What drives stock return commonalities?
Evidence from France and US using a cross-sectional approach

July 11, 2011

J.J. Lilti[a,1] and G. Mero[b]

[a]*University of Rennes 1, Center of Research in Economics and Management (UMR CNRS n°6211),11 rue Jean Macé, CS 70803, 35708 Rennes Cedex 7 France*

[b]*University of Evry, EPEE and CREST-INSEE, France, gulten.mero@ensae.fr*

In this paper, we use a cross-sectional approach to get a deeper comprehension of the common risk profile of stock returns. Instead of employing static and *ad hoc* factor selection procedure an in Fama and French (1993), we use asymptotic developments of Bai and Ng (2002, 2006) to select the relevant factors. We thus reconcile two methodologies: the statistic one and the other, founded on the observed factors. We apply our approach to French and US stock markets over the period 1999 to 2008 and test the performance of several traditional observed factors, such as credit spread and firm-characteristic variables of Fama and French (1993) and Carhart (1997). Our results put in evidence strong time and country dependencies of stock risk profile.

[1] Corresponding author. E-mail : jean-jacques.lilti@univ-rennes1.fr

I. Introduction

In equity markets, systematic risk represents common shocks that drive stock return covariances. Stock returns may vary in the same direction simply because of economic changes or other events that impact large portions of the market. A question of interest, widely documented in the literature, is how to link return commonalities to observed factors and thus facilitate economic interpretation of the covariance structure of the data. Perhaps the most well-known model aiming at explaining common risk affecting asset returns is the CAPM [Sharpe (1964), Lintner (1965) and Black (1972)]. According to the CAPM, common risk is represented by the market portfolio of all of assets in the economy, which is considered as being the only source of risk. It follows that, stock risk is measured by the covariation of the risky security with the market portfolio returns. However, the empirical results reported over the past four decades underline the lack of the CAPM to fully explain common variations of stock returns. For instance, portfolios containing stocks with relatively small capitalization and high book to market value appear to earn higher returns on average than those predicted by the CAPM[2]. The lack of empirical support of the standard CAPM motivated the development of multifactor models whose aim is to capture what is left unexplained by the static CAPM.

Two main approaches exist in the literature.

The first one consists in selecting a set of observed factors, such as economic variables or firm-characteristic factors, in order to explain stock return comovements. One category of observed factors is initially introduced with a view to capture part of common risk which is left unexplained by the CAPM. For example, Fama and French (1993) developed a three

---

[2] For example, Fama and French (1992) find that the "relation between market beta and average return is flat", and that other variables related to firm characteristics, such as market capitalization, book to market ratio and earning to price ratio, have significant additional explanatory power in explaining the cross-section of returns.

factor model which includes, in addition to the market portfolio, two well-known observed factors related to firm size and book to market ratio. These firm-characteristic variables help explaining two CAPM anomalies, the size and the book to market effects[3]. Later on, Carhart (1997) extended the three factor model of Fama and French (1993) by adding a forth one in order to take account of risk related to return persistence. On the other hand, Chen et al. (1986) use several economic variables, such as inflation rate and the GDP index, to explain common variations of stock returns. Another category of observed factors is used to deal with the variability of the market risk which is not taken into account in the standard CAPM. In fact, the CAPM is a one-period model based on the assumption that the market betas remain constant over time. As discussed by Jagannathan and Wang (1996) among others, this assumption is too restrictive since the relative risk of a firm's cash flow is likely to vary over the business cycle. The authors developed a conditional CAPM allowing for time-varying risk premium. The unconditional form of their conditional CAPM results in a two factor model including the credit spread as a proxy of the time-varying market risk premium.

The observed factors present the advantage to provide direct economic interpretation of the common variations of stock returns. However, they are not perfect proxies of the underlying fundamental factors that drive return covariances. As discussed by Shanken (1992), using observed variables that do not span the same space with the fundamental factors results in measurement errors when dealing with beta estimations.

The second approach uses latent (unobserved or statistic) factors in order to capture stock return comovements. The theoretical foundation comes from the arbitrage pricing theory (APT) [see Ross (1976) and Roll and Ross (1980)] which assumes that in the absence

[3] The CAPM anomalies, such as size, book to market and debt to equity effects, are pointed out by Banz (1981), Reinganum (1981), Basu (1983), Chan, Chen and Hseih (1985), Shanken (1985), and Bhandari (1988).

of arbitrage opportunities, a small number of factors can be used to capture common variations of a large number of asset returns. These common latent factors are directly determined by the covariance structure of the data. In practice, they are estimated using statistical methods such as factor analysis [see Lehman and Modest (1988) among others] or principal component method [see, for example, Connor and Korajzcyk (1986, 1988)]. However, these statistic factors do not provide direct economic interpretation.

In this paper, we reconcile the statistic and observed factor approaches in order to analyze and interpret the common risk profile of stock returns. The point of interest is to identify the observed factors which match the best the covariance structure of asset returns. To do so, we adapt the Darolles and Mero (2011) cross-sectional approach – initially used to assess hedge fund[4] risk exposures – to stock returns. Instead of fixing in advance the number of factors, we apply Bai and Ng (2002, 2006) criteria to estimate the risk dimension of the data and to select the relevant observed factors i.e., the factors that match the latent structure of the data. Return covariation is filtered not only from the past historical data, but also from the cross-section of returns. Then, we apply our approach to French and US stock markets and test the performance of several traditional observed factors mentioned above, such as credit spread, the firm-characteristic variables of Fama and French (1993) as well as the momentum factor of Carhart (1997).

The contribution of this paper is twofold.

i) We use subperiod analysis which suggests that the common risk profile of stock returns is time-varying and depends on either the period characteristics or the considered country. It

[4] In particular, Darolles and Mero (2011) focus on an approximate factor model framework to deal with the factor selection issue. They show that their approach outperforms the standard one consisting of an *ad hoc* factor selection as in Hasanhobvic and Lo (2006).

follows that, account should be taken not only of the beta time-variability but also of the fact that the set of relevant factors is likely to change in time as a function of market characteristics.

ii) Even if our approach is based on Bai-Ng (2002, 2006), it seems to us that we are the first to use it by differentiating crisis period and growth period and apply it to French returns. Moreover, performing international analysis allows us to provide additional insights concerning risk variability across countries.

The paper is organized as follows. In section II, we provide a brief description of recent asymptotic tests that we apply to our data in order to get a better comprehension of the common risk structure of stock returns. Section III describes the data. Section IV presents our empirical applications and summarizes the main results. Section V concludes.

II. A cross-sectional approach to assess common risk affecting stock returns

In this paper, we use a two-step methodology in order to assess common risk affecting stock returns. Instead of focusing on only one category of factors (observed or latent ones), our approach combines both sets of variables which allows us to *i)* avoid the measurement errors due to arbitrary selection procedure of the observed factors and *ii)* provide direct economic interpretation of the latent covariance structure of asset returns. To do so, we apply recent asymptotic tests developed by Bai and Ng (2002, 2006) to our data.

More precisely, in the first step of our approach we use an approximate factor framework[5] to estimate the common variations of large panels of stock returns. Instead of fixing in advance the number of latent factors to include in the analysis, we apply Bai and Ng (2002) asymptotic tests to determine their number. These tests take account of the presence of heteroscedasticity and time-series dependences in the idiosyncratic components. In the second step, we use Bai and Ng (2006) asymptotic tests to match the observed variables to the latent factors and thus provide direct economic interpretation of the covariance structure of the data.

In this section, we first summarize Bai and Ng (2002) framework that we use in the first stage of our approach in order to determine the number of latent factors. Then, we provide a brief presentation of Bai and Ng (2006) tests allowing us to match the observed variables to the latent factors (second step of our approach).

II.1 Estimating the number of latent factors using large panel data

Let X_{it} be the observed return of stock i ($i = 1,...,N$) at time t ($t = 1,...,T$). The r- factor representation is given as follows:

$$X = F\Lambda^{'} + e, \qquad (1)$$

where X is the $T \times N$ matrix of asset returns, e is the $T \times N$ matrix of idiosyncratic components[6], Λ is the $N \times r$ matrix of factor loadings, and F is the $T \times r$ matrix of common factors. The item $F\Lambda^{'}$ represents the common component of X. The idiosyncratic components are supposed to have zero mean.

[5] In the sense of Bai and Ng (2002).

[6] Note that the number of stocks N represent the cross-sectional dimension while the number of periods for each cross-section unit T gives the time-series dimension.

In the classical (strict) factor analysis, the elements of e are assumed to be i.i.d. and independent of F [see for example Roll and Ross (1980) and Andersen (1984)]. In this case, the latent factors are supposed to fully capture the common variations of stock returns and the covariance matrix of X is given by: $\Sigma_X = \Lambda\Sigma_F\Lambda^{'} + D_e$, with Σ_F being the covariance matrix of the latent factors and D_e – which represents the covariance matrix of the idiosyncratic components – being a diagonal matrix. However, this assumption appears to be quite unrealistic when dealing with large panel data. In particular, even after controlling for the latent common factors there is still some common variations left for subsets of idiosyncratic terms [see for example Connor and Korajczyk (2007)]. In addition, homoscedasticity and the absence of serial correlations in the idiosyncratic components are quite strong assumptions when dealing with stock returns which present dynamic patterns in both time-series and cross-section dimensions.

For this reason, we place our analysis in the approximate factor model framework in the sense of Bai and Ng (2002) which allows for weak time-series and cross-section dependencies and heteroscedasticity in the idiosyncratic components[7]. In particular, D_e is not restricted to be diagonal and the components of e are not supposed to be i.i.d. The idea behind the approximate factor structure in the sense of Bai and Ng (2002) is to suppose that the cross-section and time-series dependencies of the idiosyncratic terms are bounded. They show that as both T and N go to infinity the factors and their loadings can be consistently estimated using the asymptotic principal component method.

[7] Note that, Bai and Ng (2002) approximate factor structure is more general than that of Chamberlain and Rothschild (1983), which focuses only on the cross-section behavior of the model by allowing for weak cross-section dependence.

As in Bai and Ng (2002), we apply the asymptotic principal component method to equation (1) in order to estimate the factors and their loadings. This method consists in minimizing the sum across cross-section and time-series dimensions of the squared distance between the observed return of stock i and its common component[8] $\lambda_i^k F_t^{k'}$ when an arbitrary number of factors, $k = \min(T, N)$, are allowed in the analysis:

$$V(k) = \min_{\Lambda, F^k} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^k F_t^{k'})^2. \quad (2)$$

If $T > N$ ($T < N$), we can concentrate out F^k (Λ^k) and optimize the expression given in equation (2) under the normalization that $\Lambda^{k'} \Lambda^k / N = I_k$ ($F^{k'} F^k / T = I_k$). In this case, multiplying by $\sqrt{N}$ ($\sqrt{T}$) the eigenvectors corresponding to the k largest eigenvalues of the $N \times N$ ($T \times T$) covariance matrix $X^{'}X$ ($XX^{'}$) yields the estimated factor loading matrix $\tilde{\Lambda}^k$ (the estimated latent factor matrix $\tilde{F}^k$). Then, $\tilde{F}^k$ ($\tilde{\Lambda}^k$) is obtained as: $\tilde{F}^k = X\tilde{\Lambda}^k / N$ ($\tilde{\Lambda}^{k'} = \tilde{F}^{k'} X / T$).

The solution of the optimization program in (2) yields k estimated factors together with their loadings. However, $k = \min(T, N)$ is arbitrary and only a few (r) factors are statistically significant. In the Bai and Ng (2002) framework, the estimation of r is set up as a model selection problem. The higher the number of factors, the better the model fit [the lower the quadratic distance $V(k)$ as computed using equation (1)]. On the other hand, when the number of factors included in the analysis is too large, the model is overfitted which increases

[8] Note that, λ_i^k is the i th vector of length k of the $N \times k$ matrix Λ^k. On the other hand, F_t^k represents the $1 \times k$ vector corresponding to the t th time-series unit of F.

estimation errors. To deal with this trade-off between good fit and parsimony, Bai and Ng (2002) propose some information criteria of form:

$$V(k, F^k) + kg(N, T), \tag{3}$$

with $V(k, F^k)$ being the quadratic distance when k factors are included in the analysis and $g(N, T)$ being the penalty for overfitting. The criteria given in (3) is computed for different values of k; the estimated number of factors $(k = \tilde{r})$ is the one that minimizes (3). Bai and Ng (2002) show that under the approximate factor structure assumptions, as both T and N go to infinity, the estimated number of factors $\tilde{r}$ converges toward the true number[9], r. In particular, the authors propose two sets of criteria, the information criteria ($IC_i, i = 1,2,3$) and the panel information criteria ($PC_i, i = 1,2,3$). Although asymptotically equivalent, these criteria have different features in finite samples. Monte Carlo simulations performed by Bai and Ng (2002) show that IC criteria outperform PC criteria in terms of finite sample properties. In addition, Darolles and Mero (2011) use Monte Carlo simulations to assess the finite sample properties of the IC_i ($i = 1,2,3$) for data configurations similar to those considered in this paper (i.e., large cross-section but relatively small time series dimensions). Their results confirm Bai and Ng (2002) showing that IC_2 outperforms IC_1 and IC_3 criteria in inferring the appropriate number of factors. Therefore, we use IC_2 criteria to estimate the number of latent factors:

$$IC_2(k) = \ln(V(k, \tilde{F}^k)) + k\,\frac{N+T}{NT}\ln C_{NT}^2. \tag{4}$$

[9] Note that the Bai and Ng (2002) criteria depend on both dimensions. As discussed by the authors, the usual AIC and BIC criteria, which are functions of N or T alone, do not work when both dimensions of the panel are large.

In this equation, $C_{NT}. = \min(\sqrt{N}, \sqrt{T})$, $\tilde{F}^k$ is the matrix of estimated common factors when k factors are included in the analysis, $V(k, \tilde{F}^k) = N^{-1}\sum_{i=1}^{N} \hat{\sigma}_i^2$ and $\hat{\sigma}_i^2 = \tilde{e}_i'\tilde{e}_i / T$.

II.2 Economic interpretation of common variations of stock returns

The first step of our approach allows us to estimate the latent factors and their number. These factors span the covariance structure of the data which minimizes the measurement error issue. However, they do not provide direct economic explanation. It would be then useful to link these latent factors to the observed variables in order to facilitate the economic interpretation of stock return comovements. Using a set of observed factors as proxies of the latent covariance structure of returns is standard in the literature[10]. However, the choice of the observed factors is arbitrary which result in important measurement errors when estimating the market betas [see for example Shanken (1992)]. Following Darolles and Mero (2011), instead of selecting in advance which observed factors to use, we apply Bai and Ng (2006) framework in order to match the latent factors with the observed variables. This step is essential to identify which economic forces drive stock returns since it allows us to identify the relevant observed factors which match the best the covariance structure of returns.

Let G_t be the $m \times 1$ vector of observed variables at time t, with m being the number of observed variables considered in the analysis. The question of interest is if its elements are generated by (or are linear combinations of) the r latent factors F_t that underlie the common variations of stock returns. We here consider an approximate relation between the observed and latent factors:

[10]For example, in the CAPM analysis equity index returns are used as proxies of the unobserved theoretical market factor. Chen et al. (1986) use a set of macroeconomic variables to capture return covariances while Fama and French (1993) augment the standard CAPM by introducing to firm-characteristic variables related to size and book to market effects.

$$G_{jt} = \delta^{'} F_t + \varepsilon_{jt}, \qquad t = 1,...,T, \tag{6}$$

where $\varepsilon_{jt} \rightarrow N(0, \sigma_{\varepsilon}^2(j))$.

As discussed in the previous subsection, Bai and Ng (2002) show that the latent factors and their number can consistently be estimated when the simple size is large in both the cross-section and the time series dimensions[11]. Then, Bai and Ng (2006) develop some criteria to match the observed variables with the estimated latent factors. Let $\tilde{F}_t$ be the $\tilde{r} \times 1$ vector of the estimated latent factors at time t ($t = 1,...,T$), with $\tilde{r}$ being the estimated risk dimension. Equation (6) can then be rewritten as:

$$G_{jt} = \gamma^{'} \tilde{F}_t + u_{jt}, \qquad t = 1,...,T, \tag{7}$$

where $u_{jt} \rightarrow N(0, \sigma_u^2(j))$. We denote $\tilde{G}_{jt} = \tilde{\gamma}_j^{'} \tilde{F}_t$, where $\tilde{\gamma}_j$ is obtained by applying OLS regressions to equation (7). As discussed by Bai and Ng (2006), the residuals $\tilde{\varepsilon}_{jt}$ ($\hat{\varepsilon}_{jt} = G_{jt} - \tilde{G}_{jt}$) represent the measurement errors, even though they might result from systematic differences between F_t and G_{jt}.

Two statistics are proposed by Bai and Ng (2006) to assess whether the observed variables match the estimated latent factors:

$$NS(j) = \frac{\tilde{\text{var}}(\tilde{\varepsilon}(j))}{\tilde{\text{var}}(\tilde{G}(j))}, \tag{7}$$

$$R^2(j) = \frac{\tilde{\text{var}}(\tilde{G}(j))}{\tilde{\text{var}}(G(j))}, \tag{8}$$

[11]The rate of convergence and the limiting distributions for the estimated factors, factor loadings and common components, estimated by the principal component method (PCA) was developed by Bai (2003).

where a consistent estimate of $\tilde{\text{var}}(\tilde{G}(j))$ is given by:

$$\frac{1}{N}\sum_{i=1}^{N}\tilde{\gamma}_j^{'}\tilde{V}^{-1}\tilde{\Gamma}_t\tilde{V}^{-1}\tilde{\gamma}_j. \tag{9}$$

In this equation, $\tilde{V}$ is a $\tilde{r}\times\tilde{r}$ diagonal matrix consisting of the $\tilde{r}$ largest eigenvalues of sample covariance matrix $XX^{'}/NT$, and $\tilde{\Gamma}_t$ is a consistent estimate of $\Gamma_t = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}E(\lambda_i\lambda_j^{'}e_{it}e_{jt})$. To allow for serially correlated errors e_{it}, $\tilde{\Gamma}_t$ is given as follows:

$$\tilde{\Gamma}_t = \frac{1}{N}\sum_{i=1}^{N}\tilde{e}_{it}^2\tilde{\lambda}_i\tilde{\lambda}_i^{'}, \tag{10}$$

where $\tilde{\lambda}_i$ and $\tilde{e}_{it}$ are respectively the factor loadings and the idiosyncratic errors obtained in subsection II.1.

The statistics given in (7)-(8) consider one observed variable G_j at a time. The $NS(j)$ statistic can be assimilated to a noise-to-signal ratio. If $G_j = \delta^{'}F_t$, the population value of $NS(j)$ is zero. On the other hand, a large $NS(j)$ means that there is an important gap between the observed factor j and the latent factors. Concerning $R^2(j)$, the higher this statistic, the higher the adequacy between the observed and the latent factors. In the following section we show how we use these criteria to identify the relevant observed factors and apply this procedure to the French and US stock returns.

III. The data

III.1 The data panels

In this paper, we focus on the French and the US stock market samples which consist of the monthly returns of individual stocks listed in SBF250 and S&P500 indexes, respectively. The sampling period extends from December 1999 to December 2008. We apply our approach described in section II to the covariance matrix of individual stock returns as well as portfolio stock returns for each market. We create portfolios of SBF250 and S&P500 stocks sorted on market value ME and book to market equity BE/ME, as in Fama and French (1993).

More precisely, in June of each year t, we rank SBF250 (S&P500) stocks by increasing order of their ME value as measured at the end of June. Then, we allocate the stocks into four portfolios for the SBF250 and into five portfolios for the S&P500 using as breakpoints the quartile and quintile size values, respectively. Independently, we form[12] four quartile (five quintile) portfolios based on the BE/ME values of SBF250 (S&P500) stocks. The BE/ME value is obtained by dividing the book common equity for the fiscal year ending in calendar year $t-1$ by the end-of-December $t-1$ market equity value. The intersection of ME-sorted and BE/ME-sorted portfolios yields 16 size-value-sorted portfolios for the SBF250 and 25 for the S&P500. We compute the value-weighted monthly returns on these portfolios from July of t to June of $t+1$. Repeating this procedure each June, we obtain two panels of portfolio returns consisting of 16 time series of 109 observations for the French market and 25 time series of 109 observations for the US market. Market equity and book equity time series for each individual stock of the French and US market samples are extracted from Datastream.

[12] The BE/ME-sorted portfolios are constructed in the same way as the ME-sorted portfolios.

III.2 The observed factors

Concerning the observed variables, we consider:

i) The *market* factor, which is represented by value-weighted stock indexes – SBF250 and S&P500 – extracted from Datastream;

ii) The *credit spread* as measured by the difference between BAA and AAA bond returns. Concerning the US market the BAA and AAA bond yields are obtained from Bloomberg. Concerning the French market the AAA-bond yield is replaced by the 10-year government bond yield. The data for the French market are taken from Datastream;

iii) The firm-characteristic variables of Fama and French (1993) consisting of the *SMB* (small minus big) factor which is a proxy of risk related to firm size, and the *HML* (high minus low) factor which is supposed to capture risk due to book to market equity ratio. Both factors are obtained using the methodology of Fama and French (1993). More precisely, in June of each year t, we rank all SBF250 (S&P500) stocks by increasing order of their size as measured by the ME values observed at the end of June. We then split the SBF250 (S&P500) stocks into two groups, small and big (S and B) using as breakpoint the median size. Similarly, we allocate SBF250 (S&P500) into three book to market equity portfolios based on the breakpoints of the bottom 30% (*Low or L*), middle 40% (*Medium or M*), and top 30% (*High or H*) of the ranked values of BE/ME[13]. Then, the intersection of the two ME and the three BE/ME groups yields six portfolios (S/L, S/M, S/H, B/L, B/M, B/H). We compute the value-weighted monthly returns on these portfolios from July of t to June of $t+1$ and reform the portfolios in June of $t+1$. Finally, the time series of these six portfolios are used to compute

[13] The reason for constructing three BE/ME portfolios and only two ME portfolios in Fama and French (1993) comes from evidence in Fama and French (1992a) that book to market equity has a stronger role in average stock returns than size.

the SMB and HML factors. Each month, the SMB portfolio is obtained by the difference between the average of the returns on the three small-stock portfolios (S/L, S/M, S/H) and the average of the returns on the three big-stock portfolios (B/L, B/M, B/H). Similarly, the HML portfolio is measured by the average of the returns on the two high-BE/ME portfolios (S/H, B/H) less the average of the returns on the two low-BE/ME portfolios (S/L, B/L).

iv) The *UMD* factor of Carhart (1997) which captures the risk related to the momentum anomaly affecting stock returns. Each month, the stocks of the SBF250 (S&P500) index are ranked on the average returns of the past 11 months. We then allocate the SBF250 (S&P500) stocks into three portfolios using as breakpoints the bottom 30% (*Losers*), middle 40% (*Medium*) and top 30% (*Winners*) of the ranked values of the past performances. The UMD factor is then the difference between the simple average of returns on the *Winner* and *Loser* portfolios.

IV. Empirical Applications

IV.1 Estimating the number of latent factors that drive stock returns

In this section we apply Bai and Ng (2002) criteria to French and US stock markets in order to estimate the risk dimension. In particular, we consider four panels of data consisting of: *i)* the individual stock returns belonging to SBF250; *ii)* the individual stock returns listed in the S&P500; *iii)* the time series of the 16 ME-BE/ME-sorted portfolios for the French market; *iv)* the time series of the 25 ME-BE/ME-sorted portfolios for the US market.

For each of the four panels, we first begin by performing all the tests of this study using the entire sampling period extending from December 1999 to December 2008. Then, we isolate two subperiods; the first one, extending from December 1999 to September 2003 (i.e., 46 observations), corresponds to stock market meltdown period due to the collapse of the Internet and Telecom bubbles. The second subperiod goes from October 2003 to Mai 2007 (i.e., 44 observations) and represents a strong market growth due to economic recovery, low interest rates and leverage. Note that, we have excluded from the second subperiod the 19 last observations[14] (June 2007 to December 2008) in order to avoid the impact of subprime crisis in the analysis of this time interval. The subprime meltdown period is not enough large at the monthly frequency to ensure the finite sample convergence of the tests presented in section II. For this reason we focus only on the two time intervals mentioned above corresponding to market distress and growth periods, respectively.

		Overall period		Subperiod 1		Subperiod 2	
		N	T	N	T	N	T
France	Stocks	188	109	188	46	215	44
	Portfolios	16	109	16	46	16	44
USA	Stocks	448	109	448	46	476	44
	Portfolios	25	109	25	46	25	44

Table1. *T* and *N* evolutions across markets, periods and type of assets

Let T be the considered period length. Returns are standardized previously within the considered period. Let X be the T by N matrix of the asset returns for a given panel, such that the i^{th} column is the time series of the i^{th} cross section. We exclude from the analysis stocks with missing observations for a given period. Table 1 gives the different values that can be taken by T and N for a considered market at a given period and for a given type of

[14] Note that the last 19 observation points are included in the overall period analysis.

asset (individual stock or portfolio). Crossing different periods to different countries and assets yields 12 data configurations[15]. Note that, when dealing with panels of individual stock returns $N > T$ for both markets. When using portfolio returns we always get $N < T$. For each panel of data we use asymptotic principal component analysis to estimate the latent factors $\tilde{F}$ from the covariance matrix of the data, as described in subsection II.1.

	France			USA		
	Overall period	Sub-period 1	Sub-period 2	Overall period	Sub-period 1	Sub-period 2
cp1	0.28	0.26	0.21	0.25	0.21	0.16
cp2	0.07	0.09	0.05	0.08	0.11	0.08
cp3	0.03	0.05	0.04	0.05	0.06	0.06
cp4	0.03	0.04	0.03	0.03	0.04	0.04
cp5	0.03	0.03	0.03	0.03	0.03	0.04

Table 2. Eigen values of the 5 first principal components of the covariance matrix of stock returns

	France			USA		
	Overall period	Sub-period 1	Sub-period 2	Overall period	Sub-period 1	Sub-period 2
cp1	0.70	0.67	0.63	0.71	0.63	0.64
cp2	0.09	0.11	0.07	0.09	0.14	0.07
cp3	0.04	0.04	0.05	0.04	0.04	0.05
cp4	0.03	0.04	0.04	0.02	0.03	0.03
cp5	0.02	0.03	0.03	0.02	0.02	0.03

Table 3. Eigen values of the 5 first principal components of the covariance matrix of portfolio returns

[15] A given data panel is a combination of a given market (French or US) at a given period (overall period, subperiod 1, or subperiod 2), and for a given type of asset (individual stock or ME-BE/ME portfolios).

We then report in Tables 2 and 3 the eigen values of the 5 first principal components for each case. Generally speaking, the first latent factor captures an important part of common variations of stock and portfolio returns for the French as well as the US markets. Obviously, the covariance structure of asset returns is more concentrated when dealing with portfolio instead of stock returns for any considered period and for both countries. For example, for the overall period, the eigen value of the first common component varies from 0.28 (0.25) to 0.70 (0.71) for the French (US) market. As discussed by Zhang (2009), this is due to a larger degree of diversification – and thus higher common risk concentration – of portfolio returns relative to individual stock returns. In addition, the eigen values of the two first principal components appear to be higher for subperiod 1 relative to subperiod 2 for individual stock as well as portfolio returns, and for both French and US markets. For example, for US individual stocks the two most important eigen values vary from 0.21 and 0.11 to 0.16 and 0.08, respectively. As it will be discussed latter, these results suggest that the multidimensional property of common risk affecting asset returns becomes stronger in distressed period (subperiod 1) relative to the economic growth time interval (subperiod 2). However, in both subperiods, the eigen value of the second principal component is much lower as compared to the first common component when dealing with portfolio instead of individual stock returns.

	Stocks		Portfolios	
	France	USA	France	USA
Overall period	2	3	1	1
Subperiod 1	2	3	1	2
Subperiod 2	1	2	1	1

Table 4. *IC_2* values for individual stock and portfolio returns

Finally, we calculate the Bai and Ng (2002) IC_2 criterion given in equation (4) to estimate the risk dimension $\tilde{r}$ for the 12 data configurations. The results are shown in Table 4. When considering the individual stock returns for the overall period, the number of relevant latent factors in the sense of Bai and Ng (2002) is 2 for the French market and 3 for the US market. These results highlight the CAPM failure in explaining stock returns. On the other hand, the subperiod analysis points out some common risk structure heterogeneity across distressed and growth subperiods (subperiods 1 and 2, respectively); the decreasing market period is characterized by higher risk dimension for both countries [$IC_2(FR) = 2$ and $IC_2(US) = 3$] as compared to the growth time interval [$IC_2(FR) = 1$ and $IC_2(US) = 2$]. Working with portfolio returns lowers the number of the estimated latent factors for all the periods considered here and for both markets. The number of latent factors exceeds 1 only for the distressed subperiod in the US market. This result is coherent with the first principal component analysis: when dealing with diversified assets, such as ME-BE/ME-sorted portfolio returns, the systematic risk is concentrated in the first principal component.

IV.2 Economic interpretation of common variations of stock returns

Once the latent factors and their number estimated, we implement tests proposed by Bai and Ng (2006) in order to match observed risk factors with estimated latent variables. We considered the following five factors ($m = 5$): (1) *Market* represented by SBF250 (or S&P500) index; (2) *SMB* (small minus big): the spread between small and big capitalizations; (3) *HML* (high minus low): the spread between high and low book to market stocks; (4) *UMD*: the short-term reversal factor; (5) *Premium*: the credit spread between BAA and AAA bond returns.

Let G be the $T \times m$ vector of the observed factors. For each observed factor G_j ($j = 1,...,m$), we calculate $NS(j)$ and $R^2(j)$ as described in subsection II.2. Results concerning individual stocks and portfolio returns are reported in Tables 5 and 6, respectively.

		France			USA		
		Overall period	Sub-period 1	Sub-period 2	Overall period	Sub-period 1	Sub-period 2
NS(j)	Market	0.15	0.15	0.23	0.11	0.14	0.12
	SMB	7.41	6.15	6.82	3.72	5.58	2.78
	HML	0.70	0.60	67.54	0.91	0.35	2.97
	UMD	1.61	0.75	361.02	0.94	0.57	6.04
	Premium	50.53	54.31	291.39	48.00	27.75	33.80
R^2(j)	Market	0.87	0.87	0.81	0.90	0.88	0.96
	SMB	0.12	0.14	0.13	0.21	0.15	0.26
	HML	0.59	0.63	0.01	0.52	0.74	0.25
	UMD	0.39	0.57	0.00	0.52	0.64	0.14
	Premium	0.02	0.02	0.00	0.02	0.03	0.03

Table 5. *NS(j)* and *R²(j)* criteria for individual stock returns

Darolles and Mero (2011) perform Monte Carlo simulation in order to understand the finite sample properties of the $NS(j)$ criteria. Their results suggest that a given observed factor can have $NS(j)$ values up to 15 and still be a relevant factor when considering an approximate relation between the observed and the latent variables. Note that, this value is much lower that those of the irrelevant factors. Following the authors, we consider that factors having $NS(j)$ and $R^2(j)$ values inferior to approximately 15 and 0.10 respectively are relevant. Of course the lower (higher) the $NS(j)$ ($R^2(j)$) values, the higher the degree of adequacy between the j^{th} observed factor and the latent common structure of the asset returns.

On the one hand, when dealing with overall sample period the results appear to be quite homogenous between the two markets. *Market*, *SMB* and *UMD* factors are relevant for either French or US markets independently of the considered type of asset while the *Premium* factor is always rejected by the data. Moreover, the HML factor matches the covariance structure of individual stock returns in both markets but becomes irrelevant when dealing with portfolio returns, especially for the US. This suggests that risk related to value effect has stock-specific patterns in US but remains common across securities in France.

		France			USA		
		Overall period	Sub-period 1	Sub-period 2	Overall Period	Sub-period 1	Sub-period 2
NS(j)	Market	0.16	0.19	0.26	0.12	0.11	0.12
	SMB	7.95	6.05	5.85	7.91	2.50	2.86
	HML	10.63	2.41	90.34	324.49	0.41	32.52
	UMD	2.07	0.74	465.58	1.89	0.67	32.22
	Premium	45.94	45.75	184.12	53.38	3417	69.33
R^2(j)	Market	0.86	0.84	0.79	0.89	0.90	0.89
	SMB	0.11	0.14	0.15	0.11	0.29	0.26
	HML	0.09	0.29	0.01	0.00	0.71	0.03
	UMD	0.33	0.58	0.00	0.35	0.60	0.03
	Premium	0.02	0.02	0.00	0.02	0.00	0.01

Table 6 . *NS(j)* and *R^2(j)* criteria for individual stock returns

On the other hand, results concerning the subperiod analysis of individual stock returns highlight significant variations in the common risk profile across subperiods 1 and 2 for both countries. If *Market* and *SMB* remain significant all the time, *HML* and *UMD* factors are significant essentially in the distressed period (subperiod 1). In the presence of market growth, they remain significant only in the US. These results suggest that the risk related to value and momentum characteristics exhibit important crisis effects in France, while it remains significant all the time in the US. These observations seem consistent with the favorable results obtained by the Fama and French (1993) model in the US and those much less convincing when the model is tested in other countries or on specific subperiods. The *Premium* factor appears to be irrelevant.

In addition, subperiod analysis of portfolio returns yields similar results with the time interval investigation of individual stock returns. Exception is made for *HML* and *UMD* factors relative to the US market. If they are always significant when considering individual stock returns, *HML* and *UMD* exhibit important crisis effects in the portfolio analysis. It follows, that they seem to be stock specific in growth markets but become highly cross-correlated among assets in periods of market meltdown.

Finally, the results reported here highlight two main remarks:

i) The common risk affecting asset returns is multidimensional for both countries and does not depend on the period of study. At least 2 factors, *Market* and *SMB,* correctly match the covariance structure of the data independently of the considered data configuration. This confirms the failure of the CAPM in explaining stock returns.

ii) The subperiod analysis suggests that the common risk profile of stock returns is time-varying and depends on either the period characteristics or the country under study. It follows that account should be taken not only of the beta time-variability but also of the fact that the set of relevant factors is likely to change in time as a function of market characteristics. Moreover, the non-significance of the usual factors of Fama and French (1993) as well as Carhart (1997), in non-crisis periods, should prompt us to seek more robust explanatory factors.

V. Concluding Remarks

In this paper, we use a cross-sectional approach to get a deeper comprehension of the common risk profile of stock returns. Instead of employing static and *ad hoc* factor selection procedure an in Fama and French (1993), we use asymptotic developments of Bai and Ng (2002, 2006) to select the relevant factors. We thus reconcile two methodologies: the statistic one and the other, founded on the observed factors. We apply our approach to French and US stock markets over the period 1999 to 2008 and test the performance of several traditional observed factors, such as credit spread and firm-characteristic variables of Fama and French (1993) and Carhart (1997).

Moreover, focusing on French and US data allows us to provide additional insights concerning risk variability across the two countries. By using subperiod analysis to differentiate crisis period and growth period, we suggest that the common risk profile of stock returns is time-varying and depends on either the period characteristics or the studied country. It seems evident that the set of relevant factors is likely to change in time as a function of market characteristics. More precisely, this observed fact concerns classical factors, such as SMB, HML and UMD, whose risk premiums appear to be significant during the whole period but vanish when we focus on the growth subperiod. Thus, our results put in evidence that stock return commonalities exhibit dynamic patterns both, through time and across the two countries under study.

Finally, it would be interesting to extend our analysis by integrating, for several countries, additional macroeconomic factors, such as inflation rate, currency fluctuations... Then, using this extended set of observable factors, we can apply the approach herein proposed to replicate the latent common structure of stock returns. This approach captures the dynamics of asset risk profile since only the relevant observed factors for a given subperiod are included in the analysis. This point is out of the scope of this paper and is part of current research.

References

Andersen, T.W. (1984) *An Introduction to multivariate statistical analysis,* New York Wiley.

Bai, J. (2003) Inferential theory for factor models of large dimensions, *Econometrica,* **71**, 135-171.

Bai, J. and Ng, S. (2002) Determining the number of factors in approximate factor models, *Econometrica*, **70**, 191-221.

Bai, J. and Ng S. (2006) Evaluating latent and observed factors in macroeconomics and finance, *Journal of Econometrics*, **131**, 507-537.

Banz, R.W. (1981) The Relationship between return and market value of common stocks, *Journal of Financial Economics*, **9**, 3-18.

Basu, S. (1983) The Relationship between earning yield, market value and return for NYSE common stocks: further evidence, *Journal of Financial Economics,* **12**, 51-74.

Bhandari L.C. (1988) Debt/equity ratio and expected common stock returns: empirical evidence, *Journal of Finance,* **43**, 507-528.

Black, F. (1972) Capital market equilibrium with restricted borrowing, *Journal of Business*, **45**, 444-454.

Carhart, M. (1997) On persistence in mutual fund performance, *Journal of Finance*, **52**, 57-92.

Chamberlain, G. and Rothschild, M. (1983) Arbitrage, factor structure and mean-variance analysis in large asset markets, *Econometrica,* **51**, 1305-1324.

Chan, K.C., Chen, N. and Hsieh, D.A., (1985) An explanatory investigation of the firm size effect, *Journal of Financial Economics* **14**, 451-471.

Chen, N., Roll, R. and Ross, S. (1986) Economic forces and the stock market, *Journal of Business*, **59**, 383-403.

Connor, G. and Korajczyk, R.A. (1986) Performance measurement with the APT, *Journal of Financial Economics*, **15**, 373-394.

Connor, G. and Korajczyk, R.A. (1988) Risk and return in the equilibrium APT: Application to a new test methodology, *Journal of Financial Economics*, **21**, 255-289.

Connor, G. and Korajczyk, R.A. (1993) A test for the number of factors in an approximate factor model, *Journal of Finance,* **48**, 1263-1291.

Connor, G. and Korajczyk, R.A. (2007) Factor models of asset returns, *Working paper*.

Darolles, S. and Mero G. (2011) Hedge fund returns and factor models: a cross-section approach, *Bankers Markets & Investors*, **112**, 34-53.

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

Fama, E. F. and French, K. R. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, **33**, 1, 3-56.

Jagannathan, R. and Wang, Z. (1996) The conditional CAPM and the cross-section of expected returns, *Journal of Finance,* **51**, 3-54.

Lehman, B. and Modest, D. (1988) The empirical foundations of the arbitrage pricing theory, *Journal of Financial Economics,* **21**, 213-254.

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

Reingtaum, M. R. (1981) Misspecification of capital asset pricing: empirical anomalies based on earning yields and market values, *Journal of Financial Economics,* **9**, 19-46.

Roll, R. and Ross, S.A. (1980) An empirical investigation of the arbitrage pricing theory, *Journal of Finance,* **35**, 1073-1103.

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

Shanken, J. (1985) Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics*, **14**, 327-348.

Shanken, J. (1992) On the estimation of beta pricing models, *Review of Financial Studies,* **5**, 1-33.

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

Zhang, C., (2009) On the explanatory power of firm-specific variables in cross-sections of expected returns, *Journal of Empirical Finance,* **16**, 306-317.