times 1$ vector of ones ℓ_n. We focus again on the case of a single factor f to simplify the

---

[11]However, a zero covariance of both consumption factors with the currency portfolios implies that the multi-factor DCAPM is not identified.

exposition, and relegate the general multifactor case to Appendix B.

In this context, the analogue to the SDF pricing equation (1) is

$$E\left((a+bf)\mathbf{R}\right) = \ell_n \qquad (21)$$

for some real numbers (a, b). As we mentioned in the previous section, there is no longer any need to normalise (a, b), unlike in the case of excess returns.

Not surprisingly, we can equivalently write this pricing condition as a restriction on the least squares projection of $\mathbf{R}$ onto a constant and f. Specifically, if we define $\boldsymbol{\phi}$ and $\boldsymbol{\beta}$ as the vectors of intercepts and slopes in that projection, respectively, then (21) is equivalent to

$$c\boldsymbol{\phi} + d\boldsymbol{\beta} = \ell_n, \qquad (22)$$

where c is the SDF mean and d is the shadow price of the factor, or its actual price when is traded.

4.1 Traded Factors

When the factor is a gross return, it will also satisfy

$$E\left((a+bf)f\right) = 1. \qquad (23)$$

Therefore, the moments of the uncentred SDF method will be (21) and (23). Since the only parameters are (a, b), the J test will have $n-1$ degrees of freedom under correct specification.

We can also reparametrise the SDF in terms of c instead of a and rely on the moment conditions

$$E\left[\begin{array}{c} \mathbf{R}\left(c - (f-\mu)\tau\right) - \ell_n \\ f\left(c - (f-\mu)\tau\right) - 1 \\ f - \mu \end{array} \right] = \mathbf{0},$$

whose parameters are (c, τ, μ), which are analogous to (6).

Regarding the regression approach, the fact that the factor is traded implies that $d = 1$, which simplifies the least squares constraint to

$$\boldsymbol{\phi} = \kappa\left(\ell_n - \boldsymbol{\beta}\right), \qquad \kappa = 1/c,$$

when $c \neq 0$.[12] The parameter κ is usually referred to as the zero-beta return since it corresponds to the expectation of returns uncorrelated with f.

---

[12]The case $c = 0$ is such that f is equal to the minimum variance portfolio, whose return R^{**} satisfies

$$Cov\left(\mathbf{R}, R^{**}\right) = Var\left(R^{**}\right)\ell_n.$$

It is well known that this asset is the only element of the mean-variance frontier that cannot be used as a benchmark in beta pricing.

Therefore, the moments associated to the regression method will be

$$E\left[\begin{pmatrix} 1 \\ f \end{pmatrix} ((\mathbf{R} - \kappa\ell_n) - \boldsymbol{\beta}(f - \kappa))\right] = \mathbf{0},$$

with parameters $(\boldsymbol{\beta}, \kappa)$.

We show in Appendix C that there is a direct counterpart to Proposition 1 for gross returns too, so that all the approaches are numerically equivalent when implemented by single-step GMM methods.

4.2 Non-traded Factors

If the factor is not traded then we cannot rely on the additional constraint of its own pricing. The moments of the SDF method are simply (21) with parameters (a, b). Consequently, the J test has $n - 2$ degrees of freedom under correct specification.

Once again, we can also reparametrise the SDF in terms of c instead of a and rely on the moment conditions

$$E\left[\begin{array}{c} \mathbf{R}(c - (f - \mu)\tau) - \ell_n \\ f - \mu \end{array}\right] = \mathbf{0}$$

with parameters (c, τ, μ), which are analogous to (16).

On the other hand, when $c \neq 0$ we can re-express the least squares constraint as

$$\boldsymbol{\phi} = \kappa\ell_n + \boldsymbol{\beta}\varkappa, \qquad \kappa = 1/c, \qquad \varkappa = -d/c.$$

Therefore, the moments of the regression method are

$$E\left[\begin{pmatrix} 1 \\ f \end{pmatrix} ((\mathbf{R} - \kappa\ell_n) - \boldsymbol{\beta}(f + \varkappa))\right] = \mathbf{0},$$

with parameters $(\boldsymbol{\beta}, \kappa, \varkappa)$.

We show in Appendix C that there is a direct counterpart to Proposition 2 for gross returns too, so that all the approaches are numerically equivalent when implemented by single-step GMM methods.

Finally, it is worth explaining the advantages of working with gross returns in the situation discussed in section 3.3 in which the factor is uncorrelated with excess returns. For simplicity, let us assume that the reference asset used in computing $\mathbf{r}$ is a portfolio whose gross return has zero variance, so that $Cov(\mathbf{R}, f) = \mathbf{0}$ if and only if $Cov(\mathbf{r}, f) = \mathbf{0}$. As we mentioned in that section, the moment condition (1) will hold with $a = -b\mu$ regardless of $E(\mathbf{r})$, which corresponds to a valid but economically unsatisfactory SDF. In contrast, (21) cannot be satisfied unless $E(\mathbf{R})$ is proportional to ℓ_n. If this "risk neutral" condition does not hold, then the J test based on $\mathbf{R}$

will reject with probability 1 in large samples, while the rejection probability of the J test based on $\mathbf{r}$ will coincide with its size.

On the other hand, an unsatisfactory SDF that will satisfy (21) would be one in which the factor is correlated with returns but cannot discriminate across assets, so that $Cov(\mathbf{R}, f) = k\ell_n$ for some scalar constant $k \neq 0$. Similar arguments apply to $E(\mathbf{R}f)$.

5   Conclusions

There are two main approaches in empirical finance to evaluate linear factor pricing models. The oldest method relies on regressions of excess returns on factors, while the other more recent methods rely instead on the SDF representation of the model. There are two variants of the SDF approach, one that subtracts the mean of the factors and another one which does not. Given that these different procedures may lead to different empirical conclusions, it is perhaps not surprising that there has been some controversy about which approach is most adequate. In this context, our paper shows that if we use single-step methods such as CU-GMM instead of standard two-step or iterated GMM procedures, then all these procedures provide the same estimates of prices of risk, overidentifying restrictions tests, pricing errors, and Jensen's alphas irrespective of the validity of the model. Thus, we would argue that in effect the regression and SDF approaches are different representations of a single method to empirically evaluate asset-pricing models.

We prove our numerical equivalence results for both traded and non-traded factors irrespective of their number, and also for excess returns and gross returns. For the benefit of practitioners, we also develop simple, intuitive consistent parameter estimators that can be used to obtain good initial conditions for CU-GMM, and which will be efficient for elliptically distributed returns and factors.

We illustrate our results by revisiting the empirical analysis of currency returns in Lustig and Verdelhan (2007). Our findings clearly point out that the conflict among criteria for testing asset pricing models that we have previously mentioned is not only a theoretical possibility, but a hard reality. Nevertheless, such a conflict disappears when one uses single-step methods.

A different issue, though, is the interpretation of the restrictions that are effectively tested. In this sense, our results confirm Burnside's (2007) findings that the favourable empirical evidence on the asset pricing models with non-traded factors that Lustig and Verdelhan (2007) considered should be interpreted with some care because their consumption-based pricing factors seem to be uncorrelated to currency returns. We also find that if we force the models with traded factors to price those factors, and in particular the US market portfolio, then we reject the asset pricing

restrictions.

Given that our results are not specific to CU-GMM, it might be worth comparing the properties in finite samples of different members of the generalised empirical likelihood family of estimators. An alternative application of our numerical equivalence results would be the performance evaluation of mutual and hedge funds. This literature can also be divided between papers that rely on regression methods, such as Kosowski et al. (2006), and papers that rely on SDF methods, such as Dahlquist and Soderlind (1999) and Farnsworth et al. (2002). Undoubtedly, these topics constitute interesting avenues for further research.

References

Almeida, C. and R. Garcia (2009): "Assessing misspecified asset pricing models with empirical likelihood estimators", mimeo, Fundação Getulio Vargas.

Antoine, A., H. Bonnal and E. Renault (2007): "On the efficient use of the informational content of estimating equations: Implied probabilities and Euclidean empirical likelihood", *Journal of Econometrics* 138, 461-487.

Arellano, M., L.P. Hansen and E. Sentana (2009): "Underidentification?", CEMFI Working Paper 0905.

Breeden, D.T. (1979): "An intertemporal asset pricing model with stochastic consumption and investment opportunities", *Journal of Financial Economics* 7, 265-296.

Burnside, C., M. Eichenbaum, I. Kleshchelski and S. Rebelo (2006): "The returns to currency speculation," NBER Working Paper 12,489.

Burnside, C. (2007): "Empirical asset pricing and statistical power in the presence of weak risk factors", NBER Working Paper 13,357.

Chamberlain, G. (1983a): "A characterization of the distributions that imply mean-variance utility functions", *Journal of Economic Theory* 29, 185-201.

Chamberlain, G. (1983b): "Funds, factors, and diversification in arbitrage pricing models", *Econometrica* 51, 1305-1323.

Campbell, J.Y., S. Gilgio, and C. Polk (2010): "Hard Times", mimeo, Harvard University.

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

Cochrane, J.H. (2001a): *Asset Pricing*, Princeton University Press.

Cochrane, J.H. (2001b): "A rehabilitation of stochastic discount factor methodology", mimeo, GSB University of Chicago.

Dahlquist, M. and P. Söderlind (1999): "Evaluating portfolio performance with stochastic discount factors", *Journal of Business* 72, 347-83.

de Jong, R.M. and J. Davidson (2000): "Consistency of kernel estimators of heteroskedasticity and autocorrelation covariance matrices", *Econometrica* 68, 407-423.

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

Fang, K.-T., S. Kotz and K.-W. Ng (1990): *Symmetric Multivariate and Related Distributions*, Chapman and Hall.

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

Gibbons, M.R., S.A. Ross and J. Shanken (1989): "A test of the efficiency of a given portfolio", *Econometrica* 57, 1121-1152.

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

Hansen, L.P., J. Heaton and A. Yaron (1996): "Finite sample properties of some alternative GMM estimators", *Journal of Business and Economic Statistics* 14, 262-280.

Hansen, L.P. and R. Jagannathan (1991): "Implications of security market data for models of dynamic economies", *Journal of Political Economy* 99, 225-262.

Hansen, L.P. and R. Jagannathan (1997): "Assessing specification errors in stochastic discount factor models", *Journal of Finance* 52, 557-590.

Hodrick, R. J. and Zhang, X. (2001): "Evaluating the specification errors of asset pricing models", *Journal of Financial Economics* 62, 327-376.

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

Julliard, C. and A. Ghosh (2008): "Can rare events explain the equity premium puzzle?", LSE FMG Discussion Paper 610.

Kan, R. and C. Robotti (2008): "Specification tests of asset pricing models using excess returns", *Journal of Empirical Finance* 15, 816–38.

Kan, R. and C. Zhang (1999): "GMM tests of stochastic discount factor models with useless factors", *Journal of Financial Economics* 54, 103-127.

Kan, R. and G. Zhou (2000): "A critique of the stochastic discount factor methodology", *Journal of Finance* 54, 1221-1248.

Kosowski, R., A. Timmermann , R. Wermers, and H. White (2006): "Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis", *Journal of Finance* 61, 2551-2595.

Lewellen J., S. Nagel and J. Shanken (2009): "A skeptical appraisal of asset-pricing tests, *Journal of Financial Economics*, forthcoming

Lintner, J. (1965): "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics* 47, 13-37.

Lucas, R.E. (1978): "Asset prices in an exchange economy", *Econometrica* 46, 1429-1446.

Lustig, H. and A. Verdelhan (2007): "The cross section of foreign currency risk premia and consumption growth risk", *American Economic Review* 97, 89-17.

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

Magnus, J.R. and H. Neudecker (1990): *"Matrix Differential Calculus with Applications to*

Econometrics", Wiley.

Manresa, E. (2008): "Application of rank tests to linear factor asset pricing models", unpublished Master's thesis, CEMFI.

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

Newey, W.K. and R.J. Smith (2004): "Higher order properties of GMM and generalized empirical likelihood estimators", *Econometrica* 72, 219-255.

Owen, J. and R. Rabinovitch (1983): On the class of elliptical distributions and their applications to the theory of portfolio choice, *Journal of Finance* 58, 745-752.

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

Peñaranda, F. and E. Sentana (2010): "Spanning tests in portfolio and stochastic discount factor mean-variance frontiers: a unifying approach", mimeo, UPF.

Rubinstein, M. (1976): "The valuation of uncertain income streams and the pricing of options", *Bell Journal of Economics* 7, 407-425.

Ross, S.A. (1976): "The arbitrage theory of capital asset pricing", *Journal of Economic Theory* 13, 341-360.

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

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

Appendices

A   Single-step methods: Continuously Updated GMM

Let $\{\mathbf{x}_t\}_{t=1}^T$ denote a strictly stationary and ergodic stochastic process, and define $\mathbf{h}(\mathbf{x}_t;\boldsymbol{\theta})$ as a vector of known functions of $\mathbf{x}_t$, where $\boldsymbol{\theta}$ is a vector of unknown parameters. The true parameter value, $\boldsymbol{\theta}^0$, which we assume belongs to the interior of the compact set $\boldsymbol{\Theta} \subseteq \mathbb{R}^{\dim(\boldsymbol{\theta})}$, is implicitly defined by the (population) moment conditions:

$$E[\mathbf{h}(\mathbf{x}_t;\boldsymbol{\theta}^0)] = \mathbf{0},$$

where the expectation is taken with respect to the stationary distribution of $\mathbf{x}_t$. In our context of asset pricing models, $\mathbf{x}_t = (\mathbf{r}_t', \mathbf{f}_t')'$ represents data on excess returns and factors, and $\boldsymbol{\theta}$ represents the parameters of the specific model under evaluation.

GMM estimators minimise a specific norm $\bar{\mathbf{h}}_T'(\boldsymbol{\theta})\boldsymbol{\Upsilon}_T\bar{\mathbf{h}}_T(\boldsymbol{\theta})$ of the sample moments $\bar{\mathbf{h}}_T(\boldsymbol{\theta}) = T^{-1}\sum_{t=1}^T \mathbf{h}(\mathbf{x}_t;\boldsymbol{\theta})$ defined by some weighting matrix $\boldsymbol{\Upsilon}_T$. In overidentified cases such as ours, Hansen (1982) showed that if the long-run covariance matrix of the moment conditions $\mathbf{S}(\boldsymbol{\theta}^0) = avar[\sqrt{T}\bar{\mathbf{h}}_T(\boldsymbol{\theta}^0)]$ has full rank, then $\mathbf{S}^{-1}(\boldsymbol{\theta}^0)$ will be the "optimal" weighting matrix, in the sense that the difference between the asymptotic covariance matrix of the resulting GMM estimator and a GMM estimator based on any other norm of the same moment conditions is positive semidefinite. Therefore, the optimal GMM estimator of $\boldsymbol{\theta}$ will be

$$\hat{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} J_T(\boldsymbol{\theta}),$$

where

$$J_T(\boldsymbol{\theta}) = \bar{\mathbf{h}}_T'(\boldsymbol{\theta})\mathbf{S}^{-1}\left(\boldsymbol{\theta}^0\right)\bar{\mathbf{h}}_T(\boldsymbol{\theta}).$$

This optimal estimator is infeasible unless we know $\mathbf{S}(\boldsymbol{\theta}^0)$, but under additional regularity conditions, we can define an asymptotically equivalent but feasible two-step optimal GMM estimator by replacing $\mathbf{S}(\boldsymbol{\theta}^0)$ with an estimator $\mathbf{S}_T(\boldsymbol{\theta})$ evaluated at some initial consistent estimator of $\boldsymbol{\theta}^0$, $\dot{\boldsymbol{\theta}}_T$ say. There is an extensive literature on heteroskedasticity and autocorrelation consistent (HAC) estimators of long-run covariance matrices (see for example DeJong and Davidson (2000) and the references therein). In practice, we can repeat this two-step procedure many times to obtain iterated GMM estimators.

An alternative way to make the optimal GMM estimator feasible is by explicitly taking into account in the criterion function the dependence of the long-run variance on the parameter values, as in the single-step CU-GMM estimator of Hansen, Heaton and Yaron (1996), which is defined as

$$\tilde{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \tilde{J}_T(\boldsymbol{\theta}),$$

where

$$\tilde{J}_T(\boldsymbol{\theta}) = \bar{\mathbf{h}}_T'(\boldsymbol{\theta})\mathbf{S}_T^{-1}(\boldsymbol{\theta})\bar{\mathbf{h}}_T(\boldsymbol{\theta}).$$

Peñaranda and Sentana (2010) discuss how to express the CU-GMM criterion in terms of OLS output, which facilitates its optimisation. Although this estimator is often more difficult to compute than two-step and iterated estimators, particularly in linear models, an important advantage is that it is numerically invariant to normalisation, bijective reparametrisations and parameter-dependent linear transformations of the moment conditions, which will again prove useful in our context. In contrast, these properties do not necessarily hold for two-step or iterated GMM.

Newey and Smith (2004) highlight other important advantages of CU- over two-step GMM by going beyond the usual first-order asymptotic equivalence results. They also discuss alternative generalised empirical likelihood (GEL) estimators, such as empirical likelihood or exponentially-tilted methods. In fact, Antoine, Bonnal and Renault (2006) study the Euclidean empirical likelihood estimator, which is numerically equivalent to CU-GMM as far as $\boldsymbol{\theta}$ is concerned. Importantly, it is straightforward to show that these GEL methods share the numerical invariance properties of CU-GMM.

Our empirical application will consider two-step, iterated and CU-GMM. Under standard regularity conditions (see Hansen (1982)), $\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}^0)$ and $\sqrt{T}(\tilde{\boldsymbol{\theta}}_T - \boldsymbol{\theta}^0)$ will be asymptotically distributed up to first-order as the same normal random vector with zero mean and variance

$$\left[\mathbf{D}'(\boldsymbol{\theta}^0)\mathbf{S}^{-1}\left(\boldsymbol{\theta}^0\right)\mathbf{D}(\boldsymbol{\theta}^0)\right]^{-1},$$

where $\mathbf{D}(\boldsymbol{\theta}^0)$ denotes the probability limit of the Jacobian of $\bar{\mathbf{h}}_T(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}^0$. In our empirical application, we replace $\mathbf{D}(\boldsymbol{\theta}^0)$ by $\partial\bar{\mathbf{h}}_T(\hat{\boldsymbol{\theta}}_T)/\partial\boldsymbol{\theta}'$ in the case of two-step and iterated GMM estimators. In contrast, for the CU-GMM estimator $\tilde{\boldsymbol{\theta}}_T$ we compute a consistent estimator of $\mathbf{D}(\boldsymbol{\theta}^0)$ that takes into account that the weighting matrix $\mathbf{S}_T^{-1}(\boldsymbol{\theta})$ is not fixed in the criterion function. Specifically, we estimate the asymptotic variance of $\tilde{\boldsymbol{\theta}}_T$ as

$$\left[\mathcal{D}_T'(\tilde{\boldsymbol{\theta}}_T)\mathbf{S}_T^{-1}(\tilde{\boldsymbol{\theta}}_T)\mathcal{D}_T(\tilde{\boldsymbol{\theta}}_T)\right]^{-1},$$

where

$$\mathcal{D}_T(\tilde{\boldsymbol{\theta}}_T) = \frac{\partial\bar{\mathbf{h}}_T(\tilde{\boldsymbol{\theta}}_T)}{\partial\boldsymbol{\theta}'} - \frac{1}{2}\left[\bar{\mathbf{h}}_T'(\tilde{\boldsymbol{\theta}}_T)\mathbf{S}_T^{-1}(\tilde{\boldsymbol{\theta}}_T) \otimes \mathbf{I}_{\dim(\mathbf{h})}\right]\frac{\partial vec\left(\mathbf{S}_T(\tilde{\boldsymbol{\theta}}_T)\right)}{\partial\boldsymbol{\theta}'}.$$

Similarly, we estimate the asymptotic variance of $\sqrt{T}\bar{\mathbf{h}}_T(\tilde{\boldsymbol{\theta}}_T)$ as

$$\mathbf{S}_T(\tilde{\boldsymbol{\theta}}_T) - \mathcal{D}_T(\tilde{\boldsymbol{\theta}}_T)\left[\mathcal{D}_T'(\tilde{\boldsymbol{\theta}}_T)\mathbf{S}_T^{-1}(\tilde{\boldsymbol{\theta}}_T)\mathcal{D}_T(\tilde{\boldsymbol{\theta}}_T)\right]^{-1}\mathcal{D}_T'(\tilde{\boldsymbol{\theta}}_T),$$

while we replace $\mathcal{D}_T(\tilde{\boldsymbol{\theta}}_T)$ by $\partial\bar{\mathbf{h}}_T(\hat{\boldsymbol{\theta}}_T)/\partial\boldsymbol{\theta}'$ when we consider $\bar{\mathbf{h}}_T(\hat{\boldsymbol{\theta}}_T)$. Finally, $T \cdot J_T(\hat{\boldsymbol{\theta}}_T)$ and $T \cdot \tilde{J}_T(\tilde{\boldsymbol{\theta}}_T)$ will be asymptotically distributed as the same chi-square with $\dim(\mathbf{h}) - \dim(\boldsymbol{\theta})$

degrees of freedom if $E[\mathbf{h}(\mathbf{x}; \boldsymbol{\theta})] = \mathbf{0}$ holds, so that we can use those statistics to compute overidentifying restrictions (J) tests.

B Multifactor models

In what follows we represent a set of k factors by the vector $\mathbf{f}$, and their mean vector, second moment and covariance matrix by $\boldsymbol{\mu}$, $\boldsymbol{\Gamma}$ and $\boldsymbol{\Sigma} = \boldsymbol{\Gamma} - \boldsymbol{\mu}\boldsymbol{\mu}'$, respectively. In this multifactor context, the connection between the SDF and regression approaches is given by

$$E(\mathbf{r})a + E(\mathbf{r}\mathbf{f}')\mathbf{b} = E(\mathbf{r})(a + \mathbf{b}'\boldsymbol{\mu}) + Cov(\mathbf{r}, \mathbf{f})\mathbf{b} = \phi(a + \mathbf{b}'\boldsymbol{\mu}) + \mathbf{B}\boldsymbol{\mu}(a + \mathbf{b}'\boldsymbol{\mu}) + \mathbf{B}\boldsymbol{\Sigma}\mathbf{b}$$

$$= \phi(a + \mathbf{b}'\boldsymbol{\mu}) + \mathbf{B}(a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b}) = \phi c + \mathbf{B}\mathbf{d} = \mathbf{0}, \tag{B1}$$

where c is the mean of the SDF, $\mathbf{d}$ is the shadow price of $\mathbf{f}$ (or its actual price if it is a vector of traded payoffs) and

$$\mathbf{B} = \left(\begin{array}{ccc} \boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_k \end{array} \right)$$

is the $n \times k$ matrix of regression slopes. We can interpret condition (B1) as reflecting the orthogonality between the SDF and the projection of $\mathbf{r}$ onto a constant and $\mathbf{f}$ because $\phi c + \mathbf{B}\mathbf{d} = E[(a + \mathbf{b}'\mathbf{f})(\phi + \mathbf{B}\mathbf{f})] = \mathbf{0}$.

Traded factors

If the factors are excess returns themselves then condition (B1) is equivalent to $\phi = \mathbf{0}$ because $c = a + \mathbf{b}'\boldsymbol{\mu} \neq 0$ and the price of $\mathbf{f}$ must satisfy $\mathbf{d} = a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b} = \mathbf{0}$. Therefore, we can evaluate the corresponding asset pricing model by means of the SDF influence functions

$$\mathbf{h}_S(\mathbf{r}, \mathbf{f}; a, \mathbf{b}) = \left[\begin{array}{c} \mathbf{r}(a + \mathbf{f}'\mathbf{b}) \\ \mathbf{f}(a + \mathbf{f}'\mathbf{b}) \end{array} \right], \tag{B2}$$

or the regression influence functions

$$\mathbf{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B}) = \left[\begin{array}{c} \mathbf{r} - \mathbf{B}\mathbf{f} \\ vec((\mathbf{r} - \mathbf{B}\mathbf{f})\mathbf{f}') \end{array} \right]. \tag{B3}$$

The SDF functions require some normalisation in their implementation. A symmetrically normalised version of the SDF approach would use the normalisation $a^2 + \mathbf{b}'\mathbf{b} = 1$ but asymmetric normalisations are more common in empirical work. The uncentred SDF method imposes $a = 1$ and relies on the influence functions

$$\left[\begin{array}{c} \mathbf{r}(1 - \mathbf{f}'\boldsymbol{\lambda}) \\ \mathbf{f}(1 - \mathbf{f}'\boldsymbol{\lambda}) \end{array} \right]$$

with parameters $\boldsymbol{\lambda}$, while the influence functions of the centred SDF method impose $a + \boldsymbol{\mu}'\mathbf{b} = 1$, and become

$$
\begin{bmatrix}
\mathbf{r}\left(1 - (\mathbf{f} - \boldsymbol{\mu})'\boldsymbol{\tau}\right) \\
\mathbf{f}\left(1 - (\mathbf{f} - \boldsymbol{\mu})'\boldsymbol{\tau}\right) \\
\mathbf{f} - \boldsymbol{\mu}
\end{bmatrix}
$$

with parameters $(\boldsymbol{\mu}, \boldsymbol{\tau})$. The link between both sets of parameters is

$$
\boldsymbol{\mu} = \boldsymbol{\Gamma}\boldsymbol{\lambda} = \boldsymbol{\Sigma}\boldsymbol{\tau}.
$$

In all these methods the degrees of freedom of the corresponding J tests are n regardless of the number of factors k. The Jensen's alphas and pricing errors of excess returns $\mathbf{r}$ are defined by

$$
\begin{aligned}
\boldsymbol{\alpha} &= E(\mathbf{r}) - \mathbf{B}E(\mathbf{f}), \\
\boldsymbol{\pi} &= E(\mathbf{r}) - E(\mathbf{r}\mathbf{f}')\boldsymbol{\lambda}, \\
\boldsymbol{\psi} &= E(\mathbf{r}) - E(\mathbf{r}(\mathbf{f} - \boldsymbol{\mu})')\boldsymbol{\tau}.
\end{aligned}
$$

Non-traded factors

If the factors are not traded payoffs then we need $n > k$ so that there are some overidentifying restrictions to test. As in the case of traded factors, the condition (B1) holds for a valid asset pricing model. Such a constraint is equivalent to both $\boldsymbol{\phi}$ and $\mathbf{B}$ belonging to the span of some $n \times k$ matrix that we can denote as

$$
\mathbf{P} = \begin{pmatrix} \boldsymbol{\varphi}_1 & \cdots & \boldsymbol{\varphi}_k \end{pmatrix}.
$$

Assuming that $\mathbf{B}$ has full column rank, we can impose this implicit constraint on the intercepts and slopes of the regression of $\mathbf{r}$ on a constant and $\mathbf{f}$ as follows:

$$
\boldsymbol{\phi} = \mathbf{P}\mathbf{d} = \boldsymbol{\varphi}_1 d_1 + ... + \boldsymbol{\varphi}_k d_k,
$$

$$
\mathbf{B} = \begin{pmatrix} \boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_k \end{pmatrix} = \begin{pmatrix} -c\boldsymbol{\varphi}_1 & \cdots & -c\boldsymbol{\varphi}_k \end{pmatrix} = -c\mathbf{P}.
$$

Therefore, we can evaluate the corresponding asset pricing model by means of the SDF influence functions

$$
\mathbf{g}_S(\mathbf{r}, \mathbf{f}; a, \mathbf{b}) = \left[\mathbf{r}\left(a + \mathbf{f}'\mathbf{b}\right)\right], \tag{B4}
$$

or the regression influence functions

$$
\mathbf{g}_R(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}) = \begin{bmatrix} \mathbf{r} - \mathbf{P}(\mathbf{d} - c\mathbf{f}) \\ vec\left((\mathbf{r} - \mathbf{P}(\mathbf{d} - c\mathbf{f}))\mathbf{f}'\right) \end{bmatrix}. \tag{B5}
$$

The SDF functions require some normalisation in their implementation. A symmetrically normalised version of the SDF approach would use the normalisation $a^2 + \mathbf{b}'\mathbf{b} = 1$, while the regression would rely on the normalisation $c^2 + \mathbf{d}'\mathbf{d} = 1$, but asymmetric normalisations are more common in empirical work. The uncentred SDF method imposes $a = 1$ and relies on the influence functions

$$\left[\mathbf{r} \left(1 - \mathbf{f}'\boldsymbol{\lambda} \right) \right]$$

with parameters $\boldsymbol{\lambda}$. The influence functions of the centred SDF method impose $a + \boldsymbol{\mu}'\mathbf{b} = 1$ and they are

$$\begin{bmatrix} \mathbf{r} \left(1 - \left(\mathbf{f} - \boldsymbol{\mu} \right)' \boldsymbol{\tau} \right) \\ \mathbf{f} - \boldsymbol{\mu} \end{bmatrix}$$

with parameters $(\boldsymbol{\mu}, \boldsymbol{\tau})$. The usual regression imposes $c = -1$ and relies on the influence functions

$$\begin{bmatrix} \mathbf{r} - \mathbf{B} \left(\mathbf{f} + \varkappa \right) \\ vec \left(\left(\mathbf{r} - \mathbf{B} \left(\mathbf{f} + \varkappa \right) \right) \mathbf{f}' \right) \end{bmatrix},$$

with parameters $(\varkappa, \mathbf{B})$. Alternatively we can define the vector $\boldsymbol{\delta}$ of factor risk premia as $E \left(\mathbf{r} \right) = \mathbf{B}\boldsymbol{\delta}$, so that $\boldsymbol{\delta} = \varkappa + \boldsymbol{\mu}$, and add the estimation of $\boldsymbol{\mu}$. This yields

$$\begin{bmatrix} \mathbf{r} - \mathbf{B} \left(\mathbf{f} - \boldsymbol{\mu} + \boldsymbol{\delta} \right) \\ vec \left(\left(\mathbf{r} - \mathbf{B} \left(\mathbf{f} - \boldsymbol{\mu} + \boldsymbol{\delta} \right) \right) \mathbf{f}' \right) \\ \mathbf{f} - \boldsymbol{\mu} \end{bmatrix}.$$

Assuming that $\mathbf{B}$ has full column rank, the link between these parameters is

$$\boldsymbol{\delta} = \left[\boldsymbol{\Sigma} + \boldsymbol{\delta}\boldsymbol{\mu}' \right] \boldsymbol{\lambda} = \boldsymbol{\Sigma}\boldsymbol{\tau}.$$

In any case, the degrees of freedom of the J test will be $n - k$. Regarding Jensen's alphas and pricing errors, they are defined by

$$\begin{aligned} \boldsymbol{\alpha} &= E \left(\mathbf{r} \right) - \mathbf{B}\boldsymbol{\delta}, \\ \boldsymbol{\pi} &= E \left(\mathbf{r} \right) - E \left(\mathbf{r}\mathbf{f}' \right) \boldsymbol{\lambda}, \\ \boldsymbol{\psi} &= E(\mathbf{r}) - E\left(\mathbf{r} \left(\mathbf{f} - \boldsymbol{\mu} \right)' \right) \boldsymbol{\tau}. \end{aligned}$$

Evaluation with gross returns

The SDF pricing equation is

$$E \left(\left(a + \mathbf{b}'\mathbf{f} \right) \mathbf{R} \right) = \ell_n,$$

where $\mathbf{R}$ is an $n \times 1$ vector of gross returns and ℓ_n is an $n \times 1$ vector of ones. Therefore, we find

$$E \left(\mathbf{R} \right) a + E \left(\mathbf{R}\mathbf{f}' \right) \mathbf{b} = E \left(\mathbf{R} \right) c + Cov \left(\mathbf{R}, \mathbf{f} \right) \mathbf{b} = \ell_n,$$

where $c = a + \mathbf{b}'\boldsymbol{\mu}$ is the SDF mean. We can also relate the pricing of $\mathbf{R}$ to a constraint on the least squares intercepts $\boldsymbol{\phi}$ and slopes $\mathbf{B}$. Specifically,

$$E\left(\mathbf{R}\right)c + Cov\left(\mathbf{R}, \mathbf{f}\right)\mathbf{b} = \boldsymbol{\phi}c + \mathbf{B}\mathbf{d} = \ell_n,$$

where $\mathbf{d} = \boldsymbol{\mu}c + \boldsymbol{\Sigma}\mathbf{b}$ is the shadow price of the factors, or their actual prices if they are traded.

When the factors are gross returns themselves, they also satisfy

$$E\left(\left(a + \mathbf{b}'\mathbf{f}\right)\mathbf{f}\right) = \ell_k, \tag{B6}$$

where ℓ_k is a $k \times 1$ vector of ones, which implies that

$$\mathbf{d} = \boldsymbol{\mu}a + \boldsymbol{\Gamma}\mathbf{b} = \boldsymbol{\mu}c + \boldsymbol{\Sigma}\mathbf{b} = \ell_k.$$

In that case, the least squares constraint simplifies to

$$\boldsymbol{\phi}c + \mathbf{B}\ell_k = \ell_n,$$

which can also be expressed as

$$\boldsymbol{\phi} = \kappa\left(\ell_n - \mathbf{B}\ell_k\right), \qquad \kappa = 1/c,$$

when $c \neq 0$.

The influence functions of the SDF and regression methods with traded factors are

$$\mathbf{h}_S\left(\mathbf{R}, \mathbf{f}; a, \mathbf{b}\right) = \begin{bmatrix} \mathbf{R}\left(a + \mathbf{f}'\mathbf{b}\right) - \ell_n \\ \mathbf{f}\left(a + \mathbf{f}'\mathbf{b}\right) - \ell_k \end{bmatrix},$$

$$\mathbf{h}_R\left(\mathbf{R}, \mathbf{f}; \mathbf{B}, \kappa\right) = \begin{bmatrix} \left(\mathbf{R} - \kappa\ell_n\right) - \mathbf{B}\left(\mathbf{f} - \kappa\ell_k\right) \\ vec\left(\left(\left(\mathbf{R} - \kappa\ell_n\right) - \mathbf{B}\left(\mathbf{f} - \kappa\ell_k\right)\right)\mathbf{f}'\right) \end{bmatrix},$$

and the J tests have $n - 1$ degrees of freedom.

If the factors are not traded then we discard the moment condition (B6) that prices the factors. At the same time, we can re-express the least squares constraint as

$$\boldsymbol{\phi} = \kappa\ell_n + \mathbf{B}\boldsymbol{\varkappa}, \qquad \kappa = 1/c, \qquad \boldsymbol{\varkappa} = -\mathbf{d}/c.$$

when $c \neq 0$.

In this context, the influence functions of the SDF and regression approaches with non-traded factors are

$$\mathbf{g}_S\left(\mathbf{R}, \mathbf{f}; a, \mathbf{b}\right) = \left[\mathbf{R}\left(a + \mathbf{f}'\mathbf{b}\right) - \ell_n\right],$$

$$\mathbf{g}_R\left(\mathbf{R}, \mathbf{f}; \mathbf{B}, \kappa, \boldsymbol{\varkappa}\right) = \begin{bmatrix} \left(\mathbf{R} - \kappa\ell_n\right) - \mathbf{B}\left(\mathbf{f} + \boldsymbol{\varkappa}\right) \\ vec\left(\left(\left(\mathbf{R} - \kappa\ell_n\right) - \mathbf{B}\left(\mathbf{f} + \boldsymbol{\varkappa}\right)\right)\mathbf{f}'\right) \end{bmatrix},$$

and the J tests will have $n - (k + 1)$ degrees of freedom.

C Proofs

All proofs consider the multifactor context of Appendix B instead of the simplifying single factor set up in the main text. In addition, the proofs of Proposition 1 and 2 do not rely on any particular normalisation since they are irrelevant for single-step methods, even though the proposition statements in the main text refer to the usual normalisations in empirical work.

Proposition 1:

Let us define an extended regression system that adds the estimation of $(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ to $\mathbf{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B})$ defined in (B3). Specifically

$$
\mathfrak{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B}, \boldsymbol{\mu}, vech(\boldsymbol{\Gamma})) = \begin{bmatrix} \mathbf{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B}) \\ \mathbf{f} - \boldsymbol{\mu} \\ vech(\mathbf{ff}' - \boldsymbol{\Gamma}) \end{bmatrix} = \begin{bmatrix} \mathbf{r} - \mathbf{Bf} \\ vec((\mathbf{r} - \mathbf{Bf})\mathbf{f}') \\ \mathbf{f} - \boldsymbol{\mu} \\ vech(\mathbf{ff}' - \boldsymbol{\Gamma}) \end{bmatrix}.
$$

Importantly, by adding the exactly identified parameters $(\boldsymbol{\mu}, vech(\boldsymbol{\Gamma}))$, $\mathfrak{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B}, \boldsymbol{\mu}, vech(\boldsymbol{\Gamma}))$ will be numerically equivalent to $\mathbf{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B})$ in terms of both the estimates of the original parameters $\mathbf{B}$ and the J test.

If we choose $(a, \mathbf{b})$ such that

$$
a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b} = \mathbf{0},
$$

then we can carry out the following transformations of the system $\mathfrak{h}_R(\mathbf{r}, \mathbf{f}; \mathbf{B}, \boldsymbol{\mu}, vech(\boldsymbol{\Gamma}))$

$$
\begin{bmatrix} a\mathbf{I}_n & \mathbf{b}' \otimes \mathbf{I}_n & a\mathbf{B} & (\mathbf{b}' \otimes \mathbf{B})\mathcal{D} \end{bmatrix} \begin{bmatrix} \mathbf{r} - \mathbf{Bf} \\ vec((\mathbf{r} - \mathbf{Bf})\mathbf{f}') \\ \mathbf{f} - \boldsymbol{\mu} \\ vech(\mathbf{ff}' - \boldsymbol{\Gamma}) \end{bmatrix} =
$$

$$
\left[\mathbf{r}(a + \mathbf{b}'\mathbf{f})\right] - \mathbf{B}[a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b}] = \mathbf{r}(a + \mathbf{b}'\mathbf{f}),
$$

where the matrix $\mathcal{D}$ denotes the appropriate duplication matrix, that is, the matrix such that $vec(\cdot) = \mathcal{D}vech(\cdot)$ (see Magnus and Neudecker (1990)). Similarly

$$
\begin{bmatrix} \mathbf{0} & \mathbf{0} & a\mathbf{I}_k & (\mathbf{b}' \otimes \mathbf{I}_k)\mathcal{D} \end{bmatrix} \begin{pmatrix} \mathbf{r} - \mathbf{Bf} \\ vec((\mathbf{r} - \mathbf{Bf})\mathbf{f}') \\ \mathbf{f} - \boldsymbol{\mu} \\ vech(\mathbf{ff}' - \boldsymbol{\Gamma}) \end{pmatrix} =
$$

$$
\left[\mathbf{f}(a + \mathbf{f}'\mathbf{b})\right] - [a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b}] = \mathbf{f}(a + \mathbf{f}'\mathbf{b}).
$$

As we mentioned before, single-step methods are numerically invariant to normalisation, bijective reparametrisations and parameter-dependent linear transformations of the moment conditions. Therefore, for a given choice of HAC weighting matrix CU-GMM renders the extended regression system $\mathfrak{h}_R\left(\mathbf{r},\mathbf{f};\mathbf{B},\boldsymbol{\mu},vech\left(\boldsymbol{\Gamma}\right)\right)$ and the system

$$
\mathfrak{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b},\mathbf{B},vech\left(\boldsymbol{\Gamma}\right)\right) = \begin{bmatrix} \mathbf{r}\left(a+\mathbf{f}'\mathbf{b}\right) \\ \mathbf{f}\left(a+\mathbf{f}'\mathbf{b}\right) \\ vec\left(\left(\mathbf{r}-\mathbf{B}\mathbf{f}\right)\mathbf{f}'\right) \\ vech\left(\mathbf{f}\mathbf{f}'-\boldsymbol{\Gamma}\right) \end{bmatrix}
$$

numerically equivalent. In particular, the estimates of $\mathbf{B}$ and $vech\left(\boldsymbol{\Gamma}\right)$ are the same, the implied $\boldsymbol{\mu} = -\boldsymbol{\Gamma}\mathbf{b}/a$ is the same, and the J test is the same.

Given the definition of $\mathbf{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b}\right)$ in (B2), the last system can also be expressed as

$$
\mathfrak{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b},\mathbf{B},vech\left(\boldsymbol{\Gamma}\right)\right) = \begin{bmatrix} \mathbf{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b}\right) \\ vec\left(\left(\mathbf{r}-\mathbf{B}\mathbf{f}\right)\mathbf{f}'\right) \\ vech\left(\mathbf{f}\mathbf{f}'-\boldsymbol{\Gamma}\right) \end{bmatrix},
$$

where the influence functions added to $\mathbf{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b}\right)$ are exactly identified for $(\mathbf{B},vech\left(\boldsymbol{\Gamma}\right))$. Thus $\mathfrak{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b},\mathbf{B},vech\left(\boldsymbol{\Gamma}\right)\right)$ is numerically equivalent to relying on the first block $\mathbf{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b}\right)$ in terms of both the original parameters $(a,\mathbf{b})$ and the J test. Therefore, single-step methods render the systems $\mathbf{h}_R\left(\mathbf{r},\mathbf{f};\mathbf{B}\right)$ and $\mathbf{h}_S\left(\mathbf{r},\mathbf{f};a,\mathbf{b}\right)$ numerically equivalent. $\qquad\square$

Lemma 1:

We assume that the vector $\mathbf{y} = \left(\mathbf{f}',\mathbf{r}'\right)'$ follows an elliptical distribution, and denote the corresponding coefficient of multivariate excess kurtosis as κ, which is equal to $\kappa = 2/(v-4)$ in the case of Student t with v degrees of freedom, and $\kappa = 0$ under normality (see Fang, Kotz and Ng (1990) and the references therein for further details).

Let us order the estimating functions in (4) for a multifactor model as

$$
\mathbf{h}\left(\mathbf{y};\boldsymbol{\lambda}\right) = \begin{bmatrix} \mathbf{f}\left(1-\mathbf{f}'\boldsymbol{\lambda}\right) \\ \mathbf{r}\left(1-\mathbf{f}'\boldsymbol{\lambda}\right) \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1(\mathbf{f};\boldsymbol{\lambda}) \\ \mathbf{h}_2(\mathbf{r},\mathbf{f};\boldsymbol{\lambda}) \end{bmatrix}.
$$

Thus, we can define the relevant Jacobian as

$$
\mathbf{D} = E\left[\frac{\partial\mathbf{h}\left(\mathbf{y};\boldsymbol{\lambda}\right)}{\partial\boldsymbol{\lambda}'}\right] = \begin{pmatrix} \boldsymbol{\Gamma} \\ E\left(\mathbf{r}\mathbf{f}'\right) \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{pmatrix},
$$

and similarly we can decompose the relevant asymptotic covariance matrix as

$$
\mathbf{S} = avar\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{h}\left(\mathbf{y}_t;\boldsymbol{\lambda}\right)\right] = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix}.
$$

If we apply Lemma D1 in Peñaranda and Sentana (2010), then we find

$$\mathbf{S}_{11} = \omega_1 \mathbf{\Gamma} + \omega_2 \boldsymbol{\mu}\boldsymbol{\mu}',$$

$$\omega_1 = (1 - \mathcal{H})(1 + \kappa\mathcal{H}), \qquad \omega_2 = -2(1 - \mathcal{H})^2 + (3\mathcal{H}^2 - 5\mathcal{H} + 2)\kappa,$$

where $\mathcal{H} = E(\mathbf{y})' E^{-1}(\mathbf{yy}') E(\mathbf{y})$, and similarly

$$\mathbf{S}_{21} = \omega_1 E(\mathbf{rf}') + \omega_2 E(\mathbf{r})\boldsymbol{\mu}'.$$

Thus, we only need to check that condition (C1) in Lemma C1 in Peñaranda and Sentana (2010) holds, which in our context becomes

$$\mathbf{D}_2 \mathbf{D}_1^{-1} \mathbf{S}_{11} = \mathbf{S}_{21}.$$

This restriction will be satisfied because

$$\mathbf{D}_2 \mathbf{D}_1^{-1} \mathbf{S}_{11} = E(\mathbf{rf}') \mathbf{\Gamma}^{-1} \left[\omega_1 \mathbf{\Gamma} + \omega_2 \boldsymbol{\mu}\boldsymbol{\mu}' \right] = \omega_1 E(\mathbf{rf}') + \omega_2 E(\mathbf{rf}') \mathbf{\Gamma}^{-1} \boldsymbol{\mu}\boldsymbol{\mu}'$$

and, since $E(\mathbf{r}) = E(\mathbf{rf}') \mathbf{\Gamma}^{-1} \boldsymbol{\mu}$ under the null of tangency,

$$\omega_1 E(\mathbf{rf}') + \omega_2 E(\mathbf{rf}') \mathbf{\Gamma}^{-1} \boldsymbol{\mu}\boldsymbol{\mu}' = \omega_1 E(\mathbf{rf}') + \omega_2 E(\mathbf{r})\boldsymbol{\mu}' = \mathbf{S}_{21}.$$

Therefore, the linear combinations of the moment conditions in $E(\mathbf{h}(\mathbf{y};\boldsymbol{\lambda})) = \mathbf{0}$ that provide the most efficient estimators of $\boldsymbol{\lambda}$ will be given by

$$E(\mathbf{ff}'\boldsymbol{\lambda} - \mathbf{f}) = \mathbf{0}.$$

$\square$

Proposition 2:

Let us define a extended regression system that adds the estimation of $(\boldsymbol{\mu}, \mathbf{\Gamma})$ to the influence functions $\mathbf{g}_R(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d})$ defined in (B5),

$$\mathfrak{g}_R(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}, \boldsymbol{\mu}, vech(\mathbf{\Gamma})) = \begin{bmatrix} \mathbf{g}_R(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}) \\ \mathbf{f} - \boldsymbol{\mu} \\ vech(\mathbf{ff}' - \mathbf{\Gamma}) \end{bmatrix} = \begin{bmatrix} \mathbf{r} - \mathbf{P}(\mathbf{d} - c\mathbf{f}) \\ vec((\mathbf{r} - \mathbf{P}(\mathbf{d} - c\mathbf{f}))\mathbf{f}') \\ \mathbf{f} - \boldsymbol{\mu} \\ vech(\mathbf{ff}' - \mathbf{\Gamma}) \end{bmatrix}.$$

We are adding exactly identified parameters $(\boldsymbol{\mu}, vech(\mathbf{\Gamma}))$, and hence $\mathfrak{g}_R(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}, \boldsymbol{\mu}, vech(\mathbf{\Gamma}))$ is numerically equivalent to the influence functions in $\mathbf{g}_R(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d})$ in terms of both the original parameter estimates and the J test.

If we choose $(a, \mathbf{b})$ such that

$$\begin{pmatrix} a \\ \mathbf{b} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{\mu}' \\ \boldsymbol{\mu} & \boldsymbol{\Gamma} \end{pmatrix}^{-1} \begin{pmatrix} c \\ \mathbf{d} \end{pmatrix} = \begin{pmatrix} c\left(1 + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right) - \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{d} \\ \boldsymbol{\Sigma}^{-1}\left(\mathbf{d} - c\boldsymbol{\mu}\right) \end{pmatrix}$$

then we can compute the following $n \times 1$ transformation of $\mathfrak{g}_R\left(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$:

$$\begin{bmatrix} a\mathbf{I}_n & \mathbf{b}' \otimes \mathbf{I}_n & \mathbf{P}\left(ac\mathbf{I}_k + \mathbf{d}\mathbf{b}'\right) & -c\left(\mathbf{b}' \otimes \mathbf{P}\right)\mathcal{D} \end{bmatrix} \begin{bmatrix} \mathbf{r} - \mathbf{P}\left(\mathbf{d} - c\mathbf{f}\right) \\ vec\left(\left(\mathbf{r} - \mathbf{P}\left(\mathbf{d} - c\mathbf{f}\right)\right)\mathbf{f}'\right) \\ \mathbf{f} - \boldsymbol{\mu} \\ vech\left(\mathbf{f}\mathbf{f}' - \boldsymbol{\Gamma}\right) \end{bmatrix} =$$

$$\left[\mathbf{r}\left(a + \mathbf{f}'\mathbf{b}\right)\right] - \mathbf{P}\left[\mathbf{d}\left(a + \mathbf{b}'\boldsymbol{\mu}\right) - c\left(a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b}\right)\right] = \mathbf{r}\left(a + \mathbf{f}'\mathbf{b}\right),$$

where the matrix $\mathcal{D}$ denotes the corresponding duplication matrix.

Accordingly, we can also reparametrise $(c, \mathbf{d})$ in terms of the other parameters in the second block of influence functions and then construct the system

$$\mathfrak{g}_S\left(\mathbf{r}, \mathbf{f}; a, \mathbf{b}, \mathbf{P}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right) = \begin{bmatrix} \mathbf{r}\left(a + \mathbf{f}'\mathbf{b}\right) \\ vec\left(\left(\mathbf{r} - \mathbf{P}\left(\mathbf{d} - c\mathbf{f}\right)\right)\mathbf{f}'\right) \\ \mathbf{f} - \boldsymbol{\mu} \\ vech\left(\mathbf{f}\mathbf{f}' - \boldsymbol{\Gamma}\right) \end{bmatrix},$$

where

$$\begin{pmatrix} c \\ \mathbf{d} \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{\mu}' \\ \boldsymbol{\mu} & \boldsymbol{\Gamma} \end{pmatrix} \begin{pmatrix} a \\ \mathbf{b} \end{pmatrix}.$$

Under CU-GMM, and a specific choice of HAC estimator, $\mathfrak{g}_S\left(\mathbf{r}, \mathbf{f}; a, \mathbf{b}, \mathbf{P}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$ provides the same estimates and J test as $\mathfrak{g}_R\left(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$ because single-step procedures are numerically invariant to normalisation, bijective reparametrisations and parameter-dependent linear transformations of the moment conditions. As a result, the estimator of $(c, \mathbf{d})$ obtained from $\mathfrak{g}_R\left(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$ and $(a, \mathbf{b})$ from $\mathfrak{g}_S\left(\mathbf{r}, \mathbf{f}; a, \mathbf{b}, \mathbf{P}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$ coincide with their implied counterparts in the other system.

This last system can be related to the influence function $\mathbf{g}_S\left(\mathbf{r}, \mathbf{f}; a, \mathbf{b}\right)$ defined in (B4), where the influence functions that are added are exactly identified for $\left(\mathbf{P}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$ given $(a, \mathbf{b})$. Thus $\mathfrak{g}_S\left(\mathbf{r}, \mathbf{f}; a, \mathbf{b}, \mathbf{P}, \boldsymbol{\mu}, vech\left(\boldsymbol{\Gamma}\right)\right)$ is numerically equivalent to relying on $\mathbf{r}\left(a + \mathbf{b}'\mathbf{f}\right)$ in terms of both the estimates of the common parameters $(a, \mathbf{b})$ and the J test. Therefore, single-step methods render the systems $\mathbf{g}_R\left(\mathbf{r}, \mathbf{f}; \mathbf{P}, c, \mathbf{d}\right)$ and $\mathbf{g}_S\left(\mathbf{r}, \mathbf{f}; a, \mathbf{b}\right)$ numerically equivalent. $\quad\square$

Lemma 2:

We extend the results in Appendix D in Peñaranda and Sentana (2010) for elliptical distributions to the case of nontraded factors. The optimal moments are given by the linear combinations $\mathbf{D}'\mathbf{S}^{-1}\bar{\mathbf{h}}_T(\boldsymbol{\theta})$. The uncentred SDF method has the following long-run variance under the null

$$avar\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{g}_U(\mathbf{r}_t,\mathbf{f}_t;\boldsymbol{\lambda})\right] = \frac{(1+\kappa)H_1+1}{(H_2+1)^2}E(\mathbf{rr}') - \frac{\kappa H_1 + 2(1-\kappa)}{(H_2+1)^2}E(\mathbf{r})E(\mathbf{r})',$$

where $H_1 = \boldsymbol{\delta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta}$ and $H_2 = \boldsymbol{\delta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. This asymptotic variance represents a multifactor and elliptical extension of the Gaussian computations in Jagannathan and Wang (2002).

Given that $\mathbf{D} = -E(\mathbf{rf}')$ for the uncentred SDF method, the optimal moments are then proportional to the linear transformation

$$E(\mathbf{fr}')\left[E(\mathbf{rr}') - \omega E(\mathbf{r})E(\mathbf{r})'\right]^{-1}, \qquad \omega = \frac{\kappa H_1 + 2(1-\kappa)}{(1+\kappa)H_1+1}.$$

Computing the inverse, we obtain

$$E(\mathbf{fr}')\left[E^{-1}(\mathbf{rr}') + \frac{\omega}{1-\omega E(\mathbf{r})'E^{-1}(\mathbf{rr}')E(\mathbf{r})}E^{-1}(\mathbf{rr}')E(\mathbf{r})E(\mathbf{r})'E^{-1}(\mathbf{rr}')\right]$$

and imposing the null hypothesis $E(\mathbf{r}) = E(\mathbf{rf}')\boldsymbol{\lambda}$, we get

$$E(\mathbf{fr}')\left[E^{-1}(\mathbf{rr}') + \frac{\omega}{1-\omega\boldsymbol{\lambda}'E(\mathbf{fr}')E^{-1}(\mathbf{rr}')E(\mathbf{rf}')\boldsymbol{\lambda}}E^{-1}(\mathbf{rr}')E(\mathbf{rf}')\boldsymbol{\lambda}\boldsymbol{\lambda}'E(\mathbf{fr}')E^{-1}(\mathbf{rr}')\right]$$

$$=\left[\mathbf{I}_k + \frac{\omega}{1-\omega\boldsymbol{\lambda}'E(\mathbf{fr}')E^{-1}(\mathbf{rr}')E(\mathbf{rf}')\boldsymbol{\lambda}}E(\mathbf{fr}')E^{-1}(\mathbf{rr}')E(\mathbf{rf}')\boldsymbol{\lambda}\boldsymbol{\lambda}'\right]E(\mathbf{fr}')E^{-1}(\mathbf{rr}').$$

Since the $k\times k$ matrix in brackets has full rank, we can conclude that the optimal estimator of $\boldsymbol{\lambda}$ solves the sample moments

$$\frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{r}_t^+\left(1-\mathbf{f}_t'\boldsymbol{\lambda}_T\right)\right] = \mathbf{0},$$

with

$$\mathbf{r}_t^+ = E(\mathbf{fr}')E^{-1}(\mathbf{rr}')\mathbf{r}_t.$$

Finally, note that to implement this optimal estimator in practice, we need a consistent estimation of $E(\mathbf{fr}')$ and $E(\mathbf{rr}')$, which we can easily obtain from their sample counterparts. $\square$

Counterpart to Proposition 1 with gross returns:

We can obtain the SDF influence functions as a transformation and reparametrisation of the regression influence functions augmented with exactly identified moment conditions. In

particular, if we choose $(a, \mathbf{b})$ such that

$$a + \mathbf{b}'\boldsymbol{\mu} = 1/\kappa,$$

$$\boldsymbol{\mu}a + \boldsymbol{\Gamma}\mathbf{b} = \ell_k,$$

when $\kappa \neq 0$, then

$$\begin{bmatrix} a\mathbf{I}_n & \mathbf{b}' \otimes \mathbf{I}_n & a\mathbf{B} + \kappa(\ell_n - \mathbf{B}\ell_k)\mathbf{b}' & (\mathbf{b}' \otimes \mathbf{B})\mathcal{D} \end{bmatrix}$$

$$\times \begin{bmatrix} (\mathbf{R} - \kappa\ell_n) - \mathbf{B}(\mathbf{f} - \kappa\ell_k) \\ vec\left(((\mathbf{R} - \kappa\ell_n) - \mathbf{B}(\mathbf{f} - \kappa\ell_k))\mathbf{f}'\right) \\ \mathbf{f} - \boldsymbol{\mu} \\ vech\left(\mathbf{ff}' - \boldsymbol{\Gamma}\right) \end{bmatrix}$$

$$= \left[\mathbf{R}\left(a + \mathbf{b}'\mathbf{f}\right)\right] - \left[\kappa(\ell_n - \mathbf{B}\ell_k)\left(a + \mathbf{b}'\boldsymbol{\mu}\right) + \mathbf{B}\left(\boldsymbol{\mu}a + \boldsymbol{\Gamma}\mathbf{b}\right)\right]$$

$$= \left[\mathbf{R}\left(a + \mathbf{b}'\mathbf{f}\right)\right] - \left[(\ell_n - \mathbf{B}\ell_k) + \mathbf{B}\ell_k\right] = \mathbf{R}\left(a + \mathbf{b}'\mathbf{f}\right) - \ell_n,$$

where the matrix $\mathcal{D}$ denotes the appropriate duplication matrix, and

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & a\mathbf{I}_k & (\mathbf{b}' \otimes \mathbf{I}_k)\mathcal{D} \end{bmatrix} \begin{bmatrix} (\mathbf{R} - \kappa\ell_n) - \mathbf{B}(\mathbf{f} - \kappa\ell_k) \\ vec\left(((\mathbf{R} - \kappa\ell_n) - \mathbf{B}(\mathbf{f} - \kappa\ell_k))\mathbf{f}'\right) \\ \mathbf{f} - \boldsymbol{\mu} \\ vech\left(\mathbf{ff}' - \boldsymbol{\Gamma}\right) \end{bmatrix} =$$

$$\left[\mathbf{f}\left(a + \mathbf{f}'\mathbf{b}\right)\right] - \left[a\boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{b}\right] = \mathbf{f}\left(a + \mathbf{f}'\mathbf{b}\right) - \ell_k.$$

Given these relationships, we can apply similar arguments to those in the proof of Proposition 1. □

Counterpart to Proposition 2 with gross returns:

We can obtain the SDF influence functions as a transformation and reparametrisation of the regression influence functions augmented with exactly identified moment conditions. In particular, if we choose $(a, \mathbf{b})$ such that

$$a + \mathbf{b}'\boldsymbol{\mu} = 1/\kappa,$$

$$\boldsymbol{\mu}a + \boldsymbol{\Gamma}\mathbf{b} = -\varkappa/\kappa,$$

when $\kappa \neq 0$, then

$$\left[\begin{array}{cccc} a\mathbf{I}_n & \mathbf{b}' \otimes \mathbf{I}_n & a\mathbf{B} + (\kappa\ell_n + \mathbf{B}\varkappa)\,\mathbf{b}' & (\mathbf{b}' \otimes \mathbf{B})\,\mathcal{D} \end{array} \right]$$

$$\times \left[\begin{array}{c} (\mathbf{R} - \kappa\ell_n) - \mathbf{B}\,(\mathbf{f} + \varkappa) \\ vec\,(((\mathbf{R} - \kappa\ell_n) - \mathbf{B}\,(\mathbf{f} + \varkappa))\,\mathbf{f}') \\ \mathbf{f} - \boldsymbol{\mu} \\ vech\,(\mathbf{f}\mathbf{f}' - \boldsymbol{\Gamma}) \end{array} \right]$$

$$= \left[\mathbf{R}\,(a + \mathbf{b}'\mathbf{f}) \right] - \left[(\kappa\ell_n + \mathbf{B}\varkappa)\,(a + \mathbf{b}'\boldsymbol{\mu}) + \mathbf{B}\,(\boldsymbol{\mu}a + \boldsymbol{\Gamma}\mathbf{b}) \right]$$

$$= \left[\mathbf{R}\,(a + \mathbf{b}'\mathbf{f}) \right] - \left[\ell_n + \mathbf{B}\,(\varkappa/\kappa) + \mathbf{B}\,(-\varkappa/\kappa) \right] = \mathbf{R}\,(a + \mathbf{b}'\mathbf{f}) - \ell_n,$$

where the matrix $\mathcal{D}$ denotes the appropriate duplication matrix. Given this relationship, we can apply similar arguments to those in the proof of Proposition 2. □

Table 1: Descriptive statistics in annual % terms, 1953-2002

	1	2	3	4	5	6	7	8
Mean	-2.336	-0.873	-0.747	0.329	-0.151	-0.213	2.988	2.031
SD	6.346	6.628	6.614	8.415	7.443	8.121	8.090	12.417
Sharpe ratio	-0.368	-0.132	-0.113	0.039	-0.020	-0.026	0.369	0.164

Real excess returns on the 8 currency portfolios in Lustig and Verdelhan (2007) for a US investor. Portfolio 1 contains currencies with the lowest interest rates, while portfolio 8 contains currencies with the highest interest rates.

Table 2: Empirical evaluation of the CAPM

A. Risk prices and J-tests

	Uncentred SDF (λ)		Centred SDF (τ)		Regression (μ)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
MK	0.038	0.043	0.026	0.122	14.627	14.974
	(0.004)	(0.005)	(0.009)	(0.070)	(2.071)	(1.694)
J test	28.068	21.375	46.046	21.375	28.202	21.375
	(0.000)	(0.006)	(0.000)	(0.006)	(0.000)	(0.006)

B. Pricing errors

	Uncentred SDF ($\boldsymbol{\pi}$)		Centred SDF ($\boldsymbol{\psi}$)		Regression ($\boldsymbol{\alpha}$)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
1	-1.084	-0.929	-1.908**	-2.622	-1.817**	-1.369*
2	-0.342	-0.277	-0.670	-0.781	-0.511	-0.421
3	-0.609	-0.592	-0.789	-1.671	-0.882	-0.784
4	1.384	1.514	1.108	4.271	0.548	0.461
5	-0.270	-0.284	-0.260	-0.802	-0.449	-0.310
6	-0.259	-0.265	-0.283	-0.748	-0.335	-0.368
7	2.664**	2.624*	3.311**	7.404*	2.479**	2.156***
8	0.819	0.670	1.575	1.889	0.815	0.928
MK	-7.435***	-9.216***	-1.581**	-26.002***	-7.652***	-7.999***

Panel A displays estimates of "prices of risk" (λ, τ or μ) with standard errors in parenthesis, and the J test with p-values in parenthesis. We implement each method by both two-step and CU-GMM. Panel B displays the corresponding estimates of "pricing errors" (π, ψ or α) for each of the 8 currency portfolios and the pricing the factor, with their significance level (*: 10%, **: 5%, ***: 1%). The α of the traded factor is computed as the difference between its sample mean and its GMM estimate.

Table 3: Empirical evaluation of the Fama-French model

A. Risk prices and J-tests

	Uncentred SDF ($\boldsymbol{\lambda}$)		Centred SDF ($\boldsymbol{\tau}$)		Regression ($\boldsymbol{\mu}$)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
MK	0.027	0.076	0.031	0.380	12.449	9.105
	(0.005)	(0.018)	(0.009)	(0.261)	(2.261)	(2.577)
SMB	0.010	-0.014	0.005	-0.071	4.732	3.394
	(0.008)	(0.012)	(0.010)	(0.071)	(1.794)	(1.760)
HML	0.034	0.041	0.033	0.205	9.614	3.926
	(0.007)	(0.014)	(0.011)	(0.144)	(1.772)	(1.982)
J test	27.917	21.528	60.635	21.528	27.335	21.528
	(0.000)	(0.006)	(0.000)	(0.006)	(0.001)	(0.006)

B. Pricing errors

	Uncentred SDF ($\boldsymbol{\pi}$)		Centred SDF ($\boldsymbol{\psi}$)		Regression ($\boldsymbol{\alpha}$)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
1	-0.296	0.852	-1.200	4.278	-1.140	0.606
2	-0.645	0.160	-0.892	0.805	-0.352	-0.116
3	0.544	0.571	0.186	2.870	-0.169	0.804
4	1.796*	2.499*	1.869	12.551*	1.157	1.520*
5	1.125	0.536	0.905	2.691	0.154	0.764
6	1.308	0.974	1.092	4.893	0.301	0.968
7	3.034***	2.621	4.121***	13.163	2.919***	1.238
8	1.646	-0.229	2.211	-1.151	1.812	2.032*
MK	-4.278***	-20.484***	-2.601***	-102.897***	-5.474***	-2.130
SMB	-1.856***	-1.116	-0.282	-5.606	-2.300***	-0.962
HML	-2.053***	-3.487*	0.240	-17.515*	-3.940***	1.748***

Panel A displays estimates of "prices of risk" (λ, τ or μ) with standard errors in parenthesis, and the J test with p-values in parenthesis. We implement each method by both two-step and CU-GMM. Panel B displays the corresponding estimates of "pricing errors" (π, ψ or α) for each of the 8 currency portfolios and the three pricing factors, with their significance level (*: 10%, **: 5%, ***: 1%). The α of a traded factor is computed as the difference between its sample mean and its GMM estimate.

Table 4: Empirical evaluation of the (linearised) CCAPM

A. Risk prices and J-tests

	Uncentred SDF (λ)		Centred SDF (τ)		Regression (δ)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
Non-dur.	0.488	0.495	1.201	4.388	5.308	5.566
	(0.043)	(0.042)	(0.181)	(5.727)	(1.483)	(2.056)
J test	5.711	5.663	91.624	5.663	5.627	5.663
	(0.574)	(0.580)	(0.000)	(0.580)	(0.584)	(0.580)

B. Pricing errors

	Uncentred SDF ($\boldsymbol{\pi}$)		Centred SDF ($\boldsymbol{\psi}$)		Regression ($\boldsymbol{\alpha}$)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
1	-0.667	-0.645	-1.191*	-5.715	-0.337	-0.368
2	-0.992*	-0.994*	-2.272***	-8.809*	-0.885	-0.900
3	-0.450	-0.446	-0.963	-3.952	-0.909	-0.935
4	-0.108	-0.114	-0.330	-1.012	-0.548	-0.567
5	-0.689	-0.696	-1.664**	-6.166	-0.389	-0.440
6	-0.319	-0.320	-0.743	-2.838	-0.734	-0.732
7	-0.416	-0.462	-1.598**	-4.091	-0.689*	-0.743*
8	0.399	0.377	0.590	3.340	-1.230	-1.170

Panel A displays estimates of "prices of risk" (λ, τ or δ) with standard errors in parenthesis, and the J test with p-values in parenthesis. We implement each method by both two-step and CU-GMM. Panel B displays the corresponding estimates of "pricing errors" (π, ψ or α) for each of the 8 currency portfolios, with their significance level (*: 10%, **: 5%, ***: 1%).

Table 5: Empirical evaluation of the (linearised) DCAPM

A. Risk prices and J-tests

	Uncentred SDF ($\boldsymbol{\lambda}$)		Centred SDF ($\boldsymbol{\tau}$)		Regression ($\boldsymbol{\delta}$)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
Non-dur.	-0.081	-0.028	-1.134	-0.285	4.146	3.101
	(0.272)	(0.091)	(0.345)	(0.977)	(2.559)	(2.256)
Durables	0.275	0.266	0.894	2.724	9.266	8.775
	(0.131)	(0.051)	(0.230)	(2.733)	(5.464)	(5.633)
J test	2.841	4.458	16.986	4.458	5.220	4.458
	(0.829)	(0.615)	(0.009)	(0.615)	(0.516)	(0.615)

B. Pricing errors

	Uncentred SDF ($\boldsymbol{\pi}$)		Centred SDF ($\boldsymbol{\psi}$)		Regression ($\boldsymbol{\alpha}$)	
	2S-GMM	CU-GMM	2S-GMM	CU-GMM	2S-GMM	CU-GMM
1	-0.730	-0.607	-1.742***	-6.221	-0.034	-0.095
2	-0.608	-0.628*	-0.417	-6.438*	0.024	-0.010
3	-0.837	-0.803	-2.098**	-8.224	0.370	0.134
4	-0.925	-0.931	-2.740***	-9.538	0.435	-0.009
5	-0.553	-0.597	-0.610	-6.112	0.216	0.174
6	-0.796	-0.788	-2.072***	-8.076	0.672	0.252
7	-0.696	-0.934*	-0.816***	-9.567*	0.506	0.039
8	-0.348	-0.443	-1.363	-4.542	0.034	-0.189

Panel A displays estimates of "prices of risk" (λ, τ or δ) with standard errors in parenthesis, and the J test with p-values in parenthesis. We implement each method by both two-step and CU-GMM. Panel B displays the corresponding estimates of "pricing errors" (π, ψ or α) for each of the 8 currency portfolios, with their significance level (*: 10%, **: 5%, ***: 1%).

Table 6: Non-traded factor loadings and risk premia

	Zero uncentred second moments	Zero centred second moments	Zero second moments and risk premia
Non-dur.	34.851	8.390	53.039
	(0.000)	(0.396)	(0.000)
Durables	31.363	6.714	55.934
	(0.000)	(0.568)	(0.000)
Joint	49.373	14.124	77.742
	(0.000)	(0.590)	(0.000)

Each column provides the single-step GMM overidentification tests of the following null hypotheses: first column, $E(\mathbf{r}f) = \mathbf{0}$; second column, $E\left[(\mathbf{r} - \boldsymbol{\varphi})\,f\right] = \mathbf{0}$, where φ is a parameter to estimate; third column, any of the two previous hypothesis jointly with $E(\mathbf{r}) = \mathbf{0}$. The tests are computed for each consumption factor and also for both of them jointly.