Testing the new Fama and French factors with illiquidity: A panel data investigation*

François-Eric Racicot[1], William F. Rentz[2], Raymond Théoret[3]

Abstract

We investigate the new Fama-French (FF, 2015, 2016) five factors augmented with a well-known illiquidity measure (Pástor and Stambaugh, 2003), using an innovative GMM robust instrumental variables estimator casted in a panel data framework. When using OLS, the augmented FF model seems to have explanatory power regarding the FF 12-sector returns. However, our panel data framework suggests that the only consistently significant factor is the market risk factor. Nevertheless, depending on the technique we use, we find that measurement errors may be the cause of this result, thus providing some empirical evidence in support of the new FF five-factor approach. As robustness checks, we also experiment with other liquidity measures – like the Amihud (2002) ratio and the term-spread – and bond-oriented factors. Across our 12 portfolios, the results are largely unchanged. We also apply our extended model to managed portfolios – i.e., hedge fund portfolios. The returns of hedge fund strategies seem more responsive to the augmented FF five-factor model that includes illiquidity measures, especially when accounting for the subprime crisis. There is also evidence that the new FF factors embed illiquidity.

Keywords: Fama-French (2015) five factors; illiquidity; hedge funds; panel data; GMM; robust instruments.

* We would like to thank the Editor, Franck Moraux, and three anonymous referees for their insightful comments and suggestions, which greatly improved this paper. We are also very grateful to our colleague Alfred Kahl and Professor Yakov Amihud for their helpful suggestions, and to discussant Sofia Priazhkina and the seminar participants of the IAES conference held in Montreal in October 2017. We thank Conrad McCallum for editorial assistance. Financial support from the IPAG Business School is gratefully acknowledged.

1 Telfer School of Management, University of Ottawa, 55 Laurier Avenue East, Ottawa, Ontario, Canada, K1N 6N5. racicot@telfer.uottawa.ca Affiliate Research Fellow, IPAG Business School, Paris, France, Corresponding author.

2 Telfer School of Management, University of Ottawa, 55 Laurier Avenue East, Ottawa, Ontario, Canada, K1N 6N5. rentz@telfer.uottawa.ca

3 École des sciences de la gestion, Université du Québec à Montréal, 315 Ste-Catherine est, R-2915, Montréal, Québec, Canada, H2X 3X2. theoret.raymond@uqam.ca Chaire d'information financière et organisationnelle, ESG-UQAM

1. Introduction

The formulation and estimation of an asset pricing model may take many forms. Sharpe (1964), Lintner (1965), and Mossin (1966) developed what is known as the capital asset pricing model (CAPM). Jensen (1968) is credited with the development of alpha (α), which he used to investigate the performance of mutual funds via the CAPM. Black (1972) extended the theory of the CAPM to what is known as the zero-beta CAPM. Collectively, these ideas form the basis of modern portfolio management and equity valuation, and they have been widely implemented over the last 50 years by academics and practitioners. Over those same decades, there have been many attempts to extend the CAPM to a dynamic framework, such as the intertemporal CAPM (Merton, 1973) and the consumption CAPM proposed by Hansen and Singleton (1982, 1984). Later Mehra and Prescott (1985) studied the consumption CAPM to further investigate what is known as the equity premium puzzle[4]. However, another category of asset pricing models, namely, the factorial models group, focuses more on finding the empirical factors which may explain stock returns. Among them, the best known is the Fama and French three-factor model that takes into account market risk premium, size and value factors (Fama and French 1992, 1993). In a recent extension of their model, Fama and French (2015) develop a five-factor model which adds two new factors to their original: investment and profitability.

When estimating a factorial model, many issues must be addressed: identification, specification and measurement errors, multicollinearity, endogeneity, and heterogeneity, among others. Cochrane (2011, 2017) worries about the "zoo of factors". In line with his concern, Harvey *et al.* (2016) compiled a list of 316 variables discussed in the literature. Harvey (2017) and Mclean and Pontiff (2016) argue that many of these factors may be spurious. In addition to potential specification errors, some of the explanatory variables may be highly interrelated. Cochrane (1991, 2011) used a modified version of Tobin's (1969) Q theory to establish a link between asset prices and investment. Cochrane's link can be modified to express a relation between expected returns and investment[5]. Since Cochrane's Q is approximated by the market/book ratio, the FF value (*HML*) and investment

4 For a summary of these developments, see Campbell *et. al.* (1997) or Cochrane (2005, 2008). Note that the equity premium puzzle has also been depicted by Hansen and Singleton (1984).

5 See Hou, Xue, and Zhang (2015).

factors (*CMA*) are likely to be highly related. The Pástor and Stambaugh (PS, 2003) illiquidity risk factor is an example of what might be considered a generated variable because it is a parameter obtained from a regression, in this case relating stock return to its trading volume[6]. However, it is statistically indistinguishable from its original version. Although the OLS estimator may remain unbiased, constructed variables will likely increase the variance of the OLS estimator according to Pagan (1984, 1986)[7] and Shanken (1992)[8]. Thus, the resulting inference may be biased. Furthermore, Adrian et al. (2017) explain that endogeneity issues can bias traditional liquidity measures. In particular, a decrease in illiquidity – as measured, for example, by bid-ask spreads – may be not associated with an effective decrease in market illiquidity, but only with a transfer of illiquidity risk from market makers to investors. Endogeneity biases can therefore plague the OLS method.

A powerful solution to the problems of specification and measurement errors is the generalized method of moments (GMM) developed by Hansen (1982). However, the usefulness of this method is questionable when weak instruments are at play[9]. Nelson and Starz (1990a,b), Bound, Jaeger, and Baker (1995), and Hahn and Hausman (2003) show that the two-stage least squares (2SLS) estimator is inconsistent when instrumental variables are weak. Dagenais and Dagenais (1994, 1997) developed a method that creates instruments with greater robustness. These robust instruments are generated by using a Bayesian averaging approach, as recommended by Theil and Goldberger (1961). These instruments are based upon the higher moments of the explanatory variables. The approach has two principal features, namely: i) it is parsimonious in the sense that it requires minimal computational power, and; ii) it essentially minimizes a distance (d) measure. More precisely, our purpose is to propose a parsimonious approach to tackle measurement errors or the endogeneity of the explanatory variables based on the generalized method of moments. This method has the virtue of freeing the analyst from having to choose between one instrument and another. As the literature has demonstrated at length (e.g., Anderson and

6 Note that the portfolio version of this variable is the one we use. This portfolio is long on illiquid stocks and short on liquid stocks.

7 Pagan and Ullah (1988), however, find that when estimating a regression using a generated variance regressor (e.g. from GARCH), the resulting estimator is biased.

8 In the two-pass regression approach, the second step uses estimated betas. These betas may be considered as generated variables. Shanken (1992) shows that the standard error from this two-step approach should be corrected. This result appears analogous to Pagan (1984, 1986).

9 An instrument is weak when it is only slightly correlated with the explanatory endogenous variables (Greene, 2018).

Rubin, 1949, 1950; Dufour, 2003; Nelson and Startz, 1990a, b; Hahn and Hausman, 2002, 2003; Stock and Yogo, 2005; Hausman, Stock and Yogo, 2005; Olea and Pflueger, 2013), weak instruments present a perverse problem. Choosing the wrong instruments may result in increasing the problem one hoped to confront in the first place. That is, it may transform the estimator into a biased and inconsistent one. For example, it may bias the two-stage least squares estimator towards the OLS. Also, it will render the basic framework for statistical inference inappropriate (Nelson and Startz, 1990a,b; Hahn and Hausman, 2003). The robust instruments we propose deal with these issues.

We first estimate and test the new Fama and French factors model (FF factors, 2015, 2016) and an extended version that accounts for illiquidity using the Fama and French 12-sector portfolio with a panel data framework. We focus on the variables of the original FF three-factor model, on the profitability and investment factors recently introduced by FF (2015, 2016), and on the PS (2003) illiquidity factor. These risk factors appear to be the most widely recognized factors explaining the cost of equity[10]. Moreover, all of them may be represented by portfolios. If these portfolio risk factors do not span the space of the unknown state factors, then specification errors will occur. Furthermore, as noted by FF (2015, p. 2), the book/market ratio "is a noisy proxy for expected return", which implies potential measurement errors. To estimate our factorial model, we use a panel data framework allowing for specification/measurement errors. As argued by Cochrane (2005), the Fama and McBeth (1973) two-step procedure[11] to estimate an asset pricing model – which is widely used to test such models – is equivalent to a pooled regression. This is our motivation for using this approach. Note that, in the Fama and McBeth (1973) procedure, there is a bias in the estimation process for standard errors caused by the two-pass regression approach. Shanken (1992) proposes a way to correct this bias. However, as Cochrane (2005) points out, one way to confront all of these problems is to use the more powerful GMM approach. Based on this distance notion, our approach is to generalize the implicit OLS features of the Dagenais and Dagenais (1997) method to a more powerful generalized method of moments estimation framework which we refer to as

10 Pinto *et. al.* (2015, chap. 2) discuss the required knowledge and models covered by the CFA certification. Among these models are the Fama-French and the Pástor and Stambaugh ones.

11 More precisely, Fama and McBeth (1973) introduce a process for estimating cross-sectional regressions and standard errors correcting for cross-sectional correlation in a panel data framework.

GMM_d. One of the virtues of our proposed GMM_d panel[12] data framework is a systematic treatment of the previous specification errors, including the problem of measurement errors. For this reason, we cast our GMM_d in a panel framework to account for the cross-sectional and time-series variations. To the best of our knowledge, we are the first to rely on this setting to estimate an asset pricing model. The studies that are closest to ours, Li et al. (2017) and González and Jareño (2018) – they also perform the estimation of the FF five factors on sector portfolios – rely on maximum likelihood or on quantile regressions. Note that these authors use averages to report sector portfolios' exposure to the various factors. We maintain that it is always more efficient to compute such averages by relying on panel regressions with common factors[13].

The pooling method offers many advantages for the estimation of our empirical asset pricing. First, it allows to define common factors in order to focus on the most important specific factor affecting sector returns: systematic risk. In particular, we allow not only the Jensen (1968) α performance measure to vary across sectors but also the β systematic risk measure. This technique allows us to isolate the CAPM idiosyncratic risk associated with each sector and to measure its *average* exposure to the common factors. This generalization also enables us (i) to evaluate the significance of the FF five factors; and (ii) to compare this model to a six-factor model that incorporates the Pástor and Stambaugh (PS, 2003) illiquidity risk factor. This information is important for an investor wanting to know if the 12-sector portfolio we rely on makes it possible to diversify the risk associated with the common factors. Second, this empirical framework allows us to generate some new insights on the effects of unobserved heterogeneity in panel (pooled) data models that may compound measurement errors if not tackled properly. Heterogeneity may be dealt with by resorting to fixed or random effects. Fixed effects are used when there is evidence of a correlation between the factors of the model and the unobserved or omitted variables (Greene, 2018). In a portfolio model framework, the omitted variables may be macroeconomic and financial shocks impacting the factors, which are essentially the returns of mimicking portfolios spanning the space of factors. Random effects are used when there is evidence of no correlation between factors and unobserved (omitted) variables. The selection of fixed or random effects

12 In this article, we use the terms pooled estimation and panel estimation interchangeably.

13 It is also easier to compute the significance of this average by relying on pooled data.

is an empirical matter, the Hausman (1978) and Hausman et al. (2005) tests being designed to operate this choice, so we consider in our paper two GMM_d estimators: one with fixed effects and one with random effects in which all parameters may vary randomly. Third, by regrouping observations, we can rely on cross-sectional weights to tackle the heteroskedasticity linked to different sectors and to seemingly unrelated regressions (SUR) to control for the interactions between the sector innovations.

Our investigation reveals important endogeneity issues. Indeed, our results show that, using OLS in panel data for fixed or random effects models, most of the new FF risk factors are significant, although the PS illiquidity is not. When using the GMM_d approach that accounts for the non-linearity/non-normality of the data, we obtain a different picture, *viz.*, the only strongly significant risk factor is the market factor, and the illiquidity factor is weakly significant for the pooled GMM_d (fixed effects). We also find significant measurement errors for the new FF investment factor and for the PS illiquidity factor using our modified artificial regression Hausman (1978) test, referred to as the $Haus_d$ test. As a robustness check, we investigate with other liquidity measures – i.e., the Amihud (2002) ratio and the term-spread – in order to examine the multidimensional character of illiquidity. We also introduce two bond-oriented factors – i.e, the monthly change in the ten-year constant maturity yield and the credit market factor representing the monthly change in the Moody's Baa yield less ten-year Treasury constant maturity yield – to allow for the impact of macroeconomic or financial shocks[14]. Adding these factors leads to more significant results in the GMM_d estimation, especially for the value and profitability factors and the PS one. The model is also tested by dividing the sample into recessionary and expansionary periods in order to study the asymmetric behavior of sectors. We note that some sectors particularly exposed to the business cycle, such as durables, manufactured goods and financial institutions, become more risky during the subprime crisis, whereas others manage risk better, such as non durable goods and the health sector.

In the last section, we transpose our extended factor model to hedge funds. Indeed, the behavior of managed portfolios may be quite different from the one of sector portfolios. Risk management is performed by firms for sectors and by skilful portfolio managers for managed portfolios. It is

14 We thank an anonymous referee for this suggestion. These bond-oriented variables are especially used in hedge fund studies (e.g., Fung and Hsieh, 2004; Racicot and Théoret, 2016b).

also easier to diversify a managed portfolio than a sectorial one. Thus, the universe of hedge funds is a particularly relevant terrain for investigating the issues covered in this paper. Our results show that hedge funds are more active in the management of their risk exposures than industry sectors – especially when accounting for the subprime crisis.

The remainder of this article is organized as follows. Section 2 introduces an extension of the basic fixed and random effects panel data framework in the context of errors in variables in the new Fama-French (FF, 2015) five-factor model and the six-factor model that includes Pástor-Stambaugh (PS, 2003) illiquidity factor. Section 3 casts the GMM_d approach into a panel data framework and discusses both our $Haus_d$ test for measurement errors and a modified Hausman test for random versus fixed effects models. Section 4 interprets select descriptive statistics from the data obtained and presents our empirical results. Section 5 presents our robustness check. Section 6 transposes our extended model to hedge funds, while Section 7 provides our conclusions and final considerations.

2. The panel data framework for testing the new Fama-French five factors

2.1. The five- and six- risk factor models[15]

FF (2015, 2016) introduce a five-factor model[16],

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it} \tag{1}$$

where R_{it} is the return on portfolio *i*[17], R_{Ft} is the risk-free rate, $R_{Mt} - R_{FT}$ is the market risk premium, SMB_t is the size factor, HML_t is the value factor, RMW_t is the new profitability factor, and CMA_t is the new investment factor. Our goal here is to test the new FF factors and Pástor and Stambaugh (PS, 2003) liquidity factor.

15 Note that some authors have recently considered other factors besides illiquidity, like the momentum factor (see Barillas and Shanken, 2015). This factor is, however, not new and is well documented in the literature (see Carhart, 1997).

16 The data for the five FF factors and sector returns are available from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

17 Note that the results may be quite different if we estimate equation (1) on individual stocks or on factor (mimicking) portfolios as in Fama and French (2015). For more detail, see González and Jareño (2018).

Pástor and Stambaugh (PS, 2003) introduce a liquidity factor LIQ_t in the original Fama and French (1992) three-factor model. The PS liquidity factor is a constructed variable. LIQ_t is an average of the stock $\hat{\gamma}_{it}$ obtained from regression (2).

$$r_{id+1t} - r_{md+1t} = \theta_{it} + \varphi_{it} r_{idt} + \gamma_{it} sign(r_{idt} - r_{mdt}) v_{idt} + \varepsilon_{id+1t} \quad (2)$$

where r_{idt} is the return of stock i on day d in month t and v_{idt} is the dollar trading volume of stock i on day d in month t. Pagan (1984, 1986)[18] shows that generated variables may increase the variance of the OLS estimator, but the estimator remains unbiased. In this paper, we compare (1) with an augmented version of this equation that includes liquidity as a sixth factor – i.e., the PS tradable liquidity factor [19].

2.2. Fixed Effects Model

The fixed effects panel data framework including the *LIQ* factor may be written in a stacked vector format for the 12 FF sectors, which are described in Appendix 1.

$$Y = R - R_F = \sum_{i=1}^{12} \alpha_i D_i + \sum_{i=1}^{12} \beta_i D_i (R_M - R_F)$$

$$+ s\,SMB + h\,HML + r\,RMW + c\,CMA + l\,LIQ + e \quad (3)$$

$Y' = (R_{11} - R_{F1}, \cdots, R_{1T} - R_{FT}, \cdots, \cdots, R_{12,1} - R_{F1}, \cdots, R_{12,\mathrm{T}} - R_{FT})$ represents the transpose of the stacked vector Y of excess returns for each sector. $D_i' = (0, \cdots, 0, \cdots, 1, \cdots, 1, 0, \cdots, 0)$ is the transpose of the stacked dummy variable, which is 0 everywhere except for the T observations for sector i. α_i is the Jensen (1968) performance measure for sector i.

$(R_M - R_F)' = (R_{M1} - R_{F1}, \cdots, R_{MT} - R_{FT}, \cdots, R_{\mathrm{M}1} - R_{F1}, \cdots, R_{\mathrm{MT}} - R_{FT})$ is the transpose of the stacked vector of excess market returns. That is, the excess market returns are stacked 12 times, once for each sector. β_i is the sector i CAPM systematic risk beta. The other explanatory variables are

18 See also Pagan and Ullah (1988) and Shanken (1992) for more information on related matters. Adrian et al. (2017) explain that traditional liquidity measures are endogenous therefore creating an endogeneity bias when estimating via OLS.

19 The *LIQ* factor is available on Pástor's website http://faculty.chicagobooth.edu/lubos.pastor/research/ .We use the tradable *LIQ* factor and multiply it by 100 to put it in percentage form.

similarly defined. The coefficients of these other variables are 12-sector pooled coefficients. ***e*** is the stacked vector of error terms. Note that (3) is initially estimated via OLS – i.e., our least squares dummy variables (LSVD) approach – and then using the generalized method of moments (GMM) with robust instruments.

Defining the market risk premium as a specific factor of our sector portfolios – the most important factor impacting their returns – allows isolating the idiosyncratic risk of each sector in the sense of the CAPM. This idiosyncratic risk is regressed on the common factors in order to measure the *average* exposure of sectors to these factors and its degree of significance. We thus adopt a *holistic* approach that is well-suited to institutional investors – like banks, pension funds and university endowments – who want to diversify the idiosyncratic risk of their portfolio across sectors. As previously noted, similar studies seeking to measure the impact of the new FF five-factors on sector portfolios estimated the exposures of each portfolio to these factors separately, and then reported the average exposures of sectors (Li et al., 2017; González and Jareño, 2018). We contend that our method is more efficient since it computes these average exposures and their degree of significance simultaneously.

2.3. Generalized random effects model

We rely on a generalized version of the fixed effects model in (3). Thus, for comparison purposes, we use a generalized version of the random effects model where all parameters are allowed to vary randomly[20].

We can rewrite (1) in the format of the general random effects model as follows

$$R_i - R_F = [\,\mathbf{1}\; R_M - R_F\; SMB\; HML\; RMW\; CMA] \begin{bmatrix} a + v_{1i} \\ b + v_{2i} \\ s + v_{3i} \\ h + v_{4i} \\ r + v_{5i} \\ c + v_{6i} \end{bmatrix} + e_i \quad (4)$$

20 The model can be written as follows: $y_i = X_i\pi_i + e_i$ with X_i, a matrix of observations of dimension T × k, $\pi_i = \pi + v_i$ where v_i is a vector of random effects of dimension k × 1 with $E(v_i|X_i) = 0$, $E(v_i v_i'|X_i) = \Gamma$ and e_i is a vector of random errors of dimension T × 1. Thus, we can see that the π_i for an FF sector is the result of a random process with mean vector π and covariance matrix Γ.

where R_i, R_F, R_M, SMB, HML, RMW, CMA, and e_i are vectors of dimension $T \times 1$ and v_{ki} (k=1,…,6) is a random variable. Note that we assume that all of the parameters are random, not just the intercept term and the coefficient for the excess market return factor. The random effects model simplifies considerably, assuming that there is no autocorrelation or cross-sectional correlations in e_i. Following Swamy (1970), we apply GLS to (4) and obtain

$$\hat{\pi} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}y = \sum_{i=1}^{12} W_i^* b_i \tag{5}$$

where $\hat{\pi}$ is a simple weighted average of the OLS b_i. Ω must be estimated in equation (5). The technical aspects related to the estimation of the W_i^* matrix appear in Appendix 2.

Note that when multiplying the two matrices of equation (4), each parameter becomes random, that is, from the constant term to the *CMA* coefficient. This is a compact way of rewriting the FF model into a generalized random coefficient model. In the simple random coefficient model, only the constant term would have been allowed to vary randomly.

3. GMM$_d$ and Hausman$_d$ panel data framework

3.1. GMM$_d$ fixed effects model

The GMM estimator $\hat{\theta}_{GMM_d}$ for estimating the fixed effects panel data regression models is given by (Arellano and Bond, 1991; Arellano and Bover, 1995; and Arellano, 2003).

$$\hat{\theta}_{GMM_d} = \left[\left(\sum_{i=1}^{N} X_i' d_i\right)\widehat{w^{-1}}\left(\sum_{i=1}^{N} d_i' X_i\right)\right]^{-1}\left[\left(\sum_{i=1}^{N} X_i' d_i\right)\widehat{w^{-1}}\left(\sum_{i=1}^{N} d_i' Y_i\right)\right] \tag{6}$$

where $d_i = x_i - \widehat{x_i}$ is a vector of robust "distance" instruments. These new instruments – the d "distance" instruments – can be computed using a matrix-weighted average by applying GLS to a combination of two robust estimators, namely the Durbin (1954) and Pal (1980) estimators. The technical aspects related to our distance instruments appear in Appendix 3.

3.2. Implementing the panel data fixed and random effects GMM_d approaches

To implement the GMM_d approach in a fixed effects panel data framework, the first step is to create the dummy variables for each sector. Next compute the robust instruments using the algorithm described in Appendix 3. Then calculate the GMM estimators in (6) using a HAC matrix with the newly computed robust instruments and the sector dummy instruments.

To implement the GMM_d estimator in the context of the generalized random effects model, one needs to simply substitute $\hat{\theta}_{GMM_d}$ given by (6) for b_i in (12, Appendix 2). Also, the least squares variance-covariance estimator $s_i^2\left(X_i{}'X_i\right)^{-1}$ should be replaced by $s_{GMM_d,i}^2\left(X_i{}'X_i\right)^{-1}$.

3.3. Modified Hausman artificial regression test

To test whether there are measurement errors, we rely on a *modified* Hausman (1978) artificial regression, which we refer to as $Haus_d$. Each variable in the original five-factor and six-factor models has a companion variable in $Haus_d$ with its own *t* statistic that indicates whether the original variable contains measurement errors.

To implement the $Haus_d$ artificial regression, one begins by estimating the following equation using OLS:

$$Y = X\beta + \hat{\omega}\varphi + e \tag{7}$$

It is a two-stage least squares (2SLS) estimator because $\hat{\omega}$ is also obtained by OLS. The equation (7) can be rewritten as

$$Y = X\hat{\beta}_{2SLS} + \hat{\omega}\varphi + e^* \tag{8}$$

where $\varphi = \psi - \beta$ measures the under/over estimation of the OLS benchmark estimator[21].

In (8), $\hat{\omega}$ is a matrix of residuals of the regression of each explanatory variable on the instrument set. The notation $\hat{\omega}$ is commonly used in

21 See Pindyck and Rubinfeld (1998, pp. 195-197). An estimator is understated (underestimated) by OLS if its TSLS value is greater than its OLS one. It is overstated (overestimated) in the opposite case.

Hausman artificial regressions. It is equivalent to the $d_i = x_i - \widehat{x_i}$ residual that emphasizes the idea of a "distance" variable that is discussed above.

3.4. Hausman test for random vs. fixed effects

The standard approach to test whether the fixed effects model should be retained over the random effects model is usually performed via a Hausman (1978) specification test. This verifies that the quadratic distance between the fixed effects estimator is significantly different from the random effects one. The standard Hausman test can be written as follows[22]

$$H = \left(b_{FE} - \widehat{\beta}_{RE}\right)' \left[V_{FE} - V_{RE}\right]^{-1} \left(b_{FE} - \widehat{\beta}_{RE}\right) \overset{a}{\sim} \chi^2_M \tag{9}$$

which is asymptotically distributed as chi-squared with *k* degrees of freedom. More precisely, we have *k* coefficients (i.e., *k* explanatory variables excluding the constant term) in the estimator vectors b_{FE} and $\widehat{\beta}_{RE}$ for the fixed effects and random effects models, where V_{RE} can be estimated using (14) (Appendix 2) via GMM_d while V_{FE} is obtained by first running GMM_d on (3) and $\widehat{V}_{FE} = \widehat{\sigma_e^2}(\ddot{X}'\ddot{X})^{-1}$. Note that if there were only one parameter in these vectors, the square root of statistic (9) would asymptotically follow a *t* statistic and, under certain assumptions, would in fact have a normal distribution asymptotically[23]

4. Data and empirical results

4.1. Data and descriptive statistics

Our sample is composed of monthly returns of 12 indices classified by FF industry sectors[24]. The observation period ranges from January 1968 to December 2016, for a total 588 months. The panel data framework yields 12 sectors × 588 monthly observations = 7,056 total observations. The FF risk factors are drawn from French's website. The PS liquidity factor is from Pástor's website.

22 See Racicot et al. (2018) for more details.

23 See Wooldridge (2002).

24 See Appendix 1 for the description of these 12 portfolios.

Tables 1 and 2 present the descriptive statistics of the dependent and independent variables, respectively.

Table 1. Fama and French 12 sectors 1968m01 - 2016m12 (%)

	Mean	Median	Max	Min	Std. Dev.	Skewness	Kurtosis	Jarque-Bera
1 Nodur	1.08	1.08	18.88	–21.03	4.34	–0.27	5.03	107.96
2 Durbl	0.85	0.83	42.63	–32.63	6.41	0.12	7.74	551.83
3 Manuf	0.95	1.16	21.08	–28.58	5.36	–0.49	5.62	191.78
4 Enrgy	1.02	0.93	24.56	–18.33	5.59	0.04	4.13	31.51
5 Chems	0.93	1.04	20.22	–24.59	4.69	–0.22	5.18	121.14
6 Buseq	0.89	0.83	20.76	–26.03	6.60	–0.19	4.24	40.90
7 Telcm	0.95	1.16	21.36	–16.36	4.74	–0.24	4.19	40.34
8 Utils	0.88	0.93	18.84	–12.65	4.10	–0.10	3.99	24.93
9 Shops	1.03	0.97	25.86	–28.23	5.26	–0.26	5.41	149.14
10 Hlth	1.04	1.09	29.52	–20.46	4.94	0.06	5.46	148.41
11 Money	1.02	1.36	21.10	–22.10	5.58	–0.41	4.59	78.02
12 Other	0.78	1.04	19.36	–29.26	5.48	–0.48	5.17	137.35
Panel data	0.95	1.05	42.63	–32.63	5.30	–0.20	5.65	2118.32

Notes: For each sector, there are 49 years × 12 months = 588 monthly observations for a total of 12 sectors × 588 = 7,056 for the panel data set.

For all sectors, note that the Jarque-Bera (*JB*, 1980) statistic is greater than 5.99, which is the critical value of the chi-square distribution at the 5% level for 2 degrees of freedom. Thus, we reject the null hypothesis of normality for all sector returns. This empirical finding is well-known in the literature[25]. Sector 6 (Business equipment) has the highest standard deviation of 6.60. On a standalone basis in the Markowitz (1959)[26] mean-variance framework, this would indicate that Business Equipment is the riskiest sector. However, in the higher-moments framework of Rubinstein (1973) and of Jurczenko and Maillet (2006), this sector has the third lowest kurtosis. This suggests that Business Equipment is not the most risky.

25 See Mandelbrot (1963, 1972) or Haug (2007).

26 Markowitz (2012) notes that the mean-variance model still works well in the presence of moderate amounts of skewness and kurtosis.

Nine of the 12 sectors show negative skewness, which is an indicator of downside risk. Only Sector 2-Durables, Sector-4 Energy, and Sector 10-Health have the desirable positive skewness, which is an indicator of strong upside potential.

Table 2. New Fama-French (2015, 2016) and Pástor-Stambaugh risk factors 1968m01 - 2016m12 (%)

	Mean	Median	Max	Min	Std. Dev.	Skewness	Kurtosis	Jarque-Bera
R_m-R_f	0.49	0.81	16.10	–23.24	4.54	–0.51	4.79	103.88
SMB	0.19	0.04	18.27	–14.85	3.06	0.40	6.37	294.37
HML	0.38	0.30	12.90	–11.10	2.90	0.06	4.96	94.75
RMW	0.26	0.28	13.51	–18.72	2.29	–0.32	15.16	3630.35
CMA	0.35	0.22	9.58	–6.88	2.02	0.33	4.58	71.59
LIQ	0.40	0.35	11.08	–12.89	3.37	–0.02	3.94	21.82

Notes: For each sector, there are 588 observations for a total of 7,056 for the panel data set. Here the panel data contain the 588 monthly observations, repeated for each of the 12 sectors.

In Table 2, the *JB* statistics are even more indicative of non-normality for the independent variables. Among the new FF risk factors, *RMW* has an extremely high *JB* statistic, and risk factor *CMA* has the lowest *JB* statistic. Nevertheless, at 71.59, the *CMA JB* statistic is still well above the 5.99 chi-square 5% cut-off value. Even the non-FF factor *LIQ* at 21.82 is above the cut-off. The values for all the risk factors indicate that extreme events occur far more frequently than with the normal distribution. This is a reflection of the kurtosis measuring well over the normal distribution value of 3 for each of these 6 risk factors. The highest kurtosis value is for the *RMW* risk factor at 15.16, being over 5 times the normal distribution value. Only the kurtosis and *JB* statistics for *RMW* fall outside the top end of the range of the kurtosis and *JB* values from Table 1 for the sector returns.

All these results support the logic of our proposed methodology, which uses higher moments (cumulants) as instruments for the GMM estimation process. Using OLS when such strong non-normality is present–in both the dependent and explanatory variables–may lead to wrong inferences.

4.2. Empirical results and analysis

The Fixed Effects estimation of the FF new five factors

Table 3a presents our estimation results for the new FF five factors using the least squares dummy variable method (LSDV) and GMM_d approaches for the fixed effects model.

For the FF five-factor OLS pooled model of the FF twelve-sector portfolio, the coefficients of all five factors are significant at 1% except for *SMB,* which is significant at 10%. This suggests strong support for the FF five factors. However, using GMM_d pooled, only the coefficient of the market factor is significant at 1% with the *RMW* coefficient significant at 10%. Note that the beta coefficients obtained by LSDV and GMM_d are highly correlated[27], suggesting that measurement errors are low for this factor. It is also worthwhile to point out that in no case, there is a strong disagreement about whether market betas are above or below one.

From an investment performance perspective, the Jensen performance measure is negative but not significantly for both the OLS and GMM_d pooled approaches for the FF twelve-sector portfolio. For OLS, the twelve-sector portfolio appears to be weighted towards firms that are small cap (*SMB*, 0.0245), have value (*HML*, 0.0761), show robust profitability (*RMW*, 0.1806), and follow a conservative investment policy (*CMA*, 0.1091). For GMM_d, the conclusions are the same. The $Haus_d$ test suggests significant measurement errors for the *RMW* factor ($\widehat{\omega}_{RMW} = 0.3248$, $t = 2.97$).

Two recent studies have been performed on portfolios of US industry sectors. González and Jareño (GJ, 2018) analyze 10 industry sectors drawn from Bloomberg over the period stretching from November 1989 to February 2014 while the study of Li et al. (LRZM, 2017) focus on the FF 48 US industry portfolio and stretches from July 1963 to January 2017. Consistent with our study, the FF five factors are significant in the paper by LRZM when using a maximum likelihood estimation (MLE). However, the *SMB* factor is not significant in the complete model of GJ that is estimated with quantile OLS regressions including the PS illiquidity factor, while it is negative in a restricted model specification including only the FF five factors. In term of average exposures, the market risk premium is the most important factor in the three studies (including ours), the average exposure being close

27 The correlation coefficient here is almost 97%.

Table 3a. Fixed Effects Model, LSDV vs. GMMd estimation methods for the FF five-factor model by FF 12 sectors

		c	R_m-R_f	*SMB*	*HML*	*RMW*	*CMA*	$\widehat{\omega}_{R_m-R_f}$	$\widehat{\omega}_{SMB}$	$\widehat{\omega}_{HML}$	$\widehat{\omega}_{RMW}$	$\widehat{\omega}_{CMA}$	$\overline{R}^2$	*DW*
Sector	**Fama-French (2015, 2016)**													
1 NoDur	LSDV	0.1497	0.8377											
	t-stat	*1.21*	*30.75****											
	GMM_d	0.1500	0.7710											
	t-stat	*1.10*	*5.93****											
2 Durbl	LSDV	–0.2591	1.1905											
	t-stat	*–2.10***	*43.70****											
	GMM_d	–0.3574	1.3235											
	t-stat	*–2.05*	*9.44****											
3 Manuf	LSDV	–0.1338	1.1547											
	t-stat	*–1.09*	*42.39****											
	GMM_d	–0.1851	1.1922											
	t-stat	*–1.68*	*15.05****											
4 Enrgy	LSDV	0.0800	0.8584											
	t-stat	*0.65*	*31.51****											
	GMM_d	0.0905	0.7709											
	t-stat	*0.48*	*8.25****											
5 Chems	LSDV	–0.0480	0.9365											
	t-stat	*–0.39*	*34.37****											
	GMM_d	–0.0097	0.7929											
	t-stat	*–0.07*	*4.90****											
6 BusEq	LSDV	–0.2738	1.3067											
	t-stat	*–2.22***	*47.96****											
	GMM_d	–0.4397	1.5764											
	t-stat	*–1.73*	*5.15****											
7 Telcm	LSDV	0.0123	0.8410											
	t-stat	*0.10*	*30.87****											
	GMM_d	–0.0524	0.9057											
	t-stat	*–0.24*	*4.34****											

8 Utils	LSDV	*0.0809*	*0.5673*											
	t-stat	0.66	20.82***											
	GMM_d	*0.1835*	*0.2934*											
	t-stat	1.02	2.46***											
9 Shops	LSDV	*–0.0127*	*1.0534*											
	t-stat	–0.10	38.66***											
	GMM_d	*–0.0675*	*1.0981*											
	t-stat	–0.49	11.08***											
10 Hlth	LSDV	*0.0837*	*0.8776*											
	t-stat	0.68	32.21***											
	GMM_d	*0.1445*	*0.6884*											
	t-stat	0.83	4.44***											
11 Money	LSDV	*–0.0598*	*1.1271*											
	t-stat	–0.49	41.37***											
	GMM_d	*–0.0716*	*1.0847*											
	t-stat	–0.48	10.31***											
12 Other	LSDV	*–0.3286*	*1.1855*											
	t-stat	–2.67***	43.52***											
	GMM_d	*–0.3893*	*1.2422*											
	t-stat	–3.45***	17.92***											
Pooled Model														
	OLS	–0.0591[a]	0.9947[a]	0.0245	0.0761	0.1806	0.1091						0.66	1.90
	t-stat	–1.53	106.72***	1.85*	4.21***	10.15***	3.99***							
	GMM_d	*–0.0836*[a]	*0.9782*[a]	*0.0574*	*0.0164*	*0.3248*	*0.1415*						*0.66*	*1.90*
	t-stat	0.88	12.85***	0.53	0.11	1.75*	0.60							
	$Haus_d$	*–0.0914*[a]	*0.9782*[a]	*0.0575*	*0.0164*	*0.3248*	*0.1416*	*0.0120*	*–0.0311*	*0.0448*	*–0.1944*	*–0.0495*	*0.66*	*1.90*
		–1.57	22.53***	0.85	0.21	2.97***	1.13	0.27	–0.45	0.56	–1.74*	–0.38		

Notes: LSDV is the least squares dummy variable method for estimating the alpha and beta for the FF 12 sectors ranging from January 1968 to December 2016. GMM_d is the generalized method of moments method using our robust distance instruments given in (6) with the Newey-West (1987) HAC variance-covariance estimator. *** indicates significance at 1%; **, 5%; and *, 10%. $\overline{R^2}$ is the adjusted coefficient of determination, and DW is the Durbin-Watson statistic for autocorrelation of order 1. $Haus_d$ is the modified Hausman (1978) artificial regression test for measurement errors using our robust distance instruments.
[a]For the columns of the constant term (*c*) and market risk premium ($R_m - R_f$), the pooled estimates are the averages of the LSDV/GMMd estimations.

to one. Among the four remaining factors, the most important is *RMW* in the three studies, the coefficients being 0.29 (GJ), 0.26 (LRZM) and 0.18 in ours. Regarding *CMA* and *HML*, the averages exposures are equal in the LRZM and our study, the coefficients being 0.16 and 0.08, respectively. GJ find a lower coefficient for *HML* (0.06) and a higher coefficient for *CMA* (0.26). In terms of average exposures, our study is thus closer to LRZM than to GJ. However, these studies differ in terms of the average exposure of sectors to *SMB*, the coefficient being equal to 0.33 in LRZM and 0.02 in our study. For GJ, the coefficient is negative (approximately –0.30) when significant in their twelve alternative models. The GJ article also includes the PS and Amihud illiquidity factors. The coefficient of the PS factor is negative for the lower quantiles of return distributions and positive for the higher ones. The Amihud ratio provides similar results but GJ argue that the analysis is more robust when using the PS factor.

According to LSDV, we have no sectors that generate significant positive excess returns while there are 3 sectors (Durables, –0.2591; Business equipment, –0.2738; Other, –0.3286) that significantly underperform. The relative (to the market) systematic measure of risk β for all 12 sectors is significantly different from 0.

For the fixed effects GMM_d, only one sector (Other, –0.3893) has a Jensen performance measure that is significantly different from 0 and it is negative as it was using LSDV. Again, the relative systematic measure of risk is significantly different from 0 for all 12 sectors. For the five-factor model, the pooled OLS is significant for all five factors. However, when using GMM_d, only the market factor is strongly significant albeit with marginal significance for RMW.

The Fixed Effects estimation of the FF new augmented six-factor model

Table 3b presents our estimation results for the new augmented (*LIQ*) FF six-factor model using the LSDV and GMM_d approaches for the fixed effects model.

For the FF six-factor OLS pooled model for the FF twelve-sector portfolio, all of the coefficients of the five FF risk factors retain their previous significance level when the *LIQ* risk factor is added. The *LIQ* risk factor, however, is insignificant, which suggests that *LIQ* on average does not have a risk premium. This again suggests strong support for the FF five factors. However, using GMM_d pooled, only the coefficient of the market factor is significant at 1% and illiquidity may matter (15% level of significance). When testing for measurement errors using the $\mathrm{Haus}_{\mathrm{d}}$ test, the *LIQ* risk factor is significant at the 5% level and seems to be measured with significant errors ($\widehat{\omega}_{LIQ}$ $-$ 0.1077, $t = -2.20$). Furthermore, the coefficient of the *CMA* risk factor becomes significant once again at almost the 5% level (0.2216, $t = 1.95$).

From an investment performance perspective, the Jensen performance measure is negative but not significant for the OLS pooled approach for the FF twelve-sector portfolio. However, for the $\mathrm{GMM}_{\mathrm{d}}$ pooled approach, the Jensen measure is negative and almost significant at the 10% level (-0.1444, $t = -1.62$). For OLS, the twelve-sector portfolio appears to be weighted towards firms that are small cap (*SMB*, 0.0246), value (*HML*, 0.0755), robust profitability (*RMW*, 0.1807), conservative investment policy (*CMA*, 0.1097), and slightly illiquid (*LIQ*, 0.0081) and quite significant. For $\mathrm{GMM}_{\mathrm{d}}$, the twelve-sector portfolio appears to be weighted towards firms that are small cap (*SMB*, 0.0032), value (*HML*, 0.0184), robust profitability (*RMW*, 0.2764, significant at 10%), conservative investment policy (*CMA*, 0.2216), and illiquid (*LIQ*, 0.1126, significant at 15%).

According to LSDV, of course, the previous alpha and beta remain exactly the same for the individual sectors, since the coefficient of the market risk factor is the only one that varies by sector in LSDV.

Sector Analysis of the random effects models: the new FF five-factor model

Table 4a presents the OLS and $\mathrm{GMM}_{\mathrm{d}}$ results for the 12 FF sectors of the new five-factor FF model, since these results are needed in the estimation of the random effects model.

Note that the coefficient for the market factor $R_M - R_f$ is significant at 1% for all 12 FF sectors using OLS and for 10 of the 12 FF sectors using $\mathrm{GMM}_{\mathrm{d}}$, Utilities and Shops being the insignificant sectors. For *SMB*, its OLS estimated coefficients are almost all significant (at 10% or better) but for two sectors, Chemicals and Money. Turning to the $\mathrm{GMM}_{\mathrm{d}}$ estimated coefficients for *SMB*, *almost no coefficients are significant* (Durables is significant at 10%)! For the *HML* factor, Fama and French (2015) themselves conjectured that *HML* could be redundant[28] with the addition of the *RMW* and *CMA* factors. In other words, there could be multicollinearity

28 See Fama and French (2015), p. 2.

Table 3b. Fixed Effects Model, LSDV vs. GMMd estimation methods for the augmented (LIQ) FF six-factor model by FF 12 sectors

		c	R_m-R_f	SMB	HML	RMW	CMA	LIQ	$\widehat{\omega}_{R_m-R_f}$	$\widehat{\omega}_{SMB}$	$\widehat{\omega}_{HML}$	$\widehat{\omega}_{RMW}$	$\widehat{\omega}_{CMA}$	$\widehat{\omega}_{LIQ}$	$\overline{R}^2$	DW
Sector	**Fama-French (2015, 2016) and Pastor-Stambaugh (2003)**															
1 NoDur	LSDV	0,1464	0,8378													
	t-stat	*1.19*	*30.75****													
	GMM_d	0.1031	0.7631													
	t-stat	*0.73*	*5.59****													
2 Durbl	LSDV	-0.2624	1.1906													
	t-stat	*-2.13***	*43.70****													
	GMM_d	-0.4357	1.3792													
	t-stat	*-2.63****	*11.55****													
3 Manuf	LSDV	-0.1371	1.1548													
	t-stat	*-1.11*	*42.39***													
	GMM_d	-0.2507	1.2225													
	t-stat	*-2.35***	*15.93****													
4 Enrgy	LSDV	0.0767	0.8585													
	t-stat	*0.62*	*31.51****													
	GMM_d	0.0066	0.8381													
	t-stat	*0.03*	*9.05****													
5 Chems	LSDV	-0.0513	0.9366													
	t-stat	*-0.42*	*34.38****													
	GMM_d	-0.0696	0.8113													
	t-stat	*-0.53*	*5.08****													
6 BusEq	LSDV	-0.2771	1.3068													
	t-stat	*-2.25***	*47.96****													
	GMM_d	-0.4951	1.5857													
	t-stat	*-2.03***	*5.73****													
7 Telcm	LSDV	0.0090	0.8411													
	t-stat	*0.07*	*30.87****													
	GMM_d	-0.1225	0.9450													
	t-stat	*-0.60*	*4.90****													

8 Utils	LSDV	0.0776	0.5674													
	t-stat	0.63	20.83***													
	GMM_d	*0.0712*	*0.4180*													
	t-stat	0.41	3.35***													
9 Shops	LSDV	*-0.0160*	*1.0534*													
	t-stat	-0.13	38.67***													
	GMM_d	*-0.1054*	*1.0721*													
	t-stat	-0.73	9.06***													
10 Hlth	LSDV	*0.0804*	*0.8777*													
	t-stat	0.65	32.21***													
	GMM_d	*0.1156*	*0.6441*													
	t-stat	0.64	3.79***													
11 Money	LSDV	*-0.0631*	*1.1272*													
	t-stat	-0.51	41.37***													
	GMM_d	*-0.1170*	*1.0740*													
	t-stat	-0.85	12.22***													
12 Other	LSDV	*-0.3319*	*1.1856*													
	t-stat	-2.69***	43.53***													
	GMM_d	*-0.4334*	*1.2288*													
	t-stat	-3.85***	16.33***													
Pooled Model																
	OLS	-0.0624[a]	0.9948[a]	0.0246	0.0755	0.1807	0.1097	0.0081							0.69	1.95
	t-stat	-1.60	106.72***	1.94*	4.37***	10.62***	4.20***	0.77								
	GMM_d	*-0.1444*[a]	*0.9985*[a]	*0.0031*	*0.0184*	*0.2764*	*0.2216*	*0.1126*							*0.65*	*1.89*
	t-stat	-1.62	14.19***	0.03	0.12	1.65	1.19	1.54								
	$Haus_d$	*-0.0342*[a]	*0.9985*[a]	*0.0031*	*0.0184*	*0.2764*	*0.2216*	*0.1126*	*-0.0101*	*0.0245*	*0.0432*	*-0.1431*	*-0.1305*	*-0.1077*	*0.66*	*1.91*
		-0.86	25.09***	0.05	0.24	2.71***	1.95*	2.37**	-0.25	0.40	0.55	-1.37	-1.12	-2.20**		

Notes: LSDV is the least squares dummy variable method for estimating the alpha and beta for the FF 12 sectors ranging from January 1968 to December 2016. GMM_d is the generalized method of moments method using our robust distance instruments given in (6) with the Newey-West (1987) HAC variance-covariance estimator. *** indicates significance at 1%; **, 5%; and *, 10%. $\overline{R^2}$ is the adjusted coefficient of determination and DW is the Durbin-Watson statistic for autocorrelation of order 1. $Haus_d$ is the modified Hausman (1978) artificial regression test for measurement errors using our robust distance instruments.

[a] For the columns of the constant term (*c*) and market risk premium ($R_m - R_f$), the pooled estimates are the averages of the LSDV/GMM_d estimations.

Table 4a. Random Effects Model, OLS vs GMM_d estimation methods for the FF five-factor model by FF 12 sectors

		c	R_m-R_f	*SMB*	*HML*	*RMW*	*CMA*	$\overline{R}^2$	*DW*
Sector	**Fama-French (2015, 2016)**								
1 NoDur	OLS	–0.0531	0.9073	0.0722	–0.0870	0.6019	0.4240	0.78	1.93
	t-stat	*–0.59*	*42.04****	*2.35***	*–2.08***	*14.61****	*6.71****		
	GMM_d	–0.1132	0.8700	0.3334	–0.2502	0.8680	0.4834	0.73	1.96
	t-stat	*–0.88*	*9.66****	*1.77**	*–1.11*	*3.26****	*1.51*		
2 Durbl	OLS	–0.4338	1.2116	0.2168	0.5570	0.1659	–0.0419	0.71	2.09
	t-stat	*–2.90****	*33.65****	*4.24****	*7.98****	*2.41***	*–0.40*		
	GMM_d	–0.5351	1.4667	–0.3977	1.1679	–0.5301	0.0761	0.53	1.98
	t-stat	*–1.77**	*7.16****	*–0.80*	*1.43*	*–0.57*	*0.13*		
3 Manuf	OLS	–0.1848	1.1441	0.1578	0.1359	0.2609	0.0708	0.89	2.03
	t-stat	*–2.37***	*61.06****	*5.93****	*3.74****	*7.30****	*1.29*		
	GMM_d	–0.3833	1.3452	–0.0957	0.1262	0.1617	0.5731	0.84	2.02
	t-stat	*–3.05****	*13.42****	*–0.40*	*0.42*	*0.43*	*1.88**		
4 Enrgy	OLS	–0.0156	0.9352	–0.1499	0.1098	0.1548	0.3496	0.46	1.90
	t-stat	*–0.09*	*21.82****	*–2.46***	*1.32*	*1.89**	*2.79****		
	GMM_d	0.0809	0.8296	0.2713	–1.1430	1.3339	0.4808	–0.08	1.67
	t-stat	*0.19*	*2.64****	*0.47*	*–1.68**	*1.50*	*0.60*		
5 Chems	OLS	–0.1960	1.0093	–0.0395	–0.0198	0.4469	0.3679	0.79	2.04
	t-stat	*–2.11***	*45.25****	*–1.25*	*–0.46*	*10.49****	*5.63****		
	GMM_d	–0.3670	1.0456	0.0435	–0.3734	0.6373	0.9966	0.75	1.93
	t-stat	*–2.32***	*7.12****	*0.18*	*–1.36*	*1.64**	*2.77****		
6 BusEq	OLS	0.3969	1.0387	0.0667	–0.3399	–0.4512	–0.5170	0.83	2.01
	t-stat	*3.39****	*36.88****	*1.67**	*–6.23****	*–8.39****	*–6.27****		
	GMM_d	0.3973	1.1155	–0.2787	0.3841	–1.1253	–0.7229	0.71	1.92
	t-stat	*1.41*	*5.61****	*–0.72*	*0.67*	*–1.64**	*–1.36*		
7 Telcm	OLS	0.1729	0.8415	–0.2537	0.1036	–0.2546	0.0986	0.61	1.99
	t-stat	*1.34*	*27.18****	*–5.77****	*1.73**	*–4.31****	*1.09*		
	GMM_d	0.5935	0.5576	–0.2488	0.7338	–0.4035	–1.2673	0.44	1.98
	t-stat	*2.77****	*3.11****	*–0.72*	*2.16***	*–0.79*	*–2.25***		
8 Utils	OLS	–0.0569	0.6565	–0.1360	0.2267	0.1526	0.3191	0.45	1.95
	t-stat	*–0.42*	*20.10****	*–2.93****	*3.58****	*2.45***	*3.34****		
	GMM_d	0.0267	0.3057	0.4684	–0.4568	1.0399	0.3269	–0.14	1.76
	t-stat	*0.09*	*1.22*	*1.02*	*–0.87*	*1.59*	*0.45*		
9 Shops	OLS	–0.0707	1.0218	0.2648	0.0007	0.4926	0.0369	0.80	1.85
	t-stat	*–0.68*	*40.74****	*7.43****	*0.02*	*10.29****	*0.50*		

	GMM$_d$	–0.1131	1.0241	0.3110	0.2985	0.4969	–0.1973	0.78	1.78
	t-stat	*–0.53*	*5.82****	*1.20*	*0.81*	*1.20*	*–0.51*		
10 Hlth	OLS	0.1965	0.8827	–0.1736	–0.4894	0.3427	0.3847	0.65	2.12
	t-stat	*1.56*	*29.21****	*–4.05****	*–8.36****	*5.94****	*4.35****		
	GMM$_d$	–0.0803	0.8723	0.1496	–0.8984	0.8325	1.0848	0.57	2.04
	t-stat	*–0.42*	*5.41****	*0.46*	*–1.72**	*1.55*	*2.59****		
11 Money	OLS	–0.1349	1.1641	–0.0354	0.6298	0.0962	–0.2348	0.84	1.89
	t-stat	*–1.40*	*50.35****	*–1.08*	*14.05****	*2.18***	*–3.47****		
	GMM$_d$	–0.0361	1.0598	–0.0138	0.5052	0.3824	–0.4562	0.81	1.81
	t-stat	*–0.20*	*6.46****	*–0.04*	*1.03*	*0.69*	*–1.10*		
12 Other	OLS	–0.3296	1.1236	0.3043	0.0858	0.1581	0.0512	0.91	1.97
	t-stat	*–4.53****	*64.27****	*12.26****	*2.53***	*4.74****	*1.00*		
	GMM$_d$	–0.4746	1.2473	0.1474	0.1036	0.2034	0.3210	0.89	1.91
	t-stat	*–4.40****	*15.01****	*0.91*	*0.61*	*0.95*	*1.20*		
Random Effects Model : Swamy's weighted average									
	FGLS	–0.0606	0.9947	0.0246	0.0763	0.1809	0.1085	0.72	1.98
	t-stat (weighted avg)	*0.53*	*39.43****	*1.39*	*1.36*	*4.06****	*1.72**		
	t-stat (Swamy)	*–0.91*	*21.63****	*0.46*	*0.84*	*2.10***	*1.32*		
	GMM$_d$	–0.0805	0.9783	0.0579	0.0166	0.3246	0.1417	0.57	1.90
	t-stat (weighted avg)	*–0.65*	*6.87****	*0.17*	*0.21*	*0.46*	*0.42*		
	t-stat (Swamy)	*–0.83*	*10.47****	*0.74*	*0.09*	*1.57*	*0.71*		

Notes: FGLS is calculated using data for the FF 12 sectors ranging from January 1968 to December 2016 using (12, Appendix 2) for the random coefficient model. *t-stat* is calculated first as a Swamy (1970) weighted average of the OLS sector *t-stats* using (11, Appendix 2) and then using the estimated Swamy variance-covariance matrix given by (14, Appendix 2). GMM$_d$ is the generalized method of moments using our robust distance instruments given in (6) with the Newey-West (1987). HAC variance-covariance estimator for the random coefficient model. *** indicates significance at 1%; **, 5%; and *, 10%. $\overline{R^2}$ is the adjusted coefficient of determination and DW is the Durbin-Watson statistic for autocorrelation of order 1.

with its attendant problems. For OLS, the *HML* coefficient is significant at the 10%-or-better level for 9 sectors. As with *SMB*, *the GMM$_d$ estimated coefficients of the HML factor are almost all insignificant with only two being significant.*

Turning to the first new factor *RMW*, we note that all the OLS estimated coefficients are significant at the 10%-or-better level. For the GMM_d estimated *RMW* coefficients, only four sectors have coefficients that are significant at the 10%-or-better level. For the *CMA* factor, the OLS estimated coefficients are significant at the 10%-or-better level for seven sectors. The GMM_d estimated *CMA* coefficients are significant at the 10%-or-better level for three sectors. *Thus, the* GMM_d *estimation results suggest that neither new FF factors RMW nor CMA seem to have much explanatory power.*

An Investment Perspective for the random effects FF five-factor model

Turning to an investment perspective, there is only one sector (Business equipment) that generates a significant positive risk-adjusted abnormal return based on OLS (see Table 4a). However, this sector fails to generate such a return based on GMM_d.

Sector analysis of the random effects model: the augmented new FF six-factor model

Table 4b presents our estimation results for the new augmented FF six-factor model using the OLS and GMM_d approaches for the random effects model.

Business equipment and Health have positive significant alphas based on OLS. Using GMM_d, both coefficients are positive, but neither are significant. Based on OLS, the coefficients and *t* values for the new FF five factors are essentially the same in the six-factor model. This is not surprising for Business equipment, given that the liquidity coefficient is not significantly different from 0. For Health, the factor coefficients also do not change in spite of the significance of the liquidity coefficient.

LIQ is really a measure of illiquidity by design. Thus, coefficients should be positive to generate a risk premium. For example, the Durables sector has a positive sign and is significant at the 10% level for OLS. *This is consistent with the idea that durables are difficult to sell during periods of illiquidity.* However, this conclusion may not hold when using GMM_d. Only 3 of the 12 FF sectors (Health, Money, and Other) have negative coefficients for both OLS and GMM_d. Although these *LIQ* coefficients are significant at the 1% level for these sectors for OLS, only the Health sector is significant (5%) when using GMM_d.

The Random Effects estimation of the FF new five-factor model

The t statistics in Table 4a for the coefficients of random effects model are calculated using two methods. First, weighted averages of the t statistics for the 12 sectors for each coefficient are calculated using equation (11, Appendix 2). Then the t statistics are computed using the Swamy (1970) variance-covariance matrix given by equation (14, Appendix 2).

The estimation of the alpha for the random effects model is slightly negative but insignificant for both FGLS and GMM_d. The insignificance of alpha is an indicator of market efficiency. The beta coefficient for the market factor $R_M - R_f$ is close to 1 for both FGLS and GMM_d. Thus, the 12-sector portfolio has essentially the same relative market risk as the market itself and has no abnormal or superior return. This suggests that the market portfolio should be the preferred investment vehicle, as it can be cost effectively obtained from either index mutual funds or exchange traded funds (ETFs). For *SMB*, the t values are insignificant for both FGLS and GMM_d and for both methods of calculating t. For *HML*, all results are also insignificant.

Using FGLS, the new FF *RMW* factor is positive and significant at the 1% level using the weighted average t and at the 5% level using the Swamy variance-covariance matrix. For GMM_d, *RMW* is insignificant using the weighted average t and Swamy. These coefficients are 0.1809 for FGLS and 0.3246 for GMM_d. These values are much bigger than the insignificant *SMB* and *HML* values. Therefore, robust profitability firms (*RMW*) do seem to have some explanatory power for the 12-sector portfolio returns. Meanwhile, conservative firms (*CMA*) may seem to explain some of the 12-sector portfolio returns with an FGLS coefficient of 0.1085 and a t that is significant at the 10% level for the weighted average method. However, the t value is insignificant for the Swamy method, and GMM_d yields insignificant results.

The Random Effects estimation of the FF new six-factor model

For the six-factor model using FGLS, the coefficients of the FF five factors in the twelve-sector FF equally weighted portfolio are imperceptibly different from the values obtained with the five-factor model (see Tables 4a and 4b). The t values have the same levels of significance, except for the *HML* coefficient, which is now significant at the 10% level using the weighted average method for calculating t.

Table 4b. Random Effects Model, OLS vs GMM_d estimation methods for the augmented (*LIQ*) FF six-factor model by FF 12 sectors

		c	R_m-R_f	*SMB*	*HML*	*RMW*	*CMA*	*LIQ*	$\overline{R}^2$	*DW*
Sector	**Fama-French (2015, 2016) and Pastor-Stambaugh (2003)**									
1 NoDur	OLS	–0.0567	0.9074	0.0722	–0.0876	0.6021	0.4247	0.0090	0.77	1.98
	t-stat	*–0.63*	*42.01****	*2.35***	*–2.09***	*14.60****	*6.72****	*0.35*		
	GMM_d	–0.1225	0.9048	0.3001	–0.1865	0.7807	0.5682	–0.0833	0.74	1.97
	t-stat	*–0.93*	*10.03****	*1.53*	*–0.71*	*2.61****	*1.70**	*–0.75*		
2 Durbl	OLS	–0.4645	1.2124	0.2174	0.5514	0.1671	–0.0359	0.0751	0.72	2.04
	t-stat	*–3.09****	*33.73****	*4.26****	*7.91****	*2.44***	*–0.34*	*1.78**		
	GMM_d	–0.6324	1.4851	–0.4744	1.1426	–0.5728	0.1724	0.2257	0.49	1.97
	t-stat	*–2.47***	*8.42****	*–1.22*	*1.64**	*–0.76*	*0.30*	*0.98*		
3 Manuf	OLS	–0.2103	1.1448	0.1582	0.1313	0.2619	0.0758	0.0624	0.89	2.02
	t-stat	*–2.70****	*61.47****	*5.98****	*3.63****	*7.37****	*1.39*	*2.85****		
	GMM_d	–0.3866	1.3205	–0.0831	0.0521	0.2328	0.5072	0.1160	0.85	2.03
	t-stat	*–3.50****	*16.58****	*–0.53*	*0.24*	*0.91*	*2.09***	*0.81*		
4 Enrgy	OLS	–0.0506	0.9361	–0.1493	0.1034	0.1561	0.3565	0.0857	0.48	1.87
	t-stat	*–0.28*	*21.87****	*–2.46***	*1.24*	*1.91**	*2.84****	*1.70**		
	GMM_d	–0.1265	0.8716	0.1080	–1.2109	1.2541	0.6782	0.4902	–0.08	1.65
	t-stat	*–0.37*	*3.29****	*0.21*	*–1.42*	*1.30*	*0.98*	*1.29*		
5 Chems	OLS	–0.2058	1.0096	–0.0393	–0.0216	0.4472	0.3698	0.0238	0.80	2.00
	t-stat	*–2.20***	*45.25****	*–1.24*	*–0.50*	*10.50****	*5.66****	*0.91*		
	GMM_d	–0.3812	1.0867	–0.0023	–0.3267	0.5462	1.0847	–0.0568	0.75	1.91
	t-stat	*–3.01****	*9.69****	*–0.01*	*–1.44*	*1.66**	*3.46****	*–0.41*		
6 BusEq	OLS	0.3902	1.0389	0.0668	–0.3411	–0.4509	–0.5157	0.0162	0.84	1.99
	t-stat	*3.31****	*36.86****	*1.67**	*–6.24****	*–8.38****	*–6.25****	*0.49*		
	GMM_d	0.3823	1.0811	–0.2587	0.3301	–1.0515	–0.7796	0.1246	0.73	1.93
	t-stat	*1.63*	*6.50****	*–0.86*	*0.70*	*–1.99***	*–1.69**	*0.62*		
7 Telcm	OLS	0.1633	0.8417	–0.2536	0.1019	–0.2543	0.1005	0.0235	0.62	1.96
	t-stat	*1.26*	*27.18****	*–5.77****	*1.69**	*–4.30****	*1.11*	*0.64*		
	GMM_d	0.4191	0.6151	–0.3995	0.7548	–0.5444	–1.0319	0.3027	0.42	2.03
	t-stat	*1.93**	*3.04****	*–1.11*	*1.76**	*–0.97*	*–2.02***	*1.52*		
8 Utils	OLS	–0.0820	0.6571	–0.1356	0.2222	0.1536	0.3241	0.0614	0.45	1.97
	t-stat	*–0.60*	*20.14****	*–2.93****	*3.51****	*2.47***	*3.39****	*1.60*		
	GMM_d	–0.4411	0.4689	0.0501	–0.4086	0.6468	0.9671	0.8180	–0.16	1.84
	t-stat	*–1.91**	*2.31***	*0.15*	*–0.65*	*0.98*	*1.89**	*2.92****		
9 Shops	OLS	–0.0821	1.0221	0.2650	–0.0013	0.4930	0.0391	0.0279	0.80	1.87
	t-stat	*–0.78*	*40.74****	*7.44****	*–0.03*	*10.29****	*0.53*	*0.95*		

	GMM$_d$	–0.0806	1.0210	0.3349	0.3269	0.4944	–0.2140	–0.1004	0.77	1.76
	t-stat	*–0.48*	*6.66****	*1.44*	*0.78*	*1.15*	*–0.63*	*–0.57*		
10 Hlth	OLS	0.2442	0.8814	–0.1745	–0.4807	0.3408	0.3753	–0.1169	0.68	2.08
	t-stat	*1.95**	*29.42****	*–4.10****	*–8.27****	*5.96****	*4.28****	*–3.32****		
	GMM$_d$	0.1028	0.8280	0.2930	–0.8939	0.9401	0.8725	–0.3606	0.51	1.98
	t-stat	*0.51*	*5.73****	*1.04*	*–2.17***	*2.10***	*2.15***	*–2.11***		
11 Money	OLS	–0.0978	1.1632	–0.0361	0.6366	0.0948	–0.2421	–0.0909	0.85	1.89
	t-stat	*–1.02*	*50.75****	*–1.11*	*14.32****	*2.17***	*–3.61****	*–3.37****		
	GMM$_d$	–0.0769	1.0838	–0.0527	0.5648	0.3061	–0.3588	–0.0024	0.83	1.81
	t-stat	*–0.56*	*7.89****	*–0.21*	*1.29*	*0.67*	*–1.02*	*–0.01*		
12 Other	OLS	–0.2967	1.1227	0.3038	0.0918	0.1568	0.0447	–0.0803	0.91	1.99
	t-stat	*–4.11****	*65.02****	*12.39****	*2.74****	*4.76****	*0.88*	*–3.96****		
	GMM$_d$	–0.3892	1.2153	0.2214	0.0760	0.2843	0.1928	–0.1230	0.90	1.97
	t-stat	*–3.68****	*17.35****	*1.53*	*0.49*	*1.37*	*0.84*	*–1.08*		
Random Effects Model : Swamy's weighted average										
	FGLS	–0.0632	0.9950	0.0246	0.0756	0.1809	0.1100	0.0076	0.73	1.97
	t-stat (weighted avg)	*0.88*	*39.31****	*1.35*	*1.64**	*3.83****	*1.24*	*0.60*		
	t-stat (Swamy)	*–0.93*	*21.65****	*0.46*	*0.83*	*2.11***	*1.34*	*0.39*		
	GMM$_d$	–0.1461	0.9986	0.0031	0.0186	0.2767	0.2216	0.1124	0.56	1.90
	t-stat (weighted avg)	*–0.68*	*8.08****	*0.15*	*0.07*	*0.71*	*0.65*	*0.31*		
	t-stat (Swamy)	*–1.56*	*12.15****	*0.04*	*0.10*	*1.41*	*1.11*	*1.23*		

Notes: FGLS is calculated using data for the FF 12 sectors ranging from January 1968 to December 2016 using (12, Appendix 2) for the random coefficient model. *t-stat* is calculated first as a Swamy (1970) weighted average of the OLS sector *t-stats* using (11, Appendix 2) and then using the estimated Swamy variance-covariance matrix given by(14, Appendix 2). GMM$_d$ is the generalized method of moments using our robust distance instruments given in (6) with the Newey-West (1987) HAC variance-covariance estimator for the random coefficient model. *** indicates significance at 1%; **, 5%; and *, 10%. $\overline{R^2}$ is the adjusted coefficient of determination, and DW is the Durbin-Watson statistic for autocorrelation of order 1.

Looking at the investment performance of the FF twelve-sector portfolio, the performance measure is negative but insignificant even at the 20% level[29]. Using GMM$_d$, it appears that the portfolio is weighted towards stocks that

29 The *t* value is positive for the weighted average approach using FGLS even though the alpha is negative because in this particular case the weighted summation of the sectors with positive t values outweighs the magnitude of the weighted summation of the sectors with negative *t* values.

are small cap (*SMB*, 0.0031), high book to market (*HML*, 0.0186), robust profitability (*RMW*, 0.2767), conservative investment (*CMA*, 0.2216), and illiquid (*LIQ*, 0.1124). These results seem consistent with our previous Tobin Q and investment perspective. Normally, one expects large cap stocks to be liquid and hence, the *LIQ* coefficient should not be significantly different from 0 or possibly significantly negative. Here, we find that it is insignificant using GMM_d. Perhaps this is an effect of the 2007-2009 financial crisis, when even large cap stocks were somewhat illiquid.

F test for the fixed effects versus the pooled models

When testing the fixed effects model over the pooled one, the F test rejects the pooled regression approach. The standard *F* test is given by[30] $F(N-1, NT-N-k) = \dfrac{\left(R^2_{LSDV} - R^2_{Pooled}\right)/(N-1)}{\left(1 - R^2_{LSDV}\right)/(NT-N-k)}$ where R^2_{LSDV} is the coefficient of determination for the least squares dummy variables regression, R^2_{Pooled} is the coefficient of determination for the pooled regression, *N* is the number of sectors, *T* is the number of months, and *k* is the number of regressors. Table 5 provides the *F* values for the five and six-factor models using OLS and GMM_d estimation methods.

Table 5 Testing fixed effects versus random effects models

	5 factors		6 factors	
	Pool/FE	GMM_d/FE	Pool/FE	GMM_d/FE
F test	*29.44*	*13.96*	*29.56*	*14.49*
	OLS RE/FE	GMM_d RE/FE	OLS RE/FE	GMM_d RE/FE
H test	*0.0009*	*-0.61*	*-0.3671*	*-0.00004*

Notes: *F* test is a Fisher *F* test for testing the pooled versus the fixed effects models. Pool/FE designates the pooled OLS versus LSDV fixed effects models. GMM_d/FE designates the pooled GMM_d estimation method versus the fixed effects model estimated via GMM_d. H test is the Hausman test for testing fixed versus random effects models. OLS RE/FE designates the FGLS for the random effects versus the LSDV models. GMM_d RE/FE designates the GMM_d estimation method for the random effects versus the fixed effects models.

30 See, for instance, Greene (2018), p. 397. Note, however, that in our case the fixed effects model that we use is not standard per se because we also allow the beta to vary across sectors. Therefore, when computing the F test, the degrees of freedom should be adjusted compared to the standard model. One way to do that is to set N = 24.

Note that all the F tests are significant at the 1% level, which means for the five and six risk factor models, the pooled model is rejected in favor of the fixed effects model using either OLS or GMM_d estimation methods[31].

Hausman H test for the fixed effects versus the random effects models[32]

While the F tests let us draw conclusions about the fixed effects model, the data cannot help discriminate between the fixed and random effects models. The Hausman (1978) test is particularly well-suited to discriminate between models, in our case, the fixed effects versus the random effects models. The Hausman test statistic is chi-squared distributed with *k-1* degrees of freedom and is given by (9). Intuitively, the H test is a quadratic distance weighted by its variance, the distance being between the fixed and random effects estimations. Turning to our result, Table 5 shows that the Hausman test cannot reject the random effects model using either OLS or GMM_d for the five and six risk factor models[33]. Thus, the fixed effects model is rejected.

5. Robustness check

As a robustness check, we investigate whether the results of this paper depend on the choice of the liquidity factor. Indeed, liquidity risk is multi-dimensional and more than one liquidity measure may be needed to capture different aspects of liquidity risk. For instance, Goyenko et al. (2009) show that the Pástor and Stambaugh (2003) liquidity measure fails to capture the price impact of trade, while Amihud (2002) measure can be considered as a good proxy for this aspect of liquidity risk. To tackle this issue, we introduce two additional liquidity measures in our augmented Fama and French model: (i) the Amihud illiquidity ratio, and; (ii) the term spread.

The Amihud (2002) illiquidity ratio is the daily ratio of absolute stock return to its trading volume, averaged over each month, i.e.,

$$LIQ\_Amihud = \frac{1}{D_i}\sum_{i=1}^{D_i}\frac{|R_i|}{Vol_i} \tag{10}$$

31 Note that the critical value for F test in either model is 1.53.

32 Note that we did not perform the auxiliary regression version of the test. This is because we have repeated observations of the regressors, therefore rendering the test difficult to apply for this financial application. We therefore rely on the Hausman test.

33 The critical value for the chi-squared distribution of the H test is in our case 11.07 or 12.59, respectively, for the five and six risk factor models.

where D_i is the number of days of the month, R_i is the daily return on stock i and Vol_i is its corresponding trading volume. In this paper, the Amihud (2002) illiquidity measure is computed using the S&P500. The Amihud ratio quantifies the price/return response to a given size of trade[34]. According to Naes et al. (2010), this ratio is a measure of the elasticity dimension of liquidity, in the sense that it tries to capture the sensitivity of prices to trading volume. When the Amihud ratio has a high value, liquidity is low. As argued by Konstantopoulos (2016), in most research papers, the coefficients of the PS and Amihud illiquidity factors are positive. Indeed, these factors are proxies for the illiquidity premium, which is a component of returns. However, according to Acharya and Pedersen (2005), a positive shock in illiquidity predicts high future illiquidity, so contemporaneous stock prices decrease, which can result in a negative sign for the coefficients of the PS and Amihud factors.

As another liquidity measure, we rely on the term spread – i.e., the spread between the ten-year constant maturity rate on US government bonds and the 3-month T-bills rate. When there is a flight to quality – i.e., when market liquidity is low – the investors buy short-term bonds, which are less risky, and sell long-term bonds, which embed more risk. Hence, the term spread increases and market liquidity becomes scarce. We thus expect a negative sign for this variable since it is strongly countercyclical. To further develop our model, we also account for two bond-oriented factors: (1) the bond market factor, which measures the monthly change in the 10-year treasury constant maturity yield; (ii) the credit market factor, representing the monthly change in the Moody's Baa yield less 10-year treasury constant maturity yield (Fung and Hsieh, 2004; González and Jareño, 2018)[35]. A rise in the credit market factor is often associated with a deterioration in firms' financial health, and thus a decrease in stock market returns. Following this reasoning, a negative sign is expected for the credit market factor. However, the credit market factor is a proxy for the credit risk premium, a component of stock returns. An increase in this premium should result in a corresponding increase in stock returns. The sign of the credit market factor is thus an empirical matter. In other respects, a rise in

34 Price Waterhouse Corporation, August 2015, *Global financial market liquidity study.*

35 We thank an anonymous referee for his suggestion to add these factors in our analysis.

the change in the ten-year bond yield should decrease stock returns since it is associated with a rise in the funding costs of firms, among others[36].

It is also interesting to perform subperiod analysis in order to assess the time-variability of the industry portfolio risk profile. Indeed, some factors may be statistically significant only during a given subperiod, the related beta coefficients may change over time, or both (Racicot and Théoret, 2016a). When the overall period is taken into consideration, it should blur the statistical significance of one or more factors in a given subperiod. Since coefficients tend to change mainly during crises, we define two dummy variables, one which takes the value of one during crises and zero elsewhere, and the other which takes the value of one outside crises and zero elsewhere. We multiply each of these two dummies by the explanatory variables of our models. We thus obtain truncated variables which will help infer the stability of the coefficients of our model over time[37].

As in our previous experiments, the market return is the main driver of the 12 portfolio returns (Table 6). When using OLS, the four other Fama and French factors are significant at the 5% level, and *CMA*, and especially *RMW*, have the highest coefficients. In keeping with our previous results, the other variables of our model – i.e., the liquidity measures and the bond market factors – are not significant. Estimating liquidity measures one by one does not provide better results. Interestingly, when relying on our GMM_d program[38], the *RMW* factor remains significant but its coefficient is lower with GMM_d (0.16) than with OLS (0.24), which suggests that our portfolio returns behave in much the same way as stock issued by a firm with a high return on equity. When using GMM_d, two illiquidity factors become significant – i.e., the PS traded liquidity factor and the term spread – which suggests that liquidity is endogenous (Adrian et al., 2017). The PS factor has a positive sign and is significant at the 5% level, while the term spread has a negative sign and is significant at the 10% level. Estimating

36 One may question the fact that three out of the five additional factors depend on the ten-year bond yield – i.e., the change in this yield by itself, the change in the credit factor and the change in the term spread – which may be a source of multicollinearity. First, note that these three variables are expressed in first-differences, which reduces the multicollinearity problem. Second, in our sample, the only significant correlation coefficient among these three variables is the one between the change in the term spread and the change in the ten-year bond yield. At just 0.54, it is not significant.

37 Another way to analyze the stability of the coefficients of our portfolio model would be to test (ex-post) whether the factors allow us to accurately replicate out-of-sample industry portfolio returns. A case in point is Hasanhodzic and Lo (2007). The authors use a 24-month rolling window to estimate beta coefficients and then use them as weights for the replication portfolio and compute its performance for the next period.

38 Note that when we compute the GMM_d estimator in this article, we run our own designed computer program that uses the EViews programming language. This code computes our optimal robust instrumental variable approach, using a GMM algorithm. This program is available on request.

Table 6. Augmented Fama and French model with three liquidity factors and two bond market factors for 12 portfolios: OLS, GMM_d and Hausman test

	OLS				GMM_d				$\widehat{\omega}$ _$Haus_d$			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
c	–0.11**	–0.11**	–0.11**	–0.10**	–0.05**	–0.04*	–0.04**	–0.05	–0.06	–0.07	–0.07	–0.05
SMB	0.07**	0.07**	0.07**	0.07**	0.01	0.01	0.01	0.01	0.06	0.06	0.06	0.06
HML	0.08*	0.08**	0.08**	0.08**	0.12**	0.12**	0.13**	0.13**	–0.04	–0.04	–0.05	–0.05
RMW	0.24**	0.24**	0.24**	0.24**	0.16**	0.15**	0.15**	0.17**	0.11**	0.12**	0.12**	0.11**
CMA	0.12**	0.12**	0.12**	0.12**	0.06	0.04	0.04	0.05	0.09**	0.10**	0.07*	0.09**
LIQ_Pastor	0.00	0.00	–	–	0.03**	0.03**	–	–	–0.05**	–0.06**	–	–
LIQ_damihud	–0.28	–	–0.28	–	–0.06	–	–0.01	–	–0.25**	–	–0.28**	–
dtermspread	0.00	–	–	0.01	–0.14*	–	–	–0.13*	0.16**	–	–	0.18**
dr10Y	–0.02	0.15	–0.02	–0.04	0.14	0.11	0.13	0.16	–0.16**	0.04	–0.17**	–0.22**
dcreditspread	0.22	0.22	0.22	0.20	0.37**	0.37**	0.35**	0.36**	–0.18**	–0.18**	–0.15**	–0.17**
mkt_rf												
Non-durables	0.83**	0.84**	0.82**	0.86**	0.73**	0.75**	0.72**	0.73**	0.11	0.12	0.13	0.11
Durables	1.19**	1.20**	1.21**	1.19**	1.33**	1.32**	1.34**	1.31**	–0.18	–0.18	–0.19	–0.18
Manuf.	1.15**	1.17**	1.18**	1.17**	1.18**	1.17**	1.19**	1.16**	–0.04	–0.04	–0.05	–0.04
Energy	0.85**	0.86**	0.83**	0.82**	0.79**	0.80**	0.81**	0.78**	0.07	0.07	0.08	0.06
Chemicals	0.93**	0.95**	0.94**	0.93**	0.79**	0.77**	0.76**	0.75**	0.15*	0.14*	0.16*	0.15*
Bus. Eq.	1.30**	1.32**	1.29**	1.32**	1.47**	1.49**	1.46**	1.48**	–0.20**	–0.21**	–0.20**	–0.22**
Telecom.	0.84**	0.86**	0.82**	0.86**	0.87**	0.89**	0.86**	0.888**	–0.04	–0.05	–0.05	–0.04
Utilities	0.56**	0.58**	0.57**	0.56**	0.41**	0.42**	0.39**	0.42**	0.16**	0.17**	0.18**	0.16**
Shops	1.05**	1.05**	1.04**	1.08**	1.03**	1.05**	1.02**	1.02**	0.01	0.01	0.01	0.01
Health	0.87**	0.89**	0.88**	0.89**	0.61**	0.63**	0.59**	0.62**	0.29**	0.28**	0.27**	0.29**
Money	1.12**	1.15**	1.11**	1.12**	1.04**	1.05**	1.07**	1.04**	0.08	0.07	0.08	0.06
Other	1.18**	1.18**	1.19**	1.17**	1.18**	1.18**	1.19**	1.18**	–0.01	–0.01	–0.01	–0.01
R^2	0.73	0.73	0.73	0.73	0.67	0.68	0.68	0.67				
DW	2.00	2.00	2.00	2.00	1.95	2.00	2.00	2.00				

Notes: The two new liquidity factors are the Amihud (2002) ratio (*LIQ_damihud*) and the term spread (*dtermspread*) – i.e, the spread between the ten-year rate and the T-bills rate – both expressed in first-differences. The two bond factors are the ten-year constant maturity rate on US federal bonds (*dr10Y*) and the credit spread (*dcreditspread*) – i.e., the spread between the Baa rate and the ten-year rate – expressed in first-differences (Fung and Hsieh, 2004). For both estimation methods, the first column is the model estimation with the three illiquidity indicators while the three other columns provide the estimation with the three illiquidity variables taken separately. The first column of the $\widehat{\omega}$ _$Haus_d$ variables is associated with the first columns of the OLS and GMM_d estimations, and so on. *: significant at the 10% level; **: significant at the 5% level.

the illiquidity measures separately provides similar results. Our $Haus_d$ tests indicate that the coefficient of the Pástor and Stambaugh and Amihud illiquidity factors are understated in the OLS regression. It is also interesting to note, in line with our previous results, that the coefficients of the *CMA* and *RMW* factors are overstated[39].

Table 7 provides the pool estimation of our model applied to the 12 portfolios for crises and outside crises. There were many crises or recessions over our sampling period (1968-2016). However, since each crisis has its own idiosyncrasies, we resort only to the subprime crisis to truncate our explanatory variables. This crisis, by far the most important during our sampling period, lasted from June 2007 to December 2009. We note that the levels of the estimated coefficients and their significance may change from one regime to the next. For instance, some sectors – i.e., durables, manufactured goods, utilities, telecommunications and money – are more risky during the crisis, their market beta increasing substantially during this episode. Conversely, other sectors – i.e., non durables, business equipment, shops, and health – seem to bear less risk during the crisis. In other respects, among the other four factors of the augmented Fama and French model, *RMW* remains the most important factor in our regressions. Interestingly, while it is close to zero and not significant during the crisis, the coefficient of the Amihud ratio is significant at the 5% level and equal to -0.28 outside this crisis (Acharya and Pedersen, 2005). Finally, the change in the credit spread impacts positively stock returns[40]. This sign is related to the fact that the credit spread is a proxy for the credit risk premium, a component of stock returns.

6. Experiments with hedge funds

There may be differences between sectors' portfolios and managed portfolios, whose performance is related to the skills of portfolios' managers. Moreover, the transactions of most strategies are designed to be (i) nonlinear or highly nonlinear with respect to underlying assets; (ii) decorrelated from financial markets, especially when they are bearish (Fung and Hsieh, 1997, 2001, 2004). Fundamentally, this means we can expect the relationship of

39 Note that we have not included the estimated constants in Table 6 since the results are similar to the previous ones. Hence, the addition of new variables leads to similar alphas.

40 In order to limit losses in terms of "degrees of freedom," we estimate the coefficients of *dr10y* and *dcreditspread* over the whole sampling period.

Table 7. Augmented Fama and French model with three liquidity factors and two bond market factors for 12 portfolios: pool estimation during and outside the subprime crisis

	crisis				outside crisis			
c	-0.02	0.03	0.03	-0.02	-0.06**	-0.06**	-0.06**	-0.06**
SMB	0.06	0.05	0.05	0.06	0.02**	0.02**	0.02**	0.02**
HML	-0.07**	-0.06**	-0.06	-0.06**	0.08**	0.08**	0.08**	0.08**
RMW	0.17**	0.16**	0.17**	0.17**	0.18**	0.18**	0.18**	0.18**
CMA	0.13**	0.13**	0.12**	0.13**	0.10**	0.10**	0.10**	0.10**
LIQ_Pastor	-0.01	0.00			-0.00	-0.00		
LIQ_damihud	-0.01		-0.21		-0.29**		-0.28**	
dtermspread	0.38			0.37	-0.06			-0.05
dr10Y	0.02	0.03	0.04	0.02	0.02	0.03	0.04	0.02
dcreditspread	0.31	0.32	0.33*	0.31	0.31	0.32	0.33*	0.31
mkt_rf								
Non-durables	0.74**	0.73**	0.73**	0.73**	0.85**	0.85**	0.85**	0.85**
Durables	1.84**	1.83**	1.83**	1.84**	1.12**	1.12**	1.12**	1.12**
Manuf.	1.43**	1.42**	1.42**	1.43**	1.12**	1.12**	1.12**	1.12**
Energy	0.84**	0.83**	0.83**	0.84**	0.87**	0.87**	0.87**	0.87**
Chemicals	1.00**	0.99**	0.99**	0.99**	0.93**	0.93**	0.93**	0.93**
Bus. Eq.	1.15**	1.14**	1.14**	1.15**	1.32**	1.32**	1.32**	1.32**
Telecom.	1.05**	1.04**	1.04**	1.04**	0.83*	0.82**	0.82**	0.82**
Utilities	0.70**	0.70	0.69**	0.70**	0.56**	0.56**	0.56**	0.56**
Shops	0.87**	0.86**	0.86**	0.86**	1.07**	1.07**	1.07**	1.078*
Health	0.71**	0.70**	0.70**	0.71**	0.90**	0.90**	0.90**	0.90**
Money	1.43**	1.42**	1.42**	1.42**	1.10**	1.10**	1.10**	1.10**
Other	1.27**	1.26**	1.26**	1.27**	1.17**	1.17**	1.17**	1.17**
R^2	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
DW	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00

Notes: The two new liquidity factors are the Amihud (2002) ratio (*LIQ_damihud*) and the term spread (*dtermspread*) – i.e, the spread between the ten-year rate and the T-bills rate – both expressed in first-differences. The two bond factors are the ten-year constant maturity rate on US federal bonds (*dr10Y*) and the credit spread (*dcreditspread*) – i.e., the spread between the Baa rate and the ten-year rate – expressed in first-differences (Fung and Hsieh, 2004). The subprime crisis begins in June 2007 and ends in December 2009. For both regimes, the first column is the model estimation with the three illiquidity indicators while the three other columns provide the estimation with the three illiquidity variables taken separately. *: significant at the 10% level; **: significant at the 5% level.

hedge fund returns to the market index and to the other factors to be far different from the one corresponding to sector portfolios. From this perspective, hedge fund managers may try to capture the risk premia associated

Table 8. Hedge fund strategies' descriptive statistics

	Mean	Median	Maximum	Minimum	Standard-Dev.	Skewness	Kurtosis	Sharpe Ratio	CAPM Beta
Convertible	0.66%	0.84%	7.04%	-19.36%	2.17%	-4.75	43.26	0.30	0.21
Distressed	0.76%	1.00%	5.70%	-7.77%	1.75%	-0.94	6.33	0.43	0.22
Diversified Event Driven	0.76%	0.92%	8.90%	-8.20%	2.07%	-0.40	5.88	0.37	0.34
Equity Market Neutral	0.57%	0.49%	6.70%	-3.37%	1.13%	1.11	9.45	0.51	0.10
Fixed Income	0.64%	0.76%	3.47%	-8.00%	1.15%	-3.74	27.36	0.56	0.09
Futures	0.67%	0.45%	10.80%	-7.10%	2.79%	0.40	3.74	0.24	-0.06
Growth	0.80%	0.84%	18.40%	-12.20%	4.02%	0.48	5.97	0.20	0.70
Long-Short Credit	0.58%	0.78%	3.60%	-7.20%	1.22%	-1.92	12.70	0.48	0.15
Macro	0.50%	0.41%	10.50%	-9.60%	2.14%	0.56	9.36	0.23	0.22
Mergers	0.56%	0.60%	2.76%	-4.90%	0.99%	-1.36	8.21	0.57	0.12
Multi-Strategy	0.71%	0.61%	7.70%	-8.40%	2.01%	-0.28	5.46	0.35	0.32
Opportunistic	0.82%	0.80%	18.20%	-7.40%	2.72%	1.39	12.23	0.30	0.39
Short-Sellers	-0.22%	-0.41%	24.90%	-24.30%	5.14%	0.26	7.75	-0.04	-0.92
Value Index	1.24%	1.40%	11.20%	-10.10%	3.29%	-0.11	3.88	0.38	0.48
Average	0.65%	0.68%	9.99%	-9.85%	2.33%	-0.66	11.54	0.35	0.17
General Index	0.71%	0.77%	9.00%	-7.90%	2.04%	0.05	6.18	0.35	0.34
Standard & Poors 500	0.73%	1.18%	10.93%	-16.80%	4.50%	-0.65	3.93	0.16	1.00

Notes: The sample period runs from January 1995 to July 2016. The Sharpe ratio of a strategy is equal to this strategy expected excess return over its return standard deviation. The CAPM beta is computed using the simple market model.

with market illiquidity. We thus replicate our analysis of 12 portfolios on hedge fund returns. Data on these returns are drawn from the database managed by Greenwich Alternative Investment (GAI) – one of the oldest hedge fund databases, containing more than 13,500 records of individual hedge funds. The data reflects net-of-fees returns. Our dataset runs from January 1995 to July 2016, for a total of 243 observations. In addition to the general index (weighted composite index), our database includes the twelve strategies described in the Appendix 4.

6.1. Desriptive statistics

Table 8 reports the descriptive statistics of our hedge fund database. There is some heterogeneity in the historical returns and risk characteristics of hedge fund strategies. For instance, the monthly mean returns range from 0.54% for macro 1.24% for value index while the return standard deviation ranges from 1.13% for equity market neutral to 4.02% for growth. A hedge fund's market beta is generally low, the average market beta computed over all strategies being equal to 0.26. The futures strategy display a negative beta (-0.06). Selling short may thus be a dominant strategy for futures. The strategy with the highest positive beta is growth (0.70) while the strategies with the lowest positive betas are, as expected, fixed income (0.09) and equity market neutral (0.10).

The standard deviation of the general index (*gi*) return is less than the one corresponding to the S&P500 return over our sample period, the respective levels being 2.04% and 4.50%. In fact, there is evidence of a learning process at play in the hedge fund industry which is associated with a decrease in procyclicality in this sector (Racicot and Théoret, 2016a). In this context, the standard deviation of the general index return increased less during the subprime crisis than during the tech-bubble one, while the standard deviation of the S&P500 return increased much more during the subprime crisis.

Seven strategies (over fourteen) display negative skewness: convertible, fixed income, long-short credit, distressed, diversified event driven, multi-strategy, and value index (Table 8). This indicates that negative return outliers exceed positive ones for these strategies, an obvious deterrent for investors. The strategies that display the highest negative skewness are those whose business lines are greatly oriented towards credit risk or credit-related securities, like convertible and fixed income. These strategies were particularly hit by the subprime crisis that originated mainly from defaults on risky mortgages. Note that our observations are more or less in line with Chan et al. (2005) and Heuson et al. (2016), who find that most hedge fund strategies display negative skewness, which they consider an indication of tail risk. However, a more straightforward measure of tail risk is kurtosis. Most hedge fund strategies present excess kurtosis. For our hedge fund strategies, kurtosis ranges from 3.74 (futures) to 43.26 (convertible). Like the convertibles-oriented strategy, the fixed income oriented strategy was greatly hit by the subprime crisis, its kurtosis being 27.36, which is associated with catastrophic financial performance during credit crises.

6.2. Empirical results

Table 9 replicates the experiments made in the previous section for hedge funds over the period beginning in January 1995 and ending in July 2016[41]. Before analyzing this table, note that we progressively added the explanatory variables in our empirical return model, starting with the original and the new Fama and French models, and moving on to include liquidity ratios and bond market factors. In all these experiments, the alpha remained significant at the 5% level and was relatively stable at the level of 0.40%, which corresponds to an annual abnormal return equal to 4.8%. Actually, the alpha was also at about the same level before the subprime crisis (e.g., Racicot and Théoret, 2012).

According to the Hausman test, the fixed effects model is selected for hedge funds. In the OLS regression, we note that these effects are particularly high and significant for the futures, macro, growth and value index strategies, which suggests that the behavior of these strategies is quite specific. Consistent with our previous results, the market risk premium is the major driver of hedge fund returns in the OLS run. However, in the GMM_d estimation, this premium is not significant for seven strategies, which indicates an errors-in-variables issue for the risk premium, which was less the case for the 12 portfolios. Three factors – i.e., *SMB*, *HML*, and *RMW* – of the new Fama and French (2015) model are significant at the 5% level in the OLS run but only *SMB* remains significant in the GMM_d estimation including the three liquidity variables. Indeed, the *SMB* factor is an important element for hedge funds as, when there is an expansion, small cap companies tend to outperform large cap ones (Stafylas et al., 2018).

41 Statistics on hedge funds before 1995 are not reliable. Moreover, we exclude short-sellers to conduct our experiments on hedge funds because their behavior tends to run counter to the other strategies.

Table 9. Augmented Fama and French model with three liquidity factors and two bond market factors for hedge funds: OLS and GMM_d

	OLS				GMM_d				$\widehat{\omega}$_$Haus_d$			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
c	0.40**	0.39**	0.40**	0.41**	0.48**	0.37**	0.45**	0.45**				
SMB	0.08**	0.08**	0.08**	0.08**	0.10**	0.19**	0.07**	0.29**	–0.02	–0.15	0.01	–0.21**
HML	–0.04**	–0.04**	–0.04**	–0.04**	–0.01	–0.07*	–0.01	–0.10*	–0.04	0.05	–0.03	0.08
RMW	–0.02*	–0.01	–0.01	–0.01	–0.03	0.13*	–0.06	0.29**	0.01	–0.16**	0.05	–0.32**
CMA	–0.00	0.00	–0.00	–0.00	–0.07	–0.05	–0.07	–0.11**	0.07	0.07	0.06	0.12
LIQ_Pastor	2.06**	2.26**			0.40	1.62			1.68**	0.66**		
LIQ_damihud	0.33*		0.39**		2.92**		3.40**		–2.65**		–3.06**	
dtermspread	–0.62**			–0.68**	–0.46**			–2.89**	–0.16			2.28**
dr10Y	–0.03	–0.58**	–0.53**	0.08	–0.15	–0.43**	–0.60**	2.30**	0.12	–0.15	0.07	–2.27**
dcreditspread	–2.02**	–2.18**	–2.30**	–2.02**	–2.41**	–2.20**	–2.68**	–1.62**	0.42**	0.02	0.35**	–0.42**
mkt_rf												
gi	0.29**	0.29**	0.29**	0.29**	0.21**	0.24**	0.20**	0.22	0.09	0.06	0.09	0.07
conv	0.13**	0.14**	0.14**	0.14**	–0.26	–0.24	–0.30	–0.23	0.42**	0.35**	0.46**	0.37**
dist	0.15**	0.16**	0.14**	0.14**	–0.28	–0.30	–0.30	–0.35	0.44**	0.45**	0.48**	0.51**
ded	0.27**	0.27**	0.14**	0.14**	0.01	0.01	–0.01	–0.02	0.29**	0.27**	0.28**	0.32**
emn	0.05**	0.05**	0.05**	0.05**	0.15	0.20*	0.14	0.18	–0.11	–0.17	–0.12	–0.13
fi	0.04**	0.04**	0.04**	0.04**	–0.05	–0.01	–0.06	–0.03	0.09	0.05	0.10	0.07
fut	–0.09**	–0.09**	–0.09**	–0.09**	0.45	0.56	0.45	0.55	–0.56**	–0.68**	–0.54**	–0.67**
growth	0.64**	0.65**	0.65**	0.65**	0.57**	0.62**	0.55**	0.63**	0.07	0.03	0.10	0.02
lsc	0.09**	0.09**	0.09**	0.09**	0.01	0.03	–0.01	0.01	0.08	0.09	0.10	0.07
macro	0.17**	0.17**	0.17**	0.17**	0.28*	0.32*	0.28	0.29	–0.12	–0.16	–0.11	–0.13
ms	0.26**	0.26**	0.27**	0.27**	0.19*	0.23	0.18	0.20	0.07	0.03	0.08	0.06

oi	0.34**	0.34**	0.34**	0.34**	0.33**	0.37**	0.32**	0.35**	0.01	–0.03	0.04	–0.01
vi	0.42**	0.43**	0.43**	0.43**	0.22**	0.24**	0.20**	0.22**	0.20**	0.18	0.23**	0.20**
Fixed effects												
gi	–0.06	–0.07	–0.07	–0.07	–0.06	–0.06	–0.06	–0.06				
conv	0.00	0.00	0.00	0.00	0.21**	0.22**	0.22**	0.20**				
dist	0.05	0.05	0.05	0.05	0.28**	0.32**	0.29**	0.34**				
ded	0.01	0.01	0.01	0.01	0.13**	0.15**	0.13**	0.16**				
emn	–0.02	–0.02	–0.02	–0.02	–0.12**	–0.14**	–0.13**	–0.13**				
fi	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05				
fut	0.15**	0.15**	0.15**	0.15**	–0.24**	–0.29**	–0.25**	–0.30**				
growth	–0.16**	–0.16**	–0.16**	–0.16**	–0.16**	–0.17**	–0.15**	–0.19**				
lsc	–0.09	–0.09	–0.09	–0.09	–0.08	–0.08	–0.08	–0.07				
macro	–0.19**	–0.19**	–0.19**	–0.19**	–0.31**	–0.32**	–0.32**	–0.31**				
ms	–0.06	–0.06	–0.06	–0.06	–0.07	–0.06	–0.07	–0.06				
oi	0.01	0.01	0.01	0.01	–0.03	–0.04	–0.04	–0.04				
vi	0.31**	0.31**	0.31**	0.31**	0.40**	0.40**	0.40**	0.40**				
R^2	0.55	0.55	0.55	0.55	0.42	0.42	0.43	0.25				
DW	2.06	2.03	2.04	2.03	2.03	2.04	2.05	2.03				

Notes: The two new liquidity factors are the Amihud (2002) ratio (*LIQ_damihud*) and the term spread (*dtermspread*) – i.e, the spread between the ten-year rate and the T-bills rate – both expressed in first-differences. The two bond factors are the ten-year constant maturity rate on US federal bonds (*dr10Y*) and the credit spread (*dcreditspread*) – i.e., the spread between the Baa rate and the ten-year rate – expressed in first-differences (Fung and Hsieh, 2004). In addition of the return of the general index (gi), we also analyze the returns of the following strategies: convertibles (conv), distressed securities (dist), diversified event driven (ded), equity market neutral (emn), fixed income (fi), futures (fut), growth (growth), long-short credit (lsc), macro (macro), multi-strategy (ms), opportunity index (oi), and value index (vi). For both estimation methods, the first column is the model estimation with the three illiquidity indicators while the three other columns provide the estimation with the three illiquidity variables taken separately. The first column of the $\hat{\omega}$ _$Haus_d$ variables is associated with the first columns of the OLS and GMM_d estimations, and so on. *: significant at the 10% level; **: significant at the 5% level.

More importantly, the three illiquidity variables – i.e., the PS and Amihud factors, and the term spread – are significant at the 5% level in the OLS run. The coefficients of the PS and Amihud factors are positive, at 2.06 and 0.33 respectively, while the coefficient of the term spread is negative at –0.62. We obtain similar results when estimating the illiquidity variables separately. However, in the GMM_d run, the Pástor and Stambaugh factor loses its significance, and its impact is transferred to the Amihud ratio which increases from 0.33 to 2.92 when moving from OLS to GMM_d. According to the $Haus_d$ test, the coefficient of the Pástor and Stambaugh factor is thus substantially overstated in the OLS regression while, conversely, the coefficient of the Amihud ratio is greatly understated (Table 9). Moreover, the coefficient of the term spread is significant and negative in both regression models. An increase in the term spread being associated with a developing recession, hedge fund returns thus decrease after a rise in this spread. Finally, in both regressions, the coefficient of the credit spread is negative and significant at the 5% level, signalling that hedge fund returns decrease following an increase in the credit spread. Many hedge strategies are particularly exposed to credit risk – i.e., fixed income, convertibles, distressed securities, long-short credit, and multistrategy – and the estimated negative exposure of hedge funds to credit risk suggests that their performance suffered from this kind of risk during the subprime crisis.

Interestingly, when estimating the liquidity factors separately in the GMM_d run, the Amihud ratio increases from 2.92 to 3.40, while the term spread moves from –0.46 to –2.89. The illiquidity factors thus interact, suggesting that liquidity is multidimensional. When estimating the Amihud ratio separately, we also note that its estimated coefficient "absorbs" an important share of the coefficients of four of the five factors of the new FF model. This suggests that the illiquidity ratio embeds dimensions of these factors, in the sense that the corresponding mimicking portfolios include a substantial share of illiquid securities. Moreover, when estimating our model with only the term spread as illiquidity variable, all factors in the new FF model become significant, another indication that illiquidity is an important aspect of the FF factors. In this regression with only the term spread as illiquidity variable, the R^2 decreases from 0.43 to 0.25, which underlines the importance of the other illiquidity variables in explaining hedge fund returns. These results also support the endogenous character of liquidity since they are not observed in the OLS run.

Turning to the pool estimations performed during and outside the subprime crisis, we first note, as expected, that the alpha is lower during the crisis than outside it, with respective coefficients of 0.25 and 0.40, but it remains significant at the 5% level (Table 10). Second, we can distinguish two categories of strategies, according to the values of the market betas in the two regimes. For some strategies, systematic risk increased sometimes substantially during the subprime crisis. For instance, comparing the estimated beta outside and during the crisis, we note in Table 9 that the beta for convertibles increased from 0.06 to 0.57; the beta for distressed securities rose from 0.16 to 0.29; the beta for fixed income instruments, from 0.01 to 0.24, and the beta for long-short credit, from 0.08 to 0.20. All these strategies are indeed very exposed to credit risk, which drove the subprime crisis. Conversely, other strategies – i.e., growth, opportunistic index, and value index – succeeded quite well in reducing their exposure to stock markets during the crisis.

Third, the four other factors of the new Fama and French (2015) model behave quite differently during the two regimes analyzed in Table 10. As explained previously, small cap companies tend to perform better than the large cap ones during expansion periods, which explains the positive and significant exposure of hedge funds to this factor outside crisis (0.12). However, their exposure became negative during the subprime crisis, since small cap companies perform worse during crises. The case of *HML* is also quite interesting. The exposure of hedge funds to this factor is positive but low (significant at the 5% level) outside crisis but it turned negative and high in absolute value during the crisis (–0.15), growth stocks being better performers than value stocks during crises (Campbell et al, 2010). The exposure of hedge funds to *RMW* and *CMA* is low but significant outside the crisis, with a coefficient of -0.03 for both sectors, but it is not significant during the crisis. More importantly, while the coefficients of the Pástor and Stambaugh factor and Amihud ratio are not significant outside the crisis, they gain strength during the crisis, with significant and positive coefficients equal to 3.55 and 1.87, respectively. Hedge funds thus capture the illiquidity risk premium when liquidity is the most scarce – i.e., during a crisis. Outside crisis, since the premium is low, it does not significantly impact hedge fund returns. By contrast, the term spread negatively and significantly affects hedge fund returns only outside the crisis period, signalling that a decrease in the term spread – which is usually observed during normal times – results in an increase in hedge fund returns. Estimating the illiquidity variables

Table 10. Augmented Fama and French model with three liquidity factors and two bond market factors for hedge funds: pool estimation during and outside the subprime crisis

	crisis				outside crisis			
c	0.25**	0.22**	0.23**	0.26**	0.40**	0.40**	0.40**	0.40**
SMB	-0.04*	-0.05**	-0.03	-0.04*	0.12**	0.12**	0.12**	0.12**
HML	-0.15**	-0.10**	-0.19**	-0.16**	0.03**	0.03**	0.02**	0.02**
RMW	-0.06	-0.02	-0.04	0.03	-0.03**	-0.03**	-0.03**	-0.03**
CMA	0.05	0.05	0.03	-0.00	-0.03**	-0.03**	-0.03**	-0.02**
LIQ_Pastor	3.55**	7.20**			-1.07	-1.02		
LIQ_damihud	1.87**		2.18**		-0.15		-0.10	
dtermspread	-0.09			-0.29	-0.34**			-0.42**
dr10Y	-0.38**	-0.65**	-0.63**	-0.27	-0.38**	-0.65**	-0.63**	-0.63**
dcreditspread	-1.94**	-1.89**	-2.11**	-1.88**	-1.94**	-1.89**	-2.11**	-2.11**
mkt_rf								
gi	0.29**	0.28**	0.31**	0.31**	0.31**	0.31**	0.31**	0.31**
conv	0.57**	0.55**	0.59**	0.59**	0.06**	0.06**	0.06**	0.06**
dist	0.29**	0.27**	0.30**	0.30**	0.16**	0.16**	0.16**	0.16**
ded	0.33**	0.31**	0.35**	0.35**	0.29**	0.29**	0.29**	0.29**
emn	0.05*	0.04	0.07**	0.07**	0.06**	0.06**	0.06**	0.06**
fi	0.24**	0.22**	0.26**	0.26**	0.01	0.01	0.01**	0.01
fut	-0.14	-0.15*	-0.12	-0.12	-0.08	-0.08	-0.08	-0.08
growth	0.54**	0.53**	0.56**	0.56**	0.70**	0.70**	0.70**	0.70**
lsc	0.20**	0.19**	0.22**	0.22**	0.08**	0.08**	0.08**	0.08**
macro	0.14**	0.12**	0.15**	0.15**	0.18**	0.18**	0.18**	0.18**
ms	0.19**	0.18**	0.21**	0.21**	0.31**	0.31**	0.30**	0.30**
oi	0.29**	0.28**	0.31**	0.31**	0.37**	0.37**	0.37**	0.37**
vi	0.38**	0.36**	0.40**	0.40**	0.49**	0.49**	0.48**	0.48**
R^2	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56
DW	1.61	1.60	1.61	1.60	1.61	1.60	1.61	1.60

Notes: The two new liquidity factors are the Amihud (2002) ratio (*LIQ_damihud*) and the term spread (*dtermspread*) – i.e, the spread between the ten-year rate and the T-bills rate – both expressed in first-differences. The two bond factors are the ten-year constant maturity rate on US federal bonds (*dr10Y*) and the credit spread (*dcreditspread*) – i.e., the spread between the Baa rate and the ten-year rate – expressed in first-differences (Fung and Hsieh, 2004). The subprime crisis begins in June 2007 and ends in December 2009. In addition of the return of the general index (gi), we also analyze the returns of the following strategies: convertibles (conv), distressed securities (dist), diversified event driven (ded), equity market neutral (emn), fixed income (fi), futures (fut), growth (growth), long-short credit (lsc), macro (macro), multi-strategy (ms), opportunity index (oi), and value index (vi). For both regimes, the first column is the model estimation with the three illiquidity indicators while the three other columns provide the estimation with the three illiquidity variables taken separately. *: significant at the 10% level; **: significant at the 5% level.

separately provides similar results. Finally, the bond-oriented factors are also significant with the expected negative sign[42].

Discussion

Our results show that the GMM_d approach makes it possible to identify the key factors that impact the risk of a portfolio and could be used in nonlinear applications like the study of the cyclical behavior of a portfolio. For instance, using the FF 12-sector portfolio data, the GMM_d shows that the most relevant factors are *RMW* and especially the market risk premium, regardless of the method of estimation used. More precisely, to perform these experiments, the GMM_d approach was very helpful in identifying the relevant factors that impact portfolio returns, but also the endogeneity biases, especially at the level of liquidity, which, according to Adrian et al. (2017), is an endogenous variable. Adding two other liquidity variables in this section – i.e., the Amihud ratio and the term spread – GMM_d reveals that the Pástor-Stambaugh liquidity factor is understated and that the *RMW* factor is significant at the 5% level. The estimation process is thus influenced by the set of explanatory variables used to run regressions, and our GMM_d program accounts for this issue. Our findings also show that the estimated model may be very different dependent on the regime – crisis or normal times. Outside the crisis, the four factors of the new FF model (excluding the risk premium) are all significant and have positive signs, *RMW* being the most important. During the crisis, the *SMB* factor is not significant and the coefficient of the *HML* factor turns negative. Consistent with Acharya and Pedersen (2005), the Amihud ratio is significant outside the crisis with a negative sign.

Our GMM_d approach also proved very relevant for hedge funds. It contributed to identifying the key factors which impact the returns of hedge fund strategies. In that regard, GMM_d gives even more weight to the *SMB* factor than OLS – a very important risk factor in hedge fund portfolio management. Consistent with our experiments with our 12-sector portfolio, the *HML*, *RMW*, and *CMA* factors are not significant when the three illiquidity variables are included in the GMM_d regression.

However, in contrast to sector portfolios, managed portfolios should capture the illiquidity risk premium, since portfolio managers are skilled. In view of

42 In order to avoid losing to much degrees of freedom, we estimate the coefficients of *dr10y* and *dcreditspread* over the whole sampling period.

this, the loading of the Amihud ratio is positive and significant at the 5% level in both OLS and GMM_d and its value is even more important in the GMM_d run, the estimated coefficients being respectively 0.33 and 2.92. The Pástor-Stambaugh factor is only significant in the OLS regression and thus transfers its weight to the Amihud ratio in the GMM_d regression. Surprisingly, when estimating our model with only the term spread as illiquidity variable, all of the new FF model's factors become significant, which signals that illiquidity is an important dimension of the FF factors. Conversely, when estimating the model with only the Amihud ratio as illiquidity variable, the weight of four FF factors (excluding the risk premium) is largely transferred to the Amihud ratio, which once more supports the conjecture that liquidity is endogenous. As expected, the change in the cyclical behavior of hedge fund strategies' returns is more important than the one of the 12-sector portfolio. Indeed, the coefficients of *SMB* and *HML*, being positive in normal times, become negative during the subprime crisis, and the *RMW* and *CMA* factors are only significant, albeit weak, in normal times. This may explain why the coefficients of *HML*, *CMA* and *RMW* are (usually) insignificant when estimated by GMM_d which accounts for the biases present in OLS. More importantly, the liquidity measures – i.e., the Pástor-Stambaugh factor and the Amihud ratio – which are not significant in normal times gain strength during the crisis, being both significant at the 5% level. This result, clearly anticipated by our GMM_d approach, suggests that hedge fund strategies capture illiquidity risk premia in crises when these premia are actually at play.

7. Conclusion

Using LSDV estimation, we find that the new Fama and French (2015, 2016) five factors are highly significant. However, adding to this model the illiquidity factor of Pástor and Stambaugh (2003) does not provide more explanatory power to the new FF model. When applying the GMM_d approach proposed in this paper to either the FF five-factor or augmented six-factor models, a different picture emerges. In the five-factor model using the fixed effects approach, only the market risk and the profitability factors are significant, at the 1% and 10% levels, respectively. However, the Hausman auxiliary regression shows a significant measurement error for *RMW*. Turning to the random effects model, the market factor is once more significant at the 1% level, whereas, the *RMW* factor falls to the non-standard 15% level.

Adding the PS illiquidity factor to the FF five-factor model changes the conclusions in the GMM_d universe for the fixed effects model. Except for the market risk factor, none of them is significant at the standard level of significance. This result is consistent with MacKinlay (1995). The illiquidity factor, however, could be considered significant if we lower the bar to the 15% level. Note that the illiquidity factor is measured with significant error using the Hausman auxiliary regression test. Moreover, when using this test, the *CMA* factor becomes significant at 5%.

For the fixed effects model, we find that the Jensen alpha measure of performance is negative and significant at the 15% level for the FF twelve-sector, pooled, augmented six-factor model using our GMM_d approach. However, alpha is not significant for the GMM_d five-factor model. While markets may be efficient ex-ante and not ex-post, this result shows ex-post that the twelve-sector portfolio may be more or less inefficient. Therefore, as an alternative, investors would be better off holding the market portfolio. Turning to the random effects model, the alpha is also negative but insignificant for the five-factor model, for both the FGLS or GMM_d approaches. However, using GMM_d, alpha is negative and significant at the non-standard 15% level for the augmented six-factor model.

As a robustness check, we also envision the multidimensional aspects of liquidity and the time-varying dimension of the factor loadings – especially during the subprime crisis. To perform these experiments, the GMM_d approach was very helpful in identifying the relevant factors but also the endogeneity biases, especially at the level of liquidity, which is an endogenous variable (Adrian et al., 2017). Adding two other liquidity variables – i.e., the Amihud ratio and the term spread – GMM_d reveals that the Pástor-Stambaugh liquidity factor is underestimated and that the *RMW* factor is significant at the 5% level. The estimation process is thus influenced by the set of explanatory variables used to run regressions. Our findings also show that the estimated model may be very different dependent on the regime – crisis or normal times. Consistent with our GMM_d regression, the *RMW* factor remains significant during the subprime crisis and outside it, and the factor *SMB* is only significant outside the crisis while the coefficient of *HML* turns from positive to negative in the run-up to the crisis. Interestingly, the Amihud ratio is significant outside the crisis with a negative sign.

Our GMM_d approach proved also very relevant for hedge funds. It contributed to identify the key factors which impact the returns of hedge fund strategies. In this respect, the GMM_d approach gives even more weight to the *SMB* factor, which is very important for hedge fund portfolio management. By contrast, the factors *HML*, *RMW*, and *CMA* are not significant. However, contrary to sector portfolios, managed portfolios should capture the illiquidity risk premium since portfolio managers have skills. In this respect, the loading of the Amihud ratio is positive and significant at the 5% level in both OLS and GMM_d and its value is even higher in the GMM_d run, the estimated coefficients being respectively 0.33 and 2.92. The Pástor-Stambaugh factor is only significant in the OLS regression and thus transfers its weight to the Amihud ratio in the GMM_d regression. As expected, the change in the cyclical behavior of hedge fund strategies is more important than the change corresponding to the 12-sector portfolio. Indeed, the coefficient of *SMB* and *HML* turned from positive to negative when moving from normal times to the subprime crisis, and the *RMW* and *CMA* factors are only significant, albeit weakly, in normal times. More importantly, the liquidity factors – i.e., the Pástor-Stambaugh factor and the Amihud ratio – which are not significant in normal times gain strength during the crisis, becoming significant at the 5% level. This result, which was clearly identified by our GMM_d approach, suggests that hedge fund strategies capture illiquidity risk premia in crises when these premia are the most important.

References

Acharya, V.A., Pedersen, L.H., 2005. Asset pricing with liquidity risk, *Journal of Financial Economics*, 77, 375-410.

Adrian, T., Fleming, M., Shachar, O., Vogt, E., 2017. Market liquidity after the financial crisis, *Annual Review of Financial Economics*, 9, 43-83.

Aitken, A.C., 1935. On least squares and linear combinations of observations, *Proceedings of the Royal Statistical Society*, 55, 42-48.

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets*, 5, 31-56.

Anderson, T., Rubin, H., 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations, *Annals of Mathematical Statistics*, 20, 46-63.

Anderson, T., Rubin, H., 1950. The asymptotic properties of the parameters of a single equation in a complete system of stochastic equations, *Annals of Mathematical Statistics*, 21, 570-582.

Arellano, M., 2003. *Panel Data Econometrics*, Oxford University Press, New York.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and application to employment equations, *Review of Economic Studies*, 58, 277-297.

Arellano, M., Bover, O., 1995. Another look at the instrumental variables estimation of error components models. *Journal of Econometrics*, 68, 29-52.

Barillas, F., Shanken, J., 2015. Comparing asset pricing models, NBER Working Paper 21771, Cambridge, MA.

Black, F., 1972. Capital market equilibrium with restricted borrowing, *Journal of Business*, 45, 444-455.

Bound, J., Jaeger, D., Baker, R., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak, *Journal of the American Statistical Association*, 90, 443-450.

Campbell, J., Lo, A., MacKinlay, A., 1997. *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.

Campbell, J.Y., Polk, C., Tuomo, V., 2010. Growth or glamour? Fundamentals and systematic risk in stock returns, *Review of Financial Studies*, 23, 305-344.

Carhart, M., 1997. On the persistence in mutual fund performance, *Journal of Finance*, 52, 57-82.

Chan, N., Getmansky, M., Haas, S.M., Lo, A.W., 2005. Systemic risk and hedge funds. Working Paper, NBER.

Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance*, 46, 209-237.

Cochrane, J.H., 2005. *Asset Pricing*, revised ed., Princeton University Press, Princeton, NJ.

Cochrane, J.H., 2008. The dog that did not bark: A defense of return predictability, *Review of Financial Studies*, 21, 1533-1575.

Cochrane, J.H., 2011. Presidential address: Discount rates, *Journal of Finance*, 66, 1047-1108.

Cochrane, J. H., 2017. Macro-finance, *Review of Finance*, 21(3), 945-985.

Dagenais, M.G., Dagenais, D.L., 1994. GMM estimators for linear regression models with errors in the variables, Centre de recherche et développement en économique (CRDE), Working Paper 0594, University of Montreal, Montreal, QC.

Dagenais, M.G., Dagenais, D.L., 1997. Higher moments estimators for linear regression models with errors in the variables, *Journal of Econometrics*, 76, 193-221.

Driscoll, J.C., Kraay, A.C., 1998. Consistent covariance matrix estimation with spatially dependent panel data, *Review of Economics and Statistics*, 80, 546-560.

Dufour, J.M., 2003. Identification, weak instruments and statistical inference in econometrics, CIRANO, Working Paper No. 2003s-49, University of Montreal, Montreal, QC.

Durbin, J., 1954. Errors in variables, *International Statistical Review*, 22(1/3), 23-32.

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

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

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model, *Journal of Financial Economics*, 116, 1-22.

Fama, E.F., French, K.R., 2016. Dissecting anomalies with a five-factor model, *Revue of Financial Studies*, 29, 69-103.

Fama, E.F., McBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests, *Journal of Political Economy*, 81, 607-636.

Fung, W., Hsieh, D.A., 1997. Empirical characteristics of dynamic trading strategies: The case of hedge funds. Review of financial studies 10, 275-302.

Fung, W., Hsieh, D.A., 2001. The risk in hedge fund strategies: Theory and evidence from trend followers. Review of Financial Studies 14, 313-341.

Fung, W., Hsieh, D.A., 2004. Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal*, 60, 65-80.

González, M.d.l.O, Jareño, F., 2018, Testing extensions of Fama and French models: A quantile regression approach. The Quartely Review of Economics and Finance. Forthcoming.

Goyenko, R.Y., Holden, C.W., Trzcinka, C.A., 2009. Do liquidity measures measure liquidity? *Journal of Financial Economics*, 92, 153-181.

Greene, W.H., 2018. *Econometric Analysis*, 8th ed., Pearson, New York, NY.

Hahn, J., Hausman, J., 2002. A new specification test for the validity of instrumental variables, *Econometrica*, 70, 163-189.

Hahn, J., Hausman, J., 2003. Weak instruments: Diagnosis and cures in empirical economics, *American Economic Review*, 93, 118-125.

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

Hansen, L.P., Singleton, K.J., 1982. Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica*, 50, 1269-1286.

Hansen, L.P., Singleton, K.J., 1984. Generalized instrumental variables estimation of nonlinear rational expectations models: Errata, *Econometrica*, 52, 267-268.

Harvey, C.R., 2017. Presidential address: The scientific outlook in financial economics, *Journal of Finance*, forthcoming.

Harvey, C.R., Liu, Y., Zhu, H., 2016. … and the cross-section of expected returns, *Review of Financial Studies*, 29, 5-68.

Hasanhodzic, J., Lo, A., 2007. Can hedge fund returns be replicated? The linear case, *Journal of Investment Management*, 5, 5-45.

Haug, E.S., 2007. *Derivatives Models on Models*, 2nd Ed., Wiley, Chichester, England.

Hausman, J., 1978. Specification tests in econometrics, *Econometrica*, 46(6), 1251-1271.

Hausman, J., Stock, J., Yogo, M., 2005. Asymptotic properties of the Hahn-Hausman test for weak instruments, *Economics Letters*, 89, 333-342.

Heuson, A.J., Hutchinson, M.C., Kumar, A., 2016. Skewness, fund flows and hedge fund performance. Working Paper, SSRN.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach, *Review of Financial Studies*, 28, 650-705.

Jarque, C.M., Bera, A.K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters*, 6, 255-259.

Jensen, M.C., 1968. The performance of mutual funds in the period 1945-64, *Journal of Finance*, 23, 389-416.

Jurczenko, E., Maillet, B., 2006. The four-moment capital asset pricing model: Between asset pricing and asset allocation. In *Multi-Moment Asset Allocation and Pricing Models* (eds. Jurczenko, E., Maillet, B.), John Wiley & Sons, Chichester, England, Ch. 6.

Konstantopoulos, L., 2016. Illiquidity: A study on how investing in illiquid assets can be profitable. M.Sc. Thesis, Tilburg University.

Li, L., Rao, X., Zhou, W., Mizrach, B., 2017. Analysis of 48 industry portfolios with a new Fama-French 5-factor model. Apllied Mathematics 8, 1684-1702.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 46, 13-37.

MacKinlay, A.C., 1995. Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics*, 38, 3-28.

Mandelbrodt, B., 1963. The variation of certain speculative prices, *Journal of Business*, 36, 394-419.

Mandelbrodt, B., 1972. Correction of an error in "The variation of certain speculative prices", *Journal of Business*, 45, 542-543.

Markowitz, H., 1959. *Portfolio Selection: Efficient Diversification of Investments*, Cowles Foundation Monograph 16, John Wiley & Sons, New York.

Markowitz, H., 2012. The "Great Confusion" concerning MPT, *Aestimatio: The IEB International Journal of Finance*, 4, 8-27.

Mclean, R. D., Pontiff, J., 2016. Does academic research destroy stock return predictability? *Journal of Finance*, 71, 5-32.

Mehra, R., Prescott, E., 1985. The equity risk premium: A puzzle, *Journal of Monetary Economics*, 15, 145-161.

Merton, R.C., 1973. An intertemporal capital asset pricing model, *Econometrica*, 41, 867-887.

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

Naes, R., Skjeltorp, J.A., Odegaard, B.A., 2010, Stock market liquidity and the business cycle. Working Paper, Norges Bank.

Nelson, C., Startz, R., 1990a. Some further results on the exact small sample properties of the instrumental variables estimator, *Econometrica*, 58, 967-976.

Nelson, C., Startz, R., 1990b. The distribution of the instrumental variables estimator and its *t*-ratio with the instrument is a poor one, *Journal of Business*, 63, S125-S140

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-708.

Olea, J.L.M., Pflueger, C., 2013. A robust test of weak instruments, *Journal of Business and Economic Statistics*, 31, 358–369.

Pagan, A.R., 1984. Econometric issues in the analysis of regressions with generated regressors, *International Economic Review*, 25, 221-247.

Pagan, A.R., 1986. Two stage and related estimators and their applications, *Review of Economic Studies*, 53, 517-538.

Pagan, A.R., Ullah A., 1988. The econometric analysis of models with risk terms, *Journal of Applied Econometrics*, 3, 87-105.

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

Pástor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns, *Journal of Political Economy*, 111, 642-685.

Pinto, J.E., Henry, E., Robinson, T.R., Stowe, J.D., 2015. *Equity Valuation*, 3nd ed., John Wiley & Sons, New York.

Pindyck, R., Rubinfeld, D., 1998. *Econometric Models and Economic Forecasts*, 4th ed., Irwin McGraw-Hill, Boston, MA.

Racicot, F.E., 2015. Engineering robust instruments for panel data regression models with errors in variables: A note, *Applied Economics*, 47, 981-989.

Racicot, F.E., Rentz, W.F., Théoret, R., 2018. Estimating panel data fixed and random effects with application to the new Fama-French model using GMM robust instruments, Working Paper, ESG-UQAM.

Racicot, F.E., Théoret, R., 2012. Optimally weighting higher moment instruments to deal with measurement errors in financial return models, *Applied Financial Economics*, 22, 1135-1146.

Racicot, F.É., Théoret, R., 2016a. The asymmetrical behavior of hedge funds across the state of the business cycle: The q-factor model revisited, *Finance*, 37, 43-87.

Racicot, F.-É., Théoret, R., 2016b. Macroeconomic shocks, forward looking dynamics, and the behavior of hedge funds. *Journal of Banking & Finance* 62, 41-61.

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

Saunders, A., Cornett, M.M., McGraw, P.A., 2014. Financial Institutions Management: A Risk Management Approach. Fifth Canadian Edition. McGraw Hill.

Shanken, J., 1992. On the estimation of beta-pricing models, *Journal of Finance*, 47, 1-34.

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

Stafylas, D., Andersen, K., Uddin, M., 2018. Hedge fund performance attribution under various market conditions, *International Review of Financial Analysis*, 56, 221-237.

Stock, J., Yogo, M., 2005. Testing for weak instruments in linear IV regression. In *Identification and Inference in Econometrics: Festschrift in Honor of Thomas Rothenberg* (eds. Stock, J., Andrews, D.), Cambridge University Press, Cambridge, England, 80-108.

Swamy, P., 1970. Efficient inference in a random coefficient regression model, *Econometrica*, 38, 311-323.

Theil, H., Goldberger, A.S., 1961. On pure and mixed estimation in economics, *International Economic Review*, 2, 65-78.

Tobin, J., 1969. A general equilibrium approach to monetary theory, *Journal of Money, Credit and Banking*, 1, 15-29.

White, H., 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity, *Econometrica*, 48, 817-838.

Wooldridge, J.M., 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.

Appendix 1

Composition of the FF 12 sector portfolios

Portfolio	Composition	SIC	Content
1 NoDur	Consumer NonDurables -- Food, Tobacco,	0100-0999	Agricultural products
	Textiles, Apparel, Leather, Toys	2000-2399	Food and kindred products
		2700-2749	Printing, publishing & allied industries
		2770-2799	Service industries for the printing trade
		3100-3199	Leather & leather products
		3940-3989	Toys, sporting & athletic goods
2 Durbl	Consumer Durables -- Cars, TV's, Furniture,	2500-2519	Furniture and fixtures
	Household Appliances	3630-3659	Household appliances
		3710-3711	Motor vehicles
		3714-3714	Motor vehicle parts
		3716-3716	Trucks
		3900-3939	Miscellaneous manufacturing industries
3 Manuf	Manufacturing -- Machinery, Trucks, Planes,	2520-2589	Office furniture
	Office Furniture, Paper, Commercial Printing	2600-2699	Paper & allied products
		2750-2769	Commercial printing
		3000-3099	Rubber & miscellaneous plastic products
		3200-3569	Stone, clay, glass & concrete products
		3580-3629	Refrigeration & service industry machinery
		3700-3709	Transportation equipment
		3712-3713	Truck and bus bodies
		3715-3715	Truck trailers
		3717-3749	Aircraft and parts
		3752-3791	Railroad equipment

		3860-3899	Photographic equipment & supplies
4 Enrgy	Oil, Gas, and Coal Extraction and Products	1200-1399	Coal mining, oil & gas extraction
		2900-2999	Petroleum refining & related industries
5 Chems	Chemicals and Allied Products	2800-2829	Chemical & allied products
		2840-2899	Soap, detergents, cosmetics & other toilet preparation
6 BusEq	Business Equipment -- Computers, Software,	3570-3579	Computer & office equipment
	and Electronic Equipment	3660-3692	Communications equipment
		3694-3699	Electronic equipment
		3810-3829	Search systems
		7370-7379	Computer programming and data processing services
7 Telcm	Telephone and Television Transmission	4800-4899	Communications
8 Utils	Utilities	4900-4949	Electric, gas & sanitary services
9 Shops	Shops Wholesale, Retail, and Some Services	5000-5999	Wholesale trade-durable goods
	(Laundries, Repair Shops)	7200-7299	Personal services
		7600-7699	Miscellaneous repair services
10 Hlth	Healthcare, Medical Equipment, and Drugs	2830-2839	Drugs
		3840-3859	Medical instruments & supplies
		8000-8099	Health services
11 Money	Finance	6000-6999	Financial institutions
12 Other	Other – Mines, Construction, Building material,		
	Transportation, Hotels, Business Services, Entertainement		

Notes: SIC: standard industrial classification.
Source: French's website.

Appendix 2

Estimation of W_i^*

An empirical estimation of W_i^* in (5) is needed to implement this model. W_i^* in its theoretical form is given by

$$W_i^* = \left[\sum_{i=1}^{N}\left(\Delta + \sigma_{e_i}^2\left(X_i'X_i\right)^{-1}\right)^{-1}\right]^{-1}\left(\Delta + \sigma_{e_i}^2\left(X_i'X_i\right)^{-1}\right)^{-1}, \quad (11)$$

where $N = 12$ is the number of FF sectors. To estimate W_i^* Swamy (1970) estimated Δ using the empirical variance of a set of N least squares estimates for the vector b_i minus the average value of $s_i^2\left(X_i'X_i\right)^{-1}$, *viz.*

$$\widehat{\Delta} = \left[1/(N-1)\right]\left[\sum_{i=1}^{N} b_i b_i' - N\overline{b}\overline{b}'\right] - (1/N)\sum_{i=1}^{N} V_i \text{ where } \overline{b} = \left(1/N\right)\sum_{i=1}^{N} b_i$$

and $V_i = s_i^2\left(X_i'X_i\right)^{-1}$. In summary, we obtain an estimate of the vector π given by the weighted average vector $\hat{\pi}$ with an estimator of the covariance matrix Δ given by $\widehat{\Delta}$.

We can write the empirical version of (5) for the random effects (RE) model by substituting $\widehat{\Delta}$ for Δ and $s_i^2\left(X_i'X_i\right)^{-1}$ for $\sigma_{e_i}^2\left(X_i'X_i\right)^{-1}$ (i.e., $\widehat{\Omega}$ for Ω, which implies substituting $\widehat{W}_i^*$ for W_i^*) to obtain the feasible generalized least squares (FGLS),

$$\hat{\hat{\pi}}_{RE} = \left(X'\widehat{\Omega}^{-1}X\right)^{-1} X'\widehat{\Omega}^{-1}y = \sum_{i=1}^{12} \widehat{W}_i^* b_i \quad (12)$$

The asymptotic RE variance-covariance matrix of (12) is given by the standard GLS one[43],

$$V_{RE}(\hat{\hat{\pi}}) = \left(X'\widehat{\Omega}^{-1}X\right)^{-1} \quad (13)$$

which translate empirically to (Swamy, 1970)

$$V_{RE}(\hat{\hat{\pi}}) = \left[\sum_{i=1}^{N}\left(\widehat{\Delta} + s_i^2\left(X_i'X_i\right)^{-1}\right)^{-1}\right]^{-1} \quad (14)$$

As is shown above, the estimated variance-covariance matrix of the random effects model is obtained in a straightforward way, using the first part of (11).

43 See Racicot, Rentz and Théoret (2018) or Wooldridge (2002).

Appendix 3

The instruments used for GMM_d

These estimators are respectively defined, in their multivariate representation, by (Racicot, 2015),

$$\beta_D = \left(z_1{}'x\right)^{-1}\left(z_1{}'y\right) \quad \text{(Durbin)} \tag{15}$$

$$\beta_P = \left(z_2{}'x\right)^{-1}\left(z_2{}'y\right) \quad \text{(Pal)} \tag{16}$$

where $z_1 = \left[x_{ij}^2\right]$, $z_2 = z_3 - 3Diag\left(x'x\ /\ N\right)x'$, $z_3 = \left[x_{ij}^3\right]$, and $Diag(x'x/N) = x'x/N \bullet I_k$ are stacked vectors with i representing the sectors ($i = 1, \ldots, N$), k the number of explanatory variables (either 5 or 6), and t the time subscript ($t = 1,\ldots,T$). The notation $\bullet$ is the Hadamard product. The second and third power (moments) of the *de-meaned* variables (x) are then computed. This is analogous to computing the second and third moments of the explanatory variables. In short, the instruments are obtained by taking the matrix of explanatory variables (X) in deviation from its mean (x). Next, we obtain the weighted estimator (β_H) by an application of the GLS to the following combination (Racicot, 2015),

$$\beta_H = W\begin{pmatrix}\beta_D \\ \beta_P\end{pmatrix} \tag{17}$$

where $W = \left(C'S^{-1}C\right)^{-1}C'S^{-1}$ is the GLS weighting matrix, S is the covariance matrix of $\begin{pmatrix}\beta_D \\ \beta_P\end{pmatrix}$ under the null hypothesis (i.e., no measurement errors), and $C = \begin{pmatrix}I_k \\ I_k\end{pmatrix}$ is a matrix of two staked identity matrices of dimension k.

Note that this weighting approach, which relies on GLS as the weighting matrix, is *optimal* in the Aitken (1935) sense[44]. However, we opt for the GMM method to weight the Durbin and Pal's estimators. We consider this a more efficient procedure than the one used by Dagenais and Dagenais (1997) in that we rely on the asymptotic properties of the GMM estimator with respect to the correction of heteroskedasticity and autocorrelation to

44 Note that we use *W* as a weighting matrix in the GLS estimator in (17). As well-known, this matrix can be replaced by the White (1980) or the Newey-West (1987) HAC asymptotically consistent variance-covariance matrix. For the problem of cross-sectional correlation (or spatial correlation) see Driscoll and Kraay (1998).

weight the instruments obtained with GLS. Note that when using GMM, we give up some efficiency gain in order to avoid completely specifying the nature of the autocorrelation or heteroskedasticity of the innovation and the data generating process of the measurement errors (Hansen, 1982). Again, we consider this a great advantage over the GLS estimator.

Appendix 4

Description of hedge fund strategies

Strategy	Description
convertible	They take a long position in convertibles and short simultaneously the stock of companies having issued these convertibles in order to hedge a portion of the equity risk.
Distressed securities	The managers buy equity and debt at deep discounts issued by firms facing bankruptcy.
Diversified event driven	The managers follow a multistrategy event driven approach.
Equity market neutral	The managers aim at obtaining returns with low or no correlation with equity and bond markets. They exploit the pricing inefficiencies between related equity securities.
Fixed income	The managers follow a variety of fixed income strategies like exploiting relative mispricing between related sets of fixed income securities. They invest in MBS, CDO, CLO and other structured products.
Futures	The manager utilizes futures contracts to implement directional positions in global equity, interest rate, currency and commodity markets. He resorts to leveraged positions to increase his return.
Growth	The managers invest in companies experiencing strong growth in earnings per share.
Long-short credit	They take long and short positions in credit in spite of the unavailability of bonds. They invest in high-yield bonds, CDS and CDO, among others.
Macro	These funds have a particular interest in macroeconomic variables. They take positions according to their forecasts of these variables. Managers rely on quantitative models to implement their strategies.
Multi-Strategy index	The manager utilizes investment strategies from more than one of the four broad strategy group indices.
Opportunistic	The managers' investment approach changes over time to better take advantage of current market conditions and investment opportunities.
Value index	Managers invest in securities which are perceived undervalued with respect to their "fundamentals".

Sources: Greenwich Global Hedge Fund Index Construction Methodology, Greenwich Alternative Investment (2015); Saunders et al (2014).