

Cumulant Instrument Estimators for Hedge Fund Return Models with Errors in Variable

François-Éric Racicot
Raymond Théoret

WORKING PAPER
WP.2012.05

October 2012
ISSN 0701-3086

Cumulant Instrument Estimators for Hedge Fund Return Models with Errors in Variables[†]

François-Éric Racicot^{a,*}, Raymond Théoret^b

^a *Telfer School of Management, University of Ottawa, 55 Laurier Avenue East, Ottawa, Ontario; Chaire d'information financière et organisationnelle, ESG-UQAM; Laboratory for Research in Statistics and Probability, LRSP;*

^b *Université du Québec (Montréal), École des sciences de la gestion, 315 est Ste-Catherine, Montréal, Québec; Université du Québec (Outaouais); Chaire d'information financière et organisationnelle, ESG-UQAM.*

* Corresponding author. Tel: +1(613) 562-5800, ext. 4757.

E-mail addresses: Racicot@telfer.uottawa.ca (F-É. Racicot), raymond.theoret@uqam.ca (R. Théoret).

[†] **Acknowledgements.** *We would like to thank Ramazan Gençay and the participants of the Southern Finance Association held in November 2011.*

Cumulant Instrument Estimators for Hedge Fund Return Models with Errors in Variables[†]

Abstract

We revisit the factors incorporated in asset pricing models following the recent developments in financial markets, i.e. the rise of shadow banking and the change in the transmission channel of monetary policy. We propose two versions of the Fung and Hsieh (2004) hedge fund return model, especially an augmented market model which accounts for the new dynamics of financial markets and the procyclicality of hedge fund returns. We run these models with an innovative Hausman procedure tackling the measurement errors embedded in the models factor loadings. Our empirical method also allows confronting the drawbacks of the instruments used to estimate hedge fund asset pricing models.

Keywords: Cumulant instruments; Hausman test; Fung & Hsieh model; GMM.

JEL classification: C13, C19, C58, G12, G23.

1. Introduction

The first generation of hedge fund return models was based on traditional asset pricing models like the CAPM, the CCAPM or, more frequently, the Fama and French model or a multifactorial version of this model (Fama and French, 1997, 1998; Do *et al.*, 2005; Jarrow and Protter, 2011). However, these models bypass important aspects of hedge fund strategies like their involvement in options and structured products which entails non-linear payoffs better captured by contingent claim risk factors rather than by the conventional ones like the market risk premium (Heuson and Hutchinson, 2012). To tackle this problem, Fung and Hsieh (2004) proposed an extension of the Sharpe (1992) asset class factor model which, in addition to the market risk premium and a factor related to the small firm anomaly, includes factors mimicking hedge fund option-like trading strategies, essentially lookback option return factors.

In line with the evolution of financial markets and shadow banking¹ – hedge funds being a key element of this system – other factors have gained importance to explain stock returns. The term structure spread and the interest rate on long-term bonds weight more in the panoply of monetary policy indicators since the level of short-term interest rates is near its zero floor (Estrela and Hardouvelis, 1991; Adrian and Shin, 2010; Disyatat, 2010; Gambacorta and Marques-Ibanez, 2011). The term structure spread also enters in the computation of “haircuts²”, a significant dimension of hedge fund leverage and funding costs. It is also much related to the hedge fund return and leverage cycles (McGuire *et al.*, 2005). Stock return volatility is finally a very relevant factor for hedge fund return analysis since it is the key determinant of derivative prices. Even abstracting from the close link

¹ Market-oriented banking is another appellation of shadow banking. This appellation refers to the importance of market funding for the institutions composing this new form of banking, as opposed to the traditional deposit funding by banks. On the asset side, shadow bankers are also very involved in securitization and in investment in structured products.

² The term “haircut” refers to the percentage of equity required in market funding. It is the inverse of the leverage associated with a funding operation.

between stock volatility and option prices, stock market returns would be now more sensitive to market volatility than to the market risk premium (Campbell *et al.*, 2001).

To account for these developments, we propose parsimonious models of hedge fund returns. Our contribution is threefold. First, we modify the Fung and Hsieh (2004) model by combining the lookback option return variables included in their factorial model into one principal component. Second, in an augmented version of the market model³, we relate hedge fund returns to increasingly important determinants of stock returns, i.e., the term structure spread and the *VIX*, an indicator of the implicit volatility of stock returns. Third, it is obvious that hedge fund models, regardless of their precision, contain specification errors. These errors may be due to omitted variables or inherent to the measurement of their risk factors (Fama and McBeth, 1973; Chen *et al.*, 1986; Shanken, 1992; Campbell *et al.*, 1997; Lettau and Ludvigson, 2001; Cochrane, 2005). They can also be attributable to biases related to the report of hedge fund returns, like the survivorship bias⁴. To treat these problems, we propose a new version of the GMM econometric method embedding a test designed to compute directly the extent of the specification (measurement) errors.

We formulate three hypotheses in this article. First, our parsimonious model for hedge fund returns performs as well as the Fung and Hsieh model including many lookback option factors, and our principal component regrouping these lookbacks is more significant than its individual factors taken separately. Second, aside from the market risk premium, the term structure spread and the *VIX* are important determinants of the hedge fund return cyclical behaviour. Third, specification errors bias severely the estimation of the models factor loadings, especially the coefficient of the market risk premium (beta),

³ Note that the market model is the empirical counterpart of the CAPM model, which includes only one explanatory variable: the market risk premium.

⁴ Other well-known biases in the reporting of hedge fund data are the backfill, incubation and smoothing biases.

since the market risk premium is only a proxy of the market portfolio return⁵ but also because it captures some of the effects of the omitted factors.

We find that specification errors are present in the estimation process, especially for the estimation of the small firm anomaly and the market risk premium. Hedge fund returns are also cyclical, more precisely procyclical, since they are sensitive to our model cyclical variables – the term structure spread, the interest rate on long-term bonds or the *VIX*. Finally, our lookback principal component factor appears relevant for specific hedge fund strategies which are not market-oriented and/or very involved in hedging activities, a result supported by a Markov switching regime analysis.

This paper is organized as follows. Section 2 presents the empirical models. Section 3 explains the estimation methodology while section 4 reports the empirical results. Section 5 shows the performance of our new measurement errors indicator while section 6 concludes.

2. The Models

2.1. Model 1: The Fung and Hsieh (F&H) model

In this paper, we first estimate the well-known F&H (2004) model, which is:

$$R_{pt} - R_{ft} = \alpha + \beta_1(R_{mt} - R_{ft}) + \beta_2(Russel - S \& P500)_t + \beta_3BdOpt_t + \beta_4ComOpt_t + \beta_5CurOpt_t + \beta_6CredSpr_t + \beta_710Y_t + \varepsilon_t \quad (1)$$

where $(R_{pt} - R_{ft})$ represents the return of a portfolio – in our case a hedge fund portfolio – in excess of the risk-free rate (R_{ft}) , and $(R_{mt} - R_{ft})$ is the market risk premium. $(Russel - S \& P500)$ is the spread between the Russel 2000 and the S&P500 stock market returns, a proxy for the *SMB* (“small minus big”) factor in the Fama and French (1997, 1998) model, which accounts for

⁵ More precisely, the market risk premium is the spread between the market portfolio return and the risk-free rate.

the small firm anomaly. $BdOpt$, $ComOpt$, $CurOpt$ are the F&H lookback option factors which account for the hedge fund option-like trading strategies, respectively the bond-option, commodity-option and currency lookback option factors. $CredSpr$ is the credit spread, measured as the difference between the BBB and AAA corporate bond yields and $10Y$ is the Federal Reserve's ten-year constant maturity yield⁶. Equation (1) stands for Model 1.

2.2. Benchmark models: Models 2 and 3.

We also consider a more parsimonious version of equation (1) by combining the lookback option returns into one principal component, named $CpOpt$. Note that Fung and Hsieh (1997, 2001, 2004) relied on lookback straddles to study the behaviour of trend followers in the hedge fund industry. But according to these authors, there are substantial differences in trading strategies among trend following funds, so it may not be possible to find a single benchmark that can be used to monitor the performance of trend followers (Fung and Hsieh, 2001). Our empirical results show that our principal component built with the three categories of lookback straddles used in the Fung and Hsieh's studies may constitute a good benchmark to track the performance of the majority of hedge fund strategies, and not only those of trend-followers.

In other respects, we also replace the (*Russel – S & P500*) variable by its counterpart in the Fama and French (1997, 1998) model, SMB – the return of a mimicking portfolio long in small firm stocks and short in big firm stocks, size being measured by stock market capitalization. SMB appeared more significant empirically than the corresponding F&H factor used to test the small firm anomaly. Model 2 obtains:

$$R_{pt} - R_{ft} = \alpha + \beta_1 (R_{mt} - R_{ft}) + \beta_2 (SMB)_t + \beta_3 CpOpt_t + \beta_4 10Y_t + \varepsilon_t \quad (2)$$

⁶ For the sake of simplicity, in the following discussion, we refer to $10Y$ by the *ten-year interest rate*.

We finally consider another version of equation (1), i.e., Model 3, which is more in line with recent developments in the literature on asset pricing (Campbell *et al.*, 2001; Adrian and Shin, 2010, 2011).

$$R_{pt} - R_{ft} = \alpha + \beta_1(R_{mt} - R_{ft}) + \beta_2(SMB)_t + \beta_3 Spread_t + \beta_4 VIX_t + \varepsilon_t \quad (3)$$

Two variables are increasingly considered important to determine asset returns aside the market premium whose statistical impact seems to decrease. First, a variable which has gained strength in explaining returns since the advent of shadow banking is the term-structure spread ($Spread_t$). This variable has become an important indicator of monetary policy but is also a proxy for the phases of the business cycle. According to Adrian and Shin (2010), the fact that short-term interest rates are close to 0 has induced central banks to change the way they manage monetary policy. The credit channel⁷ is now partly implemented via this spread. An increase in the spread is associated with a tightening of monetary policy. In that respect, the term structure spread is an important indicator of monetary policy in the literature focusing on a new channel of the transmission of monetary policy: the risk-taking channel⁸ (e.g., Disyatat, 2010; Gambacorta and Marques-Ibanez, 2011). Moreover, the term-structure spread is a proxy of the phases of the business cycle, an increase in the spread being associated with an economic contraction. It is thus a countercyclical indicator of business conditions. As evidenced by Figure 1, the U.S. term-structure spread enters in a high volatility regime during recessions. Its first moment is also higher in recessions. We thus expect that $\beta_3 < 0$. Second, the volatility of financial market returns weights also more and more in the valuation of financial assets. Volatility is a key determinant of haircuts, a measure of institutional investors' financial leverage (Shleifer and Vishny, 2011). Volatility is also associated with financial institutions strategic

⁷ The broad credit channel regroups the traditional lending channel and the balance-sheet channel.

⁸ According to the risk-taking channel, monetary policy impacts business conditions by changing the perception of risk in the financial system. It focuses on financial frictions in the lending sector.

complementarities, an important factor in the computation of asset prices since these complementarities are an important source of systemic risk (Hirshleifer and Teoh, 2003; Barrell *et al.*, 2010). Moreover, market volatility is the driver of the Black (1976) leverage effect, according to which rising volatility is associated with decreasing returns. Finally, volatility is a key component of hedging. According to Figure 2, *VIX* high volatility regimes are associated with times of financial crises or recessions. The *VIX* first moment is also higher in recession. We expect that $\beta_4 < 0$. Equation (3) can thus be considered as an augmented market model which, aside from the market risk premium, incorporates two additional financial variables to account for the recent developments in financial markets: *Spread* and *VIX*.

Insert Figures 1 and 2 here

3. Estimation methodology

3.1. An optimal combination of Durbin and Pal's instruments

Following Fuller (1987), Dagenais and Dagenais (1994, 1997) proposed a method to tackle errors-in-variables in a cross-section setting. However, their method is easily transposable to a times series context, a situation generally characterized by the autocorrelation of errors. As suggested by Dagenais and Dagenais (1994), in the presence of autocorrelation, provided the innovation and the errors in variables are stationary and ergodic (White, 1984), the estimator is still consistent. To the best of our knowledge, we are the first to apply their method in a time series framework. Let us look more specifically to the Dagenais and Dagenais' original method and to our extensions related to the weighting of the higher moment instruments and to our

generalization of the Hausman test associated with this estimation method (Racicot and Théoret 2012).

Using the method of moments, Durbin (1954) proposes the following estimator, based on third-order moment and co-moment, to identify the parameter of a simple univariate regression model whose explanatory variable is measured with error:

$$\hat{\beta}_d = \frac{s_{xxy}}{s_x^3} \quad (4)$$

where s_{xxy} is the third-order co-moment between the dependent variable y and the explanatory variable x and s_x^3 is the third-order moment of x , defined as:

$$s_x^3 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^3 \quad (5)$$

and

$$s_{xxy} = (n-1)^{-1} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})^2 \quad (6)$$

Subsequently, Pal (1980) derives an estimator based on the fourth-order moment and co-moment of the variables of a model, defined as:

$$\hat{\beta}_p = \frac{s_{xxy} - 3\left(\sum x^2 / n\right)s_{xy}}{s_x^4 - 3\left(\sum x^2 / n\right)s_x^2} \quad (7)$$

with $s_{xy} = (n-1)^{-1} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})$ and $s_x^2 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2$. Equation (7) provides an estimator which is a ratio of cumulants, the numerator being a co-cumulant and the denominator, the fourth-order cumulant, a combination of kurtosis and variance (Stuart and Ord, 1994; Malevergne and Sornette, 2005). In the case of a multivariate regression model, the higher

moment or cumulant instrumental variables corresponding to these two estimators are respectively (Fuller, 1987; Racicot, 1993):

$$\mathbf{x} \circ \mathbf{x} \quad (8)$$

and

$$\mathbf{x} \circ \mathbf{x} \circ \mathbf{x} - 3 \mathbf{x} [\mathbf{D} \circ \mathbf{I}_k] \quad (9)$$

where \mathbf{x} is the matrix of the explanatory variables expressed in deviation from their mean, $\mathbf{D} = (\mathbf{x}^T \mathbf{x} / N)$, a diagonal matrix, and where the symbol \circ stands for the Hadamard product, an element by element matrix multiplication operator. Dagenais and Dagenais (1997) add to these two instruments other cumulants and co-cumulants which were also used previously as instruments by Durbin and Pal in order to identify the parameters of a model containing variables measured with errors. The complete list of the cumulant instrumental variables proposed by Dagenais and Dagenais (1997) is reported in Table 1.

Insert Table 1 here

To increase the robustness of these cumulant instruments⁹ and reduce their well-known instability, Dagenais and Dagenais (1997) rely on Fuller's (1987) instrumental variable (IV) estimator in order to weight them. They note that the resulting combination seems to perform better than the estimators taken separately. But Dagenais and Dagenais were not very specific on the weighting matrix to use. As shown in Racicot (1993), we can *prima facie* rely on a Generalized Least-Squares (GLS) weighting of Durbin (1954) and Pal (1980) estimators, respectively equal to $\beta_D = (\pi_1^T x)^{-1} \pi_1^T y$ and $\beta_P = (\pi_2^T x)^{-1} \pi_2^T y$, for identifying the parameters of a model containing measurement errors. In these expressions, all the variables are expressed in

⁹ Indeed, it is well known that the Durbin and Pal's instruments lack robustness (Cheng and Van Ness, 1999).

deviations from their means with $\pi_1 = [x_{ij}^2]$, $\pi_3 = \pi_2 - 3D(x^T x / N)x^T$, and $\pi_2 = [x_{ij}^3]$. We obtain the new estimator β_H :

$$\beta_H = \Phi \begin{bmatrix} \beta_D \\ \beta_P \end{bmatrix} \quad (10)$$

where Φ is the GLS weighting matrix. Note that this weighting approach, which relies on GLS as the weighting matrix, is *optimal* in the Aitken's (1935) sense¹⁰. However, we rather use the GMM method to weight the Durbin and Pal's estimators, an obviously more efficient method than Dagenais and Dagenais' procedure since we rely on the asymptotic properties of the GMM estimator with respect to the correction of heteroskedasticity and autocorrelation to weight the instruments obtained with GLS. Note that when using GMM, we give up some efficiency gain in order not to have to specify completely the nature of the autocorrelation or heteroskedasticity of the innovation nor the DGP of the measurements errors (Hansen 1982). This is also a great advantage over GLS.

More specifically, the list of instrumental variables may be extended to other moments and comoments (Table 1) but, in line with Dagenais and Dagenais, we retain the triplet $\{z_0, z_1, z_4\}$, i.e., respectively the constant and the Durbin and Pal's estimators, because the results seem more robust when using this subset of instruments rather than the whole set reported in Table 1. Since the benchmark models we estimate (Models 2 and 3) comprise four explanatory variables, the number of instruments related to cumulants thus amounts to eight, that is one z_1 and one z_4 for each explanatory variable of the benchmarks model. In the case of our model, the set of z

¹⁰ Note that we use Φ as weighting matrix in the GLS estimator (equation (10)). As well-known, this matrix can be replaced by the White (1980) asymptotically consistent variance-covariance matrix. The properties of this estimator, named β_E , are discussed in the aforementioned reference. Actually, we implicitly use an augmented version of this estimator in this article.

variables is therefore: $\mathbf{Z} = \{z_0, z_{11}, z_{12}, z_{13}, z_{14}, z_{41}, z_{42}, z_{43}, z_{44}\}$. Our GMM formulation, called GMM-z, obtains¹¹:

$$\arg \min_{\hat{\beta}} \left\{ n^{-1} \left[\mathbf{Z}^T (\mathbf{y} - \mathbf{X}\hat{\beta}) \right]^T \mathbf{W} n^{-1} \left[\mathbf{Z}^T (\mathbf{y} - \mathbf{X}\hat{\beta}) \right] \right\} \quad (11)$$

with \mathbf{W} , a weighting matrix. This method may provide a robust estimator accounting for the autocorrelation, heteroskedasticity and other usual econometric problems encountered in financial experiments.

3.2 . The Hausman artificial regression

Using a standard regression model:

$$\mathbf{Y} = \mathbf{X}^* \beta + \varepsilon \quad (12)$$

assume that \mathbf{X}^* is observed with errors. Its observed value, \mathbf{X} , is thus equal to:

$$\mathbf{X} = \mathbf{X}^* + \mathbf{v} \quad (13)$$

with \mathbf{v} being a matrix of random variables assumed to be normally distributed. Substituting (13)

in (12), we have:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon^* \quad (14)$$

with $\varepsilon^* = \varepsilon - \mathbf{v}\beta$

Obviously, \mathbf{X} is correlated to ε^* , which creates an endogeneity issue. To treat this problem, we first regress the explanatory variables \mathbf{X} on the matrix \mathbf{Z} , which contains the Durbin and Pal cumulants given by Equations (8) and (9) to obtain $\hat{\mathbf{X}}$:

$$\hat{\mathbf{X}} = \mathbf{Z}\hat{\theta} = \mathbf{Z}(\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{X} = \mathbf{P}_Z \mathbf{X} \quad (15)$$

¹¹ See Racicot and Théoret (2001), chap.11, for an introduction to the GMM and its applications in finance.

where \mathbf{P}_Z is the conventional “predicted value maker”. Having run this regression, we extract the matrix of residuals $\hat{\mathbf{w}}$:

$$\hat{\mathbf{w}} = \mathbf{X} - \hat{\mathbf{X}} = \mathbf{X} - \mathbf{P}_Z \mathbf{X} = (\mathbf{I} - \mathbf{P}_Z) \mathbf{X} \quad (16)$$

We can write:

$$\mathbf{X} = \hat{\mathbf{X}} + \hat{\mathbf{w}} \quad (17)$$

Substituting (17) in (14), we have:

$$\mathbf{Y} = \hat{\mathbf{X}}\boldsymbol{\beta} + \hat{\mathbf{w}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}^* \quad (18)$$

Since we assume measurement errors, i.e., that coefficients estimated by OLS are biased, we replace the coefficients $\boldsymbol{\beta}$ associated with $\hat{\mathbf{w}}$ by the mute coefficient vector $\boldsymbol{\theta}$:

$$\mathbf{Y} = \hat{\mathbf{X}}\boldsymbol{\beta} + \hat{\mathbf{w}}\boldsymbol{\theta} + \boldsymbol{\varepsilon}^* \quad (19)$$

To express (19) in terms of \mathbf{X} , the vector of observed variables, we replace $\hat{\mathbf{X}}$ by its value given by (17):

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \hat{\mathbf{w}}\boldsymbol{\varphi} + \boldsymbol{\varepsilon}^* \quad (20)$$

where $\boldsymbol{\varphi} = \boldsymbol{\theta} - \boldsymbol{\beta}$ (Pindyck and Rubinfeld, 1998, pp. 195-197). Then we estimate (20) with OLS. A F test on the $\boldsymbol{\varphi}$ coefficients indicates whether they are significant as a group whilst a t test on the individual coefficients indicates whether they are measured with errors. The vector $\hat{\boldsymbol{\beta}}$ computed by running OLS on (20) is identical to a TSLS estimate, that is:

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{TSLS}} = (\mathbf{X}^T \mathbf{P}_Z \mathbf{X})^{-1} \mathbf{X}^T \mathbf{P}_Z \mathbf{Y} \quad (21)$$

To estimate (20), we use our combination of Durbin and Pal instruments. After substituting the computed $\hat{\mathbf{w}}$ in (20) and running OLS, we obtain a new procedure which is a mapping from the TSLS to the Hausman artificial regression.

To complete this study on measurement errors using cumulants as instruments, a test on the magnitude of these errors is required. To build this test, we rely on the Hausman (Hausman, 1978; McKinnon, 1992; Coën and Racicot, 2007; Racicot and Théoret, 2012) artificial regression, as given by (20), which we write as:

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}}_{\text{TSLs}} + \hat{\mathbf{w}}\boldsymbol{\varphi} + \boldsymbol{\varepsilon}^* \quad (22)$$

where $\hat{\mathbf{w}}$ is the vector of the residuals of the regressions of each explanatory variable on the instrument set. As indicated in (22), the vector of estimated coefficients of the explanatory variables is identical to the one resulting from a conventional TSLs procedure using the same set of instruments (Spencer and Berk, 1981). This result, overlooked by Dagenais and Dagenais (1997) and other more recent researchers on this topic (Meng et al., 2011), increases the usefulness of Equation (22), and stands for a new way to formulate a TSLs directly embedding an Hausman errors-in-variables test on each coefficient. Relatedly, in this equation, each explanatory variable x_i has its own corresponding w_i . The associated coefficient φ_i allows detecting measurement errors on x_i . If it is significantly positive, the corresponding β_i is overstated in the OLS regression and *vice-versa* if it is significantly negative. Our estimated coefficients φ_i are thus new indicators of measurement errors which we formalize by the following empirical relationship:

$$\text{Spread}_{is} = \left(\hat{\beta}_{is,OLS} - \hat{\beta}_{is,TSLs} \right) = \hat{\pi}_0 + \hat{\pi}_1 \hat{\varphi}_{is} + \xi_{is} \quad s = 1 \text{ to } n \quad (23)$$

where s stands for a specific hedge fund strategy. To sum up, Equation (19) is another way to set up a TSLs but one may prefer this formulation to the one represented by a conventional TSLs in view of the useful information conveyed by this equation.

Insert Table 2 here

3.3. The estimation methods and the instruments

The full set of the estimation methods we use in this paper is reported in Table 2. As shown in this table, we rely essentially on three estimation methods: the Hausman method we just described, the TSLS and the GMM. To estimate these IV methods, we resort to three groups of instruments: i) the simple higher moment instruments (*hm*), which are the higher moments of the explanatory variables proposed by Fuller (1987), Cragg (1997) and Lewbel (1997); ii) the *z* instruments; iii) and the *d* instruments, or the distance variables. The *hm* instruments were originally proposed by Fuller (1987) and Lewbel (1997) in line with Dagenais and Dagenais (1997). This set of instruments is built with the higher-order moments of the dependent and explanatory variables up to power 3. These instruments are computed in deviations from their means.

The *d* instruments, which may be considered as filtered versions of the endogenous variables, are defined as follows:

$$d_{it} = x_{it} - \hat{x}_{it} \quad (24)$$

This variable removes some of the nonlinearities embedded in the x_{it} . It is thus a smoothed version of the x_{it} which might be seen as a proxy for its long-term expected value, the relevant variables in the asset pricing models being theoretically defined on the explanatory variables expected values. To compute the \hat{x}_{it} in (24), we perform the following regression using the *z* (cumulant) instruments:

$$x_{it} = \hat{\gamma}_0 + \mathbf{z}\hat{\phi} + \varsigma_t = \hat{x}_{it} + \varsigma_t \quad (25)$$

which amounts to run a polynomial adjustment on each explanatory variable.

Summarizing, we resort to three sets of instruments to estimate the benchmark models in this paper: the *hm*, the *z* (cumulants) and the *d* (distance) variables, a smoothed version of the *z*

instruments. We combine these instruments with our three estimation methods to obtain respectively HAUS-hm, HAUS-d, TSLS-hm, TSLS-z, TSLS-d, GMM-hm, GMM-z, and GMM-d¹².

4. Empirical results

4.1. Data

Our sample of hedge funds is composed of the monthly returns of 17 Van Greenwich indices classified by strategies or groups of strategies. The observation period runs from January 1995 to March 2010, for a total of 183 observations. The Fama and French risk factors, – that is the market risk premium and the mimicking portfolio *SMB* –, are drawn from the French’s website¹³. Moreover, the Fung and Hsieh lookback factors are drawn from the Hsieh’s database¹⁴. Finally, the U.S. macroeconomic and financial data come from the Fred database, which is managed by the Federal Reserve Bank of St-Louis.

[Insert Table 3 here]

Table 3 provides the descriptive statistics of our hedge fund database by strategies. Overall, the mean return computed over the strategies and the mean Sharpe index, a risk-adjusted measure of performance, are higher than the S&P500’s counterparts over the sample period. For instance, the hedge fund mean return is equal to 0.84% versus 0.46% for the market index. The corresponding Sharpe indices are respectively 0.23 and 0.04. The

¹² Note that we do not report the results associated with the methods using the *z* instruments because of their low performance in our estimations. The results related to the *d* instruments, which are a smoothed version of the *z* ones, are much better. The results associated with the methods using directly the *z* instruments are available on demand.

¹³ The address of the French’s website is: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

¹⁴ The website of the Hsieh’s database is: <http://faculty.fuqua.duke.edu/~dah7/DataLibrary/TF-FAC.xls>.

performance of hedge funds was thus quite good over our sample period even if it includes the subprime crisis. Moreover, the hedge fund return distribution is less negatively skewed than the S&P500's¹⁵ one but much more leptokurtic. Interestingly, the standard deviation of the S&P500 (4.62) is much higher than the mean standard deviation (2.41) computed over all strategies. Note that the skewness and kurtosis associated with the hedge fund weighted composite index are lower in absolute value than the S&P500's ones.

In other respects, we note that the performance of the strategies is very different from one strategy to the next over the sample period. The lowest mean return was obtained by the short-sellers strategy (0.18%) whereas the highest one was provided by the opportunistic strategy (1.22%). There is a strong correlation (0.74) between strategies return and systemic risk as measured by the CAPM beta. Interestingly, the standard deviations of the strategies returns are not a good indicator of systemic risk¹⁶, the correlation between the CAPM beta and the standard deviation being only 0.02. Surprisingly, the correlation between strategies return and kurtosis, at 0.37, is also quite low. Hence, the strategies which bear high fat-tail risk¹⁷ do not necessarily yield the best returns¹⁸.

The ranking of the strategies with respect to their return corresponds to the usual one reported by hedge fund institutes. With a mean return of 0.18%, the short sellers strategy is located at the tail of the ranking. Note that this low return is associated with the hedging capacity of short sellers, the market beta of this strategy being equal to -1.01. The futures index, with a negative beta, also provides a lower return than the average one. The macro

¹⁵ This fact may be explained by the option-like trading strategies followed by hedge funds (Heuson and Hutchinson, 2012).

¹⁶ In fact, the market beta of a stock measures its *systematic* risk, which is its non-diversifiable risk as opposed to its idiosyncratic risk which is diversifiable. In the recent banking literature, *systemic* risk is rather associated with the strategic complementarities of financial institutions, like the externalities one financial institution imposes to the others as measured by its CoVaR (Adrian and Brunnermeier, 2010; Gauthier *et al.*, 2012). Given the aim of our study, we do not make a distinction between systematic and systemic risk.

¹⁷ That is risk associated with rare events.

¹⁸ According to Heuson and Hutchinson (2012), contrary to asymmetry, there is no *a priori* reason for kurtosis to affect hedge fund performance. However, the empirical evidence on this point is quite limited.

strategy, even if it displays a higher return than the S&P500, is also particularly underperforming. In that respect, the models developed by Quants, which are a key component of the macro strategy, seem to have performed badly during our estimation period (Khandani and Lo, 2007). Furthermore, the following strategies are located at the top of the ranking: opportunistic index (1.22%), long-short (1.09%), diversified event driven (1.08%), growth (1.04%) and event driven (1.03%). Note that strategy ranking corresponding to the Sharpe index may be quite different from the ranking based on the mean return. For instance, we observe quite different strategy mean returns in the Sharpe index ranging around 1.

Insert Table 4 here

4.2. OLS estimation of the Fung and Hsieh model

Table 4 provides the OLS estimation of the F&H model (Equation (1)). As noted earlier, it is estimated using a hedge fund sample comprising 17 Van Greenwich monthly indices (strategies) observed over the period running from January 1995 to March 2010. Overall, the R^2 are consistent with previous work (Fung and Hsieh, 1997, 2001, 2004), its mean value computed over all strategies being equal to 0.52. At 1.72, the average DW statistic signals no serious autocorrelation problem in the dataset. The strategies with the highest R^2 tend to have the highest market beta. In that respect, the three strategies having the highest market beta – the growth, multi-strategy and the opportunistic strategies – also display the highest R^2 , their beta being respectively 0.72, 0.52 and 0.49 and the corresponding R^2 being 0.72, 0.81 and 0.78. Incidentally, these betas are quite close to those computed by the market model (Table 3).

The constant of this kind of regressions related to factorial asset pricing models has a particular meaning since it represents the alpha, the risk-adjusted return of hedge fund portfolios or their absolute return. Indeed, most of hedge funds studies have identified an alpha puzzle, i.e., a positive and significant alpha for the hedge fund industry, which contradicts the hypothesis of financial markets efficiency (cf. Ackermann et al. 1999)¹⁹. However, these studies were performed over the period corresponding to *the Great Moderation*. During this period, which extended from 1984 to 2007 in the U.S., GDP growth was generally positive, although at an average level lower than in the past, but, more importantly, its volatility was also much lower than previously. This period, also characterized by a better inflation control by the Federal Reserve, fostered positive returns for investors. However, in our sample which includes the subprime crisis, there seems to be no alpha puzzle. First, when a strategy's alpha is positive, it is not significant at the usual thresholds. Then, there are six strategies for which the alpha is negative and significant at least at the 10% level. Thus, hedge funds are not immunized from financial crises and the subprime crisis obviously impacted greatly their performance.

Turning to the impact of the explanatory variables, we note that the market risk premium is actually the main driver of hedge fund returns even if hedge funds tend to maintain a restricted market exposure. The other main explanatory variable is the spread between the Russell 2000 and the S&P. Its coefficient is usually positive, which suggests that hedge funds are for most of them exposed to the performance of small firms. In that respect, there seems to be some herding in the hedge fund industry toward the investment in small firms (Haiss, 2005). Our results seem to confirm the incentive of hedge funds to profit from the corresponding market anomaly.

¹⁹ More specifically the EMH, the efficient market hypothesis.

Regarding the F&H lookback option factors, we note that, except for the bond lookback which has a positive impact on most hedge fund strategies, significant at the 10% level, they are not particularly significant as a group. Moreover, the strategies where F&H factors are the most important are the ones corresponding to very specialized strategies like the futures, the arbitrage and the equity market neutral strategies. These strategies display a low beta and are quite involved in hedging or option-like activities, which may explain the importance of the lookback factors for these strategies.

Insert Figure 3 here

Interestingly, contrary to the earlier studies (Fung and Hsieh, 1997, 2001, 2004) where the long-term interest rate (*10Y*) impacts negatively hedge fund returns, this factor has a positive significant impact for most strategies over our sample period. A look at Figure 3, which plots the level of the ten-year interest rate and its Markov regimes, indicates that the long-term rate was on a downward trend over most of our estimation period, which was favourable for capturing capital gains on bonds, and that the period comprised between 2001 and 2007 was a persistent period of low regime (low volatility) for this interest rate. The earlier studies were, for most of them, done over a period of high volatility for long-term interest rates, which was difficult to hedge and thus detrimental for hedge fund performance. Over our estimation period, the long term interest rate is a more relevant proxy for the return on asset backed securities, like mortgage-backed securities, in which many hedge fund strategies are heavily invested. In this case, an increase in the long-term interest rate gives rise to an increase in hedge fund bond and asset-backed securities returns.

In that respect, note if hedge funds followed buy-and hold strategies, like mutual funds, the results would be quite different from those we obtain (Fung and Hsieh, 1997). More precisely, in a buy-and-hold strategy, an increase in long-term interest rate leads to capital losses on the holdings of fixed income securities and thus, *ceteris paribus*, to a decrease in portfolio returns. However, hedge funds follow dynamic strategies. They thus can short bonds when interest rates increase and this strategy then delivers positive returns. Being dynamic, the strategies of hedge funds change with the development of the financial system in which they operate. The modifications in the economic and financial environment over our sample period noted above seem to have altered hedge fund strategies as given by the estimated factor loadings, a result in line with the well-known Lucas'(1976) critique in macroeconometrics.

Note also that our buy-and-hold argument to which we resort to explain the link between the risk premium of a hedge fund portfolio and the long-term interest rate is formulated in a partial equilibrium framework. But this argument must be modified when we consider a general equilibrium setting. In such a case, by the dynamics of asset substitution, an increase in long-term interest rate must lead to an increase in a stock risk premium (Tobin, 1969). Otherwise, this stock would not be held. Thus, all in all, the balance is tilted toward a positive comovement between a hedge fund portfolio risk premium and an increase in long-term interest rates.

Insert Figure 4 here

Finally, the credit spread is significant at the 10% level for only four strategies: the market-neutral group, the equity market neutral, the futures and the macro strategies. Except for the equity market strategy, this factor exerts a positive impact on the return of these strategies. A look at Figure 4, which plots the credit spread, is once again instructive. As expected, the credit spread tends to increase during recessions or in times of financial crisis (e.g. the Asian crisis in 1997-1998²⁰) but the period is dominated by the big jump in the credit spread during the subprime crisis. In that respect, the figure indicates a change in regime in the spread behaviour only during the subprime crisis. Our results suggest that some hedge funds tend to short the credit spread, which explains the significant positive return associated with this factor for three strategies analyzed. In that respect, such derivatives as the credit-default-swap (CDS) are used to short the credit spread. However, the negative coefficient related to the credit spread for the equity market neutral strategy indicates that it suffered greatly from the subprime crisis, due to the rising probability of default for firms here proxied by the credit spread.

Overall, we find that the market risk premium and the return on small firm stocks are the main drivers of hedge fund returns over our sample period. Except for the bond lookback, the F&H factors have a lesser importance, so we decided to group them into a principal component factor. Moreover, the impact of the ten-year interest and the credit spread seems to be different from earlier studies, especially the F&H's ones (1997, 2001, 2004). This may be due to the relative stability of interest rates over our sample period, explained by a better control of inflation by the Federal Reserve, and the occurrence of a major financial crisis in our sample period, the subprime crisis.

²⁰ This period also overlaps the Russian and LTCM crises.

Insert Table 5 here

4.3. The IV estimations

4.3.1. The higher moment and cumulant instruments empirical performance

Table 5 provides the R^2 of the regressions of Model 2 explanatory variables on three categories of instruments: the classical, d and hm instruments. We note that the d and the hm instruments are quite comparable in terms of performance while the classical instruments, which are the predetermined values of the models explanatory variables, are very poor proxies of the model endogenous variables. As explained earlier, note also that we do not report the results related to the z instruments in our empirical work because these instruments yield results quite inferior to the other categories of instruments²¹. The d instruments, a smoothed version of the z ones, perform much better.

Insert Table 6 here

Since the d instruments are new in the literature, we present in Table 6 the regressions of the Model 2 explanatory variables on these instruments. The results suggest that they can be considered as strong instruments, the R^2 of these regressions being high. In fact, each risk factor has its own instrument to which it is related by an estimated coefficient close to 1 and is orthogonal to the other risk factors.

²¹ These results are available on demand.

4.3.2. The Hausman procedure applied to Models 2 and 3

We applied our IV estimation methods to our two benchmark models, Models 2 and 3. The results are averaged over the strategies. Note that the transposition of our general Hausman artificial regression (Equation (20)) to Model 2 (equation (2)) is:

$$R_{pt} - R_{ft} = \alpha + \beta_1 (R_{mt} - R_{ft}) + \beta_2 (SMB)_t + \beta_3 CpOpt_t + \beta_4 10Y_t + \sum_{i=1}^4 \varphi_i \hat{w}_{it} + \varepsilon_t \quad (26)$$

The estimated coefficients φ_i , which are equal to $\theta_i - \beta_i$, allow detecting measurement errors, and their estimated signs indicate whether the impact of the corresponding factor is overstated or understated in the OLS regression. Equation (26) is estimated with the *hm* and *d* instruments, respectively the Haus-hm and Haus-d methods.

Similarly, the transposition of equation (20) to Model 3 (equation (3)) is:

$$R_{pt} - R_{ft} = \alpha + \beta_1 (R_{mt} - R_{ft}) + \beta_2 (SMB)_t + \beta_3 Spread_t + \beta_4 VIX_t + \sum_{i=1}^4 \varphi_i \hat{w}_{it} + \varepsilon_t \quad (27)$$

To estimate Models 2 and 3, OLS is used as a benchmark. The other estimation methods are listed in Table 2. For the sake of simplicity, the individual strategies are not displayed, the results being averaged over the strategies.

[Insert Table 7 here]

4.3.3. Model 2 OLS and IV estimations

Table 7 provides the OLS and IV estimations of Model 2. First note that the R^2 obtained over the estimation methods with Model 2 compare quite well to the less parsimonious Model 1 – the F&H model – which includes three additional variables. For instance, the R^2 associated with the OLS estimation of Model 2, at 0.56, is even higher than the 0.52 obtained

with the OLS estimation of model 1. Moreover, the R^2 obtained with the Hausman estimations are even superior. Yet, the R^2 associated with the TSLS and GMM estimations are somewhat lower, a quite normal result. Once more, the DW statistic does not signal the presence of autocorrelation in our results.

We note that the average alpha is quite sensitive to the estimation method used. It is close to 0, although insignificant, for the Hausman regressions run with the two sets of instruments, higher moment and d instruments, but its minimum value is obtained with the GMM-hm method, at -0.8527 and significant at the 1% level. However, the sign of the alpha is robust to the estimation method. Indeed, regardless of the estimation method it remains negative, which suggests no alpha puzzle in our sample, at least in Models 1 and 2.

Insert Figure 5 here

All explanatory variables are very significant in the OLS estimation of Model 2. And once more, regardless of the estimation method, the market risk premium and the factor accounting for the small firm anomaly, here SMB , are the main drivers of hedge fund returns. In that respect, SMB is more significant than its counterpart in Model 1. Moreover, combining the three lookback option factors into one principal component leads to interesting results. Indeed, the $CpOpt$ coefficient is positive and significant at the 5% level except in the case of the Gmm-d estimation where it is positive but insignificant. This result suggests that most hedge funds benefited from their option-like trading strategies over the sample period. In that respect, Figure 5 plots the time-varying market beta associated with $CpOpt$ computed with the Kalman filter and the regimes associated with the level of this variable. The beta

behaviour shows that the portfolio related to $CpOpt$ acts as a hedge or insurance against bad times. For instance, the beta became negative on a downward tendency over the Asian crisis (1997-1998), during the U.S. 2001-2003 recession and especially during the subprime crisis, where the $CpOpt$ beta touched a low of -0.34. The figure associated with the $CpOpt$ Markov regimes shows that there is no evidence of changing regimes for this return series, in contrast with the regimes of the market portfolio return which correspond to the phases of the business cycles. This pattern is shared by the three lookbacks included in Model 1 and is related to the hedging capacity of a lookback option. Note that Fung and Hsieh lookback straddles were initially designed to study the strategies of the trend followers (Fung and Hsieh, 1997, 2001, 2004). But Figure 5 shows that we can adopt a broader justification for the introduction of lookback straddles in hedge fund return models. Indeed, the evolution of the beta²² of $CpOpt$ shows that it constitutes a *hedge factor* in the hedge fund return regressions. The estimated coefficient of $CpOpt$ thus indicates the capacity of a strategy to hedge a portfolio. For instance, a positive and significant $CpOpt$ coefficient for a specific strategy signals that the investment in this strategy can effectively hedge or insure a portfolio. Thus, the lookback option factors should not be reserved to the trend follower strategies, as in Fung and Hsieh, but to all strategies which rely on option-like trading strategies, obviously the case for most hedge funds. Finally, the coefficient associated with IOY remains positive and is significant at the 5% level for three estimation methods: OLS, GMM-hm and Haus-hm.

Insert Table 8 here

²² This time-varying beta is simulated by the Kalman filter.

Our two Hausman procedures, i.e., Haus-hm and Haus-d, allow detecting nonnegligible measurement errors. In that respect, the *SMB* factor displays the highest errors. We conjecture that this bias is related to the construction of the portfolio used to approximate the small firm anomaly²³. The coefficient associated with the w of this variable, \hat{w}_{SMB} , which is equal to -0.05 for both Hausman procedures and significant at the 5% level, indicates that the sensitivity of hedge fund returns to *SMB* is understated by OLS. This is indeed the case since the average of the *SMB* coefficient computed over the IV methods is equal to 0.1636, higher than the OLS estimate, equal to 0.1467. Note that there are nine strategies for which the coefficient associated with \hat{w}_{SMB} is significant at the 5% threshold when using Haus-hm and ten when using Haus-d.

The next factor most affected by measurement errors is given by *CpOpt*. Similarly to *SMB*, *CpOpt* is a generated portfolio which is built using a principal component of three other portfolios, which might explain the observed bias in the measurement of this variable. Note also that the F&H lookback factors are only proxies for the option-like activities of hedge funds, another source of measurement errors. In that case, the direction of the bias is unclear. Indeed, the coefficient related to \hat{w}_{CpOpt} is estimated at -10.6449 when using the Haus-hm procedure, signalling an understatement of *CpOpt* impact by OLS, whilst, in contrast, the estimate is 0.1399 when using the Haus-d procedure, signalling a small overstatement by OLS. However, the \hat{w}_{CpOpt} coefficient is not significant at the usual thresholds for both procedures. Moreover, our findings indicate that the *CpOpt* impact is significantly measured with errors for five strategies when using Haus-hm and for six strategies when using Haus-d.

²³ According to Pagan (1984, 1986), generated variables are endogenous variables which must be estimated by IV methods.

Our Hausman procedures also capture measurement errors for the market beta. We conjecture that the bias is again related to the portfolio chosen to proxy the market portfolio, which is theoretically much broader than its usual empirical counterpart (Roll, 1997). The Haus-hm procedure signals that the market beta is biased at the 5% level for four strategies and for 2 with the Haus-d one. Similarly to the results obtained for *CpOpt*, the direction of the bias is unclear for the market beta. Finally, our results indicate no significant bias for the coefficient of the ten-year interest rate for most of the strategies.

4.3.4. Model 3 OLS and IV estimations

Model 3 focuses on the new drivers of returns following the development of market-oriented banking and shadow banking: the term-structure spread (*Spread*) and the *VIX*. Overall, the performance of Model 3, as measured by the R^2 and the DW, is comparable to Model 2 (Table 8). However, in contrast to Model 2, the mean alpha is always positive in Model 3 and significant at the 5% level. At 0.7632 on a monthly basis, it is quite high since it approximates 9% on an annual basis. Thus, the alpha puzzle is present here, even after accounting for the crisis, which suggests that the selected model has a substantial impact on the estimation of the alpha. Note that these findings are more in line with the hedge fund descriptive statistics over our sample period (Table 3), where the mean returns and the Sharpe indices of the strategies are positive and even substantial for many of them. Interestingly, using the GMM-hm procedure, the cross-sectional correlation between the Sharpe indices and the alphas, two risk-adjusted measures of hedge fund returns, is equal to 0.55, significant at the 5% level²⁴.

²⁴ Note that we must be prudent here regarding the interpretation of the alpha as a risk-adjusted return in Model 3 since the term structure spread and the *VIX* are not portfolio returns.

The *Spread* variable is significant at the 10% level for all the estimation methods and displays the expected negative sign. This result simply indicates that hedge fund returns are procyclical: they have a tendency to decrease when *Spread* increases, which corresponds to economic contraction periods. This procyclicality of hedge fund returns is a point neglected in many studies. However, the impact of the *VIX* is less clear. This variable is only significant at the 5% level in the GMM-hm estimation with a coefficient equal to -0.0280. According to this regression, the *VIX* variable has the expected Black leverage effect on hedge fund returns.

Once more, our estimations permit the detection of significant measurement errors associated with the sensitivity of hedge fund returns to their risk factors. In that respect, similarly to the results of Model 2, the *SMB* factor displays the highest errors. For instance, computed over the IV estimation methods, its mean coefficient is equal to 0.1734 and significant at the 1% level. The estimated coefficients for this variable ranges from a low of 0.1451, obtained with the GMM-hm method, to a high of 0.1929, obtained with the Gmm-d and TSLS-d methods. The extent of the bias caused by measurement errors may be captured by the coefficient of the artificial variable \hat{w}_{SMB} . Using the Haus-d method, this coefficient is equal to -0.0624 and significant at the 5% level, which, according to this test, suggests an understatement of the hedge fund *SMB* coefficient by OLS, estimated at 0.1483. Note that there are eight strategies for which the \hat{w}_{SMB} coefficient is significant at the 5% threshold, with a maximum *t* statistic of 4.87 for one strategy, which suggests a serious understatement in this case. Using Haus-hm, the \hat{w}_{SMB} coefficient is, at -0.0583, close to the one obtained with Haus-d, suggesting once more that that OLS understates the sensitivity of hedge fund returns to *SMB*.

Our Hausman procedure also captures measurement errors for the market beta in Model 3. In this case, both Hausman procedures signal a systematic overstatement of the beta by the OLS method, the $\hat{w}_{r_m-r_f}$ coefficient being positive. Actually, the beta estimated with OLS is equal to 0.2109. The IV methods indicate that it is lower than this level, with a low of 0.1879 using the TSLS-hm and a high of 0.1952 with the TSLS-d and Gmm-d methods and a mean of 0.1947 over the IV methods. Actually, when using the Haus-hm method, there are seven strategies for which the coefficient of $\hat{w}_{r_m-r_f}$ is significant at the 5% level and two strategies with the Haus-d method. Another variable which seems measured with error is the *Spread* variable, five strategies having a significant \hat{w}_{Spread} coefficient using the Haus-hm method and four when relying on the Haus-d one. Finally, the *VIX* variable does not appear to be plagued by measurement errors.

Insert Table 9 here

5. The measurement error indicator

To test for the relevance of our measurement error indicator for the market beta and the coefficient associated with *SMB* for our hedge fund strategies, we rely on the estimation of equation (23). Table 9, which is built with higher moment instruments and which list strategies in increasing order of the spread, shows the high positive correlation between the strategy coefficients of measurement errors, the φ , and the spread between the coefficients estimated respectively with the OLS and Haus-hm methods. Running the cross-sectional regression of the spread over the φ , we obtain the following result:

$$spread = \left(\hat{\beta}_{r_m - r_f, OLS} - \hat{\beta}_{r_m - r_f, TSLS} \right) = 0.01 + 0.80\hat{\varphi}_s + \xi_s, \quad s = 1, \dots, 17 \quad (28)$$

(0.62) (38.97)

with a R^2 at 0.99 and the t -statistics of the estimated coefficients reported in parentheses. This equation reveals significant measurement errors at the disaggregated level for the hedge fund strategies. Figure 6, which plots equation (28) along with the observed values, shows the close relationship between the spread and φ when using higher moment instruments.

Insert Figures 6 and 7 here

We repeated the same exercise for *SMB*. Regressing the spread on the φ , we have:

$$spread = \left(\hat{\beta}_{SMB, OLS} - \hat{\beta}_{SMB, TSLS} \right) = 0.01 + 0.70\hat{\varphi}_s + \xi_s, \quad s = 1, \dots, 17 \quad (29)$$

(2.01) (71.56)

with a R^2 equal to 0.99. Figure 7 plots this tight relationship between the two variables for *SMB*. Once again, we conclude that there are significant measurement errors for the hedge fund strategies. Overall, our indicator φ captures quite well the biases in the estimation process related to measurement errors.

6. Conclusion

In this paper, we propose two new versions of hedge fund return models: a parsimonious version of the F&H model, which combines its three lookback option factors into one principal component, and an augmented version of the market model, which accounts for the on-going developments in financial markets. To estimate these models,

which may be contaminated by measurement errors and given the strong asymmetry²⁵ and leptokurtism of the hedge fund returns distributions, we resort to IV methods using higher moment instruments and a new set of instruments, the distance d instruments computed with the cumulants of the models variables.

Our results show that our two parsimonious hedge fund return models perform quite well compared to the larger F&H model. One drawback of the F&H model is that the significance of the individual lookback factors, which account for the hedge fund option-like trading strategies, is low in our sample. Combining them into a principal component factor yields better results. Moreover, the estimation of our augmented market model, which incorporates the financial variables identified as important determinants of stock returns by the on-going literature on asset pricing models – the term-structure spread and the stock market volatility – shows that these variables are very relevant to model hedge fund returns. Among others, they account for the procyclicality of hedge fund returns.

One other important contribution of our paper is that we found that higher moment and cumulant instruments may be considered as robust instruments to estimate models of financial or economic variables whose distributions are, like hedge fund returns, asymmetric and leptokurtic. The instruments used to run the regressions, the hm and d instruments, perform very well while the instruments generally used to account for endogeneity in asset pricing models display poor results. Moreover, the new set of instruments we introduce, the d instruments, are much less erratic than the z ones, an obvious defect of instruments based on cumulants. Indeed, our d instruments are based on the distance between the observed endogenous variables and their fitted values computed with our cumulant instruments, which

²⁵ In this respect, see Heuson and Hutchinson (2012). As noted before, the positive asymmetry of the return distribution of many hedge fund strategies is due to their option-like trading strategies.

we built by weighting Durbin and Pal's cumulant instruments. These instruments filter (smooth) the endogenous variables, removing their nonlinearities, which leads to more stable instruments than the z ones. These instruments display a correlation close to one with their corresponding explanatory variable and are orthogonal to the other explanatory variables of our models, obviously two essential properties of good instruments.

One other important contribution of our paper is the new procedure we propose to capture measurement errors in asset pricing models. In that respect, we combine our two sets of instruments – the hm and d instruments – with two IV estimation methods, the TSLS and the GMM. We also develop a new form of TSLS based on Hausman's works on artificial regressions, which directly embeds a measurement error test. It allows us to define a new indicator of measurement errors. Our experiments show that there may be serious measurement errors in the estimation of factorial asset pricing models.

References

- Ackermann, C.R., McEnally, R., Ravenscraft, D., 1999. The performance of hedge funds: Risk, return and incentives. *Journal of Finance* 54, 833-874.
- Adrian, T., Brunnermeier, M.K., 2010. CoVaR, Working paper, Princeton University.
- Aitken, A.C., 1935. On least squares and linear combination of observations. *Proceedings of the Royal Society of Edimburgh* 55, 42-48.
- Adrian, T., Shin, H.S., 2010. Financial intermediaries and monetary economics, Staff Report, Federal Reserve Bank of New York.
- Barrell, R., Davis, E.P., Karim D., Liadzze, I., 2010. Evaluating Off-balance sheet exposures in banking crisis determination models. Working paper, Brunel University.
- Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile.? An empirical exploration of idiosyncratic risk. *Journal of Finance*, 56, 1-43.
- Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. *The econometrics of financial markets*. Princeton University Press, Princeton.
- Chen, N.F., Roll, R., Ross, S., 1986. Economic forces and the stock market. *Journal of Business* 59, 572-621.
- Cheng, C.L., Van Ness, J.W., 1999. *Regression with Measurement Error*, Arnold.
- Cochrane, J.H., 2005. *Asset Pricing*. 2nd Edition. Princeton University Press, Princeton.
- Coën, A., Racicot, F.E., 2007. Capital asset pricing models revisited: Evidence from errors in variables. *Economics Letters* 95, 443-450.
- Cragg, J.G., 1997. Using higher moments instruments to estimate the simple errors-in-variables model. *Rand Journal of Economics* 28, S71-S91.
- Dagenais, M.G., Dagenais, D.L., 1994. GMM estimators for linear regression models with errors in the variables. Working paper, CRDE.
- Dagenais, M.G., Dagenais, D.L., 1997. Higher moment estimators for linear regression models with errors in the variables. *Journal of Econometrics* 76, 193-221.
- Disyatat, P., 2010. The bank lending channel revisited. Working paper, Bank for International Settlements.
- Do, V., Faff, R., Wickramanayake, J., 2005. An empirical analysis of hedge fund performance: The case of Australian hedge funds industry. *Journal of Multinational Financial Management* 15, 377-393.
- Durbin, J., 1954. Errors in variables. *International Statistical Review* 22, 23-32.
- Estrella, A., Hardouvelis, G.A., 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46, 555-576.
- Fama, E.F., French, K.R., 1997. Industry costs of equity. *Journal of Financial Economics* 43, 153-193.
- Fama, E.F., French, K.R., 1998. Value versus growth: The international evidence. *Journal of Finance* 53, 1995-1999.
- Fama, E., McBeth, J., 1973. Risk, returns and equilibrium: Empirical tests. *Journal of Political Economy* 81, 607-636.
- Fuller, W.A., 1987. *Measurement Error Models*, John Wiley & Sons.
- Fung, W., Hsieh, D.A., 1997. Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies* 10, 275-302.
- Fung, W., Hsieh, D.A., 2001. The risk in hedge fund strategies: Theory and evidence from trend followers. *Review of Financial Studies*, 14, 313-341.

- Fung, W., Hsieh, D.A., 2004. Hedge fund benchmarks: A risk based approach. *Financial Analysts Journal* 60, 65-80.
- Gambacorta, L., Marques-Ibanez, D., 2011. The bank lending channel: Lessons from the crisis, Working paper, Bank for International Settlements.
- Gauthier, C., Lehar, A., Souissi, M., 2012. Macroprudential capital requirements and systemic risk. *Journal of Financial Intermediation*, forthcoming.
- Haiss, P. R., 2005. Banks, herding and regulation: A review and synthesis, Working paper, Europe Institute.
- Hansen, L.P., 1982, Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* 50, 1029-1054.
- Hausman, J.A., 1978. Specification tests in econometrics. *Econometrica* 46, 1251-1271.
- Heuson, A.J., Hutchinson, M.C., 2012. Which hedge fund managers deliver alpha ? Assessing performance when returns are skewed. Working paper, University of Miami and University College Cork.
- Hirshleifer, D., Teoh, S.H., 2003. Herd-behaviour and cascading in capital markets: A review and synthesis. *European Financial Management* 9, 25-66.
- Jarrow, R., Protter, P., 2011. Positive alphas, abnormal performance and illusory arbitrage. *Mathematical Finance*, forthcoming.
- Khandani, A.E., Lo, A.W., 2007. What happened to the Quants in August 2007? Working paper, MIT.
- Lettau, M., Ludvigson, S., 2001. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109, 1238-1287.
- Lewbel A., 1997. Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D. *Econometrica* 65, 1201-1213.
- Lucas, R.E., 1976. Econometric policy evaluation: A critique. In: *The Phillips Curve and Labour Markets*, North-Holland, 19-46.
- MacKinnon, J.G., 1992. Model specification tests and artificial regressions. *Journal of Economic Literature* 30, 102-146.
- Malevergne, Y., Sornette, D., 2005. Higher-moment portfolio theory: Capitalizing on behavioural anomalies of stock markets. *Journal of Portfolio Management* 31, 49-55.
- McGuire, P., Remolona, E., Tsatsaronis, K., 2005. Time-varying exposures and leverage in hedge funds. *BIS Quarterly Review* 1, 59-72.
- Meng, J.G., Hu, G., Bai, J., 2011. Olive: A simple method for estimating betas when factors are measured with errors. *Journal of Financial Research* 34(1), 27-60.
- Pagan, A.R., 1984. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review* 25, 221-247.
- Pagan, A.R., 1986. Two stage and related estimators and their applications. *Review of Economic Studies* 53, 517-538.
- Pal, M., 1980. Consistent moment estimators of regression coefficients in the presence of errors in variables. *Journal of Econometrics* 14, 349-364.
- Pindyck, R.S., Rubinfeld, D.L., 1998. *Econometric Models and Economic Forecasts*. 4th edition. Irwin-McGraw-Hill.

- Racicot F.E., 1993. Techniques alternatives d'estimation et tests en présence d'erreurs de mesure sur les variables explicatives. M.Sc. thesis. University of Montreal: Department of economics. <https://papyrus.bib.umontreal.ca/jspui/bitstream/1866/1076/1/a1.1g638.pdf>.
- Racicot, F.E., Théoret, R., 2001. *Traité d'économétrie financière*. Presses de l'Université du Québec, Ste-Foy.
- Racicot, F.E., Théoret, R., 2012. Optimally weighting higher moment instruments to deal with measurement errors in financial return models. *Applied Financial Economics* 22, 1135-1146.
- Roll, R., 1977. A critique of asset pricing theory's test: Part 1. *Journal of Financial Economics* 4, 129-176.
- Shanken, J., 1992. On the estimation of beta-pricing models. *Review of Financial Studies* 5, 1-34.
- Sharpe, W.F., 1992. Asset allocation: Management style and performance measurement. *Journal of Portfolio Management* 18, 7-19.
- Shleifer, A., Vishny, R., 2011. Fire sales in finance and macroeconomics. *Journal of Economic Perspectives* 25, 29-48.
- Spencer, D.E., Berk, K.N., 1981. A limited information specification test. *Econometrica* 49, 1079-1085.
- Stuart, A., Ord, K., 1994. *Kendall's Advanced Theory of Statistics*, vol. 1, Arnold.
- Tobin, J., 1969. A general equilibrium approach to Monetary Theory. *Journal of Money, Credit and Banking* 1, 15-29.
- White, H., 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48, 817-838.

Tables

Table 1 List of the cumulant instrumental variables proposed by Dagenais and Dagenais (1997)

z_0	$\mathbf{1}_N$
z_1	$\mathbf{x} \circ \mathbf{x}$ (<i>Durbin</i>)
z_2	$\mathbf{x} \circ \mathbf{y}$
z_3	$\mathbf{y} \circ \mathbf{y}$
z_4	$\mathbf{x} \circ \mathbf{x} \circ \mathbf{x} - 3\mathbf{x}[\mathbf{E}(\mathbf{x}'\mathbf{x}/N) \circ \mathbf{I}_k]$ (<i>Pal</i>)
z_5	$\mathbf{x} \circ \mathbf{x} \circ \mathbf{y} - 2\mathbf{x}[\mathbf{E}(\mathbf{x}'\mathbf{y}/N) \circ \mathbf{I}_k] - \mathbf{y}\{\mathbf{1}'_k[\mathbf{E}(\mathbf{x}'\mathbf{x}/N) \circ \mathbf{I}_k]\}$
z_6	$\mathbf{x} \circ \mathbf{y} \circ \mathbf{y} - \mathbf{x}[\mathbf{E}(\mathbf{y}'\mathbf{y}/N) - 2\mathbf{y}[\mathbf{E}(\mathbf{y}'\mathbf{x}/N)$
z_7	$\mathbf{y} \circ \mathbf{y} \circ \mathbf{y} - 3\mathbf{y}[\mathbf{E}(\mathbf{y}'\mathbf{y}/N)]$

Note. $\mathbf{1}$ stands for a vector of one. \mathbf{I}_k is the identity matrix of dimension $(k \times k)$.
 \circ stands for the Hadamard product.

Table 2 List of higher moment and cumulant estimators used to correct measurement errors

Method	Instruments
HAUS-hm	higher moments: $\{x_{t-1}, x_t^2, x_t^3, y_t^2\}$
HAUS-d	the distance variables: $\{d_1, \dots, d_4\}$
TSLS-hm	higher moments: $\{x_{t-1}, x_t^2, x_t^3, y_t^2\}$
TSLS-z	$\{z_0, z_{11}, \dots, z_{14}, z_{41}, \dots, z_{44}\}$, which are the Durbin and Pal instruments (cumulants)
TSLS-d	the distance variables: $\{d_1, \dots, d_4\}$
GMM-hm	higher moments: $\{x_{t-1}, x_t^2, x_t^3, y_t^2\}$
GMM-z	$\{z_0, z_{11}, \dots, z_{14}, z_{41}, \dots, z_{44}\}$, which are the Durbin and Pal instruments (cumulants)
GMM-d	the distance variables: $\{d_1, \dots, d_4\}$

Note. The variables which enter in the computation of higher moments and cumulants are expressed in deviation from their mean. The Haus-hm and Haus-d methods are the Hausman artificial regressions using respectively as instruments the higher moments and cumulants of the explanatory variables. Three two-stage least squares methods are used, differing by their instruments: higher moments (*hm*), *z* instruments (Table 1), and distance instruments (*d*). We resort to the same procedure with GMM. Note that for the GMM estimations, the weighting matrix used to weight the moment conditions is the Newey West one.

Table 3 Descriptive statistics of the VAN Greenwich indices, 1995-2010

	Mean	Median	Max	Min	sd	Skew	Kurtosis	Sharpe index	CAPM-beta
Market neutral group	0.93	0.92	7.20	-6.06	1.48	-0.61	8.99	0.44	0.20
Equity Market Neutral	0.87	0.80	8.10	-2.53	1.41	1.33	8.95	0.42	0.08
Event driven	1.03	1.20	10.70	-6.90	2.04	-0.17	7.44	0.36	0.30
Distressed Securities	0.68	1.04	4.79	-7.44	2.06	-1.47	6.86	0.19	0.22
Diversified Event Driven	1.08	1.10	11.70	-8.00	2.38	0.01	6.74	0.34	0.36
Arbitrage Index	0.87	0.90	4.10	-8.58	1.38	-2.40	17.99	0.43	0.16
Long-Short	1.09	1.31	13.20	-9.24	3.02	0.04	5.20	0.27	0.52
Growth	1.04	1.19	20.10	-12.99	4.53	0.43	5.50	0.17	0.76
Opportunistic	1.22	1.30	21.20	-8.51	3.24	1.33	12.17	0.29	0.48
Short sellers	0.18	-0.10	11.41	-6.88	3.61	0.56	3.46	-0.03	-1.01
Value Index	0.61	1.11	5.68	-9.65	2.54	-1.21	5.94	0.13	0.56
Directional Trading Group	0.89	0.70	7.50	-6.20	2.49	0.27	2.91	0.25	0.10
Futures	0.67	0.21	7.71	-6.80	3.10	0.18	2.72	0.13	-0.08
Macro	0.55	0.66	4.00	-2.95	1.45	0.29	3.06	0.19	0.27
Speciality Strategies Group	0.88	1.00	7.90	-12.50	2.38	-1.16	8.28	0.25	0.37
Long-Short credit	0.76	0.84	4.30	-7.00	1.34	-1.25	9.98	0.36	0.13
Multi-Strategy Index	0.95	1.00	8.80	-9.60	2.53	-0.20	5.40	0.26	0.40
Mean	0.84	0.89	9.32	-7.75	2.41	-0.24	7.15	0.23	0.22
Weighted-Composite Index	0.56	0.90	4.75	-5.96	1.86	-1.01	5.18	0.15	0.37
S&P500	0.46	1.29	11.06	-18.47	4.62	-1.13	5.99	0.04	1.00

Notes. The statistics reported in this table are computed using the monthly returns of the Van Greenwich (VG) indices observed over the period running from January 1995 to March 2010. The weighted composite index is computed over the whole set of the VG indices.

Table 4 OLS estimation of the Fung and Hsieh model (Model 1) by strategy

	<i>c</i>	<i>r_m-r_f</i>	<i>Russel-S&P</i>	<i>BdOpt</i>	<i>ComOpt</i>	<i>CurOpt</i>	<i>CredSpr</i>	<i>10Y</i>	<i>R²</i>	<i>DW</i>
MNG	-1.8118 <i>-2.13</i>	0.3355 <i>14.53</i>	0.2481 <i>6.71</i>	-0.0202 <i>-2.62</i>	0.0043 <i>0.53</i>	-0.0014 <i>-0.24</i>	6.7783 <i>1.99</i>	4.6178 <i>2.99</i>	0.65	1.75
EMN	1.2150 <i>1.34</i>	0.1812 <i>7.38</i>	0.1378 <i>3.50</i>	-0.0264 <i>-3.22</i>	-0.0005 <i>-0.06</i>	-0.0061 <i>-0.99</i>	-7.6811 <i>-2.12</i>	0.0013 <i>0.00</i>	0.40	1.68
ED	1.0214 <i>0.80</i>	0.1579 <i>4.56</i>	0.0444 <i>0.80</i>	0.0404 <i>3.49</i>	0.0428 <i>3.50</i>	0.0335 <i>3.85</i>	0.1414 <i>0.03</i>	-1.0482 <i>-0.45</i>	0.28	1.96
DS	-1.0183 <i>-1.36</i>	0.2734 <i>13.46</i>	0.2177 <i>6.69</i>	-0.0218 <i>-3.22</i>	0.0032 <i>0.45</i>	-0.0019 <i>-0.37</i>	2.6545 <i>0.88</i>	3.4273 <i>2.53</i>	0.63	1.74
DED	-1.3741 <i>-1.96</i>	0.0610 <i>3.22</i>	0.1239 <i>4.08</i>	-0.0102 <i>-1.62</i>	0.0050 <i>0.74</i>	0.0047 <i>0.98</i>	-0.9470 <i>-0.34</i>	4.8918 <i>3.86</i>	0.26	1.96
AI	2.5212 <i>1.32</i>	0.0054 <i>0.10</i>	0.0169 <i>0.20</i>	0.0544 <i>3.15</i>	0.0620 <i>3.39</i>	0.0554 <i>4.25</i>	-5.4667 <i>-0.71</i>	-3.1404 <i>-0.91</i>	0.26	1.81
LS	-1.3409 <i>-1.92</i>	0.3568 <i>18.88</i>	0.2062 <i>6.81</i>	0.0000 <i>0.01</i>	0.0060 <i>0.90</i>	0.0026 <i>0.54</i>	4.1851 <i>1.50</i>	3.8549 <i>3.05</i>	0.73	1.62
GR	-1.5468 <i>-1.07</i>	0.7220 <i>18.50</i>	0.4227 <i>6.76</i>	0.0092 <i>0.71</i>	0.0095 <i>0.69</i>	0.0001 <i>0.01</i>	4.5399 <i>0.79</i>	4.1240 <i>1.58</i>	0.72	1.53
OI	-1.5327 <i>-1.84</i>	0.4962 <i>21.93</i>	0.2954 <i>8.15</i>	0.0034 <i>0.46</i>	0.0023 <i>0.29</i>	0.0019 <i>0.33</i>	4.0838 <i>1.22</i>	4.4550 <i>2.95</i>	0.78	1.55
SS	-0.0934 <i>-0.13</i>	0.1247 <i>6.49</i>	0.0571 <i>1.86</i>	-0.0173 <i>-2.70</i>	-0.0037 <i>-0.55</i>	0.0020 <i>0.42</i>	3.9940 <i>1.41</i>	0.3764 <i>0.29</i>	0.26	1.93
VI	0.6280 <i>0.33</i>	0.2989 <i>5.86</i>	0.1051 <i>1.29</i>	0.0276 <i>1.62</i>	0.0365 <i>2.03</i>	0.0034 <i>0.27</i>	4.5487 <i>0.60</i>	-1.8874 <i>-0.55</i>	0.17	1.87
DTGI	-0.9208 <i>-1.59</i>	0.1795 <i>11.43</i>	0.1393 <i>5.54</i>	-0.0118 <i>-2.25</i>	0.0011 <i>0.19</i>	-0.0021 <i>-0.53</i>	1.2014 <i>0.52</i>	3.3982 <i>3.24</i>	0.56	1.56
FUT	-1.5928 <i>-1.59</i>	0.3908 <i>14.38</i>	0.0951 <i>2.18</i>	-0.0161 <i>-1.77</i>	0.0109 <i>1.13</i>	0.0014 <i>0.21</i>	8.0520 <i>2.01</i>	3.3601 <i>1.85</i>	0.57	1.74
MI	-0.5282 <i>-0.56</i>	0.3578 <i>13.94</i>	0.1112 <i>2.70</i>	-0.0234 <i>-2.73</i>	0.0063 <i>0.70</i>	0.0018 <i>0.27</i>	6.2981 <i>1.66</i>	0.9713 <i>0.57</i>	0.57	1.55
SPEC	-2.1021 <i>-1.69</i>	0.4531 <i>13.45</i>	0.3352 <i>6.21</i>	-0.0049 <i>-0.44</i>	0.0118 <i>0.99</i>	0.0061 <i>0.72</i>	5.2732 <i>1.06</i>	5.9205 <i>2.63</i>	0.59	1.64
LSCRED	1.6302 <i>0.77</i>	-0.9820 <i>-17.06</i>	-0.4627 <i>-5.02</i>	0.0056 <i>0.29</i>	-0.0074 <i>-0.37</i>	-0.0077 <i>-0.53</i>	-9.6560 <i>-1.14</i>	-1.7266 <i>-0.45</i>	0.67	1.73
MS	-1.6494 <i>-2.00</i>	0.5266 <i>23.60</i>	0.2966 <i>8.30</i>	0.0082 <i>1.10</i>	-0.0061 <i>-0.77</i>	0.0010 <i>0.18</i>	4.5087 <i>1.37</i>	4.9069 <i>3.29</i>	0.81	1.69
Mean	-0.4997 <i>1.32</i>	0.2317 <i>12.28</i>	0.1406 <i>4.52</i>	-0.0002 <i>1.85</i>	0.0108 <i>1.02</i>	0.0056 <i>0.86</i>	1.9122 <i>1.14</i>	2.1472 <i>1.84</i>	0.52	1.72
Median	-1.0183 <i>1.36</i>	0.2989 <i>13.46</i>	0.1378 <i>5.02</i>	-0.0049 <i>1.77</i>	0.0050 <i>0.70</i>	0.0018 <i>0.42</i>	4.0838 <i>1.14</i>	3.3982 <i>1.85</i>	0.57	1.73
GI	-0.4409 <i>-0.68</i>	0.1515 <i>8.60</i>	0.0526 <i>1.86</i>	-0.0041 <i>-0.69</i>	-0.0011 <i>-0.18</i>	-0.0057 <i>-1.29</i>	0.7684 <i>0.30</i>	2.2002 <i>1.87</i>	0.36	1.41

Notes. The estimated model (Model 1) is given by equation (1). The strategies, here abbreviated, are listed in the same order as in Table 3. Coefficient t-statistics are in italics.

Table 5 Adjusted R^2 : OLS regression of Model 2 explanatory variables on instrument categories

	classical	hm	d
$r_m - r_f$	0.01	0.73	0.71
<i>smb</i>	0.01	0.54	0.71
<i>CpOpt</i>	0.03	0.63	0.54
<i>10Y</i>	0.93	0.99	0.74

Note. *hm* stands for higher moment (instruments) and *d* for distance (instruments).

Table 6 Regression of Model 2 explanatory variables on the d instruments

	$r_m - r_f$	<i>smb</i>	<i>CpOpt</i>	<i>10Y</i>
d_1	1.0064	0.0016	-0.0035	0.0002
	<i>20.03</i>	<i>0.04</i>	<i>-0.18</i>	<i>0.02</i>
d_2	0.0038	1.0010	-0.0021	0.0001
	<i>0.06</i>	<i>20.48</i>	<i>-0.09</i>	<i>0.01</i>
d_3	-0.0954	-0.0244	1.0513	-0.0033
	<i>-0.49</i>	<i>-0.16</i>	<i>14.31</i>	<i>-0.08</i>
d_4	0.0180	0.0046	-0.0097	1.0006
	<i>0.09</i>	<i>0.03</i>	<i>-0.12</i>	<i>22.63</i>

Notes. Each explanatory variable is regressed on the four d IVs built using equation (24). The coefficient t-statistics are in italics.

Table 7 Model 2 estimated by OLS and some higher moment IV methods over 17 indices, 1995-2010

	<i>c</i>	<i>r_m-r_f</i>	<i>smb</i>	<i>CpOpt</i>	<i>10Y</i>	\hat{w}_{m-r_f}	\hat{w}_{SMB}	\hat{w}_{CpOpt}	\hat{w}_{10Y}	<i>R</i> ²	<i>DW</i>
OLS	-0.2397	0.2152	0.1467	0.0665	0.1460					0.56	1.76
	<i>1.52</i>	<i>12.13</i>	<i>6.79</i>	<i>2.18</i>	<i>2.18</i>						
TSLs-hm	-0.0662	0.1934	0.1819	0.0363	0.1116					0.45	1.83
	<i>1.02</i>	<i>7.39</i>	<i>4.52</i>	<i>1.90</i>	<i>1.53</i>						
TSLs-d	-0.0691	0.2439	0.1800	0.0086	0.1069					0.48	1.81
	<i>1.23</i>	<i>5.42</i>	<i>4.13</i>	<i>1.91</i>	<i>1.41</i>						
GMM-hm	-0.8527	0.1802	0.1544	0.0514	0.2670					0.44	1.70
	<i>2.83</i>	<i>6.26</i>	<i>6.32</i>	<i>2.12</i>	<i>3.27</i>						
Gmm-d	-0.0691	0.2439	0.1800	0.0086	0.1069					0.48	1.81
	<i>1.13</i>	<i>4.53</i>	<i>5.68</i>	<i>1.40</i>	<i>1.28</i>						
Mean	-0.2594	0.2153	0.1686	0.0343	0.1477					0.48	1.78
	<i>1.55</i>	<i>7.15</i>	<i>5.49</i>	<i>1.90</i>	<i>1.93</i>						
Haus-hm	-0.0662	0.1934	0.1819	0.0363	0.1116	0.0603	-0.0502	-10.6449	0.0583	0.58	1.73
	<i>1.44</i>	<i>8.84</i>	<i>6.04</i>	<i>1.86</i>	<i>2.09</i>	<i>1.33</i>	<i>2.06</i>	<i>1.53</i>	<i>1.08</i>		
<i>t-min</i>						<i>0.02</i>	<i>0.16</i>	<i>0.26</i>	<i>0.04</i>		
<i>t-max</i>						<i>3.10</i>	<i>4.27</i>	<i>3.34</i>	<i>2.32</i>		
nb. signif. strat.						<i>4</i>	<i>9</i>	<i>5</i>	<i>1</i>		
Haus-d	-0.0478	0.2439	0.1800	0.0086	0.1069	-0.0373	-0.0509	0.1399	0.0020	0.57	1.78
	<i>0.78</i>	<i>5.61</i>	<i>4.68</i>	<i>2.06</i>	<i>1.48</i>	<i>0.86</i>	<i>2.09</i>	<i>1.31</i>	<i>0.91</i>		
<i>t-min</i>						<i>0.03</i>	<i>0.26</i>	<i>0.01</i>	<i>0.14</i>		
<i>t-max</i>						<i>3.18</i>	<i>4.65</i>	<i>3.45</i>	<i>1.96</i>		
nb. signif. strat.						<i>2</i>	<i>10</i>	<i>6</i>	<i>2</i>		

Notes. The estimated model (Model 2) is given by equation (2). The econometric methods we use are explained in Table 2. The coefficient t-statistics are reported in italics. For the Haus-hm and Haus-d estimations, we provide the minimum and the maximum t-statistics over all the strategies analyzed, and also the number of strategies for which the t-statistic of the corresponding *w* is greater than 1.96 (95% confidence level). The J-tests, not reported here for the sake of simplicity, indicate that the instrument sets used to run the GMMs are well identified. Note also that the TSLs-d and the GMM-d deliver the same sets of coefficients for the explanatory variables since both methods rely on the same variance-covariance weighting matrix. However, the coefficient t-statistics differ due to the GMM iterations. Finally, as explained in this article, note that the Haus-hm and Haus-d methods deliver the same set of estimated coefficients for the explanatory variables as, respectively, the TSLs-hm and the TSLs-d methods.

Table 8 Model 3 estimated by OLS and some higher moment IV methods over 17 indices, 1995-2010

	<i>c</i>	<i>r_m-r_f</i>	<i>smb</i>	<i>Spread</i>	<i>VLX</i>	$\hat{w}_{r_m-r_f}$	\hat{w}_{SMB}	\hat{w}_{Spread}	\hat{w}_{VLX}	<i>R</i> ²	<i>DW</i>
OLS	0.6343	0.2109	0.1483	-1.0820	-0.0005					0.52	1.71
	<i>2.24</i>	<i>11.80</i>	<i>6.76</i>	<i>1.78</i>	<i>1.60</i>						
TSLS-hm	0.7666	0.1798	0.1879	-0.7131	-0.0085					0.49	1.78
	<i>2.18</i>	<i>8.62</i>	<i>5.72</i>	<i>1.64</i>	<i>1.18</i>						
TSLS-d	0.6211	0.1952	0.1929	-1.5055	0.0026					0.45	1.75
	<i>1.25</i>	<i>2.86</i>	<i>5.37</i>	<i>1.73</i>	<i>1.36</i>						
GMM-hm	1.0378	0.1924	0.1451	-0.3663	-0.0280					0.36	1.67
	<i>3.43</i>	<i>9.67</i>	<i>7.28</i>	<i>1.68</i>	<i>2.76</i>						
Gmm-d	0.6211	0.1952	0.1929	-1.5055	0.0026					0.45	1.75
	<i>1.08</i>	<i>2.37</i>	<i>5.99</i>	<i>1.61</i>	<i>0.95</i>						
Mean	0.7362	0.1947	0.1734	-1.0345	-0.0064					0.45	1.73
	<i>2.04</i>	<i>7.06</i>	<i>6.23</i>	<i>1.69</i>	<i>1.57</i>						
Haus-hm	0.7666	0.1798	0.1879	-0.7131	-0.0085	0.0783	-0.0583	-11.1124	0.0694	0.56	1.70
	<i>2.36</i>	<i>9.08</i>	<i>6.15</i>	<i>1.74</i>	<i>1.25</i>	<i>0.08</i>	<i>-0.06</i>	<i>-11.11</i>	<i>0.07</i>		
<i>t-min</i>						<i>0.11</i>	<i>0.01</i>	<i>0.15</i>	<i>0.24</i>		
<i>t-max</i>						<i>4.22</i>	<i>4.62</i>	<i>2.87</i>	<i>1.82</i>		
nb. signif. strat.						7	9	5	0		
Haus-d	0.7211	0.1952	0.1929	-1.5055	0.0026	0.0206	-0.0624	0.5771	-0.0080	0.54	1.69
	<i>1.52</i>	<i>3.46</i>	<i>4.90</i>	<i>2.05</i>	<i>1.35</i>	<i>0.78</i>	<i>2.19</i>	<i>1.16</i>	<i>0.89</i>		
<i>t-min</i>						<i>0.07</i>	<i>0.13</i>	<i>0.09</i>	<i>0.27</i>		
<i>t-max</i>						<i>2.76</i>	<i>4.87</i>	<i>3.13</i>	<i>1.95</i>		
nb. signif. strat.						2	8	4	1		

Notes. The estimated model (Model 3) is given by equation (3). The econometric methods we use are explained in Table 2. The coefficient t-statistics are reported in italics. For the Haus-hm and Haus-d estimations, we provide the minimum and the maximum t-statistics over all the strategies analyzed, and also the number of strategies for which the t-statistic of the corresponding *w* is greater than 1.96 (95% confidence level). The J-tests, not reported here for the sake of simplicity, indicate that the instrument sets used to run the GMMs are well identified. Note also that the TSLS-d and the GMM-d deliver the same sets of coefficients for the explanatory variables since both methods rely on the same variance-covariance weighting matrix. However, the coefficient t-statistics differ due to the GMM iterations. Finally, as explained in this article, note that the Haus-hm and Haus-d methods deliver the same set of estimated coefficients for the explanatory variables as, respectively, the TSLS-hm and the TSLS-d methods.

Table 9 Hausman-d tests for the market risk premium

	OLS	haus-d	Spread	φ	$t(\varphi)$
AI	0.03	0.24	-0.21	-0.25	1.88
LSCRED	-0.93	-0.77	-0.17	-0.22	1.75
EMN	0.16	0.31	-0.15	-0.18	3.18
ED	0.16	0.22	-0.05	-0.06	0.72
DS	0.25	0.28	-0.03	-0.04	0.88
DTGI	0.16	0.19	-0.03	-0.04	1.17
SS	0.11	0.12	-0.01	0.00	0.03
MS	0.50	0.50	-0.01	-0.02	0.86
OI	0.47	0.47	0.00	-0.01	0.15
VI	0.28	0.28	0.00	-0.02	0.12
MI	0.33	0.34	0.00	-0.01	0.12
LS	0.33	0.33	0.01	0.00	0.05
MNG	0.31	0.29	0.02	0.03	0.50
SPEC	0.41	0.39	0.02	0.02	0.34
DED	0.05	0.01	0.04	0.06	1.29
FUT	0.37	0.33	0.04	0.05	0.69
GR	0.68	0.63	0.05	0.07	0.82
GI	0.14	0.21	-0.08	-0.10	2.51

The spread (measurement error) is the difference between the OLS coefficient and the corresponding Hausman coefficient resulting from the estimation of the Hausman artificial regression (Equation (23)). For each spread, we provide the coefficient φ of the corresponding artificial variable. Note the strong positive relationship between the spread and φ , the strategies being reported in increasing order of the spread.

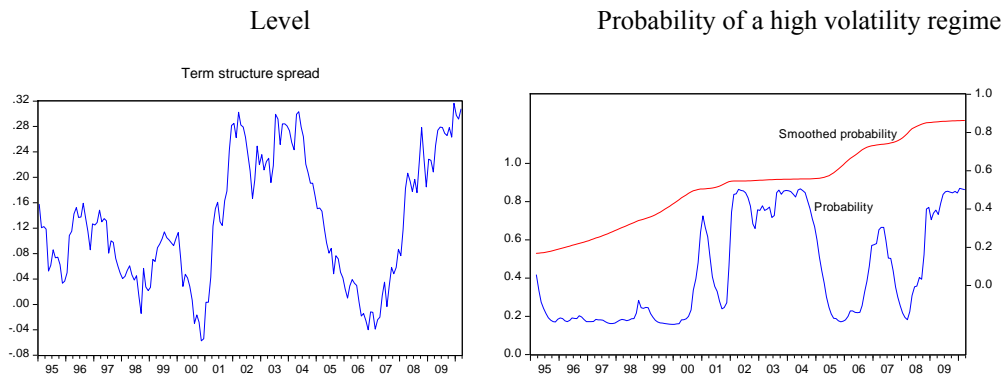
Table 10 Hausman-d tests for *SMB*

	OLS	haus-d	Spread	φ	$t(\varphi)$
SPEC	0.36	0.61	-0.25	-0.36	4.65
GR	0.44	0.62	-0.18	-0.26	2.80
DED	0.13	0.27	-0.14	-0.20	4.12
MNG	0.24	0.36	-0.12	-0.19	3.30
DS	0.21	0.34	-0.12	-0.18	3.66
EMN	0.16	0.25	-0.09	-0.12	1.86
ED	0.08	0.16	-0.08	-0.11	1.09
DTGI	0.14	0.20	-0.06	-0.10	2.47
AI	0.00	0.04	-0.04	-0.05	0.32
LS	0.22	0.26	-0.04	-0.07	1.60
OI	0.29	0.33	-0.04	-0.06	1.20
SS	0.05	0.04	0.01	0.01	0.26
FUT	0.11	0.08	0.03	0.04	0.57
MI	0.14	0.04	0.10	0.13	1.87
MS	0.27	0.17	0.10	0.12	2.09
VI	0.19	0.03	0.16	0.22	1.56
LSCRED	-0.54	-0.74	0.21	0.31	2.17
GI	0.05	0.00	0.05	0.06	1.41

The spread (measurement error) is the difference between the OLS coefficient and the corresponding Hausman coefficient resulting from the estimation of the Hausman artificial regression (Equation (23)). For each spread, we provide the coefficient φ of the corresponding artificial variable. Note the strong positive relationship between the spread and φ , the strategies being reported in increasing order of the spread.

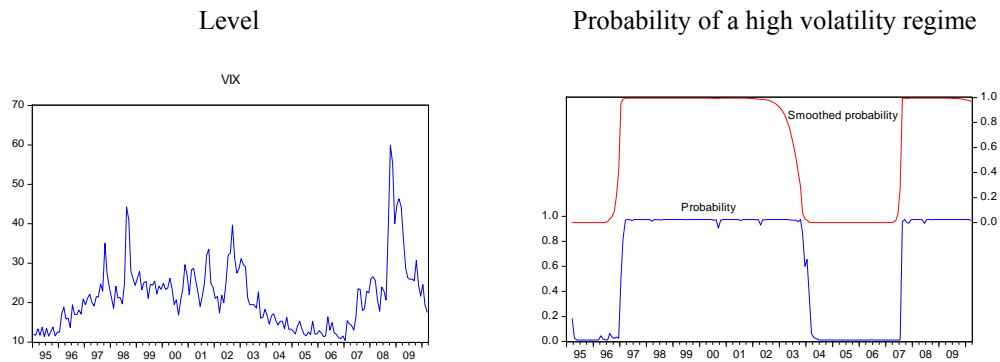
Figures

Figure 1 U.S. term-structure spread, 1995-2010



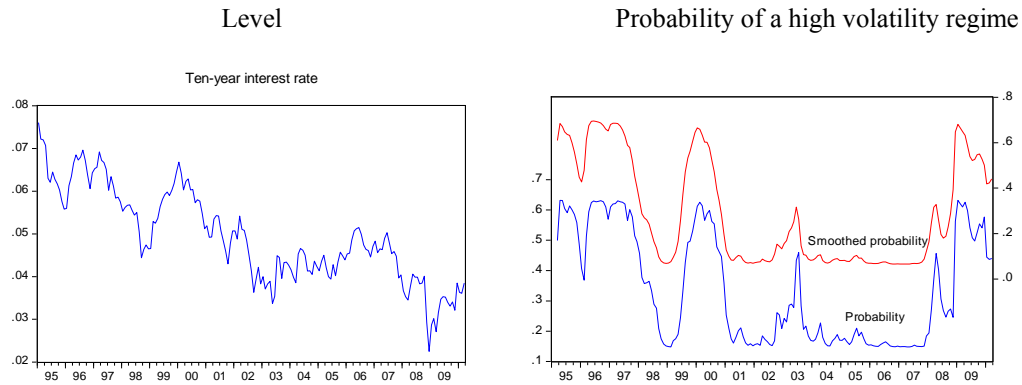
Notes. The term-structure spread is the difference between the ten-year and the three-month interest rates. The probability of a high regime is computed with a Markov switching regime algorithm.

Figure 2 VIX, 1995-2010



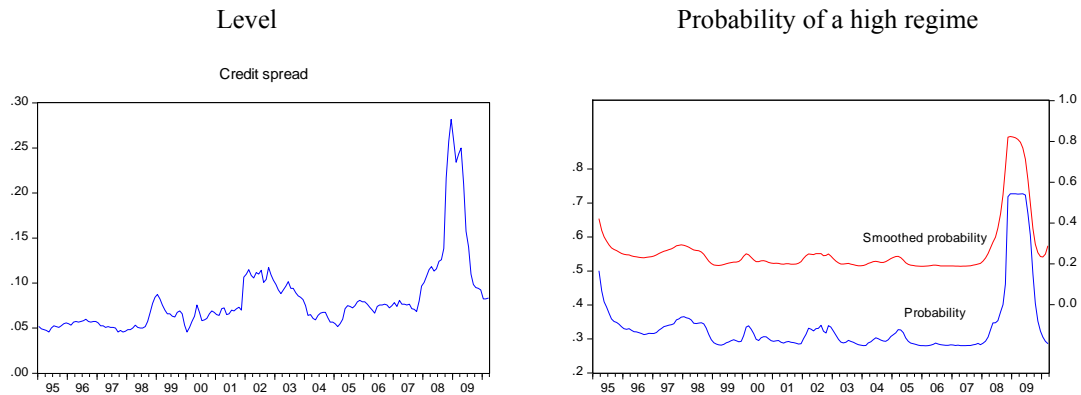
Notes. The VIX is the implicit market volatility computed using options on the S&P. The probability of a high regime is computed with a Markov switching regime algorithm.

Figure 3 U.S. ten-year interest rate, 1995-2010



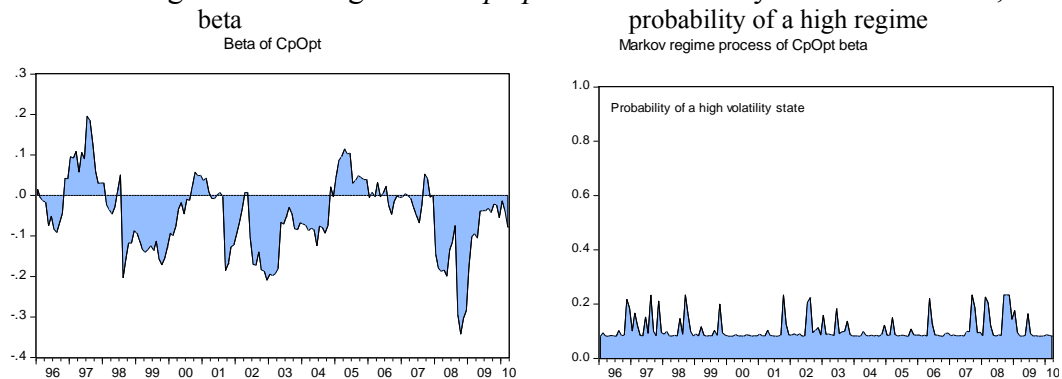
Note. The probability of a high regime is computed with a Markov switching regime algorithm.

Figure 4 U.S. credit spread, 1995-2010



Notes. The credit spread is the difference between the BBB and AAA corporate bond yields. The probability of a high regime is computed with a Markov switching regime algorithm.

Figure 5 Markov regime switching for the $CpOpt$ beta obtained by the Kalman filter, 1995-2010



Notes. $CpOpt$ is a time series obtained by computing the first principal component of the three F&H lookback option factors. The probability of a high regime is computed with a Markov switching regime algorithm.

Figure 6 Regression of spread on φ for the market risk premium

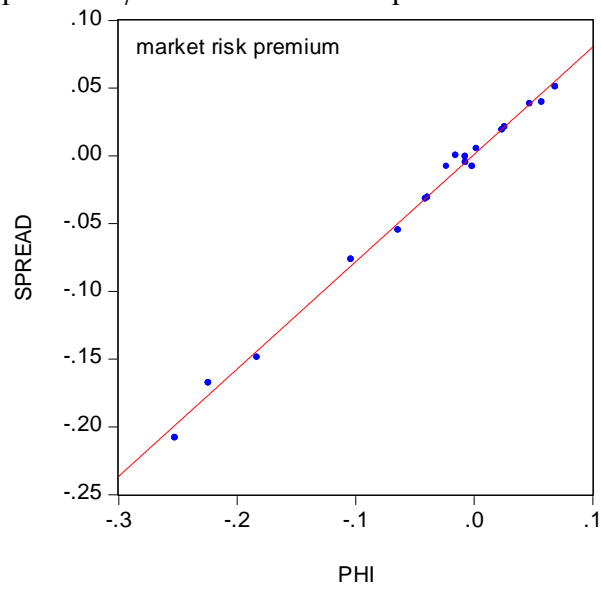


Figure 7 Regression of spread on φ for *SMB*

