ESG UQÀM

Cahier de recherche
2006-03

Chaire d'information financière et organisationnelle

Towards new empirical versions of financial and accounting models corrected for measurement errors

François-Éric Racicot[†]

Raymond Théoret[††]

Alain Coën[†††]

[†]*Professor, University of Quebec (Outaouais), (UQO), 283 Alexandre Tache Boul., Gatineau, Quebec, Canada, J8X 3X7. E-mail: francoiseric.racicot@uqo.ca, Tel. number: (819) 595-3900 ext. 1727.*
[††] *Professor, University of Quebec (Montreal), (UQAM), 315 Ste-Catherine east, Montreal, Quebec,Canada. H3X 2X2. E-mail: theoret.raymond@uqam.ca. Tel. number: (514) 987-3000 ext. 4417.*
[†††]*Professor, University of Quebec (Montreal), (UQAM). E-mail:coen.alain@uqam.ca.*

Towards new empirical versions of financial and accounting models corrected for measurement errors

Résumé

Dans cet article, nous proposons une nouvelle version empirique du modèle de Fama et French basée sur le test de spécification d'Hausman (1978) et qui vise à éliminer les erreurs de mesure dans les données. Le cadre empirique proposé est assez général pour convenir à d'autres modèles financiers ou comptables qui sont estimés à l'aide de données mesurées avec erreur. L'élimination des erreurs de mesure revêt un caractère primordial à plusieurs niveaux comme ceux de la publication de données financières, de la régie d'entreprise et de la protection des investisseurs.

Abstract

In this paper, we propose a new empirical version of the Fama and French model based on the Hausman (1978) specification test and aimed at discarding measurement errors in the variables. The proposed empirical framework is general enough to be used for correcting other financial and accounting models for measurement errors. Removing measurement errors is important at many levels as information disclosure, corporate governance and protection of investors.

Mots-clefs: *Évaluation des actifs; sélection de portefeuille; erreurs sur les variables; erreurs de mesure; moments supérieurs; variables instrumentales; test de spécification; régie d'entreprise; protection des investisseurs.*

Key words: *Asset Pricing; Portfolio Selection; Errors in variables; Measurement errors; Higher moments; Instrumental variables; Specification Test; Corporate Governance; Protection of investors.*

JEL classification: C13; C19; C49; G12; G31.

1. Introduction[1]

Until recently, the financial community was little concerned by the problem of measurement errors[2] in accounting and financial data. The financial scandals which took place at the turn of the second millennium had contributed to make the public more aware of this problem. In the case of these scandals, the «measurement errors» were caused by fraudulent manipulation of data. Saying that these data are measured with error in this situation could be an abuse of language but they are perceived as such by an uninformed public. But these errors might come obviously from many other sources: aggregation bias, bid-ask bounce, differences in timing of data, survivor bias, to name a few.

Really, neglecting errors in variables may invalidate completely the conclusions of a financial analysis or the results of the estimation of a financial model. In the econometric jargon, these errors cause a serious bias in the estimation process. For example, in this article, we will correct the Fama and French model for its problem of errors in the variables. As we know, the estimated alpha is a very important byproduct of this model: this coefficient is used to select undervalued and overvalued securities. If the model is not purged from its problem of errors in variables, the estimated alpha will give wrong signals to investors and the resulting losses might be very important. We can transpose the same reasoning to indicators used in financial analysis. If these indicators are measured erroneously, wrong decisions will result.

[1] The approach followed in this paper was previously developed by Racicot (2003) who applied it to the well-known market model. He postulates that the model resulting from the addition of artificial variables for removing errors in variables may be considered as a new model by itself. We follow a very similar approach in this article, except that we use higher moments instead of cumulants to set up the instrumental variables.

[2] In this article, we use with the same meaning «measurement errors» and «errors in variables».

If the problem of errors in variables, in spite of the consequences, is so often neglected, it seems because the methods to discard errors in variables are not so advanced as other fields of financial econometrics. In the past, it was usual to estimate financial models like CAPM or APT or accounting models, like accruals models, without correcting the problem associated to measurement errors. But this approach is no longer justifiable. We must eliminate errors in variables using empirical advances in the statistical and econometric fields.

In this article, we propose a new empirical version of the Fama and French model which includes variables correcting the exposures to the risk factors for the problem of measurement errors. We therefore derive a new measure of alpha which is purged from this problem. This alpha is likely to give better signals about the selection of securities. Our approach is general enough to be applicable to other financial and accounting models which suffer from the problem of errors in variables. Our procedure, which is based on a version of the Hausman test using artificial regressions, can detect which variables are measured with errors and can adjust consequently the coefficients of the corresponding variables. In the context of the Fama and French model, the factors which remove errors in variables are labeled: "factors of correction of risk exposures". But obviously, these correcting factors must be reinterpreted in the context of each model.

This paper is organised as follows. Section 2 outlines the context of the study and gives a brief survey on the literature of measurement errors in relation with risk analysis.

Section 3 deals with the methodology used to carry our empirical analysis. First, theoretical background of the problem caused by errors in variables and of the chosen instruments- , which consist in higher moments of the regressors of the model of Fama and French, - is discussed. Second, the method of estimation used to correct the problem of errors in variables in the model of Fama and French is presented. That will gives way to the development of our new empirical version of the Fama and French model based on the Hausman specification test which uses an artificial regression. Section 4 describes the data series on hedge funds returns used to estimate the Fama and French model and reports the empirical results. A summary and concluding remarks are given in the final section.

2. An outline of the study and a brief survey of the literature on the problem of errors in variables in relation with measures of risk

Financial returns, and especially hedge funds returns, are contaminated by errors in variables. If neglected, this problem may completely invalidate the results obtained by estimating financial or accounting models, like the CAPM, the Fama and French model or the Gordon-Shapiro one. Nevertheless, as said in the introduction, researchers seldom discard errors in variables in their empirical works, perhaps because the econometric methods related to the correction of the problem of errors in variables are not so advanced as other fields of financial econometrics. Some papers[3] have proposed correction techniques for errors in variables in a CAPM context but these studies are not numerous and much work must be done.

[3] The first study on the treatment of errors in variables in the context of asset pricing models was done by Fama and Macbeth (1973). Other following papers on this subject are those of : Shanken (1992), Ferson and Locke (1998) and Pastor and Stambaugh (1999).

This paper proposes a new empirical version of the Fama and French model based on the Hausman (1978) specification test[4]. This empirical model will incorporate correction factors to risk exposure. These factors will deal with the problem of errors in variables. Our model will also include another innovation. It will use higher moments as instruments for removing the problem of errors in variables. Really, the methods of financial econometrics were traditionally well suited to estimate financial models which, like the CAPM, postulate a linear relationship between return and risk. But this kind of relation is valid only if risk is small or under very special assumptions as Gaussian returns or quadratic utility function which is not realistic because it implies increasing absolute risk aversion. Prudence of an investor[5], which is associated to the third derivative of the utility function, and fat tail risk, which is related to the fourth derivative of the utility function, introduce nonlinearities in the relation between return and risk. In this article, we reposition the problem of correcting errors in variables in the context of a nonlinear relation between return and risk. Higher order moments and cumulants of returns become very important in this analysis of errors in variables in relation with nonlinear risk.

Following Campbell et al. (1997), the errors in variables problem in asset pricing models can be addressed in two ways. One (Fama and MacBeth (1973)) is to minimize this problem by grouping stocks into portfolios. The second approach (Shanken (1992)) is to explicitly adjust the standard errors to correct for the biases introduced by the errors in variables. But there is a more recent approach adopted by Kandel and Stambaugh (1995) which uses GLS

[4] In his Ph.D. thesis, Racicot (2003) applied a similar form of this model to the CAPM. Instead of using higher moments as instruments, he used cumulants.
[5] Prudence means "aversion to downside risk" in finance, which seems to be a universal virtue.

instead of least squares as estimator. But, in this approach, the covariance matrix required to weight the observations must be estimated, which might be a problem.

Durbin (1954), Pal (1980) and more recently Racicot (1993) and Dagenais and Dagenais (1997)[6] have proposed an estimation method to correct errors in variables based on cumulants of the error terms of the equations of variables measured with errors. These cumulants serve as instruments to remove errors in variables. In this article, we will use a variant of this method to correct the errors in variables problem which might plague the well-known augmented Fama and French model and which is based on higher moments.

As will be shown, the recourse to higher moments to eliminate errors in variables opens the doors to a synthesis between the modern asset pricing theory and the financial econometric treatment of the errors in variables. Surely, Durbin (1954), Pal (1980), Racicot (1993) and Dagenais and Dagenais (1997) did not aim to transpose their technique to asset pricing theory and risk measures. But it is well-admitted now that the first two moments of returns, i.e. the mean and the variance, are largely insufficient to measure the risk of a portfolio[7]. Huang and Litzenberger (1988), Ingersoll (1987) or either Levy (1992) had mentioned that the paradigm of portfolio selection based on the first two moments of the returns distribution maximizes the expected utility of an agent only in either of the two following situations: his utility function is quadratic or the returns distribution is normal. Of course, these two postulates are violated in the real world. Following this "constat", Samuelson (1970), Rubinstein (1973), Kraus and Litzenberger (1976), Friend and

[6] See also : Dagenais (1994) and Racicot (2003).
[7] The following remarks on risk measures are adapted from: Coën and Théoret (2004).

Westerfield (1980) and Sears and Wei (1985) have put the foundations of the approach based on higher moments for pricing financial instruments. These theoretical developments gave birth to the three-moment and the four-moment CAPM[8].

The theoretical developments related to the use of higher moments as measures of financial risk are linked to other sections of risk theory, but a synthesis is yet to do. The theory of stochastic dominance[9] has a long past. The increasing orders of stochastic dominance add more and more higher moments of the returns distribution to judge if a portfolio is superior to another in terms of stochastic dominance. This theory has produced new risk measures as the risk of shortfall which is based on the probability distribution of returns and transfers of probability mass between high wealth states and low wealth states. In the same order of ideas, Scott and Hovarth (1980) have alleged that the odd moments of the returns distribution, like a positive mean and a positive skewness, have positive marginal utility for an investor. The positive even moments, like variance and kurtosis, have for their sake negative marginal utility. These developments constitute the foundations of the modern theory of risk which is under construction.

The new empirical version of the Fama and French model based on the Hausman specification test which we propose in this paper is in line with the new literature on risk. Our contribution will be double. First, we will incorporate factors adjusting or correcting risk exposure in the Fama and French model, this correction being required by a possible problem of errors in variables which will be detected by the Hausman specification test.

[8] On the three-moment and four-moment CAPM, see: Lhabitant (2004), chap. 8; Harvey and Siddique (2000); Lim (1989); Kraus and Litzenberger (1976) and Rubinstein (1973) .
[9] See the survey of Levy (1992) on stochastic dominance.

Next, we will use as instruments for correcting errors in variables the higher moments of the predetermined regressors[10] in the equation of Fama and French. These instruments will not only serve as technical tools but also as measures of risk, what will integrate our estimation method in the modern theory of risk analysis which is essentially non linear. As we will see, our procedure is all the more relevant since the regressors (factors) of the equation of Fama and French are long-short portfolios which are consequently close to those of hedge funds. Indeed, these portfolios which mimic some "market anomalies" are long in a category of returns (e.g. returns of small firms) and short in the opposite category (e.g. returns of big firms). Consequently, these portfolios behave like portfolios of options which incorporate many nonlinearities. Moments of order higher than two are therefore required to take them into account. Taleb (1997) even suggests using moments of order higher than four to measure the risk of an option[11], odd moments being measures of asymmetry and even moments being measures of convexity[12]. The empirical work will show that our choice of instruments is judicious. Really, the correlation of a regressor with moments higher than two is often much more important than the correlation with classic inferior moments.

[10] i.e. lagged regressors.
[11] For example, the fifth moment is the asymmetry sensitivity of the fourth one. The seventh moment is the sign of the convexity change as the underlying asset moves up or down. For Taleb, moments higher than four are especially important for compound options. See: Taleb (1997), p. 202-204.
[12] For example, kurtosis measures negative convexity, which is bad from the point of view of a risk averter. In an option perspective, kurtosis represents negative gamma.

3. Methodology

3.1 The biases caused by errors-in-variables

The errors in variables problem is well-known in the econometric literature but is often overlooked in finance. Assume[13] a one independent variable regression model where observed y and x are measured with errors. The true values of these variables are y^* and x^*. The relations between the observed variables and their true counterparts are:

$$y = y^* + u \qquad (1)$$

$$x = x^* + v \qquad (2)$$

where u and v are error vectors. These error vectors have zero means and their variances are: $E(uu') = \sigma_u^2 I$ and $E(vv') = \sigma_v^2 I$. The vectors u and v are assumed orthogonal.

The exact linear relationship between the two unobserved variables is:

$$y^* = x^* \beta \qquad (3)$$

Because y^* and x^* are unobservable, we substitute equations (1) and (2) in equation (3):

$$y = x\beta + \varepsilon \qquad (4)$$

where $\varepsilon = v - u\beta$. Consequently, by equation (2), x is correlated with the error term ε , and this creates a bias. To compute it, we solve equation (4) for $\hat{\beta}$, the estimated value of parameter β , by the method of least squares:

$$\hat{\beta} = (x'x)^{-1} x'y \qquad (5)$$

Substituting equation (4) in equation (5), we obtain:

$$\hat{\beta} = \beta + (x'x)^{-1} x'e \qquad (6)$$

---

[13] This presentation follows quite closely Judge and al. (1985).

Equation (6) shows that $\hat{\beta}$ is biased. And in large samples, it is not consistent:

$$p\,lim\,\hat{\beta} = \beta\left[1 - \frac{\sigma_u^2}{\sigma_x^2}\right] \quad (7)$$

where $\sigma_x^2 = \sigma_{x*}^2 + \sigma_u^2$. We can rewrite expression (7) as:

$$p\lim\hat{\beta} = \beta(1-\lambda) \qquad (8)$$

where $\lambda < 1$.

According to equation (8), the problem of errors in variables tends to underestimate $\hat{\beta}$, the degree of underestimation being conditioned by λ. The more λ is near 1, the more the problem of errors in variables is serious. At the limit, λ is 1 and $p\lim\beta = 0$. When only y is plagued with errors of measurement in equation (1), there is no bias because x remains uncorrelated with the innovation of the equation. The problem appears when x is measured with error. That creates a correlation between x and the innovation term and therefore, a bias.

Unfortunately, if there is more than one explanatory variable in a model, we cannot know a priori the relative impact of errors in variables on the estimation process. Some parameters will be overstated and others understated. But as seen in another section, the Hausman test, and more precisely the version of this test based on artificial regressions, will not only help us to detect errors in variables but also gives us more information about the incidence of this problem on estimated parameters.

3.2 The choice of instrumental variables to estimate the augmented Fama and French (F&F) model

The augmented F&F (1992, 1993 and 1997) model is a purely empirical model which may be written as:

$$R_{pt} - R_{ft} = \alpha + \beta_1 \left(R_{mt} - R_{ft} \right) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \varepsilon_t \quad (9)$$

with:

$R_{pt} - R_{ft}$: the excess return of a portfolio, R_{ft} being the risk-free return;

$R_{mt} - R_{ft}$: the market risk premium;

SMB: a portfolio which mimics the "small firm anomaly", which is long in the returns of selected small firms and short in the returns of selected big firms;

HML: a portfolio which mimics the "income stock anomaly", which is long in returns of stocks of selected firms having a high (book value/ market value) ratio (income stocks) and short in selected stocks having a low (book value/ market value) ratio (growth stocks);

UMD: a portfolio which mimics the "momentum anomaly", which is long in returns of selected stocks having a persistent upper trend and short in stocks having a persistent downwards trend.

To explain the return of a stock or of a portfolio of stocks, the F&F model adds to the unique factor retained by the CAPM, the market risk premium, three other factors which are assumed to represent market anomalies: the small firm anomaly, the book value to market value anomaly and the momentum anomaly[14].

We postulate that the three mimicking portfolios *SMB*, *HML* and *UMD* might be measured with errors. They are thus possibly correlated with the innovation term in equation (9) and the estimators of the parameters of this equation obtained by ordinary least squares (OLS) are consequently biased and not consistent. To purge these coefficients from these biases, we must regress in a first pass the independent variables on instrumental variables. The estimated method used in this paper which is based on the Hausman test will be explained below. The problem lays in the choice of these instruments.

As we said before, it is difficult to find valuable instruments for the excess returns of the mimicking portfolios. Being long in some stocks and short in others, their cash flows are similar to those of hedge funds. Higher moments of returns, like asymmetry and kurtosis, might have a great influence on these returns. This suggests the use of higher moments of the variables on the RHS of equation (9) as instrumental variables. An econometric theory is indeed in construction on this subject. Following Durbin (1954) and Pal (1980), Dagenais and Dagenais (1997) showed that higher moments[15] of independent variables of a regression might be valid instruments to remove errors-in-variables. But instead of defining higher moments as in these papers, we will adopt a method more akin to asset pricing theory which defines higher moments of returns by powers of these returns.

---

[14] The original F&F model contained only the first two "anomalies". The momentum anomaly, which is due to Carhart (1997) and Jegadesh and Titman (1993), was introduced subsequently, to form the augmented Fama and French model.

[15] The reader will excuse ourselves for confounding at this stage higher moments and cumulants for the sake of the presentation. We will come back on the distinction between these two concepts later.

The method of asset pricing based on higher moments is not new. Rubinstein (1973) and Kraus and Litzenberger (1976) put the foundations of the three-moment and four-moment CAPM. The three-moment CAPM integrated asymmetry of returns in the analysis while the four-moment CAPM added kurtosis.

Some authors, like Kraus and Litzenberger (1976), use a general utility function to derive the moment-CAPM. Others use a Taylor expansion of the utility function which, following Samuelson (1970) and Rubinstein (1973), allows expressing utility in terms of the higher moments of returns.

Let us assume that the expected utility of wealth, $E[U(W)]$, is function of the n first moments of the distribution of wealth:

$$E[U(W)] = \varphi\left[\overline{W}, \sigma_W, skew_W, kur_W, .., sm_{nW}\right] \qquad (10)$$

with $\overline{W}$ the expected value of wealth; σ_W, the volatility of wealth; $skew_W$, its skewness; kur_W, its kurtosis and sm_{nW}, the n^{th} moment of the distribution of wealth. We incorporate moments of order 5 and more because we know that they might be important to explain the returns of long-short portfolios like mimicking portfolios which are the foundation of the F&F model.

The expected utility of end-of-period wealth is maximized over the one period horizon subject to the constraint of initial wealth, which is:

$$a_0 + \sum_{i=1}^{n} a_i = w_0 \qquad (11)$$

According to equation (11), the initial wealth w_0 is allocated between the risk-free asset a_0 and n other risky assets designated by a_i. To maximize the utility of end-of-period wealth subject to (11), we form the usual Lagrangian:

$$L = E[U(W)] - \phi\left(a_0 - \sum_{i=1}^{n} a_i - w_0\right) \qquad (12)$$

Taking the first-order conditions for a maximum, we have:

$$\frac{\partial L}{\partial a_0} = \varphi_{\overline{W}} \times \frac{\partial \overline{W}}{\partial a_0} - \phi = 0 \qquad (13)$$

$$\frac{\partial L}{\partial a_{i,i\neq 0}} = \varphi_{\overline{W}} \times \frac{\partial E(w)}{\partial a_i} + \varphi_{\sigma_W} \times \frac{\partial \sigma_W}{\partial a_i} + \varphi_{skew_W} \times \frac{\partial skew_W}{\partial a_i} + \varphi_{kur_W} \times \frac{\partial kur_W}{\partial a_i} + \ldots + \varphi_{sm_{nW}} \frac{\partial sm_{nW}}{\partial a_n} - \phi = 0$$

(14)

with $\varphi_x = \frac{\partial E[U(X)]}{\partial x}$.

End-of-period expected wealth is equal to:

$$\overline{W} = (1 + R_0)a_0 + \sum_{i=1}^{n} [(1 + E(R_i))a_i] \qquad (15)$$

We thus have:

$$\frac{\partial \overline{W}}{\partial a_0} = 1 + R_0 \qquad (16)$$

$$\frac{\partial \overline{W}}{\partial a_i} = 1 + E(R_i) \qquad (17)$$

We can therefore express the moments of wealth in terms of the moments of returns of the portfolio as:

$$\sigma_W = \sum_i a_i \beta_{ip} \sigma_p \qquad (18)$$

$$skew_W = \sum_i a_i \gamma_{ip} skew_p \qquad (19)$$

$$kur_W = \sum_i a_i \theta_{ip} kur_p \qquad (20)$$

$$sm_{nW} = \sum_i a_i \omega_{ip} sm_{np} \qquad (21)$$

with $\sigma_p = \sqrt{E\left(R_p - \overline{R}_p\right)^2}$, the volatility of the return of the risk assets; $\beta_{ip} = \frac{E\left[\left(R_i - \overline{R}_i\right)\left(R_p - \overline{R}_p\right)\right]}{\sigma_p^2}$, the beta of risk asset i with the investor's portfolio of risk assets. $skew_p = \sqrt[3]{E\left[\left(R_P - \overline{R}_p\right)^3\right]}$, the asymmetry of the return of the portfolio and $\gamma_{ip} = \frac{E\left[\left(R_i - \overline{R}_i\right)\left(R_p - \overline{R}_p\right)^2\right]}{sk_p^3}$, the gamma of risk asset *i* with the investor's portfolio of risk assets . It is easy to generate all the other variables by following this pattern.

Equating equations (13) and (14) to delete ϕ and taking into account equations (18) to (21), we arrive at an expression for E(R_i) in terms of the moments of the distribution of the return of an investor's portfolio:

$$E(R_i) - R_0 = -\left(\frac{\varphi_{\sigma_W}}{\varphi_{\overline{W}}}\right)\beta_{ip}\sigma_p - \left(\frac{\varphi_{skew_W}}{\varphi_{\overline{W}}}\right)\gamma_{ip} skew_p - \left(\frac{\varphi_{kur_W}}{\varphi_{\overline{W}}}\right)\theta_{ip} kur_p - \ldots - \left(\frac{\varphi_{sm_{nW}}}{\varphi_{\overline{W}}}\right)\omega_{ip} sm_{np} \quad (22)$$

with $\frac{\varphi_{moment}}{\varphi_{\overline{W}}}$ being an investor's marginal rate of substitution between expected wealth and a specific moment. According to Scott and Hovarth (1980), these marginal rates of substitution are positive for odd moments, like mean and positive asymmetry, and negative for even moments, like variance and kurtosis. From equation (22), odd moments have, ceteris paribus, a negative impact on expected return from the point of view of investors. Even moments have a positive impact because they represent risk.

Moving from equation (22) to the condition of market equilibrium for E(R_i) requires making, according to Kraus and Litzenberger (1976), the strong assumption of homogeneous expectations for investors. Following this assumption, equation (22) becomes, assuming that p is the market portfolio:

$$E(R_i) - R_0 = \left[\frac{d\overline{W}}{d\sigma_W}\right]\sigma_m \beta_{im} + \left[\frac{d\overline{W}}{dskew_W}\right]skew_m \gamma_{im} + \left[\frac{d\overline{W}}{dkur_w}\right]kur_m \theta_{im} + \ldots + \left[\frac{d\overline{W}}{dsm_{nW}}\right]sm_{nm}\omega_{in}$$

(23)

The terms in brackets in expression (23) are the slopes of the efficient frontiers whose arguments are expected wealth and the respective moment. We obtain finally the n-moment CAPM:

$$E(R_i) - R_f = \psi_1 \beta_{im} + \psi_2 \gamma_{im} + \psi_3 \theta_{im} + \ldots + \psi_n sm_{nm} \qquad (24)$$

We might use directly expression (24) to define our instruments for removing errors-in-variables by the methods of higher moments in the F&F model. Assume we want to correct the mimicking portfolio *SMB* for errors-in-variables. In the first pass of our regressions, we would regress this variable on the co-moments of the lagged excess return of the market portfolio. The variable *SMB* corrected for errors-in-variables would be:

$$S\hat{M}B_t = \kappa_0 + \kappa_1 \beta_{im,t-1} + \kappa_2 \gamma_{im,t-1} + \kappa_3 \theta_{im,t-1} + \ldots + \kappa_n sm_{nm,t-1} \qquad (25)$$

Really, we would have to introduce also the co-moments of the mimicking portfolios. This approach would be laborious and would require computing rolling windows of co-moments. But there is a procedure for simplifying equation (25). Kraus and Lintzenberger (1976)[16] have shown that a three-moment CAPM is consistent with the following quadratic form:

$$R_i - R_0 = \alpha_0 + \alpha_1 (R_m - R_0) + \alpha_2 (R_m - R_0)^2 \qquad (26)$$

and consequently a n-moment CAPM can be written as:

[16] See also L'Habitant (2004) on this point, chap. 8. Sometimes, the higher order moments are expressed in deviations from the mean. According to this formulation, the four-moment CAPM would be: $E(R_i) - R_0 = \beta_1[E(R_m) - R_0] + \beta_2[R_m - E(R_m)]^2 + \beta_3[R_m - E(R_m)]^3$.

$$R_i - R_0 = \alpha_0 + \alpha_1 (R_m - R_0) + \alpha_2 (R_m - R_0)^2 + \alpha_3 (R_m - R_0)^3 + ... + \alpha_{n-1} (R_m - R_0)^{n-1}$$

(27)

A test on α_2 is a test on skewness preferences in asset pricing and a test on α_3, a test on kurtosis preferences, and so on. The higher moments are consequently powers of returns in this approach. We therefore use a financial theory, the n-moment CAPM, to give an object to the method of Dagenais and Dagenais (1997) for correcting errors-in-variables. Let us return to the variable *SMB*, which we want to correct for the problem of errors-in-variables. In the first pass of the regression, this variable will be regressed on:

$$SM\hat{B}_t = f\left(F_{it-1}, F_{it-1}^2, F_{it-1}^3, ..., F_{it-1}^5, ...\right) \qquad (28)$$

where F_i are the variables in the RHS of the equation of Fama and French (equation 9) including *SMB*. They stand for the higher moments of these variables. F_{it-1}^2 stands for the skewness of factor F_i; F_{it-1}^3, for its kurtosis, and so on.The variables appearing on the RHS of equation (28) will serve as instrumental variables in the first pass of the Hausman test, as explained in the following section.

3.3 Hausman specification test and errors in variables

To detect errors in variables in our sample of hedge funds, we could use the original Hausman *h* test[17]. To explain this test, let us suppose the following classical model:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (29)$$

[17] On the Hausman test, see : Hausman (1978), Wu (1973), MacKinnon (1992), Johnston and Dinardo (1997) and Pindyck and Rubinfeld (1998). A very good presentation of the version of the Hausman test using an artificial regression in the context of correction of errors in variables may be found in Pindyck and Rubinfeld (1998). This presentation is done for one explanatory variable.

with **Y** a *(n x 1)* vector representing the dependent variable; **X**, a (*n x k*) matrix of the explicative variables; **β**, a (*k x 1*) vector of the estimates of the parameters and $\boldsymbol{\varepsilon} \sim iid(0, \sigma^2)$.

Hausman compares two sets of estimates of the parameters vector, say, $\boldsymbol{\beta}_{\mathbf{OLS}}$, the least-squares estimator (OLS), an $\boldsymbol{\beta}_{\mathbf{A}}$, and alternative estimator which can take a variety of forms but which for our purposes is the instrumental variables estimator which we designate by $\boldsymbol{\beta}_{\mathbf{IV}}$. The hypotheses to test are H_0, being in our case the absence of errors in variables and H_1, being the presence of errors in variables. The vector of estimates $\boldsymbol{\beta}_{\mathbf{IV}}$ is consistent under both H_0 and H_1 but $\boldsymbol{\beta}_{\mathbf{OLS}}$ is consistent under H_0 but inconsistent under H_1. Under H_0, $\boldsymbol{\beta}_{\mathbf{IV}}$ is indeed less efficient than $\boldsymbol{\beta}_{\mathbf{OLS}}$.

Hausman wants to verify if "endogeneity" of some variables[18], the variables measured with errors in our case, has any significant effect on the estimation of the vector of parameters. To do so, he defines the following vector of contrasts: $\boldsymbol{\beta}_{\mathbf{IV}} - \boldsymbol{\beta}_{\mathbf{OLS}}$. The test statistic may be written as follows:

$$h = \left(\hat{\boldsymbol{\beta}}_{\mathbf{IV}} - \hat{\boldsymbol{\beta}}_{\mathbf{OLS}}\right)^T \left[\mathbf{Var}\left(\hat{\boldsymbol{\beta}}_{\mathbf{IV}}\right) - \mathbf{Var}\left(\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}\right)\right]^{-1} \left(\hat{\boldsymbol{\beta}}_{\mathbf{IV}} - \hat{\boldsymbol{\beta}}_{\mathbf{OLS}}\right) \sim \chi^2(g) \qquad (30)$$

with $\mathbf{Var}\left(\hat{\boldsymbol{\beta}}_{\mathbf{IV}}\right)$ and $\mathbf{Var}\left(\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}\right)$ being consistent estimates of the covariance matrices of $\hat{\boldsymbol{\beta}}_{\mathbf{IV}}$ and $\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}$. g is the number of potentially endogenous regressors, that is the variables

[18] Therefore, the Hausman test is an orthogonality test, that is it aims to verify if plim (1/T) X'ε = 0 in large samples. But, to paraphrase Spencer and Berk (1981), the alternative hypothesis of this test is broad and diffuse. Absence of orthogonality might be caused by numerous phenomena: omission of relevant explanatory variables, errors in variables, inappropriate aggregation over time, simultaneity and incorrect functional form. But, according to these authors, it is true of all existing tests of the orthogonality assumption and it seems that the resolution of this problem requires some a priori information. For our purpose, this a priori information is potential presence of errors in variables. We want to verify if the endogeneity of some variables, the variables measured with errors, has any significant effect on the estimation of the parameters.

measured with errors in our case. H_0 will be rejected if the p-value of this test is less than α, with α being the critical threshold of the test, say 5%.

According to MacKinnon (1992), this test might run into difficulties if the matrix $\left[\mathbf{Var}\left(\hat{\boldsymbol{\beta}}_{\mathbf{IV}}\right) - \mathbf{Var}\left(\hat{\boldsymbol{\beta}}_{\mathbf{OLS}}\right)\right]$ which weights the vector of contrasts is not positive definite. Fortunately, there is an alternative way to do the Hausman test which is much easier. This test goes as follows.

Assume a two variables linear model:

$$y_t = \beta_0 + \beta_1 x_{1t}^* + \beta_2 x_{2t}^* + \varepsilon_t \qquad (31)$$

with $\varepsilon \sim N\left(0, \sigma^2\right)$.

The variables x_{1t}^* and x_{2t}^* [19] are observed with errors, that is:

$$x_{1t} = x_{1t}^* + \upsilon_{1t} \qquad (32)$$

$$x_{2t} = x_{2t}^* + \upsilon_{2t} \qquad (33)$$

with x_{1t} and x_{2t}, the corresponding observed variables which are measured with errors. By substituting equations (32) and (33) in equation (31), we have:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t^* \qquad (34)$$

with $\varepsilon_t^* = \varepsilon_t - \beta_1 \upsilon_1 - \beta_2 \upsilon_2$. As seen before, estimating coefficients of equation (34) by the OLS method gives way to biased and inconsistent coefficients because the explanatory variables are correlated with the innovation.

[19] As it is standard in econometrics, we use the asterisks for unobserved variables.

Consistent estimators can be found if we can identify an instruments vector $\mathbf{z_t}$ which is correlated with every explanatory variable but not with the innovation of equation (31). Then we regress these two explanatory variables on $\mathbf{z_t}$. We have:

$$x_{1t} = \hat{x}_{1t} + \hat{w}_{1t} = \hat{\gamma}_1 \mathbf{z_t} + \hat{w}_{1t} \qquad (35)$$

$$x_{2t} = \hat{x}_{2t} + \hat{w}_2 t = \hat{\gamma}_2 \mathbf{z_t} + \hat{w}_2 \qquad (36)$$

with $\hat{x}_{it}$, the value of x_{it} estimated with the vector of instruments and $\hat{w}_{it}$, the residuals of the regression of x_{it} on $\hat{x}_{it}$. Substituting equations (35) and (36) into equation (34), we have[20]:

$$y_t = \beta_0 + \beta_1 \hat{x}_{1t} + \beta_2 \hat{x}_{2t} + \beta_1 \hat{w}_{1t} + \beta_2 \hat{w}_{2t} + \varepsilon_t^* \qquad (37)$$

The explanatory variables of this equation are, on the one hand, the estimated values of x_{1t} and x_{2t}, obtained by regressing these two variables on the vector of instruments $\mathbf{z_t}$, and on the other hand, the respective residuals of these regressions. Equation (37) is therefore an augmented version of equation (34), which might be qualified of auxiliary or artificial regression.

We can show that:

$$p\lim \left[\frac{\sum \hat{w}_{1t} \varepsilon_t^*}{N} \right] = p\lim \left[\frac{-\beta_1 \sum x_{1t} \upsilon_{1t}}{N} \right] = -\beta_1 \sigma_{\upsilon_1}^2 \qquad (38)$$

and the same for w_{2t}. When there is no measurement error, $\sigma_{\upsilon_1}^2 = 0$ and OLS gives way to a consistent estimator for the parameter of $\hat{w}_{1t}$ in equation (37), that is β_1. When there are measurement errors, $\sigma_{\upsilon_1}^2 \neq 0$ and therefore this estimator is not consistent.

20 Racicot (2003) applies this approach to the market model. He postulates that the t test issued from the new variable $\hat{w}$ is distributed asymptotically as the normal distribution. According to Pindyck and Rubinfeld (1998), this test is adequate. Racicot (2003) also postulates in this context that the new model resulting from the addition of the artificial variables may be considered as a new model by itself, so we have a new alpha for this model. We follow a similar approach in this article.

We can therefore write the following test to detect the presence of errors in variables. As we do not know a priori if there are errors in variables, we replace the coefficients of $\hat{w}_{1t}$ and $\hat{w}_{2t}$ in equation (37) by γ_1 and γ_2 . We have:

$$y_t = \beta_0 + \beta_1 \hat{x}_{1t} + \beta_2 \hat{x}_{2t} + \gamma_1 \hat{w}_{1t} + \gamma_2 \hat{w}_{2t} + \varepsilon_t^* \quad (39)$$

But following equations (35) and (36) $\hat{x}_{1t} = x_{1t} - \hat{w}_{1t}$ and $\hat{x}_{2t} = x_{2t} - \hat{w}_{2t}$. We can therefore rewrite equation (39) as follows:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + (\gamma_1 - \beta_1)\hat{w}_{1t} + (\gamma_2 - \beta_2)\hat{w}_{2t} + \varepsilon_t^* \quad (40)$$

If there is no measurement error for both variables x_{1t} and x_{2t}, then $\gamma_1 = \beta_1$ and $\gamma_2 = \beta_2$. If there are measurement errors, $\gamma_i \neq \beta_i$ and the coefficients of the residuals terms w_{it} will not be zero.

There is more information which we can draw from equation (40). Indeed, if the estimated coefficient $(\gamma_i - \beta_i)$ is significantly positive, that indicates that the estimated coefficient of the corresponding explanatory variable x_{it} is overstated in the OLS run. Therefore, the estimated coefficient for this variable will decrease in equation (40). On the other hand, if the estimated coefficient $(\gamma_i - \beta_i)$ is significantly negative, that indicates that the estimated coefficient of the corresponding explanatory variable x_{it} is understated in the OLS run. Therefore, the estimated coefficient for this variable will increase in equation (40). These effects of errors in variables produced by equation (40) are very informative. In the next section, we will transpose these results to the F&F model.

We must notice that the coefficients β_i estimated by the equation (40) are identical to those ones produced by a two-stage least squares (TSLS) procedure using the same instruments. Equation (40) is therefore another way to do a TSLS. But in view of the useful information produced by equation (40), this equation opens the doors to new financial models. We will therefore prefer this formulation to that one represented by TSLS to estimate the augmented F&F model. And we thus have a new empirical formulation for the F&F model.

We therefore proceed as follows to test for errors in variables. First, we regress the observed explanatory variables x_{it} on the instruments vector to obtain the residuals $\hat{w}_{it}$. Then, we regress y_t on the observed explanatory variables x_{it} and on the residuals $\hat{w}_{it}$. This is an auxiliary or artificial regression. If the coefficient of the residuals of an explanatory variable is significantly different from 0, we may conclude that there is a measurement error on this explanatory variable. We may use the Wald test (F test) to see if the whole set of $(\gamma_i - \beta_i)$ coefficients is significantly different from zero.

We can generalize the preceding procedure to the case of k explanatory variables which are potentially suffering from the problem of errors in variables. Let $\mathbf{X}$ be a *(n x k)* matrix of explanatory variables which are potentially suffering from the disease of errors in variables and let $\mathbf{Z}$ be a *(n x s)* matrix of instruments ($s > k$). To perform the Hausman test based on an artificial regression, we first regress $\mathbf{X}$ on $\mathbf{Z}$ to obtain $\hat{\mathbf{X}}$, that is:

$$\hat{\mathbf{X}} = \mathbf{Z}\hat{\boldsymbol{\theta}} = \mathbf{Z}(\mathbf{Z'Z})^{-1}\mathbf{Z'X} = \mathbf{P_Z X} \qquad (41)$$

where $\mathbf{P_Z}$ is the "predicted value maker". Having performed this regression, we compute the matrix of residuals $\hat{\mathbf{w}}$:

$$\hat{\mathbf{w}} = \mathbf{X} - \hat{\mathbf{X}} = \mathbf{X} - \mathbf{P_Z X} = (\mathbf{I} - \mathbf{P_Z})\mathbf{X} \qquad (42)$$

Then we perform the following artificial regression:

$$\mathbf{y} = \boldsymbol{\beta}\mathbf{X} + \boldsymbol{\lambda}\hat{\mathbf{w}} \qquad (43)$$

An F test on the $\boldsymbol{\lambda}$ coefficients will indicate if they are significant as a group. A t test on individual coefficients will indicate if the corresponding β is understated or overstated, as discussed previously.

The vector of $\boldsymbol{\beta}$ estimated by equation (43) is identical to the TSLS estimates, that is:

$$\boldsymbol{\beta} = \boldsymbol{\beta}_{\mathbf{IV}} = (\mathbf{X'P_Z X})^{-1}\mathbf{X'P_Z y} \qquad (44)$$

To detect errors in variables in the augmented F&F model, we will run two sets of regressions. First, we will run the OLS one, that is:

$$R_{pt} - R_{ft} = \alpha + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 UMD_t + \varepsilon_t \quad (45)$$

Then, we will run the following artificial regression explained previously:

$$R_{pt} - R_{ft} = \alpha^* + \beta_1^*(R_{mt} - R_{ft}) + \beta_2^* SMB_t + \beta_3^* HML_t + \beta_4^* UMD_t + \sum_{i=1}^{4}\varphi_i \hat{w}_{it} + \varepsilon_t^* \quad (46)$$

The estimated coefficients φ_i will allow detecting errors in variables and their signs will indicate if the corresponding variable is overstated or understated in the OLS regression.

As said previously, the β^* estimated by equation (46) are equivalent to the TSLS estimates. But we prefer equation (46) because it gives more information on the problem of errors in variables. Equation (46) is thus our new empirical version of the augmented F&F

model. The φ_i are really factors of correction of the risk exposure of a Fund to the i[th] factor of risk. If φ_i is positive, that means that the exposure to the i[th] risk factor is overstated in the OLS regression. The β associated to this factor will thus decrease in the artificial regression. And vice-versa if φ_i is negative. Moreover, according to our previous developments, we expect a high positive correlation between $\left(\hat{\beta}_i - \hat{\beta}_i^*\right)$, that is the estimated error on the coefficient of factor i, and $\hat{\varphi}_i$, the estimated coefficient of the corresponding artificial variable $\left(\hat{w}_i\right)$.

4. Empirical results and analysis

Our sample of hedge funds returns comprises the monthly returns of twenty Greenwich-Van US American hedge funds indexes classified by categories or groups of categories[21]. The appendix gives the enumeration of these funds with their chosen symbol. The observation period runs from January 1995 to November 2005, for a total of 131 observations. The risk factors which appear in the F&F equation, - that is the market risk premium and the three mimicking portfolios: *SMB*, *HML* and *UMD*, - are for their part drawn from the French's website[22].

We can get a first glance at our sample of hedge funds by looking at table 1, which reports the descriptive statistics of the "average" fund. At 11.6%, the annualized mean return of this sample is quite high and its standard deviation is relatively moderate. But it is well-known that standard deviation is an acceptable measure of risk only if risk is small or

---

[21] These data can be found on the following website: http://www.hedgefund.com. Capocci (2004) gives an exhaustive directory of the hedge funds websites.

[22] The e-mail of the French's website is : http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

under very special conditions. Otherwise, we must consider the higher moments of the distributions of excess returns. The skewness of the average fund is not very different from 0 but the mean level of kurtosis, at 6.4, is quite high, the level of kurtosis associated with a normal distribution being 3. Obviously, a large number of equity oriented hedge funds strategies exhibit payoffs which look like those of a short position in a put option on the market index[23], and therefore bears significant left-tail risk, risk that is ignored by the commonly used mean-variance framework. That means that rare events are more frequent than in a normal distribution and that nonlinearities of payoffs are very present. Incidentally, there are 18 funds over 20 which have a level of kurtosis exceeding three and 17 which have non-gaussian returns according to the Jarque-Bera test.

Table 1 *Descriptive statistics of the sample of twenty hedge funds returns**

Statistics	Average	Std. Dev
Mean yield	0.0097	0.003
Median yield	0.0098	0.004
Maximum yield	0.1066	
Minimum yield	-0.0879	
Skewness	0.0959	0.817
Kurtosis	6.3812	3.304
Jarque-Bera	133.048	284.206
p-value	0.042	0.121
Number of funds with kurtosis > 3	18	
Number of funds with non-gaussian yields	17	

We must now define the instruments necessary to perform the Hausman test. Table 2 gives the correlations of the F&F factors with themselves and their chosen instruments from January 1995 to November 2005. Besides the instruments discussed before, we add other

[23] A short put option carries the risk of rare but large losses.

macroeconomic variables: the monthly and annual American inflation rate, *IPC_MENS* and *IPC_ANN*, and the monthly and annual growth rate of the American industrial production, *PROD_MENS* and *PROD_ANN*.

Table 2 *Correlation between the four factors and their instruments*

	RM_RF	SMB	HML	UMD
RM_RF	1.00			
SMB	0.23	1.00		
HML	-0.30	-0.37	1.00	
UMD	-0.23	0.04	-0.62	1.00
RM_RF(-1)	0.03	0.17	0.16	-0.11
$RM\_RF(-1)^2$	0.08	0.07	-0.02	-0.07
$RM\_RF(-1)^3$	-0.07	0.10	0.10	-0.02
$RM\_RF(-1)^4$	0.11	0.01	-0.05	-0.05
$RM\_RF(-1)^5$	-0.10	0.04	0.08	0.01
HML(-1)	-0.16	0.06	0.11	0.01
$HML(-1)^2$	-0.06	-0.23	0.17	-0.01
$HML(-1)^3$	-0.14	0.18	-0.06	0.09
$HML(-1)^4$	0.01	-0.29	0.19	-0.10
$HML(-1)^5$	-0.12	0.24	-0.11	0.11
SMB(-1)	0.08	-0.02	-0.04	0.11
$SMB(-1)^2$	-0.01	-0.19	0.20	-0.10
$SMB(-1)^3$	0.12	-0.17	0.03	0.01
$SMB(-1)^4$	0.04	-0.29	0.24	-0.15
$SMB(-1)^5$	0.13	-0.23	0.06	-0.06
UMD(-1)	0.03	-0.15	-0.07	-0.07
$UMD(-1)^2$	-0.20	-0.08	0.18	0.07
$UMD(-1)^3$	0.19	-0.06	-0.01	-0.22
$UMD(-1)^4$	-0.21	-0.12	0.16	0.15
$UMD(-1)^5$	0.22	-0.01	-0.05	-0.22
IPC_ANN	-0.03	0.00	0.10	-0.03
IPC_MENS	-0.11	-0.06	-0.01	0.06
PROD_ANN	0.12	-0.23	-0.11	0.05
PROD_MENS	-0.02	-0.04	-0.07	0.02

According to table 2, the F&F factors are more or less correlated with classic instruments like the first lag of the factor or with the macroeconomic variables. On the side of the macroeconomic variables, we observe that their correlation with the risk factors located at the head of the columns is quite low. Only the annual growth of the industrial production has a moderate correlation with three of these factors[24].

On the side of the risk factors of equation (45), we noticed before that there are many nonlinearities in these mimicking portfolios, which are similar to portfolios of hedge funds. We also noticed before that these nonlinearities might be captured by the higher moments of these factors. Corroborating this assumption, we observe at table 2 that the risk factors are generally more correlated or cross-correlated with the higher moments of the first lag of a risk factor than to the first lag itself. For example, the market risk premium is more related to the higher moments of *UMD(-1)* than to *UMD(-1)* itself. Indeed, the correlation between *RM_RF* and *UMD(-1)* is quite low: 0.03, but it is equal to 0.22 for $UMD(-1)^5$, that is the higher moment of *UMD(-1)* of order 5. The same is true for the factor *SMB* and the higher moments of *SMB(-1)*. Consequently, higher moments of lagged variables may constitute quite good instruments.

Before discussing the results, let us notice that we performed a Wald test over the whole set of the four coefficients associated to the risk factors of equation (46). For this test, the null hypothesis H_0 was:

$$\varphi_1 = \varphi_2 = \varphi_3 = \varphi_4 = 0$$

If this hypothesis is not rejected, we cannot detect errors in variables for the four factors considered as a group. This hypothesis was not rejected for any fund and consequently it seems that there is no problem of errors in variables in our sample of hedge funds when we

[24] We notice that the correlation between *SMB* and the annual growth of industrial production is quite high, being equal to -0.23. This correlation suggests that the *SMB* factor would have a positive impact on returns in recessions.

consider the four factors as a group. But as we will see, individual *t* tests reveal that in some cases, the bias caused by the presence of errors in variables may be quite serious.

Table 3 and 4 give a first grasp of the estimations of equations (45) and (46) over the whole sample of our hedge funds. These tables contain the total count of significant coefficients at a confidence level of 95% and the mean level of the *t* statistics computed over the twenty funds for both estimation methods. Table 3 indicates that the constant and the coefficients of regression of the risk premium and *SMB* are significant for the majority of the funds of our sample. The variable *UMD* is significant for approximately 50% of the sample while the variable *HML* is more problematic, being significant for only a minority of funds. Table 4, which produces the average of the *t* statistics for the constant and the four factors computed over the twenty funds, confirms those preliminary observations. Let us notice that equation (46) produces less significant coefficients than equation (45), even if this difference is not quite high. It is well known that the coefficients estimated by TSLS tend to have a larger variance than those resulting from a corresponding OLS regression. As seen previously, the artificial regression given by equation (46) is equivalent to TSLS method. On the other hand, the average R^2 of the two estimation methods are quite comparable, being 0.55 for the OLS estimation and 0.54 for the artificial regression.

Table 3 *Count of significant coefficients for the constant and the four factors for the OLS and the artificial Hausman regression**

	c	Rm-Rf	SMB	HML	UMD	Line total
OLS	13	18	17	7	9	64
Hausman	15	16	17	5	7	60

* For a variable, the table gives the number of significant coefficients for $\alpha = 5\%$. There are twenty yield indexes of hedge funds in the sample. The estimation goes from January 1995 to November 2005.

Table 4 *Average level of the t statistics for the constant and the four factors for the OLS and the artificial Hausman regression**

	c	Rm-Rf	SMB	HML	UMD	Line avera ge
OLS	3.16	8.62	5.54	1.72	1.99	4.21
Hausman	2.68	8.64	5.18	1.53	1.80	3.97

* The average is computed for a sample of twenty hedge funds indexes.

At table 5, we find the mean levels of the coefficients estimated by the OLS method and by the artificial regression for the whole set of funds. The mean level of the constant, which corresponds to the alpha, is practically the same for both regression methods. We notice that the beta seems to be overstated by the OLS regression which produces biased coefficients if there are errors in variables. The average beta resulting from the OLS estimation is 0.25 and 0.23 for the artificial one. Otherwise, the impact of *SMB* tends to be understated by the OLS regression, its average coefficient being 0.17 in the OLS regression and 0.21 in the artificial one. The average incidence of the *UMD* factor is quite similar for the two estimation methods but, as we will see, this situation hides quite a high dispersion. Finally, the influence of *HML* is quite low in both estimation methods, being moderately overstated by the OLS regression.

Table 5 *Mean level of the estimated coefficients for the constant and the four factors for the OLS and the Hausman artificial regression*

	c	Rm-Rf	SMB	HML	UMD
OLS	0.0036	0.2479	0.1709	0.0321	0.0764
Hausman	0.0039	0.2346	0.2057	-0.0126	0.0745
Spread	-0.0003	0.0133	-0.0348	0.0447	0.0019

Because the F&F model is a purely empirical one, there is no theory on the signs of the four factors of this model, except perhaps for the market index whose coefficient is generally positive according to the CAPM. But if we consider the three mimicking portfolios *SMB*, *HML* and *UMD* as factors of risk, it is reasonable to expect generally a positive sign for the estimated coefficients of these factors. Really, a fund can short a factor of risk, as it might short the market index. A hedge fund which makes short selling will have a negative beta. Another which short sells the mimicking portfolio *SMB*, say, will have a negative coefficient for this factor. But this behaviour seems to be the exception rather than the rule. Risk factors in returns equations might generally be preceded by positive signs.

Table 6 gives information on this matter for our sample. For both estimation methods, the estimated coefficients of the four factors have positive signs for the majority of funds. That comforts us in our expectations of considering the mimicking portfolios as risk factors and not as market anomalies as they were viewed in the past.

Table 6 *Count of positive signs for the constant and the four factors for both estimation methods.*

	c	Rm-Rf	SMB	HML	UMD	Line total
OLS	17	18	19	15	20	89
Hausman	17	18	18	13	17	83

To conclude these preliminary observations, we may say that the problem of errors in variables does not seem to be very important for the group of funds surveyed in this study. But average behaviour might hide a great dispersion at the individual level. As we will see in the following paragraphs, the biases caused by errors in variables might be severe for some funds.

In our sample, it was the factor *HML* which seems suffering most from errors in variables. The artificial regression reveals that its residuals were significant at the 95% confidence level for four funds. For both SMB and UMD, the residuals were significant for three funds.

We will gain a better grasp of our estimations if we look at the individuals results. Table 7 gives the estimated betas, that is the coefficient of the risk premium (R_m-R_f), for each fund and for the two estimation methods. This table also gives the corresponding coefficient φ in the artificial regression (equation 46). This coefficient is in bold when it is significant at the 95% confidence level[25].

We notice at table 7 that for five funds, the spread between the OLS and the Hausman estimate, which is the measurement error on this coefficient, is quite high. For these funds, it is overstated four times and understated one time by the OLS method. When the beta of a fund is overstated, table 7 reveals that the artificial coefficient *φ* associated to the residuals of the market risk premium in equation (46) is positive, as it must be. In the case of the "aggressive growth" fund designated by *ag*, the beta is greatly understated by the OLS

[25] It seems reasonable to consider also the φ which are significant at the 10% level.

regression and the corresponding φ is therefore negative. Let us notice that the correlation between the spreads (errors) column in our tables and the corresponding φ column is 0.98. Therefore, the association between the level of the measurement error on a coefficient and the corresponding level of φ is positive and almost perfect[26]. Moreover, the regression of the error of the ith fund, that is the spread $\left(\hat{\beta}_{i,OLS} - \hat{\beta}_{i,HAUS}\right)$, on the corresponding artificial variable $\hat{\varphi}_i$, gives the following result:

$$\left(\hat{\beta}_{i,OLS} - \hat{\beta}_{i,HAUS}\right) = -0.005 + 0{,}70\hat{\varphi}_i$$
$$(-2.45) \quad (23.96)$$

For this regression, the adjusted R^2 is 0.96. Therefore, the level of the artificial variable is a very good indicator of the measurement error of the corresponding coefficient. It thus gives precious information on the measurement error and it is why we prefer the Hausman version of the Fama and French model to the equivalent TSLS one. We will not repeat this regression for the other factors of risk of the Fama and French model because the results are very similar to those obtained for the risk premium.

[26] Notice that usually, the spread is less than φ.

Table *7 Spread (error) between the OLS and the Hausman beta for each fund of the study**

	OLS beta	Hausman beta	Error	φ
emn	0.095	0.126	-0.031	-0.039
ed	0.264	0.201	**0.063**	0.085
ds	0.144	0.082	**0.063**	0.091
ss	0.322	0.275	0.048	0.063
mna	0.082	0.069	0.014	0.018
shs	-0.908	-0.922	0.014	0.043
va	0.515	0.494	0.020	0.022
f	-0.007	-0.129	**0.122**	0.189
macro	0.303	0.349	-0.046	-0.036
mt	0.325	0.291	0.034	0.051
dtg	0.155	0.088	**0.067**	0.110
mng	0.145	0.118	0.026	0.035
lsg	0.431	0.431	0.000	0.000
ag	0.650	0.744	**-0.094**	-0.125
oi	0.368	0.360	0.008	0.013
specg	0.525	0.532	-0.008	0.009
em	0.678	0.686	-0.008	0.021
inc	0.138	0.131	0.007	0.007
msi	0.388	0.427	-0.039	-0.050
hf	0.345	0.338	0.007	0.012

** The spread (error) is the difference between the OLS coefficient and the corresponding Hausman coefficient. . For each spread, we produce the coefficient (φ) of the corresponding artificial variable. The Funds having the highest spreads in absolute value are bold-faced. The coefficient φ is bold-faced when significant. Notice the high positive relation between the error and its associated φ.*

Table 8 gives the same information as table 7 concerning the *SMB* factor. There are five funds where the measurement error is high. Concerning these funds, we notice that the coefficient of this variable is understated four times and overstated one time by the OLS regression. When the impact of *SMB* is understated, the corresponding φ is negative and when it is overstated, the corresponding φ is positive, as it might be. The *SMB* coefficient is particularly understated for the emerging markets fund (*em*) and its associated φ is therefore negative and very high in absolute value.

Table 8 *Spread (error) between the OLS and the Hausman SMB estimated coefficients for each fund of the study**

	SMB coef. OLS	SMB coef. Haus.	Spread	φ
emn	0.123	0.158	-0.035	-0.036
ed	0.225	0.252	-0.027	-0.052
ds	0.188	0.232	-0.044	-0.074
ss	0.250	0.279	-0.028	-0.056
mna	0.062	0.066	-0.004	-0.012
shs	-0.422	-0.359	**-0.063**	-0.077
va	0.331	0.272	0.058	0.101
f	0.021	-0.053	**0.073**	0.149
macro	0.199	0.381	**-0.182**	**-0.238**
mt	0.102	0.070	0.032	0.055
dtg	0.087	0.080	0.007	0.039
mng	0.133	0.150	-0.017	-0.028
lsg	0.288	0.248	0.040	0.074
ag	0.359	0.368	-0.010	-0.025
oi	0.318	0.256	**0.062**	0.140
specg	0.276	0.459	**-0.183**	**-0.293**
em	0.396	0.659	**-0.264**	**-0.423**
inc	0.105	0.145	-0.040	-0.068
msi	0.159	0.214	-0.055	-0.079
hf	0.220	0.236	-0.016	-0.021

** The spread (error) is the difference between the OLS coefficient and the corresponding Hausman coefficient. . For each spread, we produce the coefficient (φ) of the corresponding artificial variable. The Funds having the highest spreads in absolute value are bold-faced. This coefficient φ is bold-faced when significant. Notice the high positive relation between the error and its associated φ.*

Table 9 reveals that the estimation of the *HML* coefficient is quite problematic, the *φ* coefficients for this variable being very high for some funds. The OLS regression tends to overstate greatly the impact of this variable on the returns of four funds. In this study, this overstatement is explained by a possible serious problem of errors in variables.

Table 9 *Spread (error) between the OLS and the Hausman HML estimated coefficients for each fund of the study*

	HML coef. OLS	**HML coef. Haus.**	**Spread**	**φ**
emn	0.016	0.021	-0.004	0.009
ed	0.074	0.065	0.009	-0.006
ds	0.071	0.033	0.039	0.052
ss	0.084	0.093	-0.009	-0.037
mna	0.004	0.014	-0.010	-0.023
shs	0.508	0.320	**0.187**	**0.334**
va	0.029	0.019	0.010	0.024
f	0.165	-0.137	**0.302**	**0.548**
macro	0.025	-0.080	**0.106**	0.217
mt	-0.175	-0.249	**0.073**	0.131
dtg	0.037	-0.161	**0.198**	**0.362**
mng	0.025	0.026	-0.002	-0.009
lsg	-0.046	-0.072	0.026	0.058
ag	-0.297	-0.250	-0.047	-0.069
oi	-0.086	-0.185	**0.099**	**0.208**
specg	0.072	0.098	-0.026	-0.047
em	0.082	0.096	-0.014	-0.021
inc	0.086	0.105	-0.020	-0.046
msi	-0.018	0.012	-0.030	-0.046
hf	-0.014	-0.020	0.006	0.013

** The spread (error) is the difference between the OLS coefficient and the corresponding Hausman coefficient. . For each spread, we produce the coefficient (φ) of the corresponding artificial variable. The Funds having the highest spreads in absolute value are bold-faced. This coefficient φ is bold-faced when significant. Notice the high positive relation between the error and its associated φ.*

Finally, according to table 10, the overstatement of the impact of the factor *UMD* is quite high in three cases. It is moderately understated in two cases. Moreover, if we correlate the OLS and Hausman estimated coefficients for this factor, we obtain only 0.23. The corresponding coefficients for the three other factors, - R_m-R_F, *SMB* and *HML* -, are respectively 0.99, 0.91 and 0.80. Consequently, there is a great divergence between the results obtained by the OLS and Hausman regressions for the *UMD* factor, which is not observed for the other ones.

Table 10 *Spread (error) between the OLS and the Hausman UMD estimated coefficients for each fund of the study**

	UMD coef. OLS	**UMD coef. Haus.**	**Spread**	**φ**
emn	0.101	0.160	-0.060	-0.079
ed	0.086	0.083	0.003	-0.001
ds	0.019	0.009	0.009	0.016
ss	0.115	0.120	-0.005	-0.015
mna	0.029	0.036	-0.007	-0.014
shs	0.047	-0.103	**0.150**	**0.256**
va	0.045	0.054	-0.009	-0.016
f	0.187	-0.027	**0.213**	**0.352**
macro	0.150	0.204	-0.055	-0.037
mt	0.038	-0.027	**0.065**	0.103
dtg	0.139	0.024	**0.115**	**0.199**
mng	0.060	0.078	-0.017	-0.027
lsg	0.064	0.054	0.010	0.017
ag	0.059	0.082	-0.023	-0.023
oi	0.112	0.102	0.010	0.033
specg	0.075	0.176	**-0.100**	-0.136
em	0.082	0.215	**-0.133**	-0.173
inc	0.071	0.090	-0.019	-0.031
msi	0.021	0.073	-0.052	-0.072
hf	0.064	0.085	-0.021	-0.028

** The spread (error) is the difference between the OLS coefficient and the corresponding Hausman coefficient. . For each spread, we produce the coefficient (φ) of the corresponding artificial variable. The Funds having the highest spreads in absolute value are bold-faced. This coefficient φ is bold-faced when significant. Notice the high positive relation between the error and its associated φ.*

At tables 11 and 12, we find respectively the OLS and artificial regressions for the funds which seem to suffer the most from a problem of error in variables. The adjusted R^2 varies greatly for these funds, having a low of 0.02 for the futures funds and a high of 0.85 for the short selling funds. The artificial regressions (table 12) were performed by using the higher moments of the predetermined variables of our model as instruments.

Table 11 *Estimation of Funds plagued with errors in variables by the OLS method*

Fund	c	Rm-Rf	SMB	HML	UMD	adj R^2	DW
DTG	0.004	0.155	0.087	0.037	0.139	0.129	1.98
	1.80	*2.94*	*1.55*	*0.51*	*2.56*		
EM	-0.001	0.678	0.396	0.082	0.082	0.456	1.49
	-0.32	*7.38*	*4.06*	*0.65*	*0.87*		
F	0.005	-0.007	0.021	0.165	0.187	0.023	1.86
	1.61	*-0.08*	*0.23*	*1.41*	*2.13*		
MACRO	-0.001	0.303	0.199	0.025	0.150	0.361	1.93
	-0.33	*5.41*	*3.33*	*0.33*	*2.60*		
OI	0.006	0.368	0.318	-0.086	0.112	0.726	1.56
	3.76	*9.60*	*7.81*	*-1.64*	*2.85*		
SHS	0.002	-0.908	-0.422	0.508	0.047	0.854	1.91
	0.71	*-14.73*	*-6.44*	*6.01*	*0.74*		
SPECG	0.000	0.525	0.276	0.072	0.075	0.507	1.54
	-0.09	*8.43*	*4.17*	*0.85*	*1.18*		

*** In italic, we find the t statistics of the corresponding estimated coefficient.**

Table 12 *Estimation of Funds plagued with errors in variables by the Hausman method with higher moments as instruments**

Fund	c	Rm-Rf	SMB	HML	UMD	Resid_rm-rf	Resid_smb	Resid_hml	Resid_umd	adj R^2	DW
DTG	0.006	0.088	0.080	-0.161	0.024	0.110	0.039	0.362	0.199	0.148	2.03
	2.38	*0.89*	*0.91*	*-1.46*	*0.26*	*0.94*	*0.33*	*2.47*	*1.75*		
EM	-0.003	0.686	0.659	0.096	0.215	0.021	-0.423	-0.021	-0.174	0.476	1.66
	-0.63	*4.03*	*4.34*	*0.51*	*1.36*	*0.11*	*-2.13*	*-0.08*	*-0.89*		
F	0.009	-0.129	-0.053	-0.137	-0.027	0.190	0.149	0.548	0.352	0.035	1.85
	2.33	*-0.80*	*-0.37*	*-0.77*	*-0.18*	*1.00*	*0.79*	*2.30*	*1.90*		
MACRO	-0.002	0.349	0.381	-0.080	0.204	-0.036	-0.238	0.217	-0.037	0.399	2.04
	-0.70	*3.42*	*4.19*	*-0.71*	*2.16*	*-0.30*	*-2.00*	*1.44*	*-0.31*		
OI	0.006	0.360	0.256	-0.185	0.102	0.013	0.140	0.208	0.033	0.725	1.55
	3.51	*5.00*	*3.99*	*-2.32*	*1.52*	*0.15*	*1.67*	*1.95*	*0.40*		
SHS	0.003	-0.922	-0.359	0.320	-0.103	0.043	-0.077	0.334	0.256	0.854	2.02
	1.16	*-7.97*	*-3.49*	*2.49*	*-0.96*	*0.31*	*-0.57*	*1.95*	*1.93*		
SPECG	-0.002	0.532	0.459	0.098	0.176	0.009	-0.294	-0.047	-0.136	0.522	1.70
	-0.52	*4.57*	*4.42*	*0.76*	*1.62*	*0.06*	*-2.16*	*-0.27*	*-1.01*		

* The variable resid is the residuals of the regression of the attached variable, say *umd*, on the set of instruments.

Table 12 reports our preferred empirical version of the F&F model. It is a new version which has not been produced yet. Therefore, it is not only an artificial regression but a new empirical model. It includes the estimated coefficients of the residuals which produce a great amount of information about the correction of the exposures of the funds to the risk factors as we have seen before. This correcting process is required because of the problem of errors in variables. In table 13, we notice that the coefficients of the regressions performed in table 12 are identical to the coefficients obtained by a TSLS using the same instruments as equation (46). But this last estimation form is less informative than that one produced by the artificial regression which constitutes a new empirical model.

Table 13 *Estimation of Funds plagued with errors in variables by the TSLS method with higher moments as instruments*

Fund	c	Rm-Rf	SMB	HML	UMD	adj R^2	DW
DTG	0.006	0.088	0.080	-0.161	0.024	0.064	2.02
	2.27	*0.85*	*0.87*	*-1.39*	*0.25*		
EM	-0.003	0.686	0.659	0.096	0.215	0.416	1.76
	-0.59	*3.76*	*4.05*	*0.47*	*1.27*		
F	0.009	-0.129	-0.053	-0.137	-0.027	0	1.89
	2.25	*-0.77*	*-0.35*	*-0.74*	*-0.17*		
MACRO	-0.002	0.349	0.381	-0.080	0.204	0.161	1.98
	-0.60	*2.90*	*3.54*	*-0.60*	*1.82*		
OI	0.006	0.360	0.256	-0.185	0.102	0.700	1.60
	3.35	*4.78*	*3.82*	*-2.22*	*1.46*		
SHS	0.003	-0.922	-0.359	0.320	-0.103	0.833	1.87
	1.09	*-7.45*	*-3.26*	*2.33*	*-0.90*		
SPECG	-0.002	0.532	0.459	0.098	0.176	0.453	1.78
	-0.49	*4.27*	*4.13*	*0.71*	*1.52*		

To notice the relevance of higher moments as instruments, we repeated the estimations appearing in table 13 but without using higher moments. The instruments used are therefore classic predetermined variables, that is exogenous variables or lagged endogenous or exogenous variables not powered. Table 14 gives this estimation. The results are obviously bad compared to those obtained by TSLS using higher moments as instruments. Higher moments are therefore good candidates for instruments. They take into account the nonlinearities of the payoffs of hedge funds which are neglected by classic instruments.

Table 14 *Estimation of Funds plagued with errors in variables by the TSLS method with classic instruments*

Fund	c	Rm-Rf	SMB	HML	UMD	adj R^2	DW
DTG	0.008	0.018	0.064	-0.285	-0.114	0	2.07
	2.12	*0.10*	*0.29*	*-0.96*	*-0.54*		
EM	-0.011	0.849	1.014	0.483	0.811	0	2.05
	-1.35	*2.08*	*2.12*	*0.74*	*1.74*		
F	0.014	-0.231	-0.053	-0.486	-0.372	0	2.00
	2.17	*-0.72*	*-0.14*	*-0.96*	*-1.02*		
MACRO	-0.003	0.280	0.242	0.144	0.400	0.185	2.05
	-0.77	*1.38*	*1.02*	*0.44*	*1.73*		
OI	0.005	0.394	0.116	0.120	0.236	0.557	1.64
	1.46	*2.52*	*0.63*	*0.48*	*1.32*		
SHS	0.003	-0.930	-0.324	0.224	-0.032	0.818	1.80
	0.64	*-4.23*	*-1.26*	*0.64*	*-0.13*		
SPECG	-0.007	0.636	0.637	0.422	0.595	0.125	2.03
	-1.34	*2.36*	*2.02*	*0.98*	*1.94*		

5. Summary and conclusion

In this paper, we have explored a new empirical version of the Fama and French model conceived for detecting and simultaneously removing the problem of errors in variables. This model includes variables for correcting exposure to risk factors. This correction process allows removing the errors in variables disease from the estimated coefficients of the risk factors. These correction factors, based on the artificial regression of the Hausman specification test, are interesting because they give information on the understatement of overstatement of risk due to the problem of errors in variables.

Our empirical version of the model of Fama and French includes another innovation. It uses as instruments the higher moments of the distributions of the returns of the mimicking portfolios to purge the F&F model from its errors in variables. That these instruments are highly related to the risk factors is perfectly normal because the returns of the mimicking portfolios incorporate many nonlinearities. These nonlinearities cannot be captured by a classic CAPM or APT model which postulates a linear relation between the returns to be explained and their risk factors. The F&F model is also linear. But the risk factors *SMB*, *HML* and *UMD* take care for the presence of these nonlinearities. The good relation between these factors and the chosen instruments in this article, which are the higher moments of those variables, tends to demonstrate these points. Moreover, a TSLS conducted on the returns of hedge funds using classic instruments performed very poorly relatively to a TSLS using higher moments as instruments.

There are by now many articles which put in question the F&F model[27]. Our article is instead an argument in favour of this model. The three factors, *SMB*, *HML* and *UMD*, of the F&F model have their place in the explanation of risk, because they package the many nonlinearities which are present in the distribution of returns. They are not market anomalies as they were viewed in the past. They are instead the "reservoirs" of moments and co-moments risks.

Our study also puts light on the idiosyncratic behaviour of hedge funds. The beta of hedge funds is not generally high, not exceeding 0.25. The short sellers are an exception on that matter, having a mean beta near -1. For hedge funds, the three factors *SMB*, *HML* and *UMD* are really factors of risk, their signs being mostly significantly positive in the regressions of the funds returns on those factors. After the market risk premium, it is the factor *SMB* which exerts the most prominent positive impact on the returns of the hedge

[27] For example, Chung and al. (2001) observed that *SMB* and *HML* become not significant or less significant when moments or co-moments are taken into account. F&F factors might simply be good proxies for the higher moments of the return distribution. But our regression technique, which considers co-moments as instruments, suggests that they capture these co-moments but also other influences.

funds of our sample. Really, this factor allows identifying rich and cheap securities, which is one of the most important dimensions of the hedge funds industry.

Globally, the problem of errors in variables does not seem to be too serious in our sample of hedge funds. This problem gives way to an overstatement of the impact of the risk premium and to an understatement of the influence of *SMB*. While the impact of *HML* on hedge funds returns is quite low, the incidence of *UMD* is more problematic. The correlation between the coefficients corrected and uncorrected for the problem of errors in variables is low in the case of *UMD*.

Investors and regulating institutions must be more and more aware of the consequences of measurements errors in terms of false signals and financial losses. Concerning governance, financial institutions must formulate directives to their portfolio and financial managers which will incite them to take into account measurement errors in their decision process and to improve the financial data which they use and publish. The method which we propose in this article for correcting measurement errors can be transposed easily to accommodate the majority of situations involving measurement errors.

In summary, our new empirical version of the Fama and French model seems quite encouraging. One avenue of future research will be to use cumulants[28] instead of higher moments to define the instruments. In finance, we use generally moments to quantify risks but cumulants are quite promising on this matter. In his studies, Dagenais used cumulants and not higher moments. Future seems quite promising for risk analysis.

[28] On this subject, see: Malevergne and Sornette (2005).

Appendix

List of funds of the study

- *USEMN*: Equity market neutral
- *USED*: Event driven
- *USDS*: Distressed securities
- *USSS*: special situations
- *USMMNA*: Market neutral arbitrage
- *USSHS*: short selling index
- *USVA*: Value index
- *USF*: futures index
- *USMACRO*: Macro
- *USMT*: Market timing
- *USDTG*: Directional trading group
- *USMNG*: Market neutral group
- *USLSG*: long short group
- *USAG*: aggressive growth
- *USOI*: opportunity index
- *USSPECG*: special strategies
- *USEM*: emerging markets
- *USINC*: fixed income`
- *USMSI*: multi strategy index
- *USHF*: hedge funds index

References

Artzner, P., F. Delbaen, J. M. Eber and D. Heath, "Coherent measures of risk", Mathematical Finance, 9, 203-228, (1999).

Bertsimas, D., G.J. Lauprete and A. Samarov, "Shortfall as a risk measure : properties, optimization and applications", Journal of Economic Dynamics and Control, 28, 1353-1381, (2004).

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

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

Capocci, D. Introduction aux hedge funds. Economica, 2004.

Carhart, M.M., "On persistence in mutual fund performance", Journal of Finance, 52, 57-82, (1997).

Coën, A. and R. Théoret, "Vers une vision probabiliste du choix d'investissement: une application à la performance du secteur bancaire", Banque & Marchés, 72, septembre-octobre 2004, p. 32-43.

Chung, Y.P., H. Johnson and M.J. Schill, "Asset pricing when returns are nonnormal: Fama French factors vs. higher order systematic co-moments", working paper, (2001).

Dagenais, M.G., "Parameter Estimation in regression models with errors in the variables and autocorrelated disturbances", Journal of Econometrics, 64, 145-163, (1994).

Dagenais, M.G. and D.L. Dagenais, "Higher moment estimators for linear regression models with errors in the variables", Journal of Econometrics, 76, 193-221, (1997).

Durbin, J., "Errors in variables", International Statistical Review, 22, 23-32, (1954).

Fama, E.F. and K.R. French, "The cross-section of expected stock returns", Journal of Finance, 47, 427-465, (1992).

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

Fama, E.F. and K.R. French, "Industry costs of equity", Journal of Financial Economics, 43, 153-193, (1997).

Fama, E.F. and J.D. MacBeth, "Risk, Return, and Equilibrium : Empirical Tests", Journal of Political Economy, 81, 607-636, (1973).

Ferson, W.E. and D.H. Locke, "Estimating the Cost of Capital Through Time : An Analysis of the Sources of Error", Management Science, 44, 485-500, (1998).

Friend, I. and R. Westerfield, "Co-skewness and capital asset pricing", Journal of Finance, 35, 897-914, (1980).

Harvey, C.R. and A. Siddique, "Conditional skewness in asset pricing tests", Journal of Finance, 55, 1263-1295, (2000).

Hausman, J.A., "Specification Tests in Econometrics", Econometrica, 46, 1251-1271, (1978).

Johnston, J. and J. Dinardo.Econometric Methods.fourth edition.McGraw Hill, 1997.

Huang, C. F. and R.H. Litzenberger. Foundations for Financial Economics. Prentice Hall, Englewood Cliffs, 1988.

Ingersoll Jr., J. Theory of Financial Decision Making. Rowman and Littlefield Publishers, 1987.

Jegadeesh, N. and S. Titman, "Returns to buying winners and selling losers: implications for stock market efficiency", Journal of Finance, 48, 65-91, (1993).

Judge, G.G., W.E. Griffiths, R. Carter Hill, H. Lütkepohl and T.C. Lee. The Theory and Practice of Econometrics. Wiley, 1985.

Kandel S. and R. Stambaugh, "Portfolio Inefficiency and the Cross-Section of Expected Returns", Journal of Finance, 50, p. 157-184, (1995).

Kraus, A. and R. Litzenberger, "Skewness preference and the valuation of risk assets", Journal of Finance, 31, 1085-1100, (1976).

Leibowitz, M. L. and R.D. Henriksson, "Portfolio optimization with shortfall constraints: A confidence-limit approach to maintaining downside risk", Financial Analysts Journal, (March-April 1989).

Levy, H., "Stochastic dominance and expected utility: survey and analysis", Management Science, 38, 555-574, (1992).

Lhabitant, F.S. Hedge Funds . Quantitative Insights. John Wiley & Sons, 2004.

Lim, K.-G, "A New Test of the Three-Moment Capital Asset Pricing Model", Journal of Financial and Quantitative Analysis, 24, 205-216, (1989).

MacKinnon, J.G., "Model Specification Tests and Artificial Regressions", Journal of Economic Litterature, 30, 1, 102-146,(1992),.

Malevergne, Y. and Sornette, D., "Higher moments portfolio theory, capitalizing on behavioral anomalies of stock markets", Journal of Portfolio Management, 31, 49-55, (2005).

Pal, M., "Consistent moment estimators of regression coefficients in the presence of errors in variables", Journal of Econometrics, 14, 349-364, (1980).

Pastor L. and R. Stambaugh, "Costs of Equity Capital and Model Mispricing", Journal of Finance, 54, 67-121, (1999).

Pindyck, R.S. and D.L. Rubinfeld. Econometric Models and Economic Forecasts. fourth edition. Irwin-McGraw-Hill, 1998.

Racicot, F.-É., Trois essais sur l'analyse des données économiques et financières, Ph.D. thesis, chap.3, University of Quebec (Montreal) (2003).

Racicot, F-.É., Estimation et tests en présence d'erreurs de mesure sur les variables explicatives, M.Sc. thesis (economics), University of Montreal (1993).

Rubinstein, M., "The fundamental theorem of parameter-preference security valuation", Journal of Financial and Quantitative Analysis, 8, 61-69, (1973).

Samuelson, P. A., "The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments", Review of Economic Studies, 37, 537-542, (1970).

Scott, R.C. and P.A. Hovarth, "On the direction of preference for moments of higher order than the variance", Journal of Finance, 35, 915-919, (1980).

Sears, R. S. and K.C.J. Wei, "Asset pricing, higher moments, and the market risk premium: A note", Journal of Finance, 40, 1251-1253, (1985).

Shanken, J., "On the Estimation of Beta-Pricing Models", Review of Financial Studies, 5, 1-33, (1992).

Spencer, D.E. and K.N. Berk, "A Limited Information Specification Test", Econometrica, 49, 4, 1079-1085, (1981).

Taleb, N. Dynamic Hedging. John Wiley, 1997.

Wu, D. , "Alternative tests of independence between stochastic regressors and disturbances", Econometrica, 41, 733-750, (1973).