“A Robust Conditional Realized Extended 4-CAPM”

Christophe Hurlina, Patrick Kouontchoub, Bertrand Mailletc,*

*a University of Orléans (LEO)  
b Variances and University of Paris-1 (CES/CNRS)  
c A.A. Advisors-QCG (ABN AMRO), Variances and University of Paris-1 (CES/CNRS and EIF)

Abstract

In this paper we present and extend the approach of Bollerslev and Zhang (2003) for “realized” measures and co-measures of risk in some classical asset pricing models, such as the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and the Arbitrage Pricing Theory (APT) model by Ross (1976). These extensions include higher-moments asset pricing models (see Jurczenko and Maillet, 2006), conditional asset pricing models (see Bollerslev et al., 1988, and Jondeau and Rockinger, 2004). Estimations are conducted using several methodologies aiming to neutralize data measurement and model misspecification errors (see Ledoit and Wolf, 2003 and 2004), properly dealing with inter-relations between financial assets in terms of returns (see Zellner, 1962), but also in terms of higher conditional moments (see Bollerslev, 1988).

JEL Classification: C3; C4; C5; G1

Key words: Realized Loadings, CAPM, Multifactor Pricing Models, High Frequency Data, Robust Estimation

*Corresponding author. B. Maillet, MSE, CES/CNRS, 106-112 Bd de l’Hôpital F-75647 Paris Cedex 13. Tél/fax : +33 144078268/70.

Email addresses: christophe.hurlin@univ-orleans.fr (Christophe Hurlin), patrick.kouontchou@univ-paris1.fr (Patrick Kouontchou), bmaillet@univ-paris1.fr (Bertrand Maillet)

February 23, 2009
Abstract

In this paper we present and extend the approach of Bollerslev and Zhang (2003) for “realized” measures and co-measures of risk in some classical asset pricing models, such as the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and the Arbitrage Pricing Theory (APT) model by Ross (1976). These extensions include higher-moments asset pricing models (see Jurczenko and Maillet, 2006), conditional asset pricing models (see Bollerslev et al., 1988, and Jondeau and Rockinger, 2004). Estimations are conducted using several methodologies aiming to neutralize data measurement and model misspecification errors (see Ledoit and Wolf, 2003 and 2004), properly dealing with inter-relations between financial assets in term of returns (see Zellner, 1962), but also in terms of higher conditional moments (see Bollerslev, 1988).

JEL Classification: C3; C4; C5; G1

Key words: Realized Loadings, CAPM, Multifactor Pricing Models, High Frequency Data, Robust Estimation

February 23, 2009
A Robust Conditional Realized Extended 4-CAPM

1. Introduction

The Capital Asset Pricing Model (CAPM) by William Sharpe (1964) and John Lintner (1965) marks the birth of the asset pricing theory; it offers powerful and intuitive predictions about how to measure risk and the relation between expected return and risk. Unfortunately, the empirical record of the model is rather poor. Poor enough to invalidate the way it is used in applications. The CAPM’s empirical problems may reflect theoretical failings, as the result of many simplifying assumptions. But they may also be caused by difficulties in implementing valid tests of the model. In this context, due to recent market database availability, several recent research focus on high-frequency data characteristics (see Voit, 2003) and present applications of traditional low-frequency models on newly available high-frequency databases (see Bollerslev and Zhang, 2003), using robust estimation methodology (see Berkowitz and Diebold, 1998).

Indeed, financial variables exhibit strong peculiarities from leptokurticity and asymmetry, to heteroskedascity and clustering phenomenons. Most of classical financial low-frequency models are based on the close-to-normal hypothesis, that is difficult to sustain when real market conditions are under studies. That is the reason why high-frequency data financial applications deserve special research attention and precaution.

In one hand, most of authors now use information contained in the high-frequency series, because being simply closer to the real process is a valuable information (see Kunitomo, 1992), for building denoised lower-frequency estimates of the pertinent parameters that enter into the representative utility function. In the other hand, high-frequency data obviously contain pure noise that has a negative effect of the accuracy of estimations of financial model parameters (see Oomen, 2002). Whilst robust estimators of first and second moments have already been proposed in the literature (see Berkowitz and Diebold, 1998), generalizations in a four-moment world do not yet exist to our knowledge. Similarly, some attention has been recently paid to the conditional modeling of the asset dependences (see Jondeau and Rockinger, 2003) in a heterogeneous market (see Brock and Hommes, 1998 and Malevergne and Sornette, 2006). Based on these ideas, our aim on this chapter is to present an asset pricing model, encompassing some of the most important characteristics of high-frequency financial returns.

We present hereafter some estimations of “realized” measures and co-measures of risk, in the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and in
the Arbitrage Pricing Theory (APT) model by Ross (1976) using French stock high-frequency data. Bollerslev and Zhang (2003) demonstrate with US equity transaction data, that the “realized” measures and co-measures are more effective measures of the systematic risk(s) in factor models. Contrary to these standard approaches, we include higher-moments asset pricing models (see Jurczenko and Maillet, 2006), and conditional asset pricing models (see Bollerslev et al., 1988 and Jondeau and Rockinger, 2004).

The motivation for the conditional multi-moment asset pricing with heterogeneous market participants comes from three sources. First, from a theoretical perspective, financial economic considerations suggest that betas may vary with conditioning variables, an idea developed theoretically and empirically in a vast literature (that includes, among many others Berkowitz and Diebold, 1998; Andersen et al., 2002 and Bollerslev and Zhang, 2003). Second, from a different and empirical perspective, the financial econometric volatility literature (see Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2002a and Corsi, 2006) provides extensive evidence of wide fluctuations and high persistence in asset market conditional variances, and in individual equity conditional covariances with the market. Thus, even from a purely statistical viewpoint, market betas, which are ratios of time-varying conditional covariances and variances, might be expected to display persistent fluctuations. Third, all the previous contributions only assume a mean-variance strategy; the investor allocates his portfolio among some risky assets and the risk-free asset. The mean-variance criterion implicitly assumes that returns are normal, or at least that higher moments (beyond mean and variance) are not relevant for the asset allocation.

The outline of the chapter is as follows. Section 2 starts with a brief discussion of a general factor pricing model and the notion of realized factor loadings. This section also presents standard summary statistics for the monthly realized portfolio returns and factor loadings for the 43 test portfolios over the 5-year (2002-2006) sample period. Section 3 shows our conditional multi-moment asset pricing model with heterogeneous market participants in the high-frequency context. We also propose a robust estimation procedure for this model. Section 4 details the case of time-varying investment opportunities, examines the consequences of using the proposal model and provides several robustness checks of our main results. Section 5 concludes (with some suggestions for future research).

2. Factor Pricing Models and Realized Co-variations

In this section, we introduce the econometric formulations which are considered for the Realized CAPM. The basic idea is that under suitable assumptions
about the underlying return generating process, the corresponding factor loading(s) may in theory be estimated arbitrarily well through the use of sufficiently finely sampled high-frequency data (Bollerslev and Zhang, 2003). We describe the return generating process (2.1). We then derive the corresponding Realized Loadings (2.2), and finally, we discuss the setup of the actual empirical implementation (2.3).

2.1. Factor Pricing Models

Factor models are amongst the most widely used return generating processes in financial econometrics. They explain co-movements in asset returns as arising from the common effect of a (small) number of underlying variables, called factors. Following common practice in the empirical asset pricing literature, we assume that the underlying discrete-time return generating process is the \( K \)-factor model. Specifically, let us denote by \( R \) the \( N \) portfolios returns, and by \( F \) the \( K \) factors returns. If \( r_f \) is the risk free asset return, excess returns of portfolios and factors are defined respectively as: \( r = R - r_f \) and \( f = F - r_f \). Let \( r_i \) be the excess return of the \( i \)-th asset class of the entire portfolio during a specific time interval. The factor model is then specified by:

\[
    r_i = \alpha_i + \sum_{k=1}^{K} \beta_{ik} f_k + \varepsilon_i,
\]

more compactly:

\[
    r = \alpha + Bf' + e \tag{1}
\]

where \( B = (\beta_1, \beta_2, \cdots, \beta_K) \) is the \((N \times K)\) loadings, \( f = (f_1, f_2, \cdots, f_K) \) is the \((T \times K)\) vector of risk factors, \( \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_N) \) is the \((N \times 1)\) vector of intercepts and \( e = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N) \) is the \((N \times 1)\) random error with mean 0 accounts for the information not captured by the risk factors. Following equation (1), the mean of \( r \) is \( E(r) = \alpha + B[E(f)]' \) and the variance of \( r \) is \( Var(r) = B \Sigma_f B' + \sigma_e^2 \) where \( \Sigma_f \) and \( \sigma_e^2 \) are the covariance matrices of \( f \) and \( e \), respectively. Note that the components \( \beta_{ik} \) of the coefficient vector \( B \) will be zero if the \( i \)-th asset class is not exposed to the \( k \)-th risk factor.

The factor model is an extension of the well-known market model (Sharpe, 1964; Lintner, 1965), which corresponds to one-factor model (market factor, \( r_{M,t} \)) with constant loading:

\[
    r = \alpha + \beta r_M + e \tag{2}
\]

where \( \beta \) is a \((N \times 1)\) vector of market loadings and \( r_M \) is a vector of market portfolio returns.
The motivation for using the risks factors of the market is to include most possible information in the model. We can find several sources of bias which are not taken into account in one-factor model. In particular, the liquidity risk and the operational risks which can come from the heterogeneity of information available on the markets. Since there is almost no strictly independent random variable in our economic world, it is almost certain that any additional variable will, at least locally, add to the explanatory power of the CAPM. This form of factor model is most general and all other model can be derived. The factor loadings are then formally defined (see, Bollerslev and Zhang, 2003) by:

$$B = \frac{\text{cov}(r, f)}{\text{cov}(f, f)}$$ (3)

This form of estimation of factor betas is generally used for the factor model. But the empirical tests conducted with data from financial markets do not generally imply the acceptance of the model as describing correctly the range of expected returns. When regressing returns of individual assets on the market factors, $R^2$ are generally quite low. The regression of the average returns against betas is even worse. Bollerslev and Zhang (2003) propose the use of the realized loadings for estimation of factor loadings.

2.2. Realized Loadings and Co-variations

Recently, for the measure of the asset volatility, the use of high frequency data have been advocated to improve the precision of the estimation: the so-called Realized Volatility (RV) approach proposed in a series of breakthrough papers by Andersen et al. (2002) and Barndorff-Nielsen and Shephard (2002a and 2002b). As for the Realized Volatility approach, the using high frequency data in the computation of covariances and correlations between two assets leads to the analogous concept of Realized Covariance (or Covariation), Realized Correlation and Realized Betas (Bollerslev and Zhang, 2003).

Following Barndorff-Nielsen and Shephard (2002a), Bollerslev and Zhang (2003) and Morano (2007), we suppose that finest sampled $h$-period returns are available. Let $p_{t-1+ih}$ denote the $(N \times 1)$ vector of log transaction prices and let $r_{t-1+ih, h} \equiv p_{t-1+ih} - p_{t-1+(i-1)h}$ denote the $(N \times 1)$ vector of returns for the $i$-th intra-day period on day $t$, for $i = [2, 3, \cdots, 1/h]$, where $N$ is the number of stocks. We assume that $1/h$ is an integer. Then, the corresponding intra-period returns share to the same factor structure as in equation (1) with constant intra-period drift and factor loadings such as:

$$r_{t-1+ih, h} = \alpha_t + \beta_t' f_{t-1+ih, h} + \varepsilon_{t-1+ih, h}$$ (4)
where \( f_{t-1+ih,h} \) denote the \((1 \times K)\) vector of returns of the factors for the \(i\)-th intra-day period on day \( t \) and \( \varepsilon_{t-1+ih,h} \) is the random noise vector with \( E(\varepsilon_{t-1+ih,h}) = 0 \). We have \( r_t = \sum_{i=2}^{1/h} r_{t-1+ih,h} \) and \( \varepsilon_t \equiv \sum_{i=2}^{1/h} \varepsilon_{t-1+ih,h} \).

The temporally aggregated one-period returns defined by equation (4), \( r_t \), is obviously identically equal to \( r \) defined by equation (1). With this intra-day period decomposition, Andersen et al. (2002), Barndorff-Nielsen and Shephard (2002a) and Bollerslev and Zhang (2003) suggest to estimate the daily covariation by taking the outer-product of the observed high-frequency returns over the period, namely the Realized Covariance - of course, the true covariance matrix is not directly observable. However, by the theory of quadratic variations\(^1\), the corresponding Realized Covariance (denoted \( RC \)) is defined by:

\[
RC(r_t, f_t, h) = \sum_{i=2,\ldots,1/h} f_{t-1+ih,h} r_{t-1+ih,h}'
\]  

\((5)\)

Andersen et al. (2002) and Protter (2004) show that this statistic converges uniformly in probability to the true covariance matrix as \( h \to 0 \). Therefore,

\[
RC(f_t, f_t, h) = \sum_{i=2,\ldots,1/h} f_{t-1+ih,h} f_{t-1+ih,h}'
\]  

\((6)\)

is the Realized Volatility of the factors on the day \( t \).

Finally, the conditional betas, or factor loadings, defined in equations (1) and (4) can be estimated by:

\[
\hat{\beta}_{t, h} = \frac{RC(r_t, f_t, h)}{RC(f_t, f_t, h)}
\]  

\((7)\)

\( \hat{\beta}_{t, h} \) is the so-called Realized Betas or Factors Loadings (Bollerslev and Zhang, 2003) for the period \( t \) using intra-daily sample \( h \). The next subsection presents some statistics of the Realized Betas of stocks returns data on the French Stock Market.

2.3. Realized Betas on the French Stock Market

Following the definition of the realized loadings proposed in the previous subsection, we compute here the empirical counterpart on real stocks. We firstly present the characteristics of database and secondly, we present and discuss the factor loadings of some stocks in the French Stock Market.

\(^1\)See Protter, 2004, for a general discussion
The data set was obtained from Euronext and consists in transaction prices at the five-minutes sampling frequency for 50 more liquids stocks, covering the period from January 02, 2002 until December 29, 2006 (1,281 trading days). We remove stocks for which the price series start at a later date, leaving 43 stocks for the analysis\(^2\).

Table 1 gives a statistics summary of the selected stocks. These statistics cover minimum and maximum value, mean, standard deviation, skewness, kurtosis and realized volatility (mean and standard deviation of daily realized volatility). We can notice on this table that instantaneous volatilities over all the period are very weak and homogeneous for all entire assets. In the same, the average of the returns is close to zeros. We also see very asymmetrical distributions toward the left (see for example “Pernold-Ricard”, “Sodexho Alliance” and “Air Liquide”) and seldom toward the right-hand side (“Pinault Printemps” and “Vallourec”). Moreover, all the assets have very fat tails distribution (the average kurtosis is higher than 20). These reports show well that it is not possible to support the assumption of normality of the distributions of the returns in the high-frequency context.

The intra-day returns have then been employed in the computation of the daily realized regressions, as discussed in the methodological section, leading to a total of 1,281 daily observations. In order to control the different ranges of variation of the series, standardized variables have been used in realized regression estimations. For obvious reasons (errors or splits), some outliers were also removed from the intra-daily return series.

For the empirical estimation, the 5-min compounded returns on the market portfolio (denoted MKT), are constructed as the logarithmic transform of the equally-weighted percentage returns across all of the stocks. Following the approach in Fama and French (1993), the excess returns for the style portfolio (denoted HML) and the size portfolio (denoted SMB) factor are defined by the return differential between the sorted portfolios grouped according to the individual stocks book-to-market ratios and their market capitalization respectively. We also compute 9 portfolios by merging 3 portfolios according to the size with 3 portfolios according to the style (for instance the portfolio “11” will be computed by the average of the asset returns being in the 1\(^{st}\) third sorted by the market capitalization and 1\(^{st}\) third sorted by book-to-market). In Table 2, summary statistics of the monthly returns for the 9 merged portfolios are presented. We observe that all portfolios has positively average of monthly returns. Also value growth stocks (low book-to-market value) in the small size outperformed growth

\(^2\)See appendix 2 for a list of the company names used.
stocks (high book-to-market value) of the same size. Meanwhile, the average realized volatilities over the recent 5-year period are generally in line with the long-run historical sample standard deviations, with the portfolios in the lowest book-to-market exhibiting the highest average return volatility.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TF1</td>
<td>-7.3%</td>
<td>8.8%</td>
<td>0.5%</td>
<td>0.54</td>
<td>25.35</td>
<td>47%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>EADS</td>
<td>-19.1%</td>
<td>8.5%</td>
<td>0.6%</td>
<td>1.15</td>
<td>56.92</td>
<td>46%</td>
<td>8% 5%</td>
</tr>
<tr>
<td>AIR LIQUIDE</td>
<td>-9.9%</td>
<td>6.1%</td>
<td>0.4%</td>
<td>-1.33</td>
<td>46.54</td>
<td>44%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>CARREFOUR</td>
<td>-9.6%</td>
<td>6.8%</td>
<td>0.5%</td>
<td>-0.22</td>
<td>25.42</td>
<td>47%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>TOTAL FINA ELF</td>
<td>-6.1%</td>
<td>5.4%</td>
<td>0.4%</td>
<td>-0.22</td>
<td>28.05</td>
<td>44%</td>
<td>5% 3%</td>
</tr>
<tr>
<td>OREAL</td>
<td>-5.8%</td>
<td>7.7%</td>
<td>0.4%</td>
<td>0.35</td>
<td>22.97</td>
<td>44%</td>
<td>6% 3%</td>
</tr>
<tr>
<td>VALLOUREC</td>
<td>-12.1%</td>
<td>18.5%</td>
<td>0.6%</td>
<td>1.39</td>
<td>63.34</td>
<td>38%</td>
<td>8% 5%</td>
</tr>
<tr>
<td>ACCOR</td>
<td>-5.2%</td>
<td>5.6%</td>
<td>0.5%</td>
<td>-0.03</td>
<td>16.90</td>
<td>46%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>BOUYGUES</td>
<td>-14.3%</td>
<td>6.8%</td>
<td>0.5%</td>
<td>-0.87</td>
<td>53.41</td>
<td>46%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>SUEZ</td>
<td>-11.5%</td>
<td>10.1%</td>
<td>0.6%</td>
<td>0.28</td>
<td>32.22</td>
<td>46%</td>
<td>8% 6%</td>
</tr>
<tr>
<td>LAFARGE</td>
<td>-7.7%</td>
<td>6.8%</td>
<td>0.5%</td>
<td>-0.18</td>
<td>24.97</td>
<td>44%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>SANOFI SYNTHELABO</td>
<td>-7.8%</td>
<td>6.9%</td>
<td>0.5%</td>
<td>-0.51</td>
<td>26.85</td>
<td>44%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>AXA</td>
<td>-8.6%</td>
<td>9.7%</td>
<td>0.7%</td>
<td>0.10</td>
<td>26.73</td>
<td>47%</td>
<td>9% 7%</td>
</tr>
<tr>
<td>GROUPE DANONE</td>
<td>-7.7%</td>
<td>7.3%</td>
<td>0.3%</td>
<td>0.93</td>
<td>46.99</td>
<td>43%</td>
<td>5% 3%</td>
</tr>
<tr>
<td>PERNOD-RICARD</td>
<td>-20.5%</td>
<td>5.9%</td>
<td>0.4%</td>
<td>-6.25</td>
<td>368.99</td>
<td>42%</td>
<td>5% 3%</td>
</tr>
<tr>
<td>LVMO MOET VUITTON</td>
<td>-4.7%</td>
<td>6.4%</td>
<td>0.5%</td>
<td>0.03</td>
<td>16.50</td>
<td>46%</td>
<td>6% 3%</td>
</tr>
<tr>
<td>SODEXHO ALLIANCE</td>
<td>-22.2%</td>
<td>8.7%</td>
<td>0.6%</td>
<td>-2.74</td>
<td>131.81</td>
<td>47%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>MICHELIN</td>
<td>-5.4%</td>
<td>4.6%</td>
<td>0.5%</td>
<td>0.11</td>
<td>12.85</td>
<td>46%</td>
<td>6% 3%</td>
</tr>
<tr>
<td>THALES</td>
<td>-5.3%</td>
<td>6.2%</td>
<td>0.5%</td>
<td>0.24</td>
<td>15.25</td>
<td>47%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>PINAULT PRINTEMPS</td>
<td>-6.7%</td>
<td>17.4%</td>
<td>0.5%</td>
<td>2.02</td>
<td>85.11</td>
<td>44%</td>
<td>6% 5%</td>
</tr>
<tr>
<td>PEUGEOT</td>
<td>-7.5%</td>
<td>5.7%</td>
<td>0.5%</td>
<td>-0.41</td>
<td>24.42</td>
<td>47%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>ESSILOR INTL</td>
<td>-4.4%</td>
<td>7.3%</td>
<td>0.4%</td>
<td>0.46</td>
<td>18.11</td>
<td>44%</td>
<td>6% 3%</td>
</tr>
<tr>
<td>SCHNEIDER ELECTRIC</td>
<td>-5.5%</td>
<td>4.5%</td>
<td>0.4%</td>
<td>-0.08</td>
<td>16.28</td>
<td>44%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>VIVENDI ENVIRON.</td>
<td>-7.2%</td>
<td>6.3%</td>
<td>0.5%</td>
<td>0.09</td>
<td>23.49</td>
<td>47%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>SAINT-GOBAIN</td>
<td>-12.9%</td>
<td>7.5%</td>
<td>0.5%</td>
<td>-0.70</td>
<td>36.80</td>
<td>46%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>CAP GEMINI</td>
<td>-16.2%</td>
<td>15.8%</td>
<td>0.7%</td>
<td>0.09</td>
<td>42.56</td>
<td>48%</td>
<td>9% 7%</td>
</tr>
<tr>
<td>CANAL +</td>
<td>-6.4%</td>
<td>5.9%</td>
<td>0.5%</td>
<td>-0.15</td>
<td>15.71</td>
<td>32%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>VINCI</td>
<td>-3.3%</td>
<td>5.2%</td>
<td>0.4%</td>
<td>0.46</td>
<td>14.13</td>
<td>42%</td>
<td>5% 3%</td>
</tr>
<tr>
<td>CASINO GUICHARD</td>
<td>-6.6%</td>
<td>6.2%</td>
<td>0.4%</td>
<td>-0.04</td>
<td>21.24</td>
<td>43%</td>
<td>6% 3%</td>
</tr>
<tr>
<td>AGF</td>
<td>-7.3%</td>
<td>11.8%</td>
<td>0.5%</td>
<td>0.57</td>
<td>31.71</td>
<td>44%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>VIVENDI UNIVERSAL</td>
<td>-18.3%</td>
<td>13.7%</td>
<td>0.8%</td>
<td>-1.34</td>
<td>67.74</td>
<td>47%</td>
<td>9% 8%</td>
</tr>
<tr>
<td>DEXIA SICO.</td>
<td>-9.0%</td>
<td>7.6%</td>
<td>0.5%</td>
<td>-0.05</td>
<td>37.45</td>
<td>43%</td>
<td>6% 5%</td>
</tr>
<tr>
<td>STMICROELEC.SICO.</td>
<td>-7.4%</td>
<td>7.4%</td>
<td>0.6%</td>
<td>0.18</td>
<td>22.91</td>
<td>46%</td>
<td>8% 5%</td>
</tr>
<tr>
<td>ALCATEL A</td>
<td>-21.1%</td>
<td>15.4%</td>
<td>0.9%</td>
<td>-0.57</td>
<td>52.82</td>
<td>45%</td>
<td>11% 9%</td>
</tr>
<tr>
<td>LAGARDERE</td>
<td>-4.1%</td>
<td>5.7%</td>
<td>0.5%</td>
<td>0.33</td>
<td>14.34</td>
<td>44%</td>
<td>6% 4%</td>
</tr>
<tr>
<td>VALEÓ</td>
<td>-4.6%</td>
<td>7.7%</td>
<td>0.5%</td>
<td>0.63</td>
<td>18.97</td>
<td>47%</td>
<td>7% 4%</td>
</tr>
<tr>
<td>PUBLICIS GROUPE</td>
<td>-6.8%</td>
<td>8.4%</td>
<td>0.5%</td>
<td>0.30</td>
<td>20.36</td>
<td>47%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>SOCIETE GENERALE A</td>
<td>-7.4%</td>
<td>8.3%</td>
<td>0.5%</td>
<td>0.34</td>
<td>26.68</td>
<td>43%</td>
<td>7% 5%</td>
</tr>
<tr>
<td>BNP PARIBAS</td>
<td>-8.0%</td>
<td>8.0%</td>
<td>0.5%</td>
<td>0.35</td>
<td>27.00</td>
<td>45%</td>
<td>7% 4%</td>
</tr>
<tr>
<td>RENAULT</td>
<td>-6.1%</td>
<td>8.5%</td>
<td>0.5%</td>
<td>0.27</td>
<td>22.64</td>
<td>43%</td>
<td>7% 4%</td>
</tr>
<tr>
<td>FRANCE TELECOM</td>
<td>-16.3%</td>
<td>15.7%</td>
<td>0.8%</td>
<td>-0.79</td>
<td>66.70</td>
<td>47%</td>
<td>9% 8%</td>
</tr>
<tr>
<td>THOMSON</td>
<td>-10.8%</td>
<td>11.1%</td>
<td>0.6%</td>
<td>0.26</td>
<td>31.08</td>
<td>46%</td>
<td>8% 5%</td>
</tr>
<tr>
<td>CREDIT AGRICOLE</td>
<td>-15.0%</td>
<td>7.3%</td>
<td>0.5%</td>
<td>-1.36</td>
<td>63.98</td>
<td>46%</td>
<td>6% 4%</td>
</tr>
</tbody>
</table>

Source: Euronext. The table reports some main statistics of the intra-daily returns. This statistics cover minimum (Min) and maximum value (Max), standard deviation (σ), skewness (Skew.), kurtosis (Kurt.), negative returns frequency (Neg.) and realized volatility (mean and standard deviation of daily realized volatility) over the January 2002 through December 2006 sample period. The realized daily volatilities are reported in standard deviation format, and constructed from the summation of the squared 5-min returns. The mean of all stocks 5-min returns is close to 0.
Table 2: Mean, Realized Volatility for Monthly Returns

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Returns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>0.055</td>
<td>0.068</td>
<td>0.093</td>
</tr>
<tr>
<td>Middle</td>
<td>0.061</td>
<td>0.050</td>
<td>0.084</td>
</tr>
<tr>
<td>Big</td>
<td>0.064</td>
<td>0.060</td>
<td>0.079</td>
</tr>
<tr>
<td>Realized Volatility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>0.43</td>
<td>0.14</td>
<td>0.41</td>
</tr>
<tr>
<td>Middle</td>
<td>0.37</td>
<td>0.18</td>
<td>0.36</td>
</tr>
<tr>
<td>Big</td>
<td>0.78</td>
<td>0.13</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Source: Euronext. The table reports the average monthly percentage returns, realized volatility, average realized factor loadings for the 9 size and book-to-market sorted portfolios over the January 2002 through December 2006 sample period. The realized monthly volatilities are reported in standard deviation format, and constructed from the summation of the squared 5-min returns.

Although the most liquid stocks in our sample generally is traded several times within each 5-min interval, some of the less liquid stocks may not be traded for up to an hour or longer. As such, the 5-min portfolio returns are clearly under the influence of non-synchronous trading effects. It is well known that such effects may systematically bias the estimates of the factor loadings from traditional time series regressions. Several different adjustment procedures have been proposed in the empirical asset pricing literature for dealing with this phenomenon over coarser daily and monthly frequencies (see, e.g., Campbell et al., 1997). In particular, under the simplifying assumption that the trading intensity and the true latent process for the returns are independent, the one-factor CAPM adjustment procedure first proposed by Scholes and Williams (1977) works by accumulating additional leads and lags of the sample auto-covariances between the returns and the market portfolio. Adapting the Scholes-Williams procedure to a multi-factor framework, our empirical variation and covariation measures for the factors and the returns are based on the following generalizations of equations (5) and (6):

\[
RC_L(r_{t,h}, f_{t,h}) = \sum_{j=-L}^{L} \sum_{i=1}^{1/h} f_{t+i,h, f_{t+(i-j)h,h}} - 2Lh \left( \sum_{i=1}^{1/h} f_{t+i,h, h} \right) \left( \sum_{i=1}^{1/h} r_{t+(i-j)h,h} \right)'
\]

where \( L \) denotes the maximum lag length included in the adjustment. The adjusted realized factor loadings are then simply obtained by direct substitution of \( RC_L(r_{t,h}, f_{t,h}) \) and \( RC_L(f_{t,h}, f_{t,h}) \) in place of \( RC(r_{t,h}, f_{t,h}) \) and \( RC(f_{t,h}, f_{t,h}) \) respectively in equation (7)\(^3\).

\(^3\)Note that for \( L = 0 \) the expressions reduce to the standard expressions underlying the conventional unadjusted OLS regression estimates.

10
Figure 1: Average Adjusted Factor Loadings for smaller size and style Portfolio

Source: Euronext. The database includes a complete set of five-minutes transactions prices from January 02, 2002, through December 29, 2006 (1,281 trading days). The figure plots the average Scholes-Williams adjusted factor loadings for smaller size and style Portfolio as a function of the adjustment, $L = [0, 1, 2, \cdots, 10]$. SMB corresponds to the Small Minus Big size factor and HML is the High Minus Low style factor.

We choose a lag window of $L = 6$ in the empirical results reported on below. This particular choice of $L$ was motivated by visual inspection of the adjusted factor loadings, which showed little sensitivity to the choice of $L$ beyond that lag. To illustrate this feature, Figure (1) plots the average factor loadings across the full sample associated with the Market, Small Minus Big (SMB) size and High Minus Low (HML) style factors representing portfolios for the test as a function of the lag length, $L$ and Figure 2 plots the factors loadings with lag 6. The smaller size and style Portfolio consists of the stocks in the first size and first book-to-market value, and as such is among the test portfolios most prone to non-synchronous trading effects. Nonetheless, it is evident from the figure that the average loadings for all three factors stabilize fairly quickly, and is very close to the average full-day lag 59 adjusted loadings for $L$ around 6. A similar picture emerges for all of the other portfolios with the convergence to the full-day average occurring even faster (smaller values of $L$) in most other cases.
3. Conditional Asset Pricing Model

In this section we introduce the econometric formulations that are considered in this chapter. In turn, we describe the return generating process. We derive the regressions tests of the CAPM Model, and finally we propose our framework estimation.

3.1. Realized Conditional Asset Pricing Model with Higher Moments

We include in the return generating process the third moment and also the intra-daily variabilities. We first present the conditional asset pricing with three moments and after, we show the corresponding realized measures.
3.1.1. Three-moment Conditional Asset Pricing

Recall \( r \) the \((N \times T)\) matrix of excess returns of \( N \) stocks; \( r = (r_1, r_2, \cdots, r_N) \) and \( f \) the \((K \times T)\) excess return of risky factors; \( f = (f_1, f_2, \cdots, f_K) \). We have:

\[
    r_i = \alpha_i + \sum_{k=1}^{K} \beta_{ik} f_k + \sum_{k=1}^{K} \gamma_{ik} f_k^2 + \varepsilon_i, \tag{9}
\]

where \( f_k^* = F_k^2 - r_f \), is the excess returns of the square returns of factor \( k \), with \( k = [1, \cdots, K] \); \( f_k \) is the \( k \)-th common factor, \( \beta_{ik} \) is the loadings of stock \( i \) on each of the \( k \) factors, \( \gamma_{ik} \) is the sensitivity of stock \( i \) on each of the \( k \) quadratic factors \( f_k^2 \) and \( \varepsilon_i \) are assumed to be uncorrelated across assets.

More compactly:

\[
    r = \alpha + Bf + Gf^2 + e \tag{10}
\]

where \( B \) is \((N \times K)\) matrix of loadings, \( G \) \((N \times K)\) matrix of loadings of the quadratic factors and \( e \) is the \((T \times N)\) matrix of residuals with zero mean and covariance matrix \( \Sigma \).

At equilibrium, the expected return of an asset \( i \) is linearly associated with the contributions of the asset to the variance and skewness of the risk factor. The conditional version of this multi-factor model (Krauss and Litzenberger, 1976; Barone et al., 2002; Smith, 2007) is:

\[
    E(r_i) = \lambda_0 + \sum_{k=1}^{K} (\beta_{ik} \lambda_{1,k} + \gamma_{ik} \lambda_{2,k}) \tag{11}
\]

where \( \lambda_0 \) may be interpreted as the expected excess returns of a portfolio with zero covariance and coskewness with the market, \( \lambda_{1,k} \) is the price of beta risk, \( \lambda_{2,k} \) is the price of gamma risk. The value of the loading factor is given by equation 3 and:

\[
    \gamma_{i,k} = \frac{\text{Coskew}(r_i, f_k)}{\kappa_j^3}, \tag{12}
\]

is the conditional gamma of asset \( i \). The conditional coskewness of the asset \( i \) with the factor \( k \), \( \text{Coskew}(r_i, f_k) \) is given by:

\[
    \text{Coskew}(r_i, f_k) = E[(r_i - \tau_i)(f_k - \overline{f_k})^2], \tag{13}
\]

where \( \tau_i = E(r_i) \) and \( \overline{f_k} = E(f_k) \) and the conditional coskewness of the factor \( k \), \( \kappa_k^3 \) is as follows:

\[
    \kappa_k^3 = E[(f_k - \overline{f_k})^3] \tag{14}
\]

The motivation for including the square of the market returns is to fully account for coskewness with the factors. The quadratic market model (equation
10) is a direct extension of the well-known market model (Sharpe, 1964 and Lintner, 1965), which corresponds to some restrictions. With this return generating process, we assume that the individual error terms are correlated; so we have non-zero covariances between the error terms \( \varepsilon_i \). This special type of problems is called in the econometric literature the Seemingly Unrelated Regressions (SUR) Model (Zellner, 1962). We propose some basic intuition of the estimation procedure and some extensions (Parks, 1967) in the special subsection of this section.

3.1.2. Realized Loadings in Conditional Asset Pricing with Three moments

We develop and extend in this subsection the conditional multi-factor Asset Pricing Model by adding unobserved prices variabilities in the low frequency. Suppose as in the previous section that finer sampled \( h \)-period returns are available. The \( h \)-period three-moment conditional asset pricing is therefore given, in compact form similar to equation 4:

\[
\begin{align*}
\mathbf{r}_{t-1+ih, h} &= \mathbf{\mu}_t + \mathbf{\beta}_t^f \mathbf{f}_{t-1+ih, h} + \mathbf{\gamma}_t^f \mathbf{f}_{t-1+ih, h} + \varepsilon_{t-1+ih, h} \\
\end{align*}
\]

(15)

where a new parameter \( \mathbf{\gamma}_t \) is \((K \times 1)\) vector of conditional gamma at the time \( t \).

The conditional betas are defined by equation 7.

The Realized Coskewness (denoted RCoskew) of assets and risks factors with \( h \)-period sampling is defined (Beine et al., 2004) by:

\[
\text{RCoskew}(\mathbf{r}_{t, h}, \mathbf{f}_{t, h}) = \frac{1}{h} \sum_{i=2}^{1/h} \mathbf{r}_{t-1+ih, h} \mathbf{f}_{t-1+ih, h} = \frac{1}{h} \sum_{i=2}^{1/h} \mathbf{r}_{t-1+ih, h} \left( \mathbf{f}_{t-1+ih, h} - \mathbf{R}_{Ft-1+ih, h} \right)
\]

(16)

Therefore, the Realized Gamma \((\hat{\gamma}_t, h)\) for the day \( t \) is:

\[
\hat{\gamma}_t, h = \frac{\text{RCoskew}(\mathbf{r}_{t, h}, \mathbf{f}_{t, h})}{\text{RCoskew}(\mathbf{f}_{t, h}, \mathbf{f}_{t, h})}
\]

(17)

3.2. Regressions Tests of the CAPM

We describe here two types of the most popular regression tests of CAPM. The first group of tests of the Capital Asset Pricing Model is related to time-series variation (Black et al., 1972 and Gibbons et al., 1989) whilst the second is based on cross-sectional regressions (Fama and Macbeth, 1973).

3.2.1. Times-series Regressions

Recall the factor pricing model with a single factor\(^4\). All variables hereafter are considered in terms of excess returns. The loading for the asset \( i \) is given by the regression:

\[
r_i = \alpha_i + \beta_i f + \varepsilon_i
\]

(18)

\(^4\)For simplify notations, we use here a single factor model; but the results are the same for multi-factor case with vectors in place of scalars.
The model states that expected returns are linear in the betas: \( r_i = \beta_i f \). Then this means that all the regression intercepts \( \alpha_i \) should be zero. Black et al. (1972) and Gibbons et al. (1989) developed a methodology for this test. That we briefly present in the next subsection.

**The Black, Jensen and Scholes (BJS) Method**

Black et al. (1972) suggested a natural strategy for estimation and evaluation. Run time-series regressions (equation 18) for each test. The estimate of the factor risk premium is just the sample mean of the factor: \( \hat{\lambda} = \bar{f} \). Then, for the case that the regression errors in (equation 18) are uncorrelated and homoskedastic, standard Ordinary Least Square (OLS) formula for a distribution of the parameters can be used.\(^5\) We can also test the case that all the pricing errors are jointly equal to zero. With errors that are independent and identically distributed over time, homoskedastic and independent of the factors, the asymptotic joint distribution of the intercepts gives the model test statistic (\( \chi^2 \) test, see Campbell et al., 1997 and Cochrane, 2001),

\[
t_N = T \left[ 1 + (\bar{f} \hat{\sigma}^2_f)^2 \right]^{-1/2} \hat{\alpha} \hat{\Sigma}^{-1/2} \hat{\alpha}
\]

(19)

where \( \bar{f} \) denotes sample mean of factor \( f \), \( \hat{\sigma}^2_f \) denotes sample variance of factor \( f \), \( \hat{\alpha} \) is a vector of the estimated intercepts, \( \hat{\Sigma} \) is the residual covariance matrix: \( \hat{\Sigma} = E(\hat{e}\hat{e}^\prime) \). \( t_N \) converges asymptotically to \( \chi^2_N \).

For the BJS Wald-type test defined by equation (19), since \( \Sigma \) is unknown, its approximation needs the large-sample distribution theory for drawing inference (Campbell et al., 1997 and Cochrane, 2001). Gibbons et al. (1989) give a finite-sample distribution for the test.

**The Gibbons, Ross and Shanken (GRS) Method**

The candidate model for times-series regressions of Gibbons et al. (1989) is the Market Model regression equation (18). Similar to univariate F-test (which is equal to a squared \( t \)-test), its multivariate analogue looks at the ratio between the square of the estimated coefficients and their variance; in matrix notation, this involves the quadratic expression:

\[
\hat{\alpha}^\prime [Var(\hat{\alpha})]^{-1} \hat{\alpha} = T (\hat{\alpha}^\prime \hat{\Sigma}^{-1} \hat{\alpha}) \left[ 1 + (\bar{f} \hat{\sigma}^2_f)^2 \right]
\]

(20)

where \( \Sigma \) is the variance-covariance matrix of the residuals in equation (18). The finite-sample test of Gibbons et al. (1989) is achieved by an appropriate adjustment of the above expression for the degree of freedom. The test statistic \( W \) is

\(^5\)The \( t \)-tests is use to check whether the pricing errors \( \alpha \) are in fact zero.
defined as:

\[ W = (T - N - 1)N^{-1} \left[ 1 + \left( \frac{\overline{f} \sigma_f^{-1}}{\overline{q}} \right)^2 \right]^{-1} \hat{\alpha}' \Sigma^{-1} \hat{\alpha} \]

(21)

They show that, under the null hypothesis, \( W \) is unconditionally distributed as a central \( F \) with \( N \) degrees of freedom in the numerator and \( (T - N - 1) \) degrees of freedom in the denominator\(^6\), that is:

\[ W \sim F_{N, T-N-1} \]

(22)

The GRS test statistic as a useful economic interpretation since Gibbons et al. (1989) show that this statistic can be rewritten in terms of Sharpe ratios:

\[ W = (T - N - 1)N^{-1} \left[ \frac{E(q)^2}{\sigma_q^2} - \frac{\overline{f}^2}{\sigma_f^2} \right] \left\{ 1 + \left[ \frac{\overline{f} \sigma_f^{-1}}{\overline{q}} \right]^2 \right\} \]

(23)

where the portfolio denoted by \( q \) is the ex-post tangency portfolio constructed from the \( N \) assets plus the test factor \( f \). The GRS test statistic \( W \) can therefore be interpreted in terms of the difference in squared Sharpe ratios between \( q \) and \( f \). If \( f \) is ex post efficient, it has the highest Sharpe ratio of all portfolios which can be constructed and it will have the same Sharpe ratio as \( q \). Otherwise, \( q \) can only have a higher Sharpe ratio\(^7\) (Campbell et al., 1997; Cochrane, 2001).

If there are many factors that are excess returns, the quadratic form \( \hat{\alpha}' \Sigma^{-1} \hat{\alpha} \) has the distribution:

\[ \hat{\alpha}' \Sigma^{-1} \hat{\alpha} = (T - N - K)N^{-1} \left( 1 + \hat{\overline{f}' \Omega^{-1} \hat{f}} \right) \]

\[ \sim F_{N, T-N-K} \]

(24)

where \( N \) is the number of asset, \( K \) the number of factors and \( \hat{\Omega} = \frac{1}{T} \sum_{t=1}^{T} (f - \overline{f})(f - \overline{f})' \) is the sample covariance matrix of factors.

3.2.2. Cross-sectional Regressions (Fama and Macbeth, 1973)

The CAPM imposes a cross-sectional restriction, namely that expected asset returns are linearly related to their market beta; it thus seems to be straightforward to test the CAPM by running a cross-sectional regression of expected returns on betas. If we want to test specific alternatives, we could just add additional explanatory variables to the regression. In a nutshell, that is already the basic idea of cross-sectional regressions tests. But we do not know expected returns and betas, we can only estimate their values. When running cross-sectional regressions, we also need to account for potential cross-sectional correlation of

---

\(^6\)See Campbell et al. (1997) and Cochrane (2001) for more details.

\(^7\)We suppose here, for our analysis of the economic perspective of GRS test statistic, that the factor \( f \) is the Market portfolio.
residual returns. While the assumption of zero serial correlation is generally a good approximation for monthly returns, abnormal returns do considerably co-vary across assets. The cross-sectional regression model specified by Fama and MacBeth (1973) have the form for each period $t^8$:

$$ r_{i,t} = \lambda_{0,t} + \lambda_{1,t} \beta_{i,t} + \sum_{j=2}^{J} \lambda_{j,t} k_{i,j,t} + \xi_{i,t} $$

over all assets $i$, with $i = [1, \ldots, N]$, and $T$ observed periods with $t = [1, \ldots, T]$. Where $\lambda_{0,t}$ is the return in excess of the risk free rate on a fully-invested, least-variance portfolio where the effects of the others variables are zero, $\lambda_{1,t}$ is the return of a zero-investment, least-variance portfolio with a beta of one and $\lambda_{j,t}$ (for $j > 2$) is the return on a zero-investment, least-variance portfolio where the variables $k_j$ are zero$^9$.

If the CAPM holds, the expected return of $\lambda_{j,t}$ (for $j \neq 1$) should be zero, the expected return of $\lambda_{1,t}$ should be positive. Using time-series averages as estimates of expected values, we can test whether these are significantly different from zero with standard $t$-tests$^{10}$. We have for $j = [1, \ldots, N]$:

$$
\begin{align*}
\hat{\lambda}_j &= T^{-1} \sum_{t=1}^{T} \hat{\lambda}_{j,t} \\
\hat{\sigma}(\hat{\lambda}_j) &= \left\{ [T(T-1)]^{-1} \sum_{t=1}^{T} (\hat{\lambda}_{j,t} - \bar{\lambda}_j)^2 \right\}^{1/2} \\
\hat{t}(\hat{\lambda}_j) &= \hat{\lambda}_j / \hat{\sigma}(\hat{\lambda}_j)^{-1}
\end{align*}
$$

Cochrane (2001) shows how to embed the Fama-MacBeth method in a more general estimation context.

### 3.3. Estimation Procedure of the Model

For a robust estimation of the proposed model, we combine Seemingly Unrelated Regressions (SUR) Zellner (1962) and cross-sectionally correlated of Parks (1967) methodologies. This because the Seemingly Unrelated Regressions (SUR)

---

$^8$The Fama-MacBeth method allows to incorporate time-varying betas.

$^9$The relation equation (25) insures that there are no other parameters next to beta which explain the cross-section of expected returns. For instance, Fama and MacBeth (1973) test whether expected returns are also determined by variance. They postulate an alternative relationship between expected returns, beta and idiosyncratic risk $\sigma(\varepsilon_{i,t})$ of model defined in equation (18)

$^{10}$The test is formulated as:

$$
\begin{align*}
\hat{\lambda}_0 &= 0 \\
\hat{\lambda}_1 &= > 0 \\
\hat{\lambda}_j &= 0, \text{ for } j > 1
\end{align*}
$$
model, proposed by Zellner (1962), can be viewed as a special case of the
generalized regression model, but it does not share all of the features or problems
of other leading special cases (e.g., models of heteroskedasticity or serial correla-
tion). Parks (1967) proposes an estimation procedure that allows the error term
to be both serially and cross-sectionally correlated. Consider a system of \( N \) si-
multaneous equations, of which the typical \( i \)-th equation defined in equation 1.
More compactly, those latter vectors can be stacked into an \( NT \)-dimensional vec-
tor \( r \), with a corresponding arrangement for the error terms, coefficient vectors,
and regressors:

\[
\begin{pmatrix}
r_1 \\
r_2 \\
\vdots \\
r_N
\end{pmatrix}, \quad \begin{pmatrix}
e_1 \\
e_2 \\
\vdots \\
e_N
\end{pmatrix}, \quad \begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_N
\end{pmatrix},
\]

\( r \), \( e \) and \( B \) are respectively an \(( NT \times 1 )\) vector of returns, \(( NT \times 1 )\) vector of
effects terms and \(( K \times 1 )\) vector of loadings, and:

\[
F = \begin{pmatrix}
X_1 & 0 & \cdots & 0 \\
0 & X_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & X_N
\end{pmatrix},
\]

with \( K = \sum_{i=1}^{N} k_i \) and \( F \) the \(( NT \times K )\) matrix of factors. With this notation,
we have:

\[
E(r) = BF
\]  

(27)

Therefore, given that \( e_{it} \) is the error for the \( i \)-th asset in the \( t \)-th time period,
the assumption of contemporaneous disturbance correlation, but not correlation
over time, implies that the covariance matrix of this SUR system is:

\[
V(r) = \Sigma \otimes I
\]  

(28)

where:

\[
\Sigma = \begin{pmatrix}
\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\
\sigma_{21} & \sigma_{22} & \cdots & \sigma_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN}
\end{pmatrix},
\]

with \( I \) the \(( T \times T )\) identity matrix and \( \otimes \) the *Kronecker product* notation indicating
that each element of \( \Sigma \) is multiplied by an identity matrix.

With the previous notations, the classical Least Squares estimator for the vector
\( B \) is:

\[
\hat{B}_{LS} = (F'F)^{-1}F'r
\]  

(29)
and the GLS estimator (assuming $\Sigma$ is known) is:

$$\hat{B}_{\text{GLS}} = \left[ F'(\Sigma \otimes I)^{-1}F \right]^{-1} F' (\Sigma \otimes I)^{-1} r $$  \hspace{1cm} (30)

Now assume that the elements of the disturbance vector follow an $AR(1)$ process (see Parks, 1967 and Bali and Wu, 2005):

$$e_{it} = \rho_i e_{i,t-1} + \epsilon_{it},$$  \hspace{1cm} (31)

where $|\rho_i| < 1$ and $\epsilon_{it}$ are random variables with zero mean. Equation (31) indicates that the $e_{it}$ are correlated. Equation (1) for asset $i$ can then be written as:

$$r_i = b_i F_i + P_i \epsilon_i$$  \hspace{1cm} (32)

with $e_i = P_i \epsilon_i$, where $e_i$ is a $T \times 1$ random vector with $E(e_i) = 0$, $E(\epsilon_i \epsilon_i') = \sigma_{ii} I$ and:

$$P_i = \begin{bmatrix}
(1 - \rho_i^2) - 1/2 & 0 & 0 & \cdots & 0 \\
\rho_i(1 - \rho_i^2) - 1/2 & 1 & 0 & \cdots & 0 \\
\rho_i^2(1 - \rho_i^2) - 1/2 & \rho_i & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\rho_i^{T-1}(1 - \rho_i^2) - 1/2 & \rho_i^{T-2} & \rho_i^{T-3} & \cdots & 1 
\end{bmatrix}$$  \hspace{1cm} (33)

Under this setup, Parks (1967) presents a consistent and asymptotically efficient three-step estimation technique for the regression coefficients. The first step uses single equation regressions to estimate the parameters of autoregressive model. The second step uses single equation regressions on transformed equations to estimate the contemporaneous covariances. Finally, the Aitken estimator is formed using the estimated covariance,

$$\hat{B} = (F' \Omega^{-1} F)^{-1} F' \Omega^{-1} r$$  \hspace{1cm} (34)

where $\Omega = E(ee')$ denotes the general covariance matrix of the innovation. These estimators of $\Omega$ are not unbiased for true covariance matrix, but they are consistent under the usual conditions. Again imposing reasonable regularity conditions, it can be shown that this $B$ estimator is asymptotically equivalent to the infeasible Generalized Least Square (GLS) estimator which assumes $\Sigma$ is known. So inference on the parameter vector $B$ can be carried out using the approximate normality of $\hat{B}$:

$$\hat{B} \sim IN(B, \Omega)$$  \hspace{1cm} (35)

In the application, we use this mentioned methodology with the slope coefficients restricted to be the same for all assets. In particular, we use the same three-step procedure and the same covariance assumptions as in (equation 31) to (equation 33) to estimate the covariances and to generate the $t$-statistics for the parameter estimates.
4. Results and Robustness Tests

In this section, we examine the pricing efficiency of alternative models. The Fama-French Realized Portfolio Returns and Asset pricing Test are presented. We firstly present the gains associated with the use of high frequency data in the estimation of factor loadings and asset pricing. We then add a realized coskewness in the model and compare with the Bollerslev and Zhang (2003) realized loading procedure.

4.1. Fama-French Realized Portfolio Returns and Factor Loadings

The gains associated with the use of the high-frequency data for more accurately measuring and forecasting the realized factor loadings depend upon the quality of the corresponding asset pricing model. The empirical results in the literature discussed above suggest that the asset pricing anomalies related to size and book-to-market largely disappear within the context of the three-factor Fama and French (1993) model. As such, this particular model and the 9 tests portfolios sorted on the basis of firm size and book-to-market ratios provide a natural setting for evaluating the improvements entailed by the high frequency-based factor loadings\textsuperscript{11}.

The average monthly percentage loadings and standard deviations, over the 5-year (2002-2006) sample period are given in Table 3. In contrast to the earlier results, the cross-sectionally averages no longer show a clear-cut pattern in regards to the size and book-to-market characteristics. Indeed, as noted below, the model still yields an average $R^2$ in excess of 30% for explaining the cross-sectional differences in the monthly returns across the 9 portfolios more than twice the average $R^2$ for the standard one-factor CAPM over the same time period. Also, the average $R^2$ for the 9 time series regressions of the portfolio returns on the three factors exceeds 75%. These results are directly in line with the results reported in the existing literature based on longer historical samples.

The average loadings for the Small Minus Big (SMB) size factor systematically decrease with the size of the portfolio, while the average realized loadings for the High Minus Low (HML) factor increase with the book-to-market ratio. No apparent pattern is evident in the loadings association with the return on the Market portfolio (MKT). The monthly sample standard deviations reported suggest that the realized factor loadings also vary a lot through time, with the

\textsuperscript{11}The 9 portfolios is computed by merging 3 portfolios according to the size with 3 portfolios according to the style (for instance the portfolio “11” will be computed by the average of the asset returns being in the first third sorted by the market capitalization and first third sorted by book-to-market).
Table 3: Realized Factor Loadings

<table>
<thead>
<tr>
<th></th>
<th>Market</th>
<th>Small Minus Big</th>
<th>High Minus Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Middle</td>
<td>High</td>
</tr>
<tr>
<td>Market</td>
<td>1.098</td>
<td>0.708</td>
<td>0.638</td>
</tr>
<tr>
<td>Small minus Big</td>
<td>0.696</td>
<td>0.656</td>
<td>0.532</td>
</tr>
<tr>
<td>Big</td>
<td>0.973</td>
<td>0.687</td>
<td>0.576</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Middle</th>
<th>High</th>
<th>Small</th>
<th>Middle</th>
<th>High</th>
<th>Small</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.131</td>
<td>0.035</td>
<td>0.070</td>
<td>0.132</td>
<td>0.057</td>
<td>0.123</td>
<td>0.189</td>
<td>0.057</td>
<td>0.259</td>
</tr>
<tr>
<td>Middle</td>
<td>0.066</td>
<td>0.036</td>
<td>0.076</td>
<td>0.074</td>
<td>0.052</td>
<td>0.115</td>
<td>0.134</td>
<td>0.049</td>
<td>0.130</td>
</tr>
<tr>
<td>Big</td>
<td>0.096</td>
<td>0.031</td>
<td>0.071</td>
<td>0.082</td>
<td>0.050</td>
<td>0.154</td>
<td>0.148</td>
<td>0.049</td>
<td>0.123</td>
</tr>
</tbody>
</table>

Source: Euronext. The table reports the average realized factor loadings for the three-factor Fama and French (1993) Model for each of the 9 size and book-to-market sorted portfolios over the January 2002 through December 2006 sample period and the corresponding sample standard deviations. The realized factor loadings are constructed from high-frequency 5-min returns.

loadings for the largest low book-to-market portfolios. This will inflate the reported standard deviations. It would be interesting, but beyond the scope of the present chapter, to formally assess the magnitude of the measurement errors in the present context (Bai et al., 2000; Barndorff-Nielson and Shephard, 2002a and 2002b and Meddahi, 2002; all authors providing important theoretical results for the case of scalar realized volatility measures). Even though the inherent measurement errors result in somewhat erratic month-to-month realized factor loadings, an underlying systematic temporal pattern is evident for all three risk factors. Similar results are true for the other portfolios.

The squared factor is added to the three-factor Fama and French (1993) Model. The Table 4 gives the average of $R^2$ for the three-factor Fama and French (1993) Model and Fama and French (1993) Model with Squared factor returns for each of the 9 size and book-to-market sorted portfolios over the January 2002 through December 2006 sample period. The alternatives estimations are also made in realized regression and classical static regression. We observe clearly that the level of the $R^2$ is high with the squared factor that others estimations methods. This means the addition of the skewness factor makes the classical three-factor model strikingly more competitive. For example, in the large size and high style portfolio, the classical three-factor model produces an $R^2$ of 28.73 percent, the realized estimation gives 65.86 percent when the model with skewness factor produces 88.96 percent. The intercepts of the regressions for this last
Table 4: Average $R^2$ for Alternatives Models with or without Skewness Factor

<table>
<thead>
<tr>
<th></th>
<th>Realized Regression</th>
<th>Classical Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Middle</td>
</tr>
<tr>
<td>Without Skewness Factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>59.33%</td>
<td>85.12%</td>
</tr>
<tr>
<td>Middle</td>
<td>52.01%</td>
<td>85.94%</td>
</tr>
<tr>
<td>Big</td>
<td>59.59%</td>
<td>85.86%</td>
</tr>
<tr>
<td>With Skewness Factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>77.03%</td>
<td>95.34%</td>
</tr>
<tr>
<td>Middle</td>
<td>72.21%</td>
<td>96.10%</td>
</tr>
<tr>
<td>Big</td>
<td>73.65%</td>
<td>95.71%</td>
</tr>
</tbody>
</table>

Source: Euronext. The table reports the average of $R^2$ for the three-factor Fama and French (1993) Model and Fama and French (1993) Model with squared factor returns for each of the 9 size and book-to-market sorted portfolios over the January 2002 through December 2006 sample period. The estimations are carried out on realized regression and on classical static regression.

The realized factor loadings are, of course, only observable \textit{ex post}. The results in the next section demonstrate that a simple $AR(1)$ time series model for the high-frequency factor loadings effectively smoothes out the measurement error, and gives rise to more accurate forecasts and statistically significant lower pricing errors when compared to the conventional monthly rolling regression-based estimates traditionally used in the literature.

4.2. High-frequency Factor Loadings and Asset Pricing

The discussion in the previous section illustrated how high-frequency data may be used in an actual empirical setting for more accurately measuring systematic factor risk exposures. This section goes one step further and asks the question: do the high frequency- based measures give more accurate predictions of the future factor loadings and statistically significant improvements in terms of asset pricing predictions and corresponding better economic decisions? We begin by a discussion of our four different return-based forecast procedures for the factor loadings to based on monthly return observations only.

4.2.1. Predicted Factor Loadings

For the predicted factor loadings, we use several independent procedures. We first considers the factor loadings to be constant and equal to the conventional

---

12We also tested the model with the \textit{kurtosis}, but this factor is not significant and was not finally retained for the tests
regression estimates based on 5 years of monthly observations at the beginning of the out-of-sample forecast comparison period. We will refer to these Constant factor loadings by the acronym CON. We have also experimented with the loadings estimates based on the monthly return observations over the full sample. The 5-year Rolling Regression-based (RR) estimation technique advocated by Fama and MacBeth (1973) allows the betas to change every month with the new return observations. As previously noted, on using this procedure along with the actual realized factor returns, the average of the 60 monthly cross-sectional $R^2$'s (for explaining the variation in the returns across the 9 portfolios) equals 0.929, whereas the average $R^2$ from the 9 time series regressions equals 0.756.

We use, as a benchmark, a high-frequency-based procedure takes the forecast for the next month to be equal to the current month’s realized loadings. These forecasts would, of course, be optimal in a Mean Square Error sense if the time series of realized factor loadings followed a Random Walk (RW), or martingale, model. We consequently refer to this procedure as the HF-RW-based forecasts. The second high-frequency procedure is based on the one-step-ahead forecasts from a set of AR(1) models for each of the different loadings and portfolios. The estimates of the models are based on a rolling 24-months sample, leaving us with an out-of-sample forecast comparison period covering the 60 months from January 2002 through December 2006. Consistent with the significant autocorrelations reported in Table 5, the vast majority of the parameter estimates for the $3 \times 9 \times 60 = 1620$ different AR(1) models are statistically significant at standard thresholds. In addition to these rolling AR(1) models, we have also experimented with the forecasts based on a set of ARMA(1,1) models estimated over the full in-sample period\textsuperscript{13}. The basic findings, available upon request, were very similar to the ones reported below for the AR(1) models. Meanwhile, more complicated dynamic time series models, possibly allowing for asymmetries in the relationship between the past returns and the future factor loadings, may give rise to even better predictions\textsuperscript{14}. To keep the comparisons simple, we do not purpose any of these more complicated model specifications here. Of course, the relatively short 5-year time span of high-frequency data also limits the scope of such investigations.

\textsuperscript{13} The ARMA(1,1) formulation is consistent with the assumption that the true latent loadings follow an $AR(1)$ model, but that the (observed) realized loadings are equal to the true loadings plus an independent measurement error.

\textsuperscript{14} Braun et al. (1995) and Cho and Engle (1999) have documented statistically significant temporal variation and important asymmetries in CAPM betas, while Bollerslev and Zhang (2003), provide complementary evidence for asymmetric market risk exposure based on the same type of high-frequency measures utilized here.
### Table 5: Factor Loading Forecasts

| Source: Euronext. The table reports the average forecast errors for forecasting the monthly realized factor loadings across the 9 portfolios and the 5-year out-of-sample period from January 2002 through December 2006. The constant predictions take the factor loadings to be constant and equal to the conventional 5-year monthly regression based estimates at the beginning of the sample. Rolling Regression refers to the 5-year rolling regression estimates. High-frequency Random Walk assumes that the high-frequency-based realized factor loadings follow a random walk, or martingale, model. The High-frequency AutoRegressive predictions are based on a set of rolling $AR(1)$ models estimated over the previous 24 months. The rows labeled Bias, MSE, and MAE give the average errors, the average squared errors, and the average absolute errors, respectively. |

<table>
<thead>
<tr>
<th></th>
<th>Constant Regression</th>
<th>Rolling Random Walk</th>
<th>High Frequency AutoRegressive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Market</strong> Bias</td>
<td>0.03</td>
<td>-0.68</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>MSE</strong></td>
<td>0.01</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>MAE</strong></td>
<td>0.07</td>
<td>0.68</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Small Minus Big</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>-0.52</td>
<td>-0.08</td>
<td>0.00</td>
</tr>
<tr>
<td>MSE</td>
<td>0.28</td>
<td>0.12</td>
<td>0.02</td>
</tr>
<tr>
<td>MAE</td>
<td>0.52</td>
<td>0.26</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>High Minus Low</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td>-0.03</td>
<td>0.94</td>
<td>0.01</td>
</tr>
<tr>
<td>MSE</td>
<td>0.03</td>
<td>1.02</td>
<td>0.06</td>
</tr>
<tr>
<td>MAE</td>
<td>0.13</td>
<td>0.94</td>
<td>0.18</td>
</tr>
</tbody>
</table>

Table 5 reports the average bias, Mean Squared Error (MSE), and Mean Absolute Error (MAE), using each of the four procedures for forecasting the monthly realized loadings. The errors are averaged across the 9 size and book-to-market sorted portfolios and the 60-months out-of-sample period. The loadings associated with the market return, are generally the most accurate. Not surprisingly, the constant factor loadings (CON) result in the largest average Biases, MSEs and MAEs of all the procedures. The MSEs and MAEs for the RR and HF-RW procedures are fairly close. The one-step-ahead $AR(1)$ forecasts systematically result in the lowest MSEs and MAEs for all of the factor loadings. Intuitively, the HF-AR-based procedure more swiftly adjusts to changes in the underlying market conditions.
4.2.2. Tests of Pricing Performance

In order to focus directly on the potential benefits of using the high-frequency data in measuring and modeling the realized factor loadings, we rely on the actual realized excess factor returns when calculating the asset prices. The three-factor Fama-French model has been the subject of numerous empirical studies and, as discussed above, has been found to perform very well in pricing the 9 size and book-to-market portfolios analyzed here. Indeed, for none of the different factor loading forecast procedures is the hypothesis of a zero “alpha” in the three-factor model is rejected by a simple $t$-test. Specifically, let the forecast error for the returns on portfolio $i$ for month $t$ based on the loadings from procedure $m$, be denoted by $e_{i,t}(m)$. The four $t$-statistics for testing the hypotheses that the average pricing errors across the 9 portfolios and the 60-month out-of-sample period are equal to zero i.e., $\sum_i \sum_t e_{i,t}(m) = 0$.

Meanwhile, the superiority of the high frequency-based loadings is underscored by the results in Table 6, and the average absolute pricing errors, $|e_{i,t}(m)|$, reported in the first panel. Of course, the most appropriate statistical loss function for assessing the superiority of the competing forecasts will ultimately depend upon the application (see e.g., Christoffersen and Diebold, 1997 for a discussion of different loss functions in the comparison of nonlinear forecasting models). Hence, following Andersen et al. (1999), Table 6 also reports the heteroskedasticity adjusted average absolute standardized errors, $|e_{i,t}(m)/\sigma_{i,t}|$, and Table 7 reports the pseudo-$R^2$ defined by $1 - [e_{i,t}(m)/\sigma_{i,t}]^2$, where $\sigma_{i,t}$ refers to the realized volatility for portfolio $i$ for month $t$ constructed by the summation of the 5-minutes squared returns within the month augmented by the 1-day lead-lag autocovariances to account for the nonsynchronous trading effects. For the sake of parsimony, we only report the average loss across the 9 portfolios. From the results, reported in the first rows in each of the three panels, it is obvious that the HF-AR-based forecasts systematically yield the lowest absolute errors and the highest pseudo-$R^2$’s amongst the four different procedures.

To formally test for the statistical significance of these apparent differences, Table 6 complements the average loss with the values of the three relatively simple

---

15Since the errors associated with the forecasts of the factor loadings and the risk premiums are likely to be correlated, a rigorous true out-of-sample forecast comparison would also necessitate the joint modeling of the loadings and the factor risk premiums.

16The four statistics correspond to the four alternative predictions models for the factors loadings (CON, RR, HF-RW and HF-AR).

17Although the improvements may seem numerically small, the increase in the $R^2$ for the HF-AR procedure is actually somewhat higher than the improvements reported in the study by Ferson and Harvey (1999) and Bollerslev and Zhang (2003) based on explicitly modeling the temporal variation in the monthly factor loadings using exogenous explanatory variables.
pair-wise forecast comparison tests discussed by Diebold and Mariano (1995). In particular, let \( d_{i,t}(m,n) \) denote the loss differential between the forecasts based on method \( m \) and \( n \) for portfolio \( i \) and month \( t \). The \( Z_{(m,n)} \) statistic then refers to the standard \( t \)-test for the hypothesis that the mean differential is equal to zero. Alternatively, the null hypothesis of equal forecast accuracy, may also be tested by the nonparametric sign-test \( (Z_{2(m,n)}) \),

\[
Z_{2(m,n)} = \left[ \sum_i \sum_t 1I_{d_{i,t}(m,n)} - 0.5T \right] [0.25T]^{1/2},
\]

where \( 1I \) denotes the indicator function and \( T = 60 \times 9 \) refers to the sample size and \( m \) and \( n \) denote the monthly predictions models (CON, RR, HF-RW or HF-AR). Lastly, the related studentized version of the Wilcoxon’s signed-rank test is formally defined by,

\[
Z_{3(m,n)} = \left[ \sum_i \sum_t 1I_{d_{i,t}(m,n)} \times rank(|d_{i,t}(m,n)|) - 0.5T(T + 1) \right] \times [T(T + 1)(2T + 1)/24]^{-1/2}
\]

Under the null hypothesis that models \( m \) and \( n \) have the same predictive ability, all three test statistics should be asymptotically standard normally distributed. The values reported in the three lower panels in the table soundly reject the hypothesis of no superior predictive ability in favor of the high-frequency-based HF-AR procedure.\(^{18}\)

In order to further explore the performance of the HF-AR procedure, Table 8 reports the average pseudo \( R^2 \)’s for each of the 9 portfolios. Comparing the results to the \( R^2 \)’s associated with the conventional monthly rolling regression-based estimates, it follows from the second panel in the table, that the HF-AR procedure does better in an absolute sense. Interestingly however, for none of the portfolios where the RR procedure does better, is the difference statistically significant at the usual 5% level, whereas the HF-AR-based loadings result in statistically significant improvements for seven of the portfolios. Thus, in spite of the relatively short 5-year out-of-sample forecast comparison period, our various statistically tests systematically favor the asset pricing predictions based on the high-frequency factor loadings.

\(^{18}\)The results reported in Table 6 are based on the 60-month out-of-sample forecast comparison period and the simple \( AR(1) \) model.
Table 6: Average Pricing Errors and Significance Tests

<table>
<thead>
<tr>
<th></th>
<th>CON</th>
<th>RR</th>
<th>HF-RW</th>
<th>HF-AR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Absolute Errors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z1(\ldots)</td>
<td>CON</td>
<td>-31.76*</td>
<td>23.64*</td>
<td>-9.83*</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>20.35*</td>
<td>1.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HF-RW</td>
<td>-18.64*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z2(\ldots)</td>
<td>CON</td>
<td>-25.62*</td>
<td>14.85*</td>
<td>-18.30*</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>12.82*</td>
<td>20.08*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HF-RW</td>
<td>-131.98*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z3(\ldots)</td>
<td>CON</td>
<td>-14.63*</td>
<td>-12.92*</td>
<td>-15.56*</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>-13.12*</td>
<td>-14.28*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HF-RW</td>
<td>-15.67*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Panel B: Absolute Standardized Errors</strong></th>
<th>0.51</th>
<th>0.19</th>
<th>0.41</th>
<th>0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z1(\ldots)</td>
<td>CON</td>
<td>-10.73*</td>
<td>-20.65*</td>
<td>-85.49*</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>-11.04*</td>
<td>7.32*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HF-RW</td>
<td></td>
<td>11.58*</td>
<td></td>
</tr>
<tr>
<td>Z2(\ldots)</td>
<td>CON</td>
<td>-114.54*</td>
<td>-146.62*</td>
<td>-113.64*</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>-112.71*</td>
<td>54.99*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HF-RW</td>
<td></td>
<td>116.37*</td>
<td></td>
</tr>
<tr>
<td>Z3(\ldots)</td>
<td>CON</td>
<td>-15.50*</td>
<td>-15.31*</td>
<td>-16.49*</td>
</tr>
<tr>
<td></td>
<td>RR</td>
<td>-15.48*</td>
<td>-14.32*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HF-RW</td>
<td></td>
<td>-13.24*</td>
<td></td>
</tr>
</tbody>
</table>

Source: Euronext. Panel A reports the average absolute errors. The entries in the first row have been multiplied by $10^2$. Panel B gives the average absolute standardized errors obtained by standardizing each error by realized volatility for that particular portfolio and month. All of the averages are calculated over the 9 portfolios and the 5-year out-of-sample period from January 2002 through December 2006. The procedure labeled CON takes the factor loadings to be constant and equal to the conventional monthly regression estimates at the beginning of the sample. RR refers to the 5-year rolling regression estimates. HF-RW assumes that the high-frequency-based realized factor loadings follow a random walk, or martingale, model. The HF-AR predictions are based on a set of rolling $AR(1)$ models estimated over the previous 24 months. $Z_{1(m,n)}$ refers to the $t$-test for a difference in the means across the monthly predictions for models $m$ and $n$. $Z_{2(m,n)}$ gives the sign test while $Z_{3(m,n)}$ refers to the Wilcoxon’s signed rank test for the same hypothesis. Under the null hypothesis of no superior predictive ability by either of the two models, all of the test statistics should be standard normally distributed. The sign * Indicates the significance at 5% threshold.
Table 7: Average of Pseudo-$R^2$ and Significance Tests

<table>
<thead>
<tr>
<th></th>
<th>CON</th>
<th>RR</th>
<th>HF-RW</th>
<th>HF-AR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pseudo-$R^2$</strong></td>
<td>0.39</td>
<td>0.40</td>
<td>0.82</td>
<td>0.97</td>
</tr>
<tr>
<td><strong>$Z_1(\cdot,\cdot)$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CON</td>
<td>70.81*</td>
<td>103.17*</td>
<td>58.31*</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td></td>
<td>70.19*</td>
<td>-2.80*</td>
<td></td>
</tr>
<tr>
<td>HF-RW</td>
<td></td>
<td></td>
<td>-74.95*</td>
<td></td>
</tr>
<tr>
<td><strong>$Z_2(\cdot,\cdot)$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CON</td>
<td>114.54*</td>
<td>146.62*</td>
<td>113.67*</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td></td>
<td>112.71*</td>
<td>-54.99*</td>
<td></td>
</tr>
<tr>
<td>HF-RW</td>
<td></td>
<td></td>
<td>-116.37*</td>
<td></td>
</tr>
<tr>
<td><strong>$Z_3(\cdot,\cdot)$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CON</td>
<td>-13.25*</td>
<td>-12.94*</td>
<td>-13.26*</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td></td>
<td>-13.27*</td>
<td>-14.43*</td>
<td></td>
</tr>
<tr>
<td>HF-RW</td>
<td></td>
<td></td>
<td>-15.52*</td>
<td></td>
</tr>
</tbody>
</table>

Source: Euronext. Table reports the average pseudo-$R^2$ defined by one minus the squared standardized error. All of the averages are calculated over the 9 portfolios and the 5-year out-of-sample period from January 2002 through December 2006. The procedure labeled CON takes the factor loadings to be constant and equal to the conventional monthly regression estimates at the beginning of the sample. RR refers to the 5-year rolling regression estimates. HF-RW assumes that the high-frequency-based realized factor loadings follow a random walk, or martingale, model. The HF-AR predictions are based on a set of rolling AR(1) models estimated over the previous 24 months. $Z_{1(m,n)}$ refers to the $t$-test for a difference in the means across the $9 \times 60$ monthly predictions for models $m$ and $n$. $Z_{2(m,n)}$ gives the sign test while $Z_{3(m,n)}$ refers to the Wilcoxon’s signed rank test for the same hypothesis. Under the null hypothesis of no superior predictive ability by either of the two models, all of the test statistics should be standard normally distributed. The sign * Indicates the significance at 5% threshold.

4.2.3. Economic Significance of High-frequency Loadings

To highlight the economic value of the high-frequency-based loadings and corresponding asset pricing predictions, this section compares the ex post performance of the mean-variance efficient portfolios constructed from the 9 size and book-to-market sorted portfolios and the same four forecasting procedures discussed above. We construct the minimum variance portfolio as well as the portfolio that minimizes variance given a set target return, which is denoted $\mu_p$, allowing for monthly rebalancing. To be precise, we solve the following two optimization problems:

$$\min_{\omega_t} \omega_t' \Sigma_t \omega_t$$  \hspace{1cm} (38)

s.t. $\omega_t' \mu_t = \mu_p$ and $\omega_t' \iota = 1$

where $\omega_t$ is the $(N \times 1)$ vector of portfolio weights, and $\iota$ denotes an $(N \times 1)$ vector of ones. In addition, $\mu_t$ is the $(N \times 1)$ vector with conditional expected
Table 8: Pseudo $R^2$ and Significance Test of Portfolios

<table>
<thead>
<tr>
<th>Sample Mean</th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.70</td>
<td>0.48</td>
<td>0.80</td>
</tr>
<tr>
<td>Middle</td>
<td>0.90</td>
<td>0.30</td>
<td>0.81</td>
</tr>
<tr>
<td>Big</td>
<td>0.91</td>
<td>0.85</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$Z_{\text{RR, HF-AR}}$ Test</th>
<th>Small</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>-5.30*</td>
<td>-1.42</td>
<td>-2.55*</td>
</tr>
<tr>
<td>Middle</td>
<td>-3.73*</td>
<td>-0.88</td>
<td>2.67*</td>
</tr>
<tr>
<td>Big</td>
<td>6.78*</td>
<td>4.35*</td>
<td>3.68*</td>
</tr>
</tbody>
</table>

Source: Euronext. The first panel reports the pseudo-$R^2$ for each of the 9 portfolios defined by one minus the squared standardized error for the HF-AR predictions averaged across the 5-year out-of-sample period from January 2002 through December 2006. The HF-AR predictions are based on a set of rolling $AR(1)$ models estimated over the previous 24 months. The second panel reports the $t$-statistic for testing identical mean $R^2$’s between the 5-year rolling regression estimates, RR, and HF-AR based predictions. Under the null hypothesis of no superior predictive ability by either of the models, the test statistics should be standard normally distributed. * Indicates the significance at 5% level.

returns for the individual stocks, that is $\mu_t = E[r_t | I_{t-1}]$, where $I_{t-1}$ denotes the information set available at the end of day $t - 1$. Similarly, $\Sigma_t$ is the ($N \times N$) conditional covariance matrix.

The solution is of the problem is:

$$\omega_t = \frac{\mu_t \Sigma_t^{-1} \mu_t}{\mu_t^\prime \Sigma_t \mu_t}$$

In order to focus the comparisons directly on the impact of the forecasts for the factor loadings, we assume that the covariance matrix for the 9 tests portfolios underlying the mean-variance optimizations is the same across the different procedures, and equal to the realized covariance matrix for the 24-month presample period from January 2002 through December 2003. Also, we explicitly exclude the risk-free asset from the portfolio optimizations.

The first panel in Table 9 reports the resulting average monthly realized returns over the $ex \ post$ January 2004 through December 2006 sample period for the optimized portfolios with the monthly expected standard deviations indicated in the first column. For all of the monthly standard deviations ranging from 0.01 to 0.02, or 0.03 to 0.06 at an annual level, the $ex \ post$ returns are the highest for the HF-AR-based factor loadings. With the exception of the low-risk portfolio with a monthly expected standard deviation of 0.01, all of the Sharpe ratios
for the monthly ex post returns and standard deviations, reported in the second panel in Table 9, also favor the HF-AR procedure by a fairly large margin. The very high ex post Sharpe ratios reflect the unprecedented bull-run during the 5-year out-of-sample period. Although it would be unrealistic to expect these same high values to obtain over other time periods, the ordering among the different procedures are nonetheless suggestive.

Table 9: Mean-variance Efficient Portfolios

<table>
<thead>
<tr>
<th>Standard Deviation</th>
<th>Constant Roll.</th>
<th>Rolling Regression</th>
<th>High-frequency Random Walk</th>
<th>High-frequency AutoRegressive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ex post Returns</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>0.02</td>
<td>0.01</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>0.03</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
<td>0.07</td>
</tr>
<tr>
<td>0.04</td>
<td>0.05</td>
<td>0.07</td>
<td>0.04</td>
<td>0.09</td>
</tr>
<tr>
<td>0.05</td>
<td>0.07</td>
<td>0.08</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>0.06</td>
<td>0.08</td>
<td>0.12</td>
<td>0.06</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Sharpe Ratios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.25</td>
<td>3.74</td>
<td>1.75</td>
<td>8.75</td>
</tr>
<tr>
<td>0.02</td>
<td>0.38</td>
<td>1.88</td>
<td>1.38</td>
<td>2.38</td>
</tr>
<tr>
<td>0.03</td>
<td>1.25</td>
<td>1.78</td>
<td>1.08</td>
<td>2.25</td>
</tr>
<tr>
<td>0.04</td>
<td>1.19</td>
<td>1.61</td>
<td>0.94</td>
<td>2.19</td>
</tr>
<tr>
<td>0.05</td>
<td>1.35</td>
<td>1.51</td>
<td>0.95</td>
<td>1.95</td>
</tr>
<tr>
<td>0.06</td>
<td>1.26</td>
<td>1.96</td>
<td>0.96</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Source: Euronext. The table reports the average monthly ex post percentage returns and corresponding Sharpe ratios for the mean-variance efficient portfolios constructed from the 9 size and book-to-market sorted test portfolios over the 5-year out-of-sample period from January 2002 through December 2006. The expected monthly standard deviations reported in the first column are based on the pre-sample realized variance-covariance matrix constructed from 5-min returns from January 2002 through December 2003. The expected return calculations labeled Constant takes the factor loadings to be constant and equal to the conventional 5-year monthly regression-based estimates at the beginning of the sample. Rolling Regression refers to the 5-year rolling regression estimates. High-frequency Random Walk assumes that the high frequency-based realized factor loadings follow a random walk, or martingale, model. The High-frequency AutoRegressive predictions are based on a set of rolling AR(1) models estimated over the previous 24 months.

To value the performance gains associated with a particular trading strategy, we use the utility-based measure developed in Flemin et al. (2001 and 2002). This measure is based on quadratic utility as an approximation of the investor’s true utility function and assumes that relative risk aversion (γ) is constant. Under
these conditions, the average realized utility for a given portfolio \((U_p)\) is:

\[
U_p(\cdot) = W_0 \left[ \sum_{t=0}^{T-1} R_{p,t+1} - \frac{\gamma}{2(1 + \gamma)} R^2_{p,t+1} \right],
\]

(40)

where \(W_0\) is the initial wealth and \(R_{p,t+1} = R_f + \omega t r_{t+1}\) are the returns on the portfolio \((p)\). This provides a consistent estimator of the investor’s expected utility. In order to measure the value of switching from one trading strategy to another strategy, we equate their average realized utilities,

\[
\sum_{t=0}^{T-1} (R_{2,t+1} - \Delta) - \frac{\gamma}{2(1 + \gamma)} (R_{2,t+1} - \Delta)^2 = \sum_{t=0}^{T-1} R_{1,t+1} - \frac{\gamma}{2(1 + \gamma)} R^2_{1,t+1},
\]

(41)

where \(R_{1,t+1}\) and \(R_{2,t+1}\) are the returns for the two strategies. To equate the average utilities, we subtract a constant, \(\Delta\), from each of the returns for strategy 2. This represents the most that the investor could give up \((i.e.,\) a performance fee) and still have the same expected utility as under strategy 1.

<table>
<thead>
<tr>
<th>Standard Deviation</th>
<th>(\gamma = 1)</th>
<th>(\gamma = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.004</td>
<td>0.006</td>
</tr>
<tr>
<td>0.02</td>
<td>0.006</td>
<td>0.009</td>
</tr>
<tr>
<td>0.03</td>
<td>0.008</td>
<td>0.011</td>
</tr>
<tr>
<td>0.04</td>
<td>0.009</td>
<td>0.014</td>
</tr>
<tr>
<td>0.05</td>
<td>0.010</td>
<td>0.018</td>
</tr>
<tr>
<td>0.06</td>
<td>0.015</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Source: Euronext. The table reports the monthly fees (in percentage points) that an investor with a quadratic utility function and constant relative risk aversion \((\gamma)\) would be willing to pay to switch from the 5-year rolling regression-based return forecasts (RR) to the high-frequency-based AR(1) forecasts for the realized factor loadings (HF-AR) for the 9 size and book-to-market sorted test portfolios. The fees are based on the monthly returns for the 5-year out-of-sample period from January 2002 through December 2006, with the monthly expected standard deviations for both methods determined by the pre-sample realized covariance matrix from January 2002 through December 2003.

The Table 10 reports the monthly fees that an investor with a quadratic utility function and constant relative risk aversion equal to \(\gamma\) would be willing to pay to switch from the traditional RR 5-year monthly rolling regression-based forecasts to the high-frequency-based HF-AR forecasts for the 9 portfolios returns. The fees in the table are expressed in monthly percentage points and determined
empirically as the value which equalizes the *ex post* utility for the two different procedures when evaluated on the basis of the monthly returns\(^{19}\). To facilitate comparisons, we again fix the expected covariance matrix for both procedures at the realized covariance matrix for the January 2002 through December 2003 pre-sample period. Fixing the expected standard deviation at 6%, an investor with constant aversion risk \(\gamma = 10\) would be willing to pay 2.3% per month, to switch from the Rolling Regression strategy to the high-frequency autoregressive strategy monthly return-based forecasts.

5. Conclusion

Following Bollerslev and Zhang (2003) using “Realized” measures and co-measures of risk in some classical asset pricing models, this chapter extends their approach in an Extended Realized Conditional Asset Pricing Model framework. These extensions have included third-moment asset pricing models (see Jurczenko and Maillet, 2006) and conditional asset pricing models. We also make use of several methodologies aiming to neutralize data measurement and model misspecification errors (see Ledoit and Wolf, 2003 and 2004), trying to properly deal with inter-relations between financial assets in term of returns (see Zellner, 1962).

For the empirical investigation, high frequency data for some stocks on the French stock market have been used. Firstly, this chapter shows how high-frequency intra-day transaction prices for the different portfolios and the underlying factor return lead to more effective measures and models of the realized co-variations. As in Bollerslev and Zhang (2003), our empirical analysis are focused on a 1-month return horizon and the Fama-French test portfolios. In the French stock market on a limited sample, cross-sectional tests of the single-factor an three-factor model of Fama and French (1993) have shown that systematic risk measured by the covariance (or beta) with the market and other classical factors does not satisfactorily explain the cross-sectional variation in expected excess returns. We show also that the unexplained component in expected excess returns does not vary across portfolios and it is modest in magnitude.

This finding implies that additional variables representing portfolios characteristics we consider have no explanatory power for expected excess returns when coskewness is taken in account. This result cannot be obtained if coskewness is neglected. In addition to that, our results have implications for testing methodologies, since they show that neglecting coskewness risk can cause misleading inference. Indeed, we find that coskewness is positively correlated with size. This

\(^{19}\)The fee \(\Delta\) refer, in the table, to the *ex post* monthly returns for the mean-variance efficient portfolio with the same benchmark expected standard deviations and expected returns determined by the HF-AR and RR loading forecasts, respectively.
suggests that a possible justification for the anomalous explanatory power of size in the cross-section of expected returns, is that it proxies for omitted coskewness risk.

Some tracks of improvement are however possible. It would be interesting to formally evaluate the importance of the errors of measurement of the sensitivities to the factors in the context of realized co-variations for the CAPM. Bai et al. (2000), Barndorff-Nielsen and Shephard (2002a, b), Meddahi (2002) or recently Griffin and Oomen (2008), provide theoretical results for the case of realized volatility. Instantaneous volatility is then broken up into volatility and autocorrelation term. Griffin and Oomen (2008) show theoretically and empirically that the bias of autocorrelation decreases with the sampling rate of the prices to be close to zeros after 25 minutes. A similar work could thus be made on realized beta and realized skewness. In the same way, by using a model with independent factors, the various relations between risks factors could be simplified. Indeed, by extraction of the common factors by an Independent Component Analysis with noises filtering (Kopriva and Seršić, 2008), all the cumulants matrixes will be diagonal (because of property of independence of the factors). So the estimate of the parameters with higher moments will be able.

Lastly, inspired by work of Brock and Hommes (1998) or Malevergne and Sornette (2006), the generalization of the previous approaches within a heterogeneous framework - where market participants confront their price- could constitute a natural extension of this study. For that, we can consider several sources of heterogeneity. Indeed, we can suppose that the market participants have different points of view about anticipations of the prices (Brock and Hommes, 1998).

References


Appendix

Appendix 1: Assumptions and Properties of Realized Moments

Following Beine et al. (2004), we present here assumptions and properties of Realized Moments. When computing the realized moments for closing prices (24h GMT), daily returns are defined as: \( r_t = \sum_{i=1}^{K} r_{i, t} \) where \( K \) is the number of intervals per day (e.g. 24) and \( r_{i, t} \) is the \( i \)-th intra-day return of day \( t \).

A1.1: Realized Volatility

Consider two measures of the daily volatility \( V(r_t) \): \( V_{1, t} \equiv r_{t}^2 \); i.e. daily squared returns, and \( V_{2, t} \equiv \sum_{i=1}^{K} r_{i, t}^2 \); i.e. the so-called realized volatility.

**Assumption A1:** \( r_{i, t} \) is a martingale difference sequence, i.e. \( E(r_{i, t} | \Omega_{i-1, t}) = 0 \), where \( \Omega_{i-1, t} \) denotes a filtration (information set) including past information on \( r \) up to (and including) the point in time \( i-1 \) on day \( t \).

Implications of Assumption A1:
1. \( E(r_{i, t}) = 0 \);
2. \( Cov(r_{i, t}, r_{j, t'}) | \Omega_{\max(i-1, j-1), \max(t, t')} = 0 \), which implies \( Cov(r_{i, t}, r_{j, t'}) = 0 \);
3. \( E(r_{i, t} r_{j, t'}^k | \Omega_{i-1, t}) = 0 \) if \( i > j \) and \( t', k \in 1, 2, \ldots \).

The implications follow immediately when we use the law of iterated expectations. Note that Assumption A1 does not imply that \( E(r_{i, t} | r_{j, t'}^k | \Omega_{j-1, t'}) = 0 \) if \( i \leq j \) and \( t < t' \).

**Proposition 1:** Under A1, \( V_{1, t} \) and \( V_{2, t} \) are unbiased estimators of \( V(r_t) \).

**Proof.**

\[
E(V_{1, t}) = E(V_{2, t}) \overset{A1}{=} V(r_t),
\]

where \( \overset{A1}{=} \) indicates the implication of the assumption A1.
\[ V(r_t) \overset{A_1}{=} E(r_t^2) \]
\[ = E \left[ \left( \sum_{i=1}^{K} r_{i, t} \right)^2 \right] \]
\[ = E \left( \sum_{i=1}^{K} r_{i, t}^2 + 2 \sum_{i=1}^{K} \sum_{j=1}^{K} r_{i, t} r_{j, t} \right) \]
\[ \overset{A_1}{=} E \left( \sum_{i=1}^{K} r_{i, t}^2 \right) \]
\[ = E(V_{1, t}). \]

**Assumption A2:** \( E(r_{i, t} r_{j, t}^3 | \Omega_{j-1, t}) = 0, \forall i < j. \)
This assumption imply the zero skewness and symmetry of the same conditional distributions.

**Proposition 2:** Under \( A_1 - A_2 \), \( V_{2, t} \) is a more efficient estimate of \( V(r_t) \) than \( V_{1, t} \).

**Proof.** For the sake of simplicity, let us consider the case where \( K = 2 \).

\[
E(V_{1, t}^2) = E \left[ (r_{1, t} + r_{2, t})^4 \right] = E(r_{1, t}^4 + r_{2, t}^4 + 4r_{1, t}^3 r_{2, t} + 6r_{1, t}^2 r_{2, t}^2) \]
\[ \overset{A_1 - A_2}{=} E(r_{1, t}^4 + 6r_{1, t}^2 r_{2, t}^2) \]

More generally, it follows that:

\[
E(V_{1, t}^2) = E \left( \sum_{i=1}^{K} r_{i, t}^4 + 3 \sum_{i=1}^{K} \sum_{j \neq i} r_{i, t}^2 r_{j, t}^2 \right) \]

A similarly, for \( V_{2, t} \),

\[
E(V_{2, t}^2) = E \left[ (r_{1, t}^2 + r_{2, t}^2)^2 \right] = E(r_{1, t}^4 + r_{2, t}^4 + 2r_{1, t}^2 r_{2, t}^2) \]

or more generally:

\[
E(V_{2, t}^2) = E \left( \sum_{i=1}^{K} r_{i, t}^4 + \sum_{i=1}^{K} \sum_{j \neq i} r_{i, t}^2 r_{j, t}^2 \right) \]

Since \( (r_{i, t}^2 r_{j, t}^2) \geq 0 \), it follows directly that \( E(V_{1, t}^2) \geq E(V_{2, t}^2) \) and since \( E(V_{1, t}) = E(V_{2, t}) = V(r_t), V(V_{1, t}) > V(V_{2, t}) \).
**A1.2: Realized Skewness**

Recall that the skewness is defined as: $E[(r_t - E(r_t))^{3}] E(r_t^2)^{-3/2} = E(r_t^3) E(r_t^2)^{-3/2}$.

For this reason, a first measure of the daily cube returns are naturally $S_{1,t} \equiv r_t^3$. Extending the idea of realized volatility to the third moment would suggest the estimator $S_{2,t} \equiv \sum_{i=1}^{K} r_{i,t}^3$ is an unbiased estimator of $E(r_t^3)$.

**Assumption** $A_3$: $E(r_{i,t}^3, r_{j,t} \mid \Omega_{j-1,t}) = 0$, $\forall i > j$.

**Proposition 3:** Under $A_1, A_3, S_{2,t}$ is an unbiased estimator of $E(r_t^3)$.

**Proof.** For the sake of simplicity, let us consider the case where $K = 2$.

$$E(r_t^3) = E[(r_{1,t} + r_{2,t})^3] = E(r_{1,t}^3 + r_{2,t}^3 + 3 r_{1,t} r_{2,t}^2 + 3 r_{1,t}^2 r_{2,t} + 3 r_{1,t} r_{2,t}^2) \overset{A_1, A_3}{=} E(r_{1,t}^3 + r_{2,t}^3) = E(S_{2,t}).$$

The better estimators dominates in terms of efficiency depends on sign and size of higher moments as the simple case when $K = 2$ shows.

**Proof.**

$$E(S_{1,t}^3) = E[(r_{1,t} + r_{2,t})^6] \overset{A_3}{=} E(r_{1,t}^6 + r_{2,t}^6 + 6 r_{1,t} r_{2,t}^5 + 15 r_{1,t}^2 r_{2,t}^4 + 20 r_{1,t}^3 r_{2,t}^3 + 15 r_{1,t}^4 r_{2,t}^2 + 15 r_{1,t}^3 r_{2,t}^3) \overset{A_1, A_3}{=} E(r_{1,t}^6 + r_{2,t}^6) = E(S_{2,t}^6),$$

$$E(S_{1,t}^3) - E(S_{2,t}^3) = E(18 r_{1,t}^2 r_{2,t}^4 + E(6 r_{1,t}^3 r_{2,t}^3) + E(15 r_{1,t}^4 r_{2,t}^2 + 15 r_{1,t}^3 r_{2,t}^3) > 0,$$

Assuming independence over time of the intradaily returns, which implies Assumption $A_3$, and assuming a symmetric distribution around the mean of these returns, we get that $E(r_{1,t}^3, r_{2,t}^3) = 0$ and $E(r_{1,t} r_{2,t}^2, r_{1,t}^2 r_{2,t}^2) = 0$. In that case, $E(S_{1,t}^3) - E(S_{2,t}^3) = E(15 r_{1,t}^4 r_{2,t}^2 + 15 r_{1,t}^3 r_{2,t}^3) > 0$, which implies that $S_{2,t}$ is more efficient than $S_{1,t}$.
Appendix 2: Decomposition of the Portfolio Moments

The expected return, the variance, the skewness and the kurtosis of a portfolio $p$, composed by $N$ risky assets are respectively given by:

$$
\begin{align*}
E(R_p) &= w_{p0} R_f + \sum_{i=1}^{N} w_{pi} R_i \\
\sigma^2(R_p) &= E\left\{[R_p - E(R_p)]^2\right\} \\
s^3(R_p) &= E\left\{[R_p - E(R_p)]^3\right\} \\
\kappa^4(R_p) &= E\left\{[R_p - E(R_p)]^4\right\}
\end{align*}
$$

with:

$$
\begin{align*}
R_p &= w_{p0} R_f + \sum_{i=1}^{N} w_{pi} R_i \\
\sum_{i=1}^{N} w_{pi} &= (1 - w_{p0})
\end{align*}
$$

where $R_p$, $R_i$ and $R_f$ are respectively the portfolio return, the return of the risky asset $i$, with $i = [1, \ldots, N]$ and the return of the non risky asset; $w_{pi}$ represent the weight of the risky asset in the portfolio $p$ and $w_{p0}$ is the non risky asset weight.

While rearranging the terms, it comes:

$$
\begin{align*}
E(R_p) &= w_{p0} R_f + \sum_{i=1}^{N} w_{pi} E(R_i) \\
\sigma^2(R_p) &= \sum_{i=1}^{N} \sum_{j=1}^{N} w_{pi} w_{pj} \sigma_{ij} \\
s^3(R_p) &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} w_{pi} w_{pj} w_{pk} s_{ijk} \\
\kappa^4(R_p) &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} \kappa_{ijkl}
\end{align*}
$$

with:

$$
\begin{align*}
\sigma_{ij} &= E\{[R_i - E(R_i)][R_j - E(R_j)]\} \\
s_{ijk} &= E\{[R_i - E(R_i)][R_j - E(R_j)][R_k - E(R_k)]\} \\
\kappa_{ijkl} &= E\{[R_i - E(R_i)][R_j - E(R_j)][R_k - E(R_k)][R_l - E(R_l)]\}
\end{align*}
$$

where $(\sigma_{ij})_{i,j}$, $(s_{ijk})_{i,j,k}$ and $(\kappa_{ijkl})_{i,j,k,l}$ represent, respectively, the covariance of the returns of the risky assets $i$ and $j$, the co-skewness of the returns of the assets $i$, $j$ and $k$, and the cokurtosis of the returns of the assets $i$, $j$, $k$ and $l$, with $(i,j,k,l) = (IN^*)^4$.

What can be expressed, by using tensorial notations, in the following way:

$$
\begin{align*}
E(R_p) &= w_{p0} R_f + w_p^T E \\
\sigma^2(R_p) &= w_p^T \Omega w_p \\
s^3(R_p) &= w_p^T \Sigma (w_p \otimes w_p) \\
\kappa^4(R_p) &= w_p^T \Gamma (w_p \otimes w_p \otimes w_p)
\end{align*}
$$

with:

$$
\begin{align*}
\Omega &= \sum_{i=1}^{N} \sum_{j=1}^{N} w_{pi} w_{pj} \\
\Sigma &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} w_{pi} w_{pj} w_{pk} \\
\Gamma &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl}
\end{align*}
$$
\[ w_p' \mathbf{1} = (1 - w_{p0}) \]

where \( w_p \) is the \((N \times 1)\) vector of portfolio weights of risky assets and \( w_p' \) its is transposed; \( \mathbf{1} \) is the \((N \times 1)\) vector of one; \( \mathbf{E} \) is \((N \times 1)\) vector of the expected return of the \( N \) risky assets; \( \Omega \) is the \((N \times N)\) variance-covariance matrix of returns; \( \Sigma \) represents the \((N \times N^2)\) skewness-coskewness matrix and \( \Gamma \) is the \((N \times N^3)\) kurtosis-cokurtosis matrix of the return of the \( N \) risky assets, defined by:

\[
\begin{align*}
(\Sigma)_{(N\times N^2)} &= (\Sigma_1 \Sigma_2 \cdots \Sigma_N) \\
(\Gamma)_{(N\times N^3)} &= (\Gamma_{11} \Gamma_{12} \cdots \Gamma_{1N} | \Gamma_{21} \Gamma_{22} \cdots \Gamma_{2N} | \cdots | \Gamma_{N1} \Gamma_{12} \cdots \Gamma_{NN})
\end{align*}
\]

where \( \Sigma_k \) and \( \Gamma_{kl} \) are the \((N \times N)\) submatrices associated with \( \Sigma \) and \( \Gamma \), with elements \((s_{ijk})_{i,j=1,\ldots,N}\) and \((\kappa_{ijkl})_{i,j=1,\ldots,N}\) and the sign \( \otimes \) is the symbol of Kronecker product.

We can rewrite them in the form:

\[
\begin{align*}
\sigma^2 (R_p) &= w_p' \Omega w_p \\
s^3 (R_p) &= w_p' \Sigma_p \\
\kappa^4 (R_p) &= w_p' \Gamma_p
\end{align*}
\]

with (by using the previews notations):

\[
\begin{align*}
\Sigma_p &= \Sigma (w_p \otimes w_p) \\
\Gamma_p &= \Gamma (w_p \otimes w_p \otimes w_p)
\end{align*}
\]

and:

\[ w_p' \mathbf{1} = (1 - w_{p0}) \]

where \( \Omega w_p \) represents the \((N \times 1)\) vector of covariances of the return of the \( N \) risky assets with the portfolio return; \( \Sigma_p \) is the \((N \times 1)\) vector of co-skewness of the return of the \( N \) risky assets with the portfolio return, \( \Gamma_p \) is the \((N \times 1)\) vector of cokurtosis of the return of the \( N \) risky assets with the portfolio return; \( \mathbf{1} \) is the \((N \times 1)\) vector of one; and \( w_p, w_p', \Omega, \Sigma, \Gamma \) and the sign \( \otimes \) correspond respectively to the vector of weights of the \( N \) risky assets in the portfolio \( p \), to the weights vector transposed, to the matrices of variance-covariance, skewness-coskewness and cokurtosis-cokurtosis of the \( N \) risky assets and to the Kronecker product symbol.

From the previews relations, we can find a finer decomposition of the moments (Skewness and Kurtosis) of the portfolio.
5.1. Decomposition of the Skewness

The Skewness of a portfolio \( p \) is given by:

\[
s^3 (R_p) = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} w_{pi} w_{pj} w_{pk} s_{ijk}
\]  

with:

\[
s_{ijk} = E \{ [R_i - E (R_i)] [R_j - E (R_j)] [R_k - E (R_k)] \}
\]  

By developing and rearranging the equation (47), we have:

\[
s_{ijk} = E (R_i R_j R_k) - E (R_i R_j R_k) - E (R_i R_j R_k) + E (R_i R_j R_k) - E (R_i R_j R_k)
\]

where, by replacing in the equation (46) \( s_{ijk} \) by the relation (48), we obtain:

\[
s^3 (R_p) = Q_{ijk} - (Q_{i\bar{j}k} + Q_{i\bar{k}j} + Q_{\bar{i}j\bar{k}}) + (Q_{i\bar{j}k} + Q_{i\bar{k}j} + Q_{\bar{i}j\bar{k}}) - Q_{\bar{i}j\bar{k}}
\]

with:

\[
\begin{align*}
Q_{ijk} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} w_{pi} w_{pj} w_{pk} E (R_i R_j R_k) \\
Q_{i\bar{j}k} &= \sum_{k=1}^{N} w_{pk} \bar{R}_k \sum_{i=1}^{N} \sum_{j=1}^{N} w_{pi} w_{pj} E (R_i R_j) \\
Q_{i\bar{k}j} &= \sum_{i=1}^{N} w_{pi} w_{pk} R_i \sum_{j=1}^{N} w_{pj} E (R_j) \\
Q_{\bar{i}j\bar{k}} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} w_{pi} w_{pj} w_{pk} E (\bar{R}_i \bar{R}_j \bar{R}_k)
\end{align*}
\]

i.e., by synthesizing the developments:

\[
s^3 (R_p) = \mathbf{w}'_p (\mathbf{\Sigma}^* - \mathbf{\bar{\Sigma}}^*) (\mathbf{w}_p \otimes \mathbf{w}_p) - 3\mathbf{w}'_p \mathbf{E} \mathbf{w}'_p (\mathbf{\Omega}^* - \mathbf{\bar{\Omega}}^*) \mathbf{w}_p
\]

with:

\[
\begin{align*}
\mathbf{\Omega}^* &= (\sigma^*_{ij})_{i,j=1,\ldots,N} \quad \text{with } \sigma^*_{ij} = E (R_i R_j) \\
\mathbf{\bar{\Omega}}^* &= (\bar{\sigma}^*_{ij})_{i,j=1,\ldots,N} \quad \text{with } \bar{\sigma}^*_{ij} = E (\bar{R}_i \bar{R}_j) \\
\mathbf{\Sigma}^* &= (\mathbf{\Sigma}_1^* \mathbf{\Sigma}_2^* \cdots \mathbf{\Sigma}_N^*) \text{ with } \mathbf{\Sigma}_k^* = (s^*_{ijk})_{i,j,k=1,\ldots,N} \quad \text{and } s^*_{ijk} = E (R_i R_j R_k) \\
\mathbf{\bar{\Sigma}}^* &= (\mathbf{\bar{\Sigma}}_1^* \mathbf{\bar{\Sigma}}_2^* \cdots \mathbf{\bar{\Sigma}}_N^*) \text{ with } \mathbf{\bar{\Sigma}}_k^* = (\bar{s}^*_{ijk})_{i,j,k=1,\ldots,N} \quad \text{and } \bar{s}^*_{ijk} = E (\bar{R}_i \bar{R}_j \bar{R}_k)
\end{align*}
\]

where \( \mathbf{W}_p \) are the \((N \times 1)\) vector of the risky assets weights and \( \mathbf{W}'_p \) its are transposed; \( \mathbf{E} \) are the \((N \times 1)\) vector of the expected return of the \(N\) risky assets; \( \mathbf{\Omega}^* \) are the order 2 covariation \((N \times N)\) matrix of the \(N\) assets and \( \mathbf{\Sigma} \) represent the order 3 covariation \((N \times N^2)\) matrix.
5.2. Decomposition of the Kurtosis

The Kurtosis of a portfolio \( p \) is given by:

\[
\kappa^4 (R_p) = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} \kappa_{ijkl} = (51)
\]

with:

\[
\kappa_{ijkl} = E \{ [R_i - E (R_i)] [R_j - E (R_j)] [R_k - E (R_k)] [R_l - E (R_l)] \} = (52)
\]

By developing equation (52), we have:

\[
\begin{align*}
\kappa_{ijkl} &= E \{ [R_i - E (R_i)] [R_j - E (R_j)] [R_k - E (R_k)] [R_l - E (R_l)] \} \\
&= (53)
\end{align*}
\]

Therefore the equation (51) becomes:

\[
\begin{align*}
\kappa^4 (R_p) &= Q_{ijkl} - (Q_{ijkl} + Q_{ijkl} + Q_{ijkl} Q_{ijkl}) \\
&= (Q_{ijkl} + Q_{ijkl} + Q_{ijkl} + Q_{ijkl} + Q_{ijkl} + Q_{ijkl}) \\
&= (Q_{ijkl} + Q_{ijkl} + Q_{ijkl} + Q_{ijkl} + Q_{ijkl} + Q_{ijkl}) \\
&= (54)
\end{align*}
\]

with:

\[
\begin{align*}
Q_{ijkl} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} E (R_i R_j R_k R_l) \\
Q_{ijkl} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} E (R_i R_j R_k R_l) \\
Q_{ijkl} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} E (R_i R_j R_k R_l) \\
Q_{ijkl} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} E (R_i R_j R_k R_l) \\
Q_{ijkl} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} E (R_i R_j R_k R_l) \\
Q_{ijkl} &= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} w_{pi} w_{pj} w_{pk} w_{pl} E (R_i R_j R_k R_l) \\
\end{align*}
\]

finally:

\[
\begin{align*}
\kappa^4 (R_p) &= w_p' (\Gamma^* - \tilde{\Gamma}^*) (w_p \otimes w_p \otimes w_p) \\
&= -4 w_p' E w_p' (\Sigma^* - \tilde{\Sigma}^*) (w_p \otimes w_p) \\
&= +6 \left( w_p' \Omega^* w_p \right) (w_p' \tilde{\Omega}^* w_p) = (55)
\end{align*}
\]
with:

\[
\begin{align*}
\Omega^* &= \begin{pmatrix} \sigma^*_{ij} \end{pmatrix}_{i,j=1,\ldots,N} \quad \text{with } \sigma^*_{ij} = E(R_iR_j) \\
\bar{\Omega}^* &= \begin{pmatrix} \bar{\sigma}^*_{ij} \end{pmatrix}_{i,j=1,\ldots,N} \quad \text{with } \bar{\sigma}^*_{ij} = E(\bar{R}_i\bar{R}_j) \\
\Sigma^* &= (\Sigma^*_1 \Sigma^*_2 \cdots \Sigma^*_N) \quad \text{with } \Sigma^*_k = \begin{pmatrix} s^*_{ijk} \end{pmatrix}_{i,j=1,\ldots,N} \quad \text{and } s^*_{ijk} = E(R_iR_jR_k) \\
\bar{\Sigma}^* &= (\bar{\Sigma}^*_1 \bar{\Sigma}^*_2 \cdots \bar{\Sigma}^*_N) \quad \text{with } \bar{\Sigma}^*_k = \begin{pmatrix} \bar{s}^*_{ijk} \end{pmatrix}_{i,j=1,\ldots,N} \quad \text{and } \bar{s}^*_{ijk} = E(\bar{R}_i\bar{R}_j\bar{R}_k) \\
\Gamma^* &= (\Gamma^*_{11} \Gamma^*_{12} \cdots \Gamma^*_{1N} \mid \Gamma^*_{21} \Gamma^*_{22} \cdots \Gamma^*_{2N} \mid \cdots \mid \Gamma^*_{N1} \Gamma^*_{12} \cdots \Gamma^*_{NN}) \\
\text{where } \Gamma^*_{kl} &= \begin{pmatrix} \kappa^*_{ijkl} \end{pmatrix}_{i,j=1,\ldots,N} \quad \text{and } \kappa^*_{ijkl} = E(R_iR_jR_kR_l) \\
\bar{\Gamma}^* &= (\bar{\Gamma}^*_{11} \bar{\Gamma}^*_{12} \cdots \bar{\Gamma}^*_{1N} \mid \bar{\Gamma}^*_{21} \bar{\Gamma}^*_{22} \cdots \bar{\Gamma}^*_{2N} \mid \cdots \mid \bar{\Gamma}^*_{N1} \bar{\Gamma}^*_{12} \cdots \bar{\Gamma}^*_{NN}) \\
\text{where } \bar{\Gamma}^*_{kl} &= \begin{pmatrix} \bar{\kappa}^*_{ijkl} \end{pmatrix}_{i,j=1,\ldots,N} \quad \text{and } \bar{\kappa}^*_{ijkl} = E(\bar{R}_i\bar{R}_j\bar{R}_k\bar{R}_l)
\end{align*}
\]

where \( W_p \) are the \((N \times 1)\) vector of the risky assets weights and \( W_p' \) its are transposed; \( E \) are the \((N \times 1)\) vector of the expected return of the \( N \) risky assets; \( \Omega^* \) are the order 2 covariation \((N \times N)\) matrix of the \( N \) assets; \( \Sigma \) represent the order 3 covariation \((N \times N^2)\) matrix and \( \Gamma^* \) are the order 4 covariation \((N \times N^3)\) matrix of the \( N \) risky assets.
Appendix 3: List of Firms

This table gives list of the 43 firms included in the analysis. For these stocks, the database includes a complete set of five-minutes transactions prices from January 02, 2002, through December 29, 2006 (1,281 trading days).

<table>
<thead>
<tr>
<th>Code</th>
<th>Names</th>
<th>Code</th>
<th>Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>3232</td>
<td>TFI</td>
<td>8058</td>
<td>SCHNEIDER ELECTRIC</td>
</tr>
<tr>
<td>3340</td>
<td>EADS</td>
<td>8088</td>
<td>VIVENDI ENVIRON.</td>
</tr>
<tr>
<td>7886</td>
<td>AIR LIQUIDE</td>
<td>8131</td>
<td>SAINT-GOBAIN</td>
</tr>
<tr>
<td>7894</td>
<td>CARREFOUR</td>
<td>8146</td>
<td>CAP GEMINI</td>
</tr>
<tr>
<td>7903</td>
<td>TOTAL FINA ELF</td>
<td>8155</td>
<td>CANAL +</td>
</tr>
<tr>
<td>7908</td>
<td>OREAL</td>
<td>8157</td>
<td>VINCI</td>
</tr>
<tr>
<td>7910</td>
<td>VALLOUREC</td>
<td>8159</td>
<td>CASINO GUICHARD</td>
</tr>
<tr>
<td>7914</td>
<td>ACCOR</td>
<td>8181</td>
<td>AGF</td>
</tr>
<tr>
<td>7924</td>
<td>BOUYGUES</td>
<td>8194</td>
<td>VIVENDI UNIVERSAL</td>
</tr>
<tr>
<td>7926</td>
<td>SUEZ</td>
<td>8217</td>
<td>DEXIA SICO.</td>
</tr>
<tr>
<td>7927</td>
<td>LAFARGE</td>
<td>8279</td>
<td>STMICROELEC.SICO.</td>
</tr>
<tr>
<td>7931</td>
<td>SANOFI SYNTHELABO</td>
<td>8287</td>
<td>ALCATEL A</td>
</tr>
<tr>
<td>7935</td>
<td>AXA</td>
<td>8301</td>
<td>LAGARDERE</td>
</tr>
<tr>
<td>7937</td>
<td>GROUPE DANONE</td>
<td>8312</td>
<td>VALEO</td>
</tr>
<tr>
<td>7942</td>
<td>PERNOD-RICARD</td>
<td>8335</td>
<td>PUBLICIS GROUPE</td>
</tr>
<tr>
<td>7971</td>
<td>LVMH MOET VUITTON</td>
<td>8350</td>
<td>SOCIETE GENERALE A</td>
</tr>
<tr>
<td>7990</td>
<td>SODEXHO ALLIANCE</td>
<td>8356</td>
<td>BNP PARIBAS</td>
</tr>
<tr>
<td>7994</td>
<td>MICHELIN</td>
<td>8421</td>
<td>RENAULT</td>
</tr>
<tr>
<td>8000</td>
<td>THALES</td>
<td>8546</td>
<td>FRANCE TELECOM</td>
</tr>
<tr>
<td>8013</td>
<td>PINAULT PRINTEMPS</td>
<td>11149</td>
<td>THOMSON</td>
</tr>
<tr>
<td>8015</td>
<td>PEUGEOT</td>
<td>84569</td>
<td>CREDIT AGRICOLE</td>
</tr>
<tr>
<td>8029</td>
<td>ESSILOR INTL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Euronext. This table reports a list of the 43 firms included in the analysis. For these stocks, the database includes a complete set of five-minutes transactions prices from January 02, 2002, through December 29, 2006 (1,281 trading days).