Original Article

Hedge fund return specification with errors-in-variables

Received (in revised form): 20th June 2010

Alain Coën
is Associate Professor of Finance at the Graduate School of Business (ESG) of the University of Quebec in Montreal (UQAM). He obtained his PhD in Finance from the University of Grenoble, and his PhD in Economics from the University of Paris I Panthéon-Sorbonne. He holds a Master of Arts in Economics with major in Macroeconomics from Laval University and an Accreditation to supervise research (HDR) from Paris-Dauphine University. His research interests focus on asset pricing, international finance, hedge funds, business cycles and financial econometrics. A part of the present paper was revised when he was at HEC-ULg: he would like to thank the KBL Chair in Fund industry during this period. He has published in several international journals and has written a book in financial management.

Georges Hübner
(PhD, INSEAD) is the Deloitte Professor of Financial Management, co-chair of the Finance Department and Board member of HEC-Liège. He is also Associate Professor of Finance at Maastricht University and Affiliate Professor at EDHEC (Lille/Nice). Georges regularly provides preparation seminars for the Global Association of Risk Professionals (GARP) and Chartered Alternative Investment Analyst (CAIA) certifications. He is also co-founder and managing director of Gambit Financial solutions, a spin-off company of HEC Management School – University of Liège that produces sophisticated software solutions for investor profiling, portfolio optimization and risk management. Georges has published numerous books and research articles about credit risk, hedge funds and derivatives. He obtained the best paper award 2002 of the *Journal of Banking and Finance*, and the *Operational Risk & Compliance Achievement Award 2006* for the best academic paper. He is the inventor of the Generalized Treynor Ratio, a popular portfolio performance measure.

Aurélie Desfleurs
is Associate Professor of Finance at the Department of Accounting of the University of Quebec in Outaouasis (UQO). She is graduated from EDHEC School of Management, received her MBA and PhD in Finance from Laval University and teaches investment and corporate finance. Her research interests include asset pricing and earnings forecasting. She has published several papers in the areas of multinational finance and financial forecasting.

Correspondence: Alain Coën, Department of Finance, Graduate School of Business (Ecole des Sciences de la Gestion: ESG), University of Quebec in Montreal (UQÀM), Succursale centre-ville, Case postale 6192, Montréal, QC, Canada H3C 4R2
E-mail: coen.alain@uqam.ca

ABSTRACT **In linear models for hedge fund returns, errors-in-variables may significantly alter the measurement of factor loadings and the estimation of abnormal performance. The higher moment estimator (HME) introduced by Dagenais and Dagenais (1997) effectively deals with these issues. Results on individual funds show that the HME specification does not uncover systematic performance biases, but can modify estimated alphas in most cases and identifies relative persistence for directional funds in bearish market conditions.**

Overall, the risk premia calculated with HME remain relatively stable when compared to ordinary least squares specifications.

Journal of Derivatives & Hedge Funds (2010) **16,** 22–52. doi:10.1057/jdhf.2009.19

Keywords: errors-in-variables; measurement errors; hedge fund performance; asset pricing models

INTRODUCTION

Early research on hedge fund performance, such as that by Fung and Hsieh[1] and Liang,[2] attempted to decompose the required returns on hedge fund strategies through linear combinations of returns on different asset classes such as stocks, bonds and commodities. Regardless of the strategy under study, these traditional sets of factors could not produce significance levels higher than 80 per cent, even with as many as 11 risk premia.

To improve on the static factor model, two streams have emerged in the literature. The first approach is to relax the assumption of constant exposures to the risk factors. This can be achieved by measuring regime-switching betas, as in the study by Edwards and Caglayan[3] and Capocci *et al*,[4] who measure different betas in up and down markets, or by allowing for time-varying betas. This approach is advocated by Fung and Hsieh[5] and Posthuma and Van der Sluis[6] in a Kalman filtering approach.

The second approach aims at capturing the nonlinear risk exposures of hedge fund returns through the factor premia themselves, leaving the betas constant. Mitchell and Pulvino[7] introduce piecewise linear regressions to account for nonlinearities in hedge fund returns for risk arbitrage strategies. Fung and Hsieh[5,8] introduce risk premia accounting for option straddles for trend-following funds, as these strategies tend to exhibit payoffs that resemble long lookback option positions. In a similar vein, Agarwal and Naik[9] adapt a framework proposed by Glosten and Jagannathan[10] to estimate the returns of exchange-traded calls and puts. Finally, Chan *et al*[11] apply a regime-switching multi-moments model to hedge fund returns.

Irrespective of the approach considered, a highly non-normal behavior of returns is likely to flaw statistical inference if there is evidence of measurement errors in the explanatory variables. These errors influence the point estimators of the risk factor loadings. Errors in the variables may lead to the non-convergence of the ordinary least squares (OLS) estimator, casting doubt on the results. Paradoxically, few theoretical or applied efforts have been made to reduce this important bias. Fama and MacBeth[12] try to reduce measurement errors, grouping equities in portfolios. Shanken[13] suggests adjusting standard errors to correct biases induced by errors-in-variables (EIV). Kandel and Stambaugh[14] use a Generalized Least Squares method, but this requires a tedious estimation of covariance matrix.

The use of estimators based on moments of order higher than two proposed by Cragg,[15,16] Dagenais and Dagenais[17] and Lewbel[18] is well suited to dealing with the bias resulting from measurement error. In the context of hedge funds, Coën and Hübner[19] have shown that Dagenais and Dagenais's[17] higher moment estimator (HME), which belongs to the class of instrumental variables (IV) estimators, can account for the nonlinearity of exposures. Thus, the higher moment (HM) technique appears to be particularly well suited for hedge funds. We apply it to a set of equally weighted hedge fund indexes, as well as to a set of individual funds, both for the analysis of risk-adjusted performance and for the persistence of this performance. The analysis is carried out in bullish and bearish sub-periods.

Unfortunately, the Dagenais and Dagenais[17] method displays a practical drawback. For each independent variable used in the original return-generating process, the corrected estimator generates a new regressor that specifically accounts for the estimated measurement error. This mechanical adjustment doubles the number of variables, and may lead to a serious model over-specification, especially for small databases. We propose a simple heuristic method to circumvent this issue by setting up a recursive regression algorithm that gradually unfolds the set of independent variables. This procedure allows us to ensure proper care of the measurement errors while limiting the inflation in the variables to what is strictly necessary to optimize the significance level of the regression.

We apply this approach to a set of hedge fund indexes constructed with equally weighted hedge fund portfolios.[20] Unlike the investable indexes that can be retrieved by data providers, this database most closely resembles portfolio returns that researchers typically use in empirical studies. Using this database, we can assess the impact of our procedure for research purposes. We estimate the return-generating process with a mix of asset- and option-based explanatory factors. Applying our procedure to a set of individual funds, we find that alphas are impacted by the presence of measurement errors, but without overwhelming evidence of a systematic bias in either direction. Furthermore, the HM technique allows us to uncover some mild evidence of persistence in risk-adjusted performance.

Our results indicate that the correction for measurement errors has a significant impact on the performance measurement of hedge fund strategies, especially when option-based strategies are considered. We also propose a heuristic approach to practically reduce the number of variables in most cases, leaving the economic interpretation of the risk premia unchanged, while reducing the number of instruments in the return-generating process.

The article is organized as follows. The next section provides the econometric framework for the EIV correction. The section after that contains the description of data, the risk factors (asset- and option-based) and the model specifications. In the subsequent section, we analyze the results of the HME regression estimations for equally weighted hedge indexes and individual hedge funds. The penultimate section analyzes the empirical properties of different model specifications with the application of a heuristic reduction algorithm, and discusses the stability of risk premia. The final section concludes.

ESTIMATION METHODS

In the fundamental work of Frisch,[21] the analysis of measurement errors in the independent variables[22] of a regression, also called EIV, played a central role in the early days of econometrics. Although the economic literature acknowledges that EIV leads to inconsistent linear OLS estimators, this issue has only resurfaced with the work of Cragg[15] and Dagenais and Dagenais.[17] The reason for this general neglect of EIV arises from an oversimplification of the problem. The analysis of measurement error in finance has mostly been focused on the one-factor linear model. In this case, measurement error creates an 'attenuation effect' that biases the slope coefficient toward zero (see Cragg[15]): EIV seem at worst to give rise to conservative estimates. But in the context of multiple regressions, Cragg[15,16] emphasizes a 'contamination effect' that is much more difficult to handle.

Indeed, the bias of any given parameter not only depends on its own error (the attenuation effect), but also on the errors in all others variables (the contamination effect). Thus, every parameter of a multiple regression is inconsistent provided at least one variable is measured with error.

The problem of EIV can be illustrated through the estimation of the following K-factor model of asset return, R_t:

$$R_t = \alpha + \tilde{F}_t'\beta + u_t \tag{1}$$

where α is a constant, $\tilde{F}_t$ is a column vector representing the realization of the K factors in period t, β is the vector of the k factor loadings and u_t is a residual idiosyncratic risk.

Suppose all factors are unobserved and the relationships between true and observed factors are additive.

$$F_t = \tilde{F}_t + v_t \tag{2}$$

where F_t, $\tilde{F}_t$ and v_t are $K \times 1$ column vectors holding the observed factors, the true factors and the measurement errors, respectively. Consider further that v_t has mean zero and variance matrix equal to Σ_{VV}. The measurement errors (v_t) are moreover assumed to be independent through time and uncorrelated[23] with the true unobserved variables, $\tilde{F}_t$, and the residual idiosyncratic risk, u_t. Substituting (1) in (2) yields:

$$R_t = \alpha + F_t \cdot \beta + u_t - v_t \cdot \beta \tag{3}$$

The OLS estimates $\hat{\alpha}$ and $\hat{\beta}$ are inconsistent because the compound error in (3), $u_t - v_t \cdot \beta$, is correlated with regressor F_t through the measurement error v_t.[24] In the presence of EIV, factor loading estimates are contaminated by the attenuation and contamination biases mentioned above.

Without further assumptions, the parameters of EIV models (1) and (2) are not identified. The standard solution suggested in the literature is to introduce additional moment conditions. If there exist IV correlated with the true regressors but not with measurement errors, then adding these moments can help to solve the identification problem. Specifically, if the distributions of factors $\tilde{F}_t$ are skewed and exhibit non-zero excess kurtosis, then Cragg[16] and Dagenais and Dagenais[17] show that own and cross third and fourth moments of these explanatory variables are valid instruments that can be used as additional moment restrictions to consistently estimate parameters α and β.

Following Durbin[25] and Pal,[26] Dagenais and Dagenais[17] introduce new unbiased HME exhibiting *considerably smaller standard errors*. Under the hypothesis of no measurement error in the variables, the estimators introduced by Durbin[25] and Pal[26] are unbiased. But, as demonstrated by Kendall and Stuart[27] and Malinvaud,[28] these HME have higher standard errors than the corresponding least squares estimators, and may be described as more erratic. Taking into account this feature, Dagenais and Dagenais[17] develop a new instrumental variable estimator, β^{HM}, which is a linear matrix combination of the generalized version of β_d, Durbin's estimator, and β_p, Pal's estimator. More precisely, the instruments are $z_1 = f * f$, $z_2 = f * f * f - 3f[E(f'f/N) * I_K]$ *and a constant.* f_{ij} *are the elements of the matrix* f *and* $f = AF$ *where* $A = I_N - ii'/N$. *The matrix* f *is the* $T \times K$ *matrix* F *calculated in mean deviation, standing for the matrix of* K *factor loadings* *where* T *is here the number of observations. The symbol* **is the Hadamard element-by-element matrix multiplication operator.*

Dagenais and Dagenais's estimator can most easily be computed by means of artificial regressions, as suggested by Davidson and MacKinnon.[29] The first step consists in constructing estimates $\hat{F}_{kt}$ of the true regressors. This is carried out with K artificial regressions having F_{kt} as dependent variables and third and fourth moments (own and cross moments) of F_{kt} as independent variables. These are then used to construct measures of the EIV $\hat{w}_{kt} = F_{kt} - \hat{F}_{kt}$. The latter are then introduced as additional regressors, called the adjustment variables, in equation (4) as follows:

$$R_t = \alpha + \sum_{k=1}^{K} \beta_k^{HM} \cdot F_{kt} + \sum_{k=1}^{K} \psi_k \cdot \hat{w}_{kt} + \varepsilon_t \qquad (4)$$

With this new model specification, we can test the null hypothesis (H_0: $\hat{w}_{tk} = 0$, $k = 1, \ldots, K$) that there are no EIV applying a Durbin-Wu-Hausman type test. Furthermore, the HM estimator automatically corrects to take into account this bias owing to measurement error. If there is no EIV, then the model collapses to OLS.

DATA AND EMPIRICAL METHODS

Hedge funds data

We use the Barclay Group database with monthly net returns on 2617 funds belonging to 11 strategies: Event Driven (EDR), Funds of Funds (FOF), Global (GLO), Global Emerging Markets (GEM), Global International Markets (GIN), Global Macro (GMA), Global Regional Established (GES), Long Only Leveraged (LOL), Market Neutral (MKN), Sector (SEC) and Short Selling (SHO) for the period January 1994 – December 2002. Of these funds, 1589 were still alive at the end of the period and 1028 had ceased reporting before the end of the time window. Funds that reported fewer than one consecutive year of returns have been removed from the database.[30] Data from the same period were used by Cappoci *et al*[4] with the Managed Account Reports (MAR) database, and have been found to be relatively reliable in returns of survivorship and instant return history biases.

The series of dependent variables in our regression are built by computing the equally weighted average monthly returns of all funds, either living or dead, that follow a particular strategy during a given month. We also use individual hedge fund data to assess the relative significance of risk-adjusted performance measures (alphas) obtained with alternative specifications, as well as the persistence of performance under varying market conditions.

Risk factors

To implement the estimation procedure and the recursive regression algorithm, we use an extended version of the Fama-French[31]–Carhart[32] linear asset pricing model.

We start the implementation with the four-factor model proposed by Carhart,[32] supposedly achieving better significance levels than the Fama and French[31] specification for hedge fund returns (see Agarwal and Naik and[9] Capocci and Hübner[33]). This market model is taken as the benchmark of a correctly specified model. Its equation is:

$$R_t = \alpha + \beta_m MKT_t + \beta_s SMB_t + \beta_h HML_t + \beta_u UMD_t + \varepsilon_t \qquad (5)$$

where R_t is the hedge fund return in excess of the 13-week T-Bill rate, MKT_t is the excess return on the market index proposed by Fama and French,[31] SMB_t is the factor-mimicking portfolio for size (*small minus big*), HML_t is the factor-mimicking portfolio for the book-to-market effect (*high minus low*) and $UMD_t =$ the factor-mimicking portfolio for the momentum effect (*up minus down*). Factors are extracted from French's website.

This specification typically achieves significance levels that can easily be improved with style-based indexes. Among them, Capocci and Hübner[33] show that an additional factor accounting for the emerging bond market investment strategy triggers a major shift in the explanatory power of the hedge fund return regressions. Consequently, we choose this particular asset-based index as the fifth independent variable.

Finally, we introduce an option-based factor as the sixth regressor. To make sure that the way this variable is constructed does not unduly alter the analysis, we propose two alternative characterizations.[34]

First, we construct monthly returns from index options with a procedure similar to that put forward by Agarwal and Naik[9] to build two series of actual returns of at-the-money (ATM) put and call options. As options are never perfectly ATM, we approximate each option closing price on the last trading day of the month with a linear interpolation of the closest in-the-money (ITM) and out-of-the-money (OTM) option prices. The next month, we use the same technique to obtain the closing price. This method ensures the time consistency of the series of options used. We apply a similar technique for OTM puts and calls, where the strike price is 5 per cent away from the current value of the index. The choice of this degree of moneyness is consistent with the results empirically derived by Diez de los Rios and Garcia.[35] The variables corresponding to these series of options are called ACr, OCr, APr and OPr for ATM and OTM calls and ATM and OTM puts, respectively.

Second, we compute artificial option returns with a procedure that refines that used by Glosten and Jagannathan.[10] Each month, we identify the value of the S&P500 index. We then construct four sets of synthetic options with 1 month to maturity: an ATM put, an ATM call, an OTM put and an OTM call. The initial price of these options is proxied by using the Black-Scholes formula with the continuously compounded 1-month T-bill rate (risk-free rate), the historical volatility on the S&P500 (volatility) and the contemporaneous value of the S&P500 index multiplied by 0.95 (for the OTM puts), by 1 (for the ATM options) and by 1.05 (for OTM calls) as the strike prices. We call ACa, OCa, APa, OPa the series of realized returns on these artificial strategies.

The most comprehensive specification is a six-factor model depicted in equation (6).

$$\begin{aligned} R_t = {} & \alpha + \beta_m MKT_t + \beta_s SMB_t + \beta_h HML_t \\ & + \beta_u UMD_t + \beta_e EMB_t + \beta_o Opt_t + \varepsilon_t \end{aligned} \quad (6)$$

where Opt is the option-based factor among ACr, OCr, APr, OPr, ACa, OCa APa and OPa that provides the highest level of information in the regression. The estimated regression coefficients will be noted $\hat{\alpha}^{OLS}$ and $\hat{\beta}_k^{OLS}$ for $k \in \{m, s, h, u, e, o\}$.

Similarly, the HM specification used to estimate the same model has the following form:

$$\begin{aligned} R_t = {} & \alpha + \beta_m MKT_t + \beta_s SMB_t \\ & + \beta_h HML_t + \beta_u UMD_t + \beta_e EMB_t \\ & + \beta_o Opt_t + \psi_m \hat{w}_{mt} + \psi_s \hat{w}_{st} + \psi_h \hat{w}_{ht} \\ & + \psi_u \hat{w}_{ut} + \psi_e \hat{w}_{et} + \psi_o \hat{w}_{ot} + \varepsilon_t \end{aligned} \quad (7)$$

Table 1: Descriptive statistics of hedge fund strategies

	No. of funds	*Living*	*Dead*	*Mean*	*SD*	*t(mean)*	*Median*	*Min*	*Max*	*M. exc.*	*t(m. exc.)*	*Sharpe*
EDR	226	154	72	0.92	1.88	5.02	1.09	−8.45	4.96	0.55	3.01	0.49
FOF	599	410	189	0.70	1.76	4.06	0.61	−6.71	6.45	0.33	1.90	0.40
GLO	156	1	155	0.41	3.98	1.05	0.54	−28.39	13.90	0.04	0.10	0.10
GEM	147	100	47	1.09	4.95	2.25	1.75	−21.85	14.35	0.72	1.49	0.22
GIN	71	50	21	0.84	2.44	3.55	0.84	−6.80	8.92	0.47	1.99	0.35
GMA	129	52	77	0.81	2.21	3.74	0.53	−4.06	7.05	0.44	2.03	0.37
GES	467	306	161	1.23	3.23	3.90	1.07	−9.96	12.24	0.86	2.73	0.38
LON	32	19	13	0.88	5.89	1.52	1.37	−17.44	13.28	0.51	0.88	0.15
MKN	574	365	209	1.19	1.16	10.47	1.16	−3.49	3.86	0.82	7.20	1.02
SEC	182	111	71	1.63	4.44	3.76	2.06	−13.11	19.90	1.26	2.91	0.37
SHO	34	21	13	0.98	4.50	2.23	0.69	−13.63	13.24	0.61	1.38	0.22

This table shows the mean returns, *t*-stats for mean=0, standard deviations, medians, minima, maxima, mean excess returns, *t*-stats for mean excess return=0, Sharpe ratios for the individual hedge funds in our database following 11 active strategies for the sample period 1994:02–2002:12. Sharpe ratio is the ratio of excess return and standard deviation. No. of funds represent the number of funds following a particular strategy, living funds and dead funds represent the number of surviving and dead funds (in December 2002, without considering the new funds established in 12:2002). We calculate the mean excess return and the Sharpe ratio considering the Ibbotson Associates 1-month T-bills. Returns in the table are in percentages. EDR=Event Driven, FOF=Funds of Funds, GLO=Global, GEM=Global Emerging Markets, GIN=Global International Markets, GMA=Global Macro, GES=Global Regional Established, LON=Long Only Leveraged, MKN=Market Neutral, SEC=Sector, SHO=Short Selling.

where $\hat{w}_{kt}$ are the adjustment variables for $k \in \{m, s, h, u, e, o\}$. Again, the estimated regression coefficients will be noted $\hat{\alpha}^{HM}$, $\hat{\beta}_k^{HM}$ and $\hat{\psi}_k$, obtained by applying Dagenais and Dagenais's[17] artificial regression technique.

We can note that the HM estimator gives the same result as the OLS estimator in perfect absence of EIV. This point can be interpreted as an illustration of the superiority of HM, and empirically confirms Dagenais and Dagenais' numerical and simulated results: *The relative performance of HM estimators is always superior to that of OLS estimators, when there are errors in variables* (Dagenais and Dagenais,[17] p. 209).

Thus, one could stick to an OLS specification only when there is no evidence of enhancement of model quality, thanks to the HM specification. This enhancement is assessed through the test of statistical significance of the presence of EIV. We suggest the use of the following decision rule: if presence of EIV is statistically significant, use the HM estimator; if not, the OLS estimator can be used.

Of course, the consequences of wrong decisions based on linear asset pricing models are straightforward in the financial industry. Furthermore, thanks to this new approach, we shed a new light on absolute returns, comparing

Table 2: Descriptive statistics of risk factors

	Mean	*SD*	*t(mean)*	*Median*	*Min*	*Max*	*M. exc.*	*t(m. exc.)*	*Sharpe*	*Skewness*	*Kurtosis*
MKT	0.76	4.83	1.61	1.54	−15.69	8.33	0.39	0.83	0.16	−0.70	0.30
SMB	0.01	4.47	0.01	−0.41	−16.26	21.38	−0.36	−0.83	0.00	0.88	5.65
HML	0.64	4.19	1.57	0.77	−8.91	13.67	0.27	0.67	0.15	0.47	1.10
UMD	1.14	5.82	2.01	1.32	−25.13	18.21	0.77	1.36	0.20	−0.73	4.46
EMB	−0.60	7.35	−0.83	0.40	−34.65	12.71	−0.97	−1.35	−0.08	−1.12	3.56
ACr	0.01	0.82	0.06	−0.21	−0.98	1.88	−0.36	−4.58	0.01	0.46	−1.02
OCr	−0.01	1.28	−0.10	−0.58	−0.99	4.79	−0.38	−3.07	−0.01	1.76	2.95
APr	−0.17	0.89	−1.95	−0.59	−0.94	2.55	−0.54	−6.20	−0.19	1.33	0.77
OPr	−0.26	1.13	−2.32	−0.78	−0.97	5.17	−0.63	−5.66	−0.23	2.75	8.86
ACa	0.35	1.45	2.47	0.00	−1.00	4.18	−0.02	−0.14	0.24	0.68	−0.87
OCa	0.39	1.62	2.45	−0.26	−1.00	4.93	0.02	0.12	0.24	0.84	−0.62
APa	0.06	1.86	0.34	−1.00	−1.00	10.27	−0.31	−1.70	0.03	2.45	8.20
OPa	1.24	5.02	2.52	−1.00	−1.00	30.10	0.87	1.77	0.25	3.85	18.23

This table shows the mean returns (in per cents), *t*-stats for mean=0, standard deviations, medians, minima, maxima, mean excess returns, *t*-stats for mean excess return=0, Sharpe ratios, skewness and kurtosis for the premia for the sample period 1994:02–2002:12. Sharpe ratio is the ratio of excess return and standard deviation. We calculate the mean excess return and the Sharpe ratio considering the Ibbotson Associates 1-month T-bills. Numbers in the table are in percentages. MKT=the market premium, SMB=Small Minus Low which is the difference between the returns on a portfolio of small stocks and a portfolio of large stocks, HML=High Minus Low, which is the difference between the returns on a portfolio of high-book-to-market-equity stocks and a portfolio of low-book-to-market-equity stocks, UMD=Momentum factor (see Carhart[32]), EMB=MSCI emerging market index; ACr=return of a true ATM index call, OCr=return of a true 5% OTM index call; APr=return of a true ATM index put, OPr=return of a true 5% OTM index put; ACa=return of an artificial ATM index call; OCa=return of an artificial 5% OTM index call; APa=return of an artificial ATM index put; OPa=return of an artificial 5% OTM index put. The underlying index for all options is the S&P500.

$\hat{\alpha}^{OLS}$ with $\hat{\alpha}^{HME}$, given that we know that the HM specification is at least as good as classical OLS.

Descriptive statistics

The descriptive statistics of our sample are given in Tables 1 and 2. Our database includes a substantially higher number of dead funds (+446) than the MAR database used by Capocci *et al*,[4] especially for the GES (+97 dead funds), FOF (+77) and MKN (+72) strategies.

Consistent with previous studies, some strategies appear to achieve extremely favorable performance for all measures. SEC, GES, GEM and MKN strategies exhibit average monthly returns greater than 1 per cent. The Sharpe ratio of MKN funds is up to eight times greater than that of the market proxy. EDR, SEC, GES, GMA and FOF strategies also obtain Sharpe ratios more than twice higher than the market proxy. Thus, a classical model-free performance measure suggests that there might be significant abnormal performance present in hedge fund returns (Table 3).

As acknowledged by a growing literature, the two first moments of returns are insufficient to provide a good description of risk. Descriptive statistics reported in Table 2 confirm this view, and cast doubt on the normality of the returns and the risk factor loadings. If the risk factor loadings exhibit non-Gaussian higher moments, HM estimators could be used to obtain consistent estimators, and thus to correct for EIV. To test for the normality of the distributions, we use a series of tests (Jarque-Bera, Lilliefors, Cramer-von Mises, Watson and Anderson-Darling). Results are conclusive: we can reject the hypothesis of normality for all strategies (except SHO) and risk premia, with the exception of the HML factor. This indicates that higher moments (than the variance) of the regressors are highly likely to influence hedge funds performance measurement.

The correlation between and among hedge and among risk factors is reported in Tables 4 and 5. The correlations between the regressors and the hedge fund returns do not exceed 0.80, except LOL and GES, which display high correlations with the market proxy. The correlation among the asset-based regressors is low, thereby raising no serious concern about multi-colinearity. Nevertheless, the introduction of an option-based factor induces a high correlation with the market excess return variable (MKT), especially when an ATM option is used. Furthermore, these mutli-colinearity problems also exist among option-based factors. This feature of optional factors suggests that one should be particularly cautious when interpreting regression coefficients arising from a specification using several option-based factors, such as in the studies by Agarwal and Naik[9] or Bailey *et al*.[36]

MULTI-FACTOR MODEL AND RESULTS

Analysis of equally weighted fund indexes

The HM estimation procedure entails that the regression results are directly comparable with the OLS results for each original asset pricing specification. Thus, we run OLS on our four-, five- and six-factor models depicted above, and compare the significance levels achieved with the HME procedure. We use a standard F-test to detect the presence of EIV (Durbin-Wu-Hausman test): we test for $\psi_k = 0$ for each k. All F-stats are statistically significant at the 1 per cent level, highlighting the presence of EIV in all regressions. OLS estimates are biased. The results are presented in Table 6.

We split this table into three panels: Panel A displays strategies for which the lowest significance levels are achieved ($\bar{R}^2$ below 70 per cent for the 6-factor OLS model). In Panel B, results are displayed for $\bar{R}^2$ between 70 per cent and 80 per cent. Finally, strategies achieving the highest significance levels ($\bar{R}^2 > 80$ per cent) are reported in Panel C.

Panel A reports a consistent result regarding the additional information brought by HME for the MKN and three internationally driven

Table 3: Normality tests of hedge funds strategies and risk premia

	Jarque-Bera		*Lilliefors*		*Cramer-v. Mises*		*Watson*		*Anderson-Darling*	
	Stat.	*P-value*	*Stat.*	*P-value*	*Stat.*	*P-value*	*Stat.*	*P-value*	*Stat.*	*P-value*
MKN	11.439	0.003	0.052	0.100	0.035	0.762	0.035	0.718	0.307	0.558
GLO	2943.61	0.000	0.131	0.000	0.607	0.000	0.590	0.000	4.143	0.000
GIN	9.453	0.008	0.077	0.100	0.089	0.159	0.086	0.146	0.642	0.092
GMA	4.100	0.128	0.097	0.016	0.155	0.020	0.141	0.021	0.817	0.034
FOF	42.027	0.000	0.072	0.100	0.122	0.055	0.120	0.043	0.892	0.022
GEM	58.384	0.000	0.064	0.100	0.090	0.152	0.086	0.145	0.622	0.102
EDR	111.499	0.000	0.061	0.100	0.079	0.214	0.070	0.244	0.622	0.103
SHO	2.045	0.359	0.049	0.100	0.037	0.727	0.036	0.703	0.285	0.621
LON	3.214	0.200	0.073	0.100	0.078	0.215	0.065	0.281	0.466	0.248
SEC	30.779	0.000	0.074	0.100	0.095	0.130	0.095	0.107	0.653	0.086
GES	10.630	0.004	0.057	0.100	0.050	0.507	0.050	0.467	0.435	0.295
MKT	8.557	0.013	0.105	0.006	0.182	0.009	0.150	0.015	1.071	0.008
SMB	137.387	0.000	0.091	0.031	0.176	0.011	0.162	0.011	1.312	0.002
HML	7.980	0.018	0.071	0.100	0.079	0.211	0.071	0.231	0.567	0.139
UMD	85.77	0.000	0.131	0.000	0.470	0.000	0.465	0.000	2.480	0.000
EMB	70.119	0.000	0.116	0.001	0.217	0.003	0.189	0.004	1.154	0.005

Table 3 *continued*

	Jarque-Bera		*Lilliefors*		*Cramer-v. Mises*		*Watson*		*Anderson-Darling*	
	Stat.	*P-value*	*Stat.*	*P-value*	*Stat.*	*P-value*	*Stat.*	*P-value*	*Stat.*	*P-value*
ACr	8.238	0.016	0.125	0.000	0.407	0.000	0.380	0.000	2.699	0.000
APr	32.192	0.000	0.211	0.000	1.456	0.000	1.275	0.000	8.062	0.000
OCr	85.549	0.000	0.221	0.000	0.138	0.000	1.194	0.000	7.999	0.000
OPr	436.781	0.000	0.285	0.000	2.576	0.000	2.320	0.000	13.167	0.000
ACa	11.207	0.003	0.214	0.000	0.978	0.000	0.916	0.000	6.161	0.000
APa	365.228	0.000	0.325	0.000	2.802	0.000	2.582	0.000	14.201	0.000
OCa	13.735	0.001	0.226	0.000	1.300	0.000	1.207	0.000	7.707	0.000
OPa	1563.87	0.000	0.328	0.000	3.529	0.000	3.303	0.000	17.665	0.000

This table reports the normality tests of the variables for the sample period 1994:02–2002:12: Jarque-Bera, Lilliefors, Cramer-von Mises, Watson and Anderson-Darling. EDR=Event Driven, FOF=Funds of Funds, GLO=Global, GEM=Global Emerging Markets, GIN=Global International Markets, GMA=Global Macro, GES=Global Regional Established, LON=Long Only Leveraged, MKN=Market Neutral, SHO=Short Selling, SEC=Sector; MKT=the market premium, SMB=Small Minus Low which is the difference between the returns on a portfolio of small stocks and a portfolio of large stocks, HML=High Minus Low which is the difference between the returns on a portfolio of high-book-to-market-equity stocks and a portfolio of low-book-to-market-equity stocks, UMD=Momentum factor (see Carhart[32]), EMB=MSCI emerging market index; ACr=return of a true ATM index call, OCr=return of a true 5% OTM index call; APr=return of a true ATM index put, OPr=return of a true 5% OTM index put; ACa=return of an artificial ATM index call; OCa=return of an artificial 5% OTM index call; APa=return of an artificial ATM index put; OPa=return of an artificial 5% OTM index put. The underlying index for all options is the S&P500.

Table 4: Correlations among hedge funds and between hedge funds and risk premia

	EDR	*FOF*	*GLO*	*GEM*	*GIN*	*GMA*	*GES*	*LON*	*MKN*	*SEC*	*SHO*
EDR	1.00	—	—	—	—	—	—	—	—	—	—
FOF	0.85	1.00	—	—	—	—	—	—	—	—	—
GLO	0.70	0.67	1.00	—	—	—	—	—	—	—	—
GEM	0.70	0.80	0.76	1.00	—	—	—	—	—	—	—
GIN	0.78	0.88	0.60	0.76	1.00	—	—	—	—	—	—
GMA	0.74	0.88	0.50	0.63	0.81	1.00	—	—	—	—	—
GES	0.87	0.87	0.61	0.69	0.83	0.84	1.00	—	—	—	—
LON	0.83	0.80	0.60	0.67	0.72	0.78	0.91	1.00	—	—	—
MKN	0.81	0.81	0.57	0.62	0.73	0.70	0.72	0.66	1.00	—	—
SEC	0.82	0.82	0.55	0.62	0.75	0.79	0.93	0.87	0.65	1.00	—
SHO	−0.68	−0.63	−0.51	−0.55	−0.61	−0.68	−0.83	−0.82	−0.50	−0.84	1.00
MKT	0.72	0.65	0.56	0.59	0.68	0.68	0.86	0.86	0.57	0.79	−0.78
SMB	0.51	0.47	0.30	0.33	0.36	0.40	0.47	0.42	0.36	0.57	−0.51
HML	−0.34	−0.31	−0.31	−0.32	−0.32	−0.37	−0.51	−0.55	−0.17	−0.51	0.57
UMD	−0.05	0.17	−0.06	−0.03	0.05	0.21	0.03	−0.08	0.04	0.13	−0.04
EMB	0.68	0.70	0.74	0.84	0.75	0.63	0.72	0.72	0.54	0.64	−0.60

Table 4 *continued*

	EDR	*FOF*	*GLO*	*GEM*	*GIN*	*GMA*	*GES*	*LON*	*MKN*	*SEC*	*SHO*
ACr	0.53	0.48	0.41	0.41	0.51	0.56	0.65	0.69	0.45	0.54	−0.59
APr	0.41	0.38	0.35	0.31	0.40	0.47	0.52	0.54	0.39	0.42	−0.47
OCr	−0.64	−0.56	−0.50	−0.48	−0.58	−0.59	−0.73	−0.76	−0.51	−0.65	0.63
OPr	−0.62	−0.52	−0.56	−0.47	−0.54	−0.50	−0.65	−0.63	−0.47	−0.61	0.61
ACa	0.50	0.47	0.36	0.41	0.48	0.56	0.66	0.67	0.42	0.57	−0.60
APa	0.47	0.44	0.34	0.39	0.45	0.53	0.62	0.63	0.40	0.54	−0.57
OCa	−0.73	−0.64	−0.67	−0.57	−0.63	−0.59	−0.76	−0.75	−0.60	−0.69	0.65
OPa	−0.64	−0.57	−0.63	−0.51	−0.55	−0.51	−0.65	−0.63	−0.53	−0.58	0.54

This table reports the correlations among hedge fund returns and between hedge fund returns and risk premia for the sample period 1994:02–2002:12. EDR=Event Driven, FOF=Funds of Funds, GLO=Global, GEM=Global Emerging Markets, GIN=Global International Markets, GMA=Global Macro, GES=Global Regional Established, LON=Long Only Leveraged, MKN=Market Neutral, SHO=Short Selling, SEC=Sector; MKT=the market premium, SMB=Small Minus Low which is the difference between the returns on a portfolio of small stocks and a portfolio of large stocks, HML=High Minus Low which is the difference between the returns on a portfolio of high-book-to-market-equity stocks and a portfolio of low-book-to-market-equity stocks, UMD=Momentum factor (see Carhart[32]), EMB=MSCI emerging market index; ACr=return of a true ATM index call, OCr=return of a true 5% OTM index call; APr=return of a true ATM index put, OPr=return of a true 5% OTM index put; ACa=return of an artificial ATM index call; OCa=return of an artificial 5% OTM index call; APa=return of an artificial ATM index put; OPa=return of an artificial 5% OTM index put. The underlying index for all options is the S&P500.

Table 5: Correlations among risk premia

	MKT	*SMB*	*HML*	*UMD*	*EMB*	*ACr*	*OCr*	*APr*	*OPr*	*ACa*	*OCa*	*APa*	*OPa*
MKT	1.00	—	—	—	—	—	—	—	—	—	—	—	—
SMB	0.14	1.00	—	—	—	—	—	—	—	—	—	—	—
HML	−0.57	−0.26	1.00	—	—	—	—	—	—	—	—	—	—
UMD	−0.19	0.19	0.09	1.00	—	—	—	—	—	—	—	—	—
EMB	0.73	0.28	−0.45	−0.21	1.00	—	—	—	—	—	—	—	—
ACr	0.82	−0.13	−0.43	−0.18	0.50	1.00	—	—	—	—	—	—	—
APr	0.66	−0.16	−0.32	−0.13	0.37	0.90	1.00	—	—	—	—	—	—
OCr	−0.90	−0.03	0.48	0.21	−0.64	−0.76	−0.55	1.00	—	—	—	—	—
OPr	−0.79	−0.08	0.40	0.18	−0.61	−0.58	−0.41	0.88	1.00	—	—	—	—
ACa	0.81	−0.07	−0.45	−0.14	0.49	0.93	0.84	−0.68	−0.51	1.00	—	—	—
APa	0.78	−0.10	−0.43	−0.14	0.47	0.91	0.84	−0.63	−0.47	1.00	1.00	—	—
OCa	−0.86	−0.15	0.43	0.14	−0.70	−0.61	−0.43	0.92	0.88	−0.54	−0.49	1.00	—
OPa	−0.70	−0.15	0.36	0.06	−0.60	−0.49	−0.34	0.78	0.74	−0.42	−0.38	0.91	1.00

This table reports the correlations between risk premia for the sample period 1994:02–2002:12. MKT=the market premium, SMB=Small Minus Low which is the difference between the returns on a portfolio of small stocks and a portfolio of large stocks, HML=High Minus Low which is the difference between the returns on a portfolio of high-book-to-market-equity stocks and a portfolio of low-book-to-market-equity stocks, UMD=Momentum factor (see Carhart[32]), EMB=MSCI emerging market index; ACr=return of a true ATM index call, OCr=return of a true 5% OTM index call; APr=return of a true ATM index put, OPr=return of a true 5% OTM index put; ACa=return of an artificial ATM index call; OCa=return of an artificial 5% OTM index call; APa=return of an artificial ATM index put; OPa=return of an artificial 5% OTM index put. The underlying index for all options is the S&P500.

(GLO, GIN and GMA) strategies. The 6-factor HME specification always dominates the corresponding 6-factor OLS, with an increase in the $\bar{R}^2$ ranging between 0.1 per cent and 5.4 per cent, despite the fact that all variables are duplicated with the HM characterization, increasing the number of coefficients from 7 to 13. Significance levels are only close for the GIN strategy, as indicated by the insignificance of each coefficient of the adjustment variables. For the other three strategies displayed in the panel, there are between one and four significant loadings for the adjustment variables.

Some coefficients that used to be significant with OLS might not be any longer under HME. This phenomenon is particularly noteworthy for the MKN strategy, where out of the five significant OLS coefficients only two (the HML and EMB coefficients) remain significant with HME. The association of adjustment variables with the original factors may thus induce a dilution effect among the variables.

Table 6: OLS and HME regressions on hedge fund strategies

	$\bar{R}^2$	α	Factors							F-stat	Adjustment variables					
			MKT	*SMB*	*HML*	*UMD*	*EMB*	*Option*			w_m	w_s	w_h	w_u	w_e	w_o
Panel A: Strategies with low significance levels ($\bar{R}^2 < 0.7$)																
MKN	0.448	0.681***	0.164***	0.086***	0.078***	0.018	—	—	—	—	—	—	—	—	—	—
	0.468	0.713***	0.124***	0.073***	0.077***	0.023	0.038**	—	—	—	—	—	—	—	—	—
	0.487	0.752***	0.065*	0.073***	0.068***	0.021	0.030*	APa	−0.189**	—	—	—	—	—	—	—
	0.486	0.743***	0.187***	0.087***	0.146**	−0.054*	—	—	—	16.05***	−0.023	0.005	−0.050	0.107***	—	—
	0.504	0.863***	0.100*	0.040	0.093	−0.003	0.095***	—	—	12.76***	0.029	0.047	0.001	0.054*	−0.076*	—
	0.522	0.855***	0.090	0.046	0.123**	−0.010	0.095**	ACr	0.044	10.01***	−0.047	0.057	−0.035	0.061*	−0.072	0.687
GLO	0.342	−0.197	0.473***	0.207***	0.068	0.004	—	—	—	—	—	—	—	—	—	—
	0.547	0.121	0.077	0.080	0.063	0.062	0.378***	—	—	—	—	—	—	—	—	—
	0.604	0.338	−0.254**	0.079	0.014	0.047	0.334***	APa	−1.060***	—	—	—	—	—	—	—
	0.395	−0.210	0.456***	0.241*	0.476*	−0.330***	—	—	—	12.89***	0.095	0.012	−0.334	0.461***	—	—
	0.617	0.811***	−0.251	−0.169*	0.031	0.082	0.787***	—	—	8.98***	0.391*	0.349***	0.101	0.036	−0.512***	—
	0.658	1.335***	−0.861***	−0.131	−0.152	0.085***	0.606***	APa	−1.721	8.54***	0.903***	0.320***	0.276	0.012	−0.341***	1.675**
GIN	0.540	0.202	0.378***	0.147***	0.095**	0.055*	—	—	—	—	—	—	—	—	—	—
	0.674	0.359***	0.184***	0.085***	0.092**	0.083***	0.186***	—	—	—	—	—	—	—	—	—
	0.677	0.391***	0.108	0.106***	0.094***	0.081***	0.192***	ACr	0.478	—	—	—	—	—	—	—
	0.530	0.181	0.443***	0.071	0.163	0.050	—	—	—	24.57***	−0.062	0.112	−0.055	0.008	—	—
	0.665	0.265	0.257***	0.038	0.130	0.090**	0.147**	—	—	19.57***	−0.082	0.078	−0.029	−0.019	0.049	—
	0.678	0.417*	0.307***	0.025	0.152	0.085**	0.156***	APa	0.065	13.76***	−0.158	0.102	−0.053	−0.021	0.051	0.377
GMA	0.632	0.144	0.341***	0.129***	0.046	0.115***	—	—	—	—	—	—	—	—	—	—
	0.666	0.218*	0.249***	0.100***	0.045	0.128***	0.088***	—	—	—	—	—	—	—	—	—
	0.690	0.275**	0.112*	0.139***	0.047***	0.125***	0.099***	ACr	0.858***	—	—	—	—	—	—	—
	0.636	0.201	0.239***	0.200***	0.001	0.053	—	—	—	22.33***	0.139*	−0.075	0.076	0.093*	—	—
	0.660	0.314*	0.233***	0.137***	−0.093	0.136***	0.059	—	—	18.98***	0.000	−0.042	0.161	0.000	0.037	—
	0.701	0.419***	−0.247	0.300***	−0.061	0.119***	0.154***	ACr	2.137***	13.28***	0.437***	−0.193***	0.112	0.020	−0.051	−1.366*

Panel B: Strategies with average significance levels $(0.7<\bar{R}^2<0.8)$																
FOF	0.624	0.097	0.266***	0.144***	0.071**	0.069***	—	—	—	—	—	—	—	—	—	—
	0.711	0.188*	0.153***	0.108***	0.069***	0.086***	0.108***	—	—	—	—	—	—	—	—	—
	0.718	0.216**	0.087*	0.127***	0.070***	0.084***	0.113***	ACr	0.414*	—	—	—	—	—	—	—
	0.623	0.116	0.235***	0.160***	0.101	0.014	—	—	—	23.39***	0.054	−0.005	−0.008	0.080*	—	—
	0.704	0.300***	0.095	0.083	0.013	0.097	0.153	—	—	19.44***	0.068	0.041	0.076	−0.003	−0.056	—
	0.727	0.380***	0.145	0.067	0.043	0.091***	0.146***	ACr	−0.192	13.57***	−0.097	0.085	0.041	0.007	−0.044	1.617***
GEM	0.393	0.374	0.624***	0.293***	0.107	0.027	—	—	—	—	—	—	—	—	—	—
	0.736	0.881***	−0.006	0.091	0.098	0.119***	0.602***	—	—	—	—	—	—	—	—	—
	0.739	0.788***	−0.119	0.120*	0.105	0.114***	0.614***	OCa	0.373	—	—	—	—	—	—	—
	0.396	0.290	0.635***	0.212	0.398	−0.167	—	—	—	21.05***	0.078	0.180	−0.211	0.267	—	—
	0.736	0.804***	0.038	−0.020	0.148	0.077	0.568***	—	—	18.42***	0.044	0.211	0.029	0.081	0.038	—
	0.759	0.989***	0.636**	−0.224*	0.182	0.078	0.466***	ACr	−2.418**	11.79***	−0.890***	0.484***	−0.002	0.088	0.147	5.636***
EDR	0.709	0.378***	0.298***	0.194***	0.093***	−0.002	—	—	—	—	—	—	—	—	—	—
	0.727	0.425***	0.240***	0.175***	0.092***	0.006	0.056***	—	—	—	—	—	—	—	—	—
	0.748	0.488***	0.144***	0.175***	0.078***	0.002	0.043**	APa	−0.308***	—	—	—	—	—	—	—
	0.732	0.461***	0.349***	0.153***	0.162**	−0.069**	—	—	—	18.61***	−0.048	0.068	−0.037	0.107***	—	—
	0.765	0.633***	0.223***	0.084**	0.092	−0.004	0.136***	—	—	11.79***	0.008	0.134***	0.033	0.032	−0.103**	—
	0.780	0.511***	0.329***	0.056	0.155***	−0.015	0.107**	ACr	−0.425	8.58***	−0.212*	0.176***	−0.037	0.042	−0.072	0.820
SHO	0.786	0.910***	−0.664******	−0.360***	0.091	−0.087***	—	—	—	—	—	—	—	—	—	—
	0.787	0.953***	−0.717***	−0.377***	0.090	−0.079**	0.051	—	—	—	—	—	—	—	—	—
	0.791	0.862***	−0.874***	−0.359***	0.086	−0.077**	0.053	APr	−0.903*	—	—	—	—	—	—	—
	0.797	1.209***	−0.806***	−0.362***	−0.395***	0.043	—	—	—	20.61***	0.180	0.037	0.570***	−0.125	—	—
	0.796	1.275***	−1.038***	−0.362***	−0.355**	0.017	0.143	—	—	16.52***	0.403***	0.025	0.536***	−0.082	−0.103	—
	0.803	1.112***	−1.098***	−0.336***	−0.364***	0.008	0.105	OPa	−0.060	12.21***	0.445***	−0.008	0.540***	−0.069	−0.074	−0.212
Panel C: Strategies with high significance levels $(\bar{R}^2>0.8)$																
LON	0.816	0.108	0.987***	0.384***	−0.023	0.022	—	—	—	—	—	—	—	—	—	—
	0.819	0.177	0.901***	0.357***	−0.025	0.035	0.082	—	—	—	—	—	—	—	—	—
	0.832	0.291	0.630***	0.433***	−0.019	0.028	0.104**	ACr	1.697***	—	—	—	—	—	—	—
	0.815	0.102	0.939***	0.435***	0.121	−0.122	—	—	—	24.20***	0.087	−0.044	−0.116	0.197	—	—

Table 6 *continued*

	$\bar{R}^2$	α	*Factors* MKT	SMB	HML	UMD	EMB	Option		*F-stat*	*Adjustment variables* w_m	w_s	w_h	w_u	w_e	w_o
	0.811	0.234	1.000***	0.350***	−0.013	−0.007	0.035	—	—	20.40***	−0.133	0.016	0.008	0.065	0.053	—
	0.825	0.505	0.652**	0.459***	−0.060	−0.002	0.112	ACr	1.331	16.62***	−0.069	−0.031	0.049	0.066	−0.012	1.579
SEC	0.873	0.795***	0.716***	0.424***	0.021	0.150***	—	—	—	—	—	—	—	—	—	—
	0.872	0.802***	0.707***	0.421***	0.020	0.152***	0.008	—	—	—	—	—	—	—	—	—
	0.871	0.830***	0.756***	0.415***	0.022	0.151***	0.007	APr	0.280	—	—	—	—	—	—	—
	0.886	0.491***	0.718***	0.518***	0.288**	0.108**	—	—	—	18.62***	−0.026	−0.167**	−0.365***	0.001	—	—
	0.882	0.471***	0.890***	0.498***	0.251**	0.128***	−0.096	—	—	15.43***	−0.247*	−0.151*	−0.326***	−0.017	0.128	—
	0.886	0.580***	0.804***	0.499***	0.250**	0.125***	−0.116	OPr	−0.548*	11.32***	−0.103	−0.170**	−0.320***	−0.009	0.151*	1.562**
GES	0.876	0.526***	0.577***	0.248***	0.045	0.070***	—	—	—	—	—	—	—	—	—	—
	0.880	0.568***	0.525***	0.232***	0.044	0.077***	0.049**	—	—	—	—	—	—	—	—	—
	0.884	0.604***	0.439***	0.256***	0.046	0.075***	0.056***	ACr	0.543**	—	—	—	—	—	—	—
	0.878	0.469***	0.604***	0.232***	0.198**	−0.006	—	—	—	23.09***	−0.013	0.033	−0.147	0.098**	—	—
	0.880	0.503***	0.641***	0.198***	0.153*	0.029	0.002	—	—	18.90***	−0.122	0.056	−0.099	0.065	0.053	—
	0.894	0.747***	0.734***	0.148***	0.167**	0.033	0.001	OCr	−0.304	12.48***	−0.347***	0.131**	−0.106	0.059	0.059	1.520***

This table reports the regression results of hedge fund returns for the sample period 1994:02–2002:12. The OLS specifications take the form of equation $R_t = \hat{\alpha}^{OLS} + \sum_{k=1}^{K} \hat{\beta}_k^{OLS} \cdot F_{kt} + \upsilon_t$ with four, five and six risk factors. The HME specifications take the form $R_t = \hat{\alpha}^{HM} + \sum_{k=1}^{K} \hat{\beta}_k^{HM} \cdot F_{kt} + \sum_{k=1}^{K} \hat{\psi}_k \cdot \hat{w}_{kt} + \varepsilon_t$ with four, five and six risk factors and the corresponding adjustment variables. For both the OLS and HME specifications, we adopt as sixth risk premium the option-based factor that obtains the best value for the Akaike information criterion. The alpha is expressed in per cents. F is a standard F-test to detect EIV by testing for ψ_k=0 for each k. *** Significant at the 1% level, ** significant at the 5% level and * significant at the 10% level.

In contrast, some coefficients that are insignificant with OLS may become significant with HME. This phenomenon occurs with the UMD coefficient for the GLO strategy and with the MKT coefficient for the GMA strategy. In such cases, the corresponding adjustment variables pick up most of the coefficient variance and the factor loading is estimated with greater precision.

Finally, some OLS coefficients are seen to lose their significance because the measurement error is responsible for the effect. This is the case with the option-based factor coefficient for the GLO strategy and with the MKT coefficient for the GMA strategy. In these cases, the signs of the coefficients of the original and adjustment factors are opposite.

In Panel B, we obtain qualitatively similar results for those strategies that already had fairly high significance levels with OLS. The significance gains are limited (between 0.9 per cent and 3.2 per cent) owing to the fact that OLS performed relatively well already. We observe that with the option-based factor coefficient for GEM and with the HML coefficient for Short Sales, the coefficient is insignificant with OLS, but the coefficients for both the original factor and the adjustment variable become strongly significant, although with opposite signs, with the HME specification. For these strategies, our results suggest that the OLS coefficient hides two opposite effects, one for raw factor risk and one for measurement risk, that tend to compensate each other if the exposures are not separated.

Panel C displays a significant result regarding the LOL strategy: it is the only one for which the OLS specification dominates HME. Adjustment variable coefficients are insignificant, while some OLS coefficients lose their significance under HME. As this strategy most closely resembles long portfolios held by mutual funds whose market exposures are relatively well under control, such a result is not very surprising. For the other two strategies, the gains from HME are not very large, but are positive. For SEC, as for the GEM strategy in Panel B, the OLS coefficient of the HML variable is not significantly different from zero, but both corresponding coefficients under HME are significant and of opposite signs.

When considering Table 6 globally, one also obtains some useful insight in terms of strategy performance. Aside from the LOL strategy where OLS dominates, the account for measurement errors in the HME specification appears to generate higher alphas for all but the SEC strategy, where it decreases by 25 bps per months. Yet, the level of alpha gains is limited, as they range from 2.6 bps (for GIN) to 25 bps (for Short Sales).[37] This finding reflects the underlying interpretation of the interference of measurement errors in the original OLS specification. Once their effect is removed and transferred in the adjustment variables, the sources of risk exposures are magnified and the generation of performance can be properly isolated.

Some variables also appear to be more prone to corrections for measurement errors than others. The coefficients of the adjustment variable for the MKT, SMB and option-based factors are significant for six out of 11 strategies. The other three variables (HML, UMD and EMB) trigger a significant loading for the adjustment variable in no more than two cases. For the HML variable, the adjustment variable coefficient is highly significant for the GMA and SEC strategies regardless of the number of factors chosen. For the other nine strategies, this coefficient is consistently insignificant.

Performance and persistence on individual funds

We use and adapt a methodology initiated by Capocci and Hübner[33] to test the performance and persistence of individual funds. To emphasize the economic consequences of HME on performance attribution, we isolate two sub-periods in our sample window: a bearish sub-period (01:1994–12:1997) and a bullish sub-period (07:1996–06:2000). Using a five-factor model (MKT, SMB, HML, UMD and EMB), we select only living individual hedge funds (dead funds are not considered) for each sub-period and compare the OLS alphas with HME alphas. The results for Non-Directional strategies (EDR, MKN: 119 funds for the first sub-period and 206 funds for the second sub-period), Directional strategies (GLO, GEM, GIN, GMA, GES, LON, SEC, SHO: 298 funds for the bearish period and 291 funds for the bullish period) and FOF (FOF: 114 funds for the bearish period and 197 funds for the bullish period) are reported in Table 7.

The examination of the columns reporting significant alpha differences suggests a very different pattern for the first and the second sub-periods. During the earlier period, the number of funds for which $\alpha_{HME} < \alpha_{OLS}$ is much larger than those for which the opposite relation holds, while the opposite holds for the second sub-period, although in a less obvious fashion. This effect is particularly pronounced for the non-directional strategies. Nevertheless, there is no overwhelming evidence of a systematic bias of the OLS alpha over the (better-specified) HME alpha. This seems to indicate that accounting for EIV might result in hardly predictable alterations in performance estimation at the fund level.

To study the persistence in alphas, we use an approach similar to that of Capocci *et al.*[4] that adapts that originally proposed by Carhart[32] on mutual funds. For the two sub-periods ((01:1994–12:1997) and (07:1996–06:2000)) selected above, we constructed five portfolios for each broad strategy (Directional, Non-Directional and FOF) based on alpha OLS and alpha HME performance. We then analyzed the performance of each portfolio during the subsequent 30 months. Our results are reported in Table 8.

The results for both the bearish and bullish sub-period indicate that neither the OLS nor the HME specification emphasizes systematic persistence in hedge fund returns, especially during the bullish period. Nonetheless, there appears to be an economically as well as statistically significant difference in the top and bottom alphas for directional funds during the bearish period. This difference is only emphasized when the HME is used, in contrast with the OLS alphas that are insignificant, a finding consistent with those of Capocci *et al.*[4]

EMPIRICAL ANALYSIS OF MODEL SPECIFICATIONS

The previous section displays results that are globally in favor of the HM estimation method, but suffer from the inflation in the number of variables. We present a heuristic algorithm that aims at mitigating this drawback by gradually reducing the number of variables.

We assess the quality of this procedure in two ways. First, we review the optimal number of adjustment variables to drop and observe the gains in overall regression significance. Next, we

Table 7: Performance during the sub-period (01:1994–12:1997) sub-period (07:1996–06:2000)

	Sub-Period 01: 1994–12:1997				*Sub-Period 07:1996–06:2000*			
	N	$\alpha_{HME} < \alpha_{OLS}$ (%)	$\alpha_{HME} = \alpha_{OLS}$ (%)	$\alpha_{HME} > \alpha_{OLS}$ (%)	N	$\alpha_{HME} < \alpha_{OLS}$ (%)	$\alpha_{HME} = \alpha_{OLS}$ (%)	$\alpha_{HME} > \alpha_{OLS}$ (%)
Persistence during the period (01:1994–06:2000)								
Panel A: Non-directional strategies								
EDR	45	55.56	22.22	22.22	78	12.82	60.25	26.93
MKN	74	48.65	27.03	24.32	128	10.16	57.03	32.81
Panel B: Directional strategies								
GLO	23	73.91	9.70	17.39	—	—	—	—
GEM	17	35.29	23.53	41.18	40	27.50	45.00	27.50
GIN	15	26.67	53.33	20.00	23	13.04	52.18	34.78
GMA	18	50.00	16.67	33.33	31	38.71	54.84	6.45
GES	194	45.88	28.87	25.25	140	23.57	45.00	31.43
LON	2	0	50.00	50.00	8	12.50	75.00	12.50
SEC	19	26.32	47.37	26.32	42	14.29	57.14	28.57
SHO	10	50.00	20.00	30.00	7	0	85.71	14.29
Panel C: Funds of funds								
FOF	114	45.61	22.81	31.58	197	17.74	55.84	28.42

This table reports the intercept of the regression results of individual hedge fund returns with a full data history for the sub-sample periods of 3 years 01:1994–12:1997 and 07:1996–06:2000. EDR=Event Driven, FOF=Funds of Funds, GLO=Global, GEM=Global Emerging Markets, GIN=Global International Markets, GMA=Global Macro, GES=Global Regional Established, LON=Long Only Leveraged, MKN=Market Neutral, SHO=Short Selling, SEC=Sector; MKT=the market premium, SMB=Small Minus Low which is the difference between the returns on a portfolio of small stocks and a portfolio of large stocks, HML=High Minus Low, which is the difference between the returns on a portfolio of high-book-to-market-equity stocks and a portfolio of low-book-to-market-equity stocks, UMD=Momentum factor (see Carhart[32]), EMB=MSCI emerging market index. For each strategy, each individual fund is priced using a five factor-model with factors MKT, SMB, HML, UMD and EMB, with and without the HM estimators. Robust t-stats are used to compare $\alpha_{\mathbf{HME}}$ with $\alpha_{\mathbf{OLS}}$. Significance is assessed at the 5% confidence level. N stands for the number of funds for each strategy.

Table 8: Persistence during the sub-period (01:1998–06:2000) sub-period (07:2000–12:2002)

Quintile	*Bullish period (01:1998–06:2000)*								
	Directional Return			*Non-Directional Return*			*Funds of funds Return*		
	N(194)	α_{OLS}	α_{HME}	*N(119)*	α_{OLS}	α_{HME}	*N(114)*	α_{OLS}	α_{HME}
Panel A									
Q1	38	1.94	1.70	23	1.23	1.30	22	1.21	0.70
Q2	39	1.47	1.06	24	0.71	0.72	23	0.74	1.11
Q3	39	1.22	0.90	24	0.80	0.60	23	0.92	0.83
Q4	39	1.36	1.32	24	0.61	0.98	23	1.00	1.07
Q5	39	0.26	1.41	24	0.95	0.81	23	0.69	0.80
Test- *t*-stat Q5 vs Q1	—	−1.29	−0.24	—	−0.31	−0.45	—	−0.71	0.13

Quintile	*Bearish period (07:2000–12:2002)*								
	Directional Return			*Non-Directional Return*			*Funds of funds Return*		
	N (260)	α_{OLS}	α_{HME}	*N(206)*	α_{OLS}	α_{HME}	*N(197)*	α_{OLS}	α_{HME}
Panel B									
Q1	52	−1.18	−1.83	41	0.47	0.26	39	0.12	0.05
Q2	52	−0.26	−0.29	41	0.34	0.42	39	0.19	0.12
Q3	52	0.05	−0.05	41	0.58	0.60	39	0.25	0.29
Q4	52	−0.18	−0.31	41	0.39	0.43	40	0.35	0.33
Q5	52	−0.33	0.49	42	0.44	0.50	40	0.01	0.13
Test- *t*-stat Q5 vs Q1	—	0.76	2.01⋆⋆	—	−0.05	0.53	—	−0.32	0.24

This table reports the persistence of alphas of quintile portfolios of individual hedge fund during the 30-month bullish sub-period 01:1998–06:2000 (Panel A) and the 30-month bearish sub-period 07:2000–12:2002.
The estimation is done on sub-sample periods of 3 years 01:1994–12:1997 (Panel A) and 07:1996–06:2000 (Panel B) for hedge funds exhibiting a full data history. For each strategy, each individual fund is priced using a five factor-model with factors MKT, SMB, HML, UMD and EMB, with and without the HM estimators. *N* stands for the number of funds for each portfolio. All funds are ranked based on their previous period's alpha. Portfolios are equally weighted and weights are readjusted whenever a fund disappears. Funds with the lowest previous alpha go into portfolio Q1 and funds with the highest go into portfolio Q5. ⋆⋆⋆ Significant at the 1% level, ⋆⋆ significant at the 5% level and ⋆ significant at the 10% level.

verify the evolution of the risk premia associated with each source of risk under the OLS, the HME and the heuristic hybrid specification.

Heuristic reduction of instruments

The expanded regression model displayed in equation (4) accurately transforms the linear asset pricing model to account for measurement errors. Unfortunately, the cost of this operation is a considerable inflation in the number of independent variables. As each regressor is flanked by a twin variable, a K-factor model becomes a $2K$-modified model. Although this can be econometrically justified, the lack of significance of some adjustment variables might hinder the economic relevance of the expanded model.

We propose a heuristic recursive method to reduce the number of variables. The principle of the algorithm is to detect the adjustment variable that exhibits the lowest significance level. If the corresponding original variable were not prone to measurement error, it would have been more effective to use the OLS instead. As a heuristic check, we subtract the OLS term corresponding to this independent variable (regression coefficient times the observation) from the value of the dependent variable, and define a new dependent variable equal to this difference. We then run the HM estimation again on this new variable with the remaining variables. The procedure stops when the significance level of the new model becomes lower than the former specification, which corresponds to an empirical stopping rule based on a trade-off between model parsimony and accuracy.

To estimate the significance level of this new regression equation with D removed variables ($D<K$), we get the unadjusted R^2 by simply computing $R^2=1-\hat{\sigma}_{\vartheta}^2/\hat{\sigma}_R^2$, where $\hat{\sigma}_{\vartheta}^2$ is the variance of residuals from regression (5) and $\hat{\sigma}_R^2$ is the variance of the original returns. The pseudo-adjusted R^2 is then computed as Ps. $\bar{R}^2=\frac{1-(2K-D)}{T-(2K-D)}+\frac{T-1}{T-(2K-D)}R^2$ where T is the number of observations, K is the number of original risk factors and D is the number of adjustment variables removed from the model.

The application of the recursive regression algorithm on the 11 hedge fund strategies yields the results presented in Table 9, using the same types of panels as in Table 6.

For the strategies with low significance levels (Panel A), the algorithm brings some improvement for the GLO and GIN return indexes. The gains in significance, measured with the Pseudo-adjusted R-squared (Ps. $\bar{R}^2$), are 5.2 per cent and 6.5 per cent, respectively. Nevertheless, the sources of these gains are qualitatively very different. For the GLO strategy, two adjustment variable coefficients are insignificant in the HME: they naturally fade away with the algorithm, leaving only the adjustment variables that account for a priced measurement error. For the GIN strategy, the HME globally (slightly) improves over the OLS, but without any significant loading for the new variables. Thus, it is likely that when they are taken individually, they are considered superfluous. The remaining coefficients after three runs of the algorithm are not even significant yet. For these two strategies, the total improvement over the OLS $\bar{R}^2$ is 10.6 per cent (for GLO) and 6.6 per cent (for GIN).

In Panel B, the algorithm increases the Pseudo-adjusted R-squared of the asset pricing specification for FOF (+3.9 per cent with respect to HME, +4.8 per cent with respect to OLS), GEM (+2.8 per cent with respect to HME, +4.8 per cent with respect to OLS) and EDR (+4.1 per cent with respect to HME,

Table 9: Regression coefficients using the recursive regression algorithm

	Specification	Ps. $\bar{R}^2$	α	Factors						Adjustment variables					
				Mkt	SMB	HML	UMD	EMB	OPT	w_m	w_s	w_h	w_u	w_e	w_o
Panel A: Strategies with low significance levels ($\bar{R}^2 < 0.7$)															
MKN	OLS	0.487	0.752***	0.065*	0.073***	0.068***	0.021	0.030*	−0.189**	—	—	—	—	—	—
	HME	0.522	0.855***	0.090	0.046	0.123**	−0.010	0.095**	0.044	−0.047	0.057	−0.035	0.061*	−0.072	0.687
	1st fact. (MKT)	0.417	0.911***	—	0.037	0.094*	−0.004	0.117***	−0.022	—	0.070	−0.004	0.049	−0.093***	0.759**
	2nd fact. (HML)	0.497	0.926***	—	0.033	—	0.002	0.118***	−0.074	—	0.070	—	0.043	−0.094***	0.801***
	3rd fact. (UMD)	0.480	0.925***	—	0.022	—	—	0.127***	−0.120	—	0.091**	—	—	−0.101***	0.828***
GLO	OLS	0.604	0.338	−0.254**	0.079	0.014	0.047	0.334***	−1.060***	—	—	—	—	—	—
	HME	0.658	1.335***	−0.861***	−0.131	−0.152	0.085	0.606***	−1.721***	0.903***	0.320***	0.276	0.012	−0.341***	1.675**
	1st fact. (UMD)	0.669	1.494***	−0.903***	−0.120	−0.161	—	0.640***	−1.709***	0.894***	0.311**	0.263		−0.371***	1.706**
	2nd fact. (HML)	0.710	1.707***	−1.063***	−0.098	—	—	0.895***	−1.551***	1.055***	0.238**	—	—	−0.642***	1.719***
	3rd fact. (SMB)	0.703	1.957***	−1.081***	—	—	—	0.851***	−1.716***	1.102***	—	—	—	−0.580***	2.362***
GIN	OLS	0.677	0.391***	0.108	0.106***	0.094***	0.081***	0.192***	0.478	—	—	—	—	—	—
	HME	0.678	0.417*	0.307***	0.025	0.152	0.085**	0.156***	0.065	−0.158	0.102	−0.053	−0.021	0.051	0.377
	1st fact. (UMD)	0.686	0.439*	0.295**	0.026	0.144	—	0.164**	0.068	−0.122	0.085	−0.037	—	0.033	0.358
	2nd fact. (HML)	0.734	0.386*	0.325***	0.020	—	—	0.103	0.019	−0.166	0.093	—	—	0.097	0.337
	3rd fact. (APa)	0.743	0.244	0.308***	0.030	—	—	0.110	—	−0.114	0.076	—	—	0.084	—
GMA	OLS	0.690	0.275**	0.112*	0.139***	0.047	0.125***	0.099***	0.858***	—	—	—	—	—	—
	HME	0.701	0.419***	−0.247	0.300***	−0.061	0.119***	0.154***	2.137***	0.437***	−0.193***	0.112	0.020	−0.051	−1.366
	1st fact. (UMD)	0.680	0.271	−0.063	0.263***	−0.005	—	0.087	1.691***	0.214	−0.159*	0.039	—	0.029	−1.029
	2nd fact. (EMB)	0.504	0.326**	−0.053	0.248***	−0.005	—	—	1.537***	0.209	−0.131	0.039	—	—	−0.505
	3rd fact. (HML)	0.565	0.204	−0.096	0.235***	—	—	—	1.901***	0.301***	−0.111	—	—	—	−1.174

Panel B: Strategies with average significance levels ($0.7<\bar{R}^2<0.8$)															
FOF	*OLS*	*0.718*	*0.216***	*0.087**	*0.127****	*0.070****	*0.084****	*0.113****	*0.414**	—	—	—	—	—	—
	HME	0.727	0.380***	0.145	0.067	0.043	0.091***	0.146***	−0.192	−0.097	0.085	0.041	0.007	−0.044	1.617***
	1st fact. (UMD)	0.721	0.311*	0.236	0.057	0.075	—	0.111	−0.396	−0.212	0.092	−0.002	—	−0.004	1.748***
	2nd fact. (HML)	**0.766**	**0.342*****	**0.202**	**0.057**	—	—	**0.127**	**−0.331**	**−0.168**	**0.094**	—	—	**−0.019**	**1.659*****
	3rd fact. (EMB)	0.525	0.324***	0.237***	0.048	—	—	—	−0.451	−0.200**	0.098*	—	—	—	1.612***
GEM	*OLS*	*0.739*	*0.788****	*−0.119*	*0.120**	*0.105*	*0.114****	*0.614****	*0.373*	—	—	—	—	—	—
	HME	0.759	0.989***	0.636**	−0.225*	0.182	0.078	0.466***	−2.418**	−0.890***	0.484***	−0.002	0.088	0.147	5.636***
	1st fact. (HML)	0.774	1.085***	0.552*	−0.224*	—	0.109	0.506***	−2.385**	−0.862***	0.459***	—	0.039	0.107	5.588***
	2nd fact. (UMD)	**0.787**	**1.070*****	**0.633***	**−0.153**	—	—	**0.486*****	**−2.662****	**−0.982*****	**0.446*****	—	—	**0.114**	**5.863*****
	3rd fact. (EMB)	0.057	1.131***	0.309	−0.153	—	—	—	−1.589	−0.604**	0.346**	—	—	—	3.489**
EDR	*OLS*	**0.748**	**0.488*****	**0.144*****	**0.175*****	**0.078*****	**0.002**	**0.043****	**−0.308*****	—	—	—	—	—	—
	HME	0.780	0.511***	0.329***	0.061	0.155***	−0.015	0.107**	−0.425	−0.212	0.176***	−0.037	0.042	−0.072	0.820
	1st fact. (HML)	0.816	0.587***	0.268***	0.066	—	0.000	0.129***	−0.375	−0.167	0.164***	—	0.024	−0.093*	0.827
	2nd fact. (UMD)	**0.821**	**0.610*****	**0.252****	**0.213*****	—	—	**0.143*****	**−0.358**	**−0.149**	**0.170*****	—	—	**−0.107***	**0.810**
	3rd fact. (Mkt)	0.686	0.753***	—	0.060	—	—	0.213***	−0.426*	—	0.179***	—	—	−0.189***	1.189***
SHO	*OLS*	*0.791*	*0.862****	*−0.874****	*−0.359*	*0.086*	*−0.077*	*0.053*	*−0.903*	—	—	—	—	—	—
	HME	**0.803**	**1.112*****	**−1.098*****	**−0.233**	**−0.364**	**0.008**	**0.105**	**−0.060**	**0.445*****	**−0.008**	**0.540*****	**−0.069**	**−0.074**	**−0.212**
	1st fact. (SMB)	0.723	1.042***	−1.010***	—	−0.217	−0.006	0.068	−0.083	0.330*	—	0.367*	−0.054	−0.018	−0.120
	2nd fact. (EMB)	0.756	0.997***	−0.981***	—	−0.274	−0.007	—	−0.086	0.253	—	0.343*	−0.069	—	−0.144
	3rd fact. (0/P)	0.775	1.197***	−0.880***	—	−0.274	0.014	—	—	0.157	—	0.400**	−0.092	—	—
Panel C: Strategies with high significance levels ($\bar{R}^2>0.8$)															
LON	**OLS**	**0.832**	**0.291**	**0.630*****	**0.433*****	**−0.019**	**0.028**	**0.104****	**1.697*****	—	—	—	—	—	—
	HME	0.825	0.505	0.652**	0.439***	−0.060	−0.002	0.112	1.331	−0.069	−0.031	0.049	0.066	−0.012	1.579
	1st fact. (EMB)	0.793	0.527*	0.720***	−0.029	−0.035	−0.031	—	0.993	−0.144	−0.011	0.024	0.091	—	1.847

Table 9 *continued*

	Specification	*Ps.* $\bar{R}^2$	α	*Factors*						*Adjustment variables*					
				Mkt	*SMB*	*HML*	*UMD*	*EMB*	*OPT*	w_m	w_s	w_h	w_u	w_e	w_o
	2nd fact. (SMB)	0.764	0.566*	0.722***	—	−0.096	−0.087	—	0.990	−0.132	—	0.040	0.169	—	1.592
	3rd fact. (HML)	0.764	0.543**	0.700***	—	—	−0.096	—	1.122	−0.109	—	—	0.174		1.407
SEC	*OLS*	*0.871*	*0.830****	*0.756****	*0.415****	*0.022*	*0.151****	*0.007*	*0.280*	—	—	—	—	—	—
	HME	0.886	0.580***	0.804***	0.499***	0.415***	0.125***	−0.116	−0.548*	−0.103	−0.170**	−0.320***	−0.009	0.151*	1.562**
	1st fact. (UMD)	**0.886**	**0.451***	**0.879****	**0.482****	**0.250****		**−0.161***	**−0.562***	**−0.183**	**−0.171****	**−0.353****	—	**0.199****	**1.542****
	2nd fact. (Mkt)	0.640	0.505*	—	0.254*	0.288**	—	−0.089	−0.822	—	−0.166**	−0.323**	—	0.128	1.700***
	3rd fact. (EMB)	0.636	0.543**	—	0.485***	0.276**	—	—	−0.446*	—	−0.164**	−0.353***	—	—	0.868*
GES	*OLS*	*0.884*	*0.604****	*0.439****	*0.256****	*0.046*	*0.075****	*0.056****	*0.543***	—	—	—	—	—	—
	HME	**0.894**	**0.747****	**0.734****	**0.148****	**0.256****	**0.033**	**0.001**	**−0.304**	**−0.347****	**0.131****	**−0.106**	**0.059**	**0.059**	**1.520****
	1st fact. (EMB)	0.867	0.794***	0.637***	0.157***	0.167*	0.039	—	−0.190	−0.240**	0.113*	−0.103	0.047	—	1.232***
	2nd fact. (HML)	0.877	0.846***	0.564***	0.178***	—	0.055	—	−0.094	−0.159*	0.081	—	0.036	—	1.019**
	3rd fact. (UMD)	0.880	0.862***	0.605***	0.178***	—	—	—	−0.177	−0.223**	0.090	—	—	—	1.129***

This table reports the regression results of hedge fund returns for the sample period 1994:02–2002:12. The OLS specifications take the form of equation $R_t = \hat{\alpha}^{OLS} + \sum_{k=1}^{K} \hat{\beta}_k^{OLS} \cdot F_{kt} + \upsilon_t$ with six risk factors. The HME specifications take the form $R_t = \hat{\alpha}^{HM} + \sum_{k=1}^{K} \hat{\beta}_k^{HM} \cdot F_{kt} + \sum_{k=1}^{K} \hat{\psi}_k \cdot \hat{w}_{kt} + \varepsilon_t$ with six risk factors and the corresponding adjustment variables. For both the OLS and HME specifications, we adopt as sixth risk premium the option-based factor that obtains the best value for the Akaike information criterion. For each iteration of the recursive algorithm, we remove the OLS risk premium corresponding to the least significant coefficient $\hat{\psi}_i$ and re-estimate equation

$$R_t - \hat{\beta}_i^{OLS} \cdot F_{it} = \hat{\alpha}^{HM,1} + \sum_{\substack{k=1 \\ k \neq i}}^{K} \hat{\beta}_k^{HM,1} \cdot F_{kt} + \sum_{\substack{k=1 \\ k \neq i}}^{K} \hat{\psi}_k^1 \cdot \hat{w}_{kt} + \vartheta_t.$$

Ps. $\bar{R}^2 = \frac{1-(2K-D)}{T-(2K-D)} + \frac{T-1}{T-(2K-D)} R^2$ where R^2 is the unadjusted coefficient of determination. For OLS and HME, Ps. $\bar{R}^2 \equiv \bar{R}^2$. The alpha is expressed in percentage. *** Significant at the 1% level, ** significant at the 5% level and * significant at the 10% level

Note: Best specifications based on Ps. $\bar{R}^2$ are in bold; OLS estimates are italicized.

+7.3 per cent with respect to OLS). For each strategy, the same two coefficients (HML and UMD) cancel out. The results are much less interesting for Panel C, as the original specification (OLS for LOL, HME for the other two) does not appear to appreciate thanks to the application of the algorithm.

We do not witness any large variation in alphas when moving from HME to the hybrid specification. It increases in three cases (+37.2 bps for GLO, +8.1 bps for GEM, +9.9 bps for EDR) and decreases in three cases (−17.3 bps for GMA, −3.8 bps for FOF, −12.9 bps for SEC). As to the significance levels of the individual regression coefficients, they remain very stable for the original factors with two exceptions. For GIN, the significance levels of the MKT and EMB coefficients drop when the algorithm is applied, and for GEM the (weakly) significant SMB coefficient becomes insignificant. In both cases, however, this adverse effect is compensated through the replacement of an insignificant coefficient under HME with a significant OLS coefficient: for GIN, the HML coefficient of 0.152 is replaced with the corresponding highly significant value of 0.094 under OLS; for GEM, the UMD coefficient of 0.078 is swapped with the highly significant value of 0.114 under OLS.

Stability of risk premia

Our results show a sensible gain in the model quality in several cases. This gain results from the existence of measurement errors in factor risk loadings. By definition, such errors imply economically important uncertainty about factor risk premia. Hence, we now have to consider whether the economic substance of the model is not altered by the passage from OLS to HME, and then from HME to the hybrid model specification.

We meet this objective by assessing, for every strategy, the stability of mean total excess return attributable to each primary source of risk, whether captured by the original observed factor or by the adjustment variable. For the OLS specification, the mean total risk premium of factor k for the whole sample period is only measured by the product of the estimated loading with the average factor value: $\text{PremTot}_k = \hat{\beta}_k^{OLS} \cdot \bar{F}_k$ for $k \in \{m, s, h, u, e, o\}$. For HME, the mean total risk premium is defined as $\text{PremTot}_k = \hat{\beta}_k^{HM} \cdot \bar{F}_k + \hat{\psi}_k \cdot \bar{\hat{w}}_k$. Of course, whenever the specification with highest adjusted significance is a hybrid, the mean total risk premium for each factor is either that obtained with OLS or that of HME, depending of whether or not the adjustment variable has been removed from the regression equation.

Table 10 compares mean total risk premia obtained with the OLS regression technique with those generated by HM estimators and those generated by the application of the recursive algorithm, if applicable.

As follows from the average increase in alphas, the risk premia associated with the factors decrease on average when migrating from the OLS to the HME. Yet, the evolution is not homogenous from one strategy to another or from one factor to another.

When individual hedge fund strategies are considered, two strategies experience dramatic changes in risk premia from the OLS to the HME specification: GLO and Short Sellers. For both of these, several risk premium experience large swings: MKT (+37.4 bps), HML (−20.6 bps), EMB (−52.1 bps) and OPT (−72.5 bps) for GLO, and MKT (+20.8 bps), HML (−49.1 bps), UMD (+13.2 bps), EMB

Table 10: Risk premia corresponding to the risk factors

	Specif.	α	*Mkt*			*SMB*			*HML*			*UMD*			*EMB*			*OPT*		
			factor	*adjust.*	*total*	*factor*	*adjust.*	*total*	*factor*	*adjust.*	*total*	*factor*	*adjust.*	*total*	*factor*	*adjust.*	*total*	*factor*	*adjust.*	*total*
MKN	*OLS*	**0.752**	*0.025*	—	**0.025**	*0.001*	—	**0.001**	*0.050*	—	**0.050**	*0.025*	—	**0.025**	*−0.026*	—	**−0.026**	*0.002*	—	**0.002**
	HME	**0.855**	0.035	−0.008	**0.027**	0.000	0.008	**0.009**	0.079	0.007	**0.086**	−0.011	−0.007	**−0.018**	−0.057	−0.036	**−0.092**	0.000	−0.051	**−0.050**
GLO	*OLS*	**0.338**	*−0.099*	—	**−0.099**	*0.000*	—	**0.000**	*0.009*	—	**0.009**	*0.053*	—	**0.053**	*−0.199*	—	**−0.199**	*−0.065*	—	**−0.065**
	HME	**1.335**	−0.336	0.611	**0.275**	−0.001	0.041	**0.040**	−0.097	−0.099	**−0.197**	0.098	−0.004	**0.094**	−0.361	−0.359	**−0.720**	−0.106	−0.683	**−0.790**
	Hybrid	**1.707**	−0.415	0.752	**0.337**	−0.001	0.036	**0.036**	*0.009*	—	**0.009**	*0.053*	—	**0.053**	−0.534	−0.742	**−1.276**	−0.096	−0.733	**−0.829**
GIN	*OLS*	**0.391**	*0.080*	—	**0.080**	*0.000*	—	**0.000**	*0.061*	—	**0.061**	*0.096*	—	**0.096**	*−0.112*	—	**−0.112**	*0.004*	—	**0.004**
	HME	**0.417**	0.120	−0.107	**0.013**	0.000	0.013	**0.013**	0.097	0.019	**0.116**	0.097	0.007	**0.104**	−0.093	0.054	**−0.039**	0.004	−0.154	**−0.150**
	Hybrid	**0.244**	0.120	−0.111	**0.009**	0.000	0.009	**0.009**	*0.061*	—	**0.061**	*0.096*	—	**0.096**	−0.066	0.116	**0.051**	*0.004*	—	**0.004**
GMA	*OLS*	**0.275**	*0.044*	—	**0.044**	*0.001*	—	**0.001**	*0.030*	—	**0.030**	*0.143*	—	**0.143**	*−0.059*	—	**−0.059**	*0.004*	—	**0.004**
	HME	**0.419**	−0.096	0.074	**−0.022**	0.002	−0.028	**−0.027**	−0.039	−0.021	**−0.061**	0.137	−0.002	**0.134**	−0.092	−0.026	**−0.117**	0.011	0.101	**0.111**
FOF	*OLS*	**0.216**	*0.034*	—	**0.034**	*0.001*	—	**0.001**	*0.045*	—	**0.045**	*0.096*	—	**0.096**	*−0.068*	—	**−0.068**	*0.002*	—	**0.002**
	HME	**0.380**	0.056	−0.016	**0.040**	0.000	0.012	**0.013**	0.028	−0.008	**0.020**	0.104	−0.001	**0.103**	−0.087	−0.022	**−0.109**	−0.001	−0.119	**−0.120**
	Hybrid	**0.342**	0.079	−0.042	**0.037**	0.000	0.026	**0.026**	*0.045*	—	**0.045**	*0.096*	—	**0.096**	−0.076	−0.014	**−0.090**	−0.002	−0.128	**−0.130**
GEM	*OLS*	**0.788**	*−0.055*	—	**−0.055**	*0.001*	—	**0.001**	*0.065*	—	**0.065**	*0.132*	—	**0.132**	*−0.366*	—	**−0.366**	*0.004*	—	**0.004**
	HME	**0.989**	0.248	−0.152	**0.097**	−0.001	0.071	**0.070**	0.117	0.000	**0.117**	0.089	−0.010	**0.079**	−0.278	0.073	**−0.205**	−0.012	−0.415	**−0.428**
	Hybrid	**1.070**	0.247	−0.247	**0.000**	−0.001	0.123	**0.122**	*0.065*	—	**0.065**	*0.132*	—	**0.132**	−0.290	0.085	**−0.205**	−0.013	−0.453	**−0.466**
EDR	*OLS*	**0.488**	*0.064*	—	**0.064**	*0.001*	—	**0.001**	*0.060*	—	**0.060**	*0.005*	—	**0.005**	*−0.037*	—	**−0.037**	*0.002*	—	**0.002**
	HME	**0.511**	0.128	−0.036	**0.092**	0.000	0.026	**0.026**	0.100	0.007	**0.107**	−0.017	−0.005	**−0.022**	−0.064	−0.036	**−0.099**	−0.002	−0.060	**−0.063**
	Hybrid	**0.610**	0.098	−0.038	**0.061**	0.000	0.047	**0.047**	*0.060*	—	**0.060**	*0.005*	—	**0.005**	−0.086	−0.080	**−0.166**	−0.002	−0.063	**−0.064**
SHO	OLS	**0.862**	−0.294	—	**−0.294**	−0.002	—	**−0.002**	0.055	—	**0.055**	−0.095	—	**−0.095**	−0.026	—	**−0.026**	−0.066	—	**−0.066**
	HME	**1.112**	−0.429	0.343	**−0.086**	−0.002	−0.001	**−0.003**	−0.234	−0.203	**−0.436**	0.009	0.028	**0.037**	−0.062	−0.083	**−0.145**	−0.074	0.203	**0.129**

LON	*OLS*	**0.291**	*0.246*	—	**0.246**	*0.003*	—	**0.003**	*−0.012*	—	**−0.012**	*0.032*	—	**0.032**	*−0.062*	—	**−0.062**	*0.009*	—	**0.009**
SEC	*OLS*	**0.830**	*0.268*	—	**0.268**	*0.002*	—	**0.002**	*0.012*	—	**0.012**	*0.174*	—	**0.174**	*−0.004*	—	**−0.004**	*0.032*	—	**0.032**
	HME	**0.580**	0.314	−0.056	**0.258**	0.003	−0.035	**−0.032**	0.160	0.101	**0.261**	0.143	0.001	**0.144**	0.069	0.142	**0.211**	0.140	−0.300	**−0.160**
	Hybrid	**0.451**	0.343	−0.090	**0.253**	0.003	−0.021	**−0.018**	0.185	0.101	**0.285**	*0.174*	—	**0.174**	0.096	0.173	**0.270**	0.144	−0.295	**−0.152**
GES	*OLS*	**0.604**	*0.185*	—	**0.185**	*0.001*	—	**0.001**	*0.028*	—	**0.028**	*0.086*	—	**0.086**	*−0.033*	—	**−0.033**	*−0.003*	—	**−0.003**
	HME	**0.747**	0.287	−0.155	**0.132**	0.001	0.017	**0.017**	0.107	0.029	**0.137**	0.037	−0.013	**0.024**	0.000	0.048	**0.047**	0.004	−0.246	**−0.242**

This table reports the value of the total mean risk premia (in percents) attributable to each risk factor for the sample period 1994:02–2002:12. For each specification, the alpha is the intercept of the regression. With OLS, the mean total risk premium of factor k is equal to $\hat{\beta}_k^{OLS}\bar{F}_k$. With HME, the mean total risk premium of factor k is equal to $\hat{\beta}_k^{HM}\bar{F}_k + \hat{\psi}_k\bar{\hat{w}}_k$. Under the hybrid specification, the mean total risk premium of factor k is either the OLS risk premium if the variable has been removed from the HME estimation, or the mean total risk premium of the last pass of the HME.

Note: Best specifications based on Ps. $\bar{R}^2$ are in bold; OLS estimates are italicised.

(−11.9 bps) and OPT (+19.5 bps) for Short Sellers. The explanation of these two series of returns seems to suffer from significant alterations from the change in specifications. For the GLO strategy, this could be reasonably explained by the very small number of live funds at the end of the sample period (only one live fund on December 2002) that weakens the economic significance of the strategy returns. For Short Sellers, the sample also suffers from a small number of funds, and this may explain the instability of the risk premium.

For the other eight strategies (excluding LOL, for which HME does not dominate OLS), the stability of the first five risk premia is quite high. The average difference in mean total risk premium between HME and OLS is equal to 1.5 bps with a standard deviation of 7.7 bps (40 observations). Such evidence contrasts with the large decline in the option-based risk premium: from OLS to HME, it decreases on average by 14.3 bps, with only one positive value (GMA) and a standard deviation of 15.7 bps. Thus, accounting for measurement errors in option-based factors appears to decrease the average risk premium of these eight hedge fund strategies by a substantial yearly 1.7 per cent.

The economic relevance of the recursive regression algorithm can also be assessed by considering the difference in risk premia between the HME and the hybrid specifications. For the five strategies (excluding the GLO strategy) for which the HME and hybrid specification differ (GIN, FOF, GEM, EDR, SEC), the average difference between the HME risk premium and the new risk premium (calculated with HME) of the hybrid specification is as low as 0.2 bp, with a standard deviation of 4.0 bps (20 observations). Hence, for our sample, we find no evidence that the application of the algorithm significantly alters the risk premium associated with each factor. This finding holds provided that the strategy return index features a sufficient number of funds, as shown by the inconclusive results for the GLO and, to a lesser extent, the Short Sellers strategies.

CONCLUSION

Hedge fund return data are characterized by small samples of usable data, with large nonlinearities in risk exposures, and with a clear presence of option-based factors to explain the pattern of returns. In this context, this article shows that the application of the Dagenais and Dagenais[17] HME is a natural and logical choice. Yet, the price to pay for an accurate account for EIV is a substantial inflation in the number of coefficients to estimate. In parallel with the estimation, we have therefore developed a new heuristic algorithm to circumvent one of the weaknesses of the proposed estimator.

The empirical test on a sample of hedge fund data provides informative evidence on the applicability of the procedure. The results suggest that the use of HME emphasizes a statistically significant presence of EIV, and therefore this estimation approach dominates OLS specifications for most series of returns. The performance of hedge fund strategies is enhanced, on average, when measurement errors are properly taken into account. The use of HME does not significantly alter the risk premia attributable to each source of risk, except for the optional factor, where we find that the OLS tends to overestimate the exposure to option-based risk factors.

Our persistence study on individual hedge funds, which is performed on two sub-periods that correspond to bullish and bearish overall market conditions, emphasizes the ability of our instrumental approach to detect persistence in a more consistent way than an OLS specification. This finding is hardly surprising, as the purpose of an instrumental approach is precisely to improve the quality of the estimation, and thus detect more accurately the source of performance. This finding opens up clear avenues for further research on relative persistence of hedge fund performance.

We can reduce the number of instruments through a heuristic reduction algorithm, with no reduction in overall significance. We find evidence of very small changes in the factor risk premia with respect to HME. This procedure enables us to associate the rigor of HME with the parsimony of the OLS specification.

The scope of application of this approach seems to be very large. Given the fact that many data samples are too small to lend themselves easily to nonlinear estimation such as the use of dynamic or conditional betas, HME together with the recursive regression algorithm might serve as a very credible alternative. In the context of hedge fund research, the inflation in the number of candidate variables to explain hedge fund returns will probably soon call for a solution that reconciles robustness and parsimony. We view our contribution as a step in this direction. Finally, the specific study of which variables are likely to trigger contamination or attenuation issues could prove to be of strong economic significance. This type of investigation features in our ongoing research agenda.

ACKNOWLEDGEMENTS

We thank Benoît Carmichael, Jean-François L'Her, François-Éric Racicot and participants at the Australasian Finance and Banking Conference in Sydney (2006), at the French

Finance Association Conference in Paris (2006) and at the European Financial Management Association Conference in Vienna (2007) for helpful comments. A part of this article was revised when the first author was at HEC School of Management of the University of Liège: he thanks the KBL Chair in Fund industry during this period. Georges Hübner thanks Deloitte (Luxembourg) for financial support. All errors are ours.

REFERENCES AND NOTES

1 Fung, W. and Hsieh, D.A. (1997) Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies* 10: 275–302.
2 Liang, B. (1999) On the performance of hedge funds. *Financial Analysts Journal* 55(4): 72–85.
3 Edwards, F.R. and Caglayan, M.O. (2001) Hedge fund and commodity fund investments in bull and bear markets. *Journal of Portfolio Management* 27: 97–108.
4 Capocci, D., Corhay, A. and Hübner, G. (2005) Hedge fund performance in bull and bear markets. *European Journal of Finance* 5: 361–392.
5 Fung, W. and Hsieh, D.A. (2004) Hedge fund benchmarks: A risk-based approach. *Financial Analysts Journal* 60(5): 65–81.
6 Posthuma, N. and Van der Sluis, P.J. (2005) Analyzing style drift in hedge funds. In: G. Gregoriou, G. Hübner, N. Papageorgiou and F. Rouah (eds.) *Hedge Funds. Insights in Performance Measurement, Risk Analysis, and Portfolio Allocation*. New York: Wiley & sons, pp. 83–104.
7 Mitchell, M. and Pulvino, T. (2001) Characteristics of risk and return in risk arbitrage. *Journal of Finance* 56: 2135–2175.
8 Fung, W. and Hsieh, D.A. (2000) Performance characteristics of hedge funds and commodity funds: Natural versus spurious biases. *Journal of Financial and Quantitative Analysis* 35: 291–307.
9 Agarwal, V. and Naik, N.Y. (2004) Risks and portfolio decisions involving hedge funds. *Review of Financial Studies* 17: 63–98.
10 Glosten, L. and Jagannathan, R. (1994) A contingent claims approach to performance evaluation. *Journal of Empirical Finance* 1: 133–160.
11 Chan, N., Getmansky, M. and Haas, S.M. (2006) Systemic risk and hedge funds. In: M. Carey and R. Stulz (eds.) *The Risks of Financial Institutions and the Financial Sector*. Chicago, IL: University of Chicago Press.
12 Fama, G. and MacBeth, J. (1973) Risk, return and equilibrium: Empirical test. *Journal of Political Economy* 81: 607–636.
13 Shanken, W.F. (1992) On the estimation of beta-pricing models. *Review of Financial Studies* 5: 1–34.
14 Kandel, S. and Stambaugh, R. (1995) Portfolio inefficiency and the cross-section of expected returns. *Journal of Finance* 50: 157–184.
15 Cragg, J.G. (1994) Making good inferences from bad data. *Canadian Journal of Economics* 27: 776–800.
16 Cragg, J.G. (1997) Using higher moments to estimate simple error-in-variable model. *Rand Journal of Economics* 28(0): S71–S91.
17 Dagenais, M.G. and Dagenais, D.L. (1997) Higher moment estimators for linear regression models with errors in the variables. *Journal of Econometrics* 76: 193–221.
18 Lewbel, A. (2006) Using Heteroskedasticity to Identify and Estimate Mismeasurement and Endogenous Regressor Models. Boston College, Working paper.
19 Coën, A. and Hübner, G. (2009) Risk and performance estimation in hedge funds: Evidence from errors in variables. *Journal of Empirical Finance* 16: 112–125.

20 Hedge fund models are essentially empirical and lack of theoretical foundations compared to the CAPM or APT model. Following Fung and Hsieh[8] or Agarwal and Naik,[9] the main idea is to explain returns by *ad hoc* linear asset pricing models. Thus, it is quite difficult to define *ex ante* the sources of errors. These models are good candidates for EIV and measurement errors.

21 Frisch, R.A. (1934) *Statistical Confluence Analysis by Means of Complete Regression System*. Oslo, Norway: University Institute of Economics.

22 There is no bias when only the dependent variable is plagued with measurement errors (see, for example, Davidson and MacKinnon.[27]
23 The measurement error is said to be 'classical' when it is uncorrelated with the true unobserved factor.
24 The nature and extent of the bias is obtained by computing the asymptotic values of the OLS estimates, which converge to their true value if true factors are observable (see Carmichael and Coën[38]).

25 Durbin, J. (1954) Errors in variables. *International Statistical Review* 22: 23–32.
26 Pal, M. (1980) Consistent moment estimators of regression coefficients in the presence of errors in variables. *Journal of Econometrics* 14: 349–364.
27 Kendall, M.G. and Stuart, A. (1963) *The Advanced Theory of Statistics*, 2nd edn. London: Charles Griffin.
28 Malinvaud, E. (1978) *Méthodes Statistiques de l'Économétrie*, 3rd edn. Paris: Dunod.

29 Davidson, R. and MacKinnon, J.G. (1993) *Estimation and Inference in Econometrics*. New York: Oxford University Press.
30 This treatment explains why we have a lower number of funds remaining than in Capocci *et al.*[4]
31 Fama, E.F. and French, K.R. (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56.
32 Carhart, M. (1997) On persistence in mutual funds performance. *Journal of Finance* 52: 57–82.
33 Capocci, D. and Hübner, G. (2004) Analysis of hedge fund performance. *Journal of Empirical Finance* 11: 55–89.
34 We would like to mention that the Dagenais and Dagenais HME was developed for linear regressions. Thus, we have chosen to ignore nonlinear regressions and add optional factors as suggested by Agarwal and Naik.[9] Estimation for nonlinear models would be different and this is not the purpose of our study.
35 Diez de los Rios, A. and Garcia, R. (2005) Assessing and Valuing the Nonlinear Structure of Hedge Funds' Returns. Working Paper. CIRANO.
36 Bailey, W., Li, H. and Zhang, X. (2004) Hedge Fund Performance Evaluation: A Stochastic Discount Factor Approach. Cornell University. Working Paper.
37 The peaking monthly 1.335 per cent versus 0.338 per cent for the Global strategy is a clear outlier with respect to the rest of the table. This is probably due to the fact that this strategy mostly consists of dead funds, as the funds belonging to this strategy have been reshuffled to the other 'Global-based' strategies since 1999 (see Capocci and Hübner[33]).
38 Carmichael, B. and Coën, A. (2008) Asset pricing models with errors-in-variables. *Journal of Empirical Finance* 15: 778–788.