Robust reverse engineering of cross-sectional returns and improved portfolio allocation performance using the CAPM

XIAOHUI NI, YANNICK MALEVERGNE, DIDIER SORNETTE, AND PETER WOEHRMANN

XIAOHUI NI

is a Ph. D candidate in the Research Center for Econophysics, East China University of Science and Technology, Shanghai, China.
nicoson@mail.ecust.edu.cn

YANNICK MALEVERGNE

is a professor of finance at the Université de Lyon – Université de Saint-Etienne, France, and deputy director of Coactis E.A. 4161. He is also Affiliate professor of finance at EM LYON Business School, France, and Associate researcher at ETH Zürich, Switzerland.
ymalevergne@ethz.ch

DIDIER SORNETTE

holds the chair of Entrepreneurial Risks in the department of Management, Technology and Economics at ETH Zürich, Zürich, Switzerland and is professor of finance at the Swiss Finance Institute, c/o University of Geneva, 40 blvd. Du Pont d'Arve, CH 1211 Geneva 4, Switzerland
dsornette@ethz.ch

PETER WOEHRMANN

is a visiting assistant professor in the department of management science and engineering at Stanford University in Palo Alto, CA.
peterwl@stanford.edu

Keywords: CAPM, mean-variance portfolio optimization, constrained optimization, Fama-French, value-size portfolios, dynamical allocation, expected returns

JEL: G11, G12

Explaining the cross-sectional characteristics of asset returns is the grail of financial economists. Using their properties to devise efficient asset allocations is the quest of professional managers. The mean-variance approach [Markovitz, 1952] and the CAPM (capital asset pricing model) were the first integration of these two goals into a fully coherent self-consistent framework [Treynor, 1961; Sharpe 1964; Lintner, 1965; Mossin, 1966]: because rational investors all choose to invest in the optimal tangent Markowitz portfolio, the expected return of each asset has to adjust at equilibrium, i.e., under the balance of supply and demand, to exhibit a risk premium that is simply proportional to the expected return of the global market portfolio, which is itself the tangent Markowitz portfolio. This kills two birds with one stone, as it provides a full explanation of financial returns and makes portfolio managers redundant since the optimal investment strategy is to hold a mixture of the market portfolio and the risk-free asset (supposedly a short-term US government bond). In practice, it is well-known that the CAPM does not work (and cannot even be properly tested [Roll, 1977]), so that a wealth of extensions and alternative theories have been concerned with refining, extending, or generalizing altogether the CAPM, leading for instance to the arbitrage pricing theory [Ross, 1976], or to the phenomenological three factor model of Fama and French [1993] that accounts for size and book-to-market effect or the four factor model of Cahart [1997] that introduces the momentum effect among many others. The emphasis has been and still is on the many pricing anomalies, on the non-Gaussian and long memory properties of financial returns and on the behavioral bounded rationality of investors and how they impact the cross-sectional characteristics of asset returns and the art of investment management.

Recently, Levy and Roll [2010] proposed a radically different perspective and suggested that the inadequacy of the CAPM could just result from statistical errors. Markowitz' portfolio optimization requires the knowledge of both the expected returns and of the covariance matrix of the assets. It is well known that the optimum portfolio weights are very sensitive to return expectations, which are very difficult to determine. For instance, historical returns are bad predictors of future returns [Siegel, 2007]. Estimating covariance matrices is a delicate statistical challenge that requires sophisticated methods (see for instance [Ledoit and Wolf, 2004]). It is fair to state that, due to the large statistical errors of the inputs of Markowitz' portfolio optimization, its results are not reliable and should be considered very cautiously. This led Levy and Roll [2010] to turn the usual approach on its head and posit that the market portfolio is the efficient tangent Markowitz portfolio, i.e., it is mean-variance efficient, and ask how much the inputs should be tinkered with to allow this. Remarkably, they found that only minor adjustments of the input parameters are needed, well within the statistical uncertainties. They showed in addition that the systematic risk $\hat{\beta}_i$ of the *i*th asset derived from their procedure is consistent with the β_i calculated directly from CAPM.

Here, we present a series of tests and applications. We apply first the Levy-Roll procedure to the 25 Fama-French portfolios sorted by sizes and book-to-market values. We check the consistency of the Levy-Roll approach by investigating how the adjusted stock returns of specific stocks are modified when varying the basket of stocks they belong to. We test the dynamical performance of the Levy-Roll procedure over the period from January 1992 to December 2009. Finally, we show how to exploit the method for better portfolio allocations.

THE LEVY-ROLL PROCEDURE

Levy and Roll have found that applying small variations on the asset expected returns and covariance within their estimation error bounds make the market portfolio proxy mean-variance efficient. Following their results, a new estimation procedure of these return parameters can be proposed. The key idea is to find the return parameter vector μ and standard deviation parameter vector σ , which are closest to their sample counterparts $(\mu, \sigma)^{\mathbf{sam}}$, and at the same time ensure that the market proxy *m* is mean-variance efficient. Then, standard statistical tests are applied to qualify that one cannot reject at the usual confidence level of 95% that the obtained vectors μ and σ are generated from the same distribution as the data. Specifically, we numerically solve the optimization problem of EXHIBIT 1.

EXHIBIT 1

Optimization Problem 1:

Minimize with respect to (μ, σ):

$$\mathbf{D}\left((\boldsymbol{\mu}, \boldsymbol{\sigma}), (\boldsymbol{\mu}, \boldsymbol{\sigma})^{\mathbf{sam}}\right) = \sqrt{\alpha \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\mu_i - \mu_i^{sam}}{\sigma_i^{sam}} \right)^2 + (1 - \alpha) \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\sigma_i - \sigma_i^{sam}}{\sigma_i^{sam}} \right)^2}$$

Subject to:

$$\begin{bmatrix} \sigma_1 & 0 & \mathrm{L} & 0 \\ 0 & \sigma_2 & & \mathrm{M} \\ \mathrm{M} & & \mathrm{O} & 0 \\ 0 & \mathrm{L} & 0 & \sigma_N \end{bmatrix} \begin{bmatrix} & & \\ & \rho^{sam} & \\ & & \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \mathrm{L} & 0 \\ 0 & \sigma_2 & & \mathrm{M} \\ \mathrm{M} & & \mathrm{O} & 0 \\ 0 & \mathrm{L} & 0 & \sigma_N \end{bmatrix} \begin{bmatrix} x_{m1} \\ x_{m2} \\ \mathrm{M} \\ x_{mN} \end{bmatrix} = q \cdot \begin{bmatrix} \mu_1 - r_f \\ \mu_2 - r_f \\ \mathrm{M} \\ \mu_N - r_f \end{bmatrix}$$

In EXHIBIT 1, the function D(μ, σ) is defined as the distance between any parameter set (μ, σ) and the sample parameter set $(\mu, \sigma)^{\mathbf{sam}}$. The parameter α weights the relative impact of deviations of the returns versus the standard deviations from their sample counterparts. Scanning all possible values of α from 0.1 to 0.9, we find that the results are robust for α in the range from 0.5 to 0.75 (this last value being that chosen by Levy and Roll (2010)), with perhaps more stable results for α=0.6, the value corresponding to the results presented below.

The constraint in the optimization problem 1 is nothing but the mean-variance portfolio equation, where ρ^{sam} is the sample correlation matrix, x_m is the vector of market proxy portfolio weights, $q = \dfrac{\sigma_m^2}{R_m - r_f}$, σ_m is the market standard deviation, R_m is the market expected returns, and r_f is the risk free returns.

In order to calculate the covariance matrices, we apply the shrinkage method of Ledoit and Wolf [2003, 2004], which addresses the issues of singular covariance matrices occurring when the number of data points per asset (here 168 monthly returns) is smaller than the number of assets.

APPLICATION TO THE 25 FAMA-FRENCH PORTFOLIOS

We first investigate how this procedure performs on the Fama-French portfolio benchmarks [Fama and French, 1993]. Consider the 25 portfolios constructed from the size and book-to-market quintiles of the US market portfolio, as described on K. French [2010]'s website, based on the stocks in the NYSE, Amex and NASDAQ stock markets. These 25 portfolios are the canonical supporters of the three-factor Fama-French factor model. Is it possible that the Levy-Roll procedure could equally well explain the excess returns of these 25 portfolios, with just the single market factor without the need for the two additional size and book-to-market factors?

We proceed as follows. We first use the NYSE market. Using the data available at the end of June of each year *t,* the NYSE stocks are sorted by size (*ME*, market equity defined as the price times the number of outstanding shares) and (independently) by book-to-market equity (*BE*/*ME*, the ratio of the book value of a firm's common stock (*BE*) to its market value). For the list of stocks sorted by book-to-market ratio, the *ME* of a given stock is defined as the market equity at the end of December of the former year $(t-1)$, and the *BE* is the common book equity for the fiscal year ending in calendar year $(t-1)$. The stocks of the NYSE are then split into 5 size quintiles and 5 book-to-market quintiles, and the breakpoints of the size quintile s_i and of the book-to-market quintile b_i is recorded for the next step. Then, all stock data of the three markets (NYSE, Amex and NASDAQ) are included and split into 5 size quintiles and 5 book-to-market quintiles by using the NYSE breakpoints s_i and b_i. The 25 portfolios are then obtained as the intersections of the size and *BE*/*ME* quintiles. Applying the Levy-Roll procedure described in Exhibit 1 on each of the 25 Fama-French portfolios, we obtain 25 sets (μ, σ), on which we apply two statistical tests, one for the vector μ , and one for the vector σ . For 23 portfolios out of 25, we find that all estimated expected stock returns are located in the 95% confidence interval derived from the *t*-value of the mean estimator. For the two remaining portfolios (Size 3 - *BE*/*ME* 2 and Size Big - *BE*/*ME* 2), the μ estimates of 97% stocks are located in the 95% confidence interval.

In order to qualify the statistical significance of the estimated σ 's, the ratio $\left((\sigma_i^*)^2 / (\sigma_i^{sam})^2\right)$ is calculated. Under the null hypothesis that the obtained standard deviations are generated by the same distribution as the sample estimates, we have $(n-1)\sigma^{sam^2} / \sigma^2 \sim \chi^2(n-1)$, where n is the number of observations. For 24 portfolios out of 25, we find that all estimated standard deviations are located in the 95% confidence interval defined above. For the remaining portfolio (Size 4 - *BE*/*ME* 2), the σ estimators of 97% stocks are found to be within the 95% confidence interval. From the two tests, it can be concluded that most of the estimated parameters are well within 95% confidence interval of the null hypothesis that these values could have been generated by the same process as the data sample.

EXHIBIT 2 lists the number of stocks in each of these 25 Fama-French like portfolios constructed from 949 stocks in our database on the year 2009. We restrict our analysis to the 949 stocks for which we could ascertain data quality and completeness (see Appendix A).

Applying the Levy-Roll procedure described in Exhibit 1 on each of the 25 Fama-French portfolios, we obtain 25 sets (μ, σ), on which we apply two statistical tests, one for the vector μ, and one for the vector σ. For 23 portfolios out of 25, we find that all estimated expected stock returns are located in the 95% confidence interval derived from the *t*-value of the mean estimator. For the two remaining portfolios (Size 3 - *BE*/*ME* 2 and Size Big - *BE*/*ME* 2), the μ estimates of 97% stocks are located in the 95% confidence interval.

In order to qualify the statistical significance of the estimated σ 's, the ratio $\left((\sigma_i^*)^2 / (\sigma_i^{sam})^2\right)$ is calculated. Under the null hypothesis that the obtained standard deviations are generated by the same distribution as the sample estimates, we have $(n-1)\sigma^{sam^2} / \sigma^2 \sim \chi^2(n-1)$, where n is the number of observations. For 24 portfolios out of 25, we find that all estimated standard deviations are located in the 95% confidence interval defined above. For the remaining portfolio (Size 4 - *BE*/*ME* 2), the σ estimators of 97% stocks are found to be within the 95% confidence interval. From the two tests, it can be concluded that most of the estimated parameters are well within 95% confidence interval of the null hypothesis that these values could have been generated by the same process as the data sample.

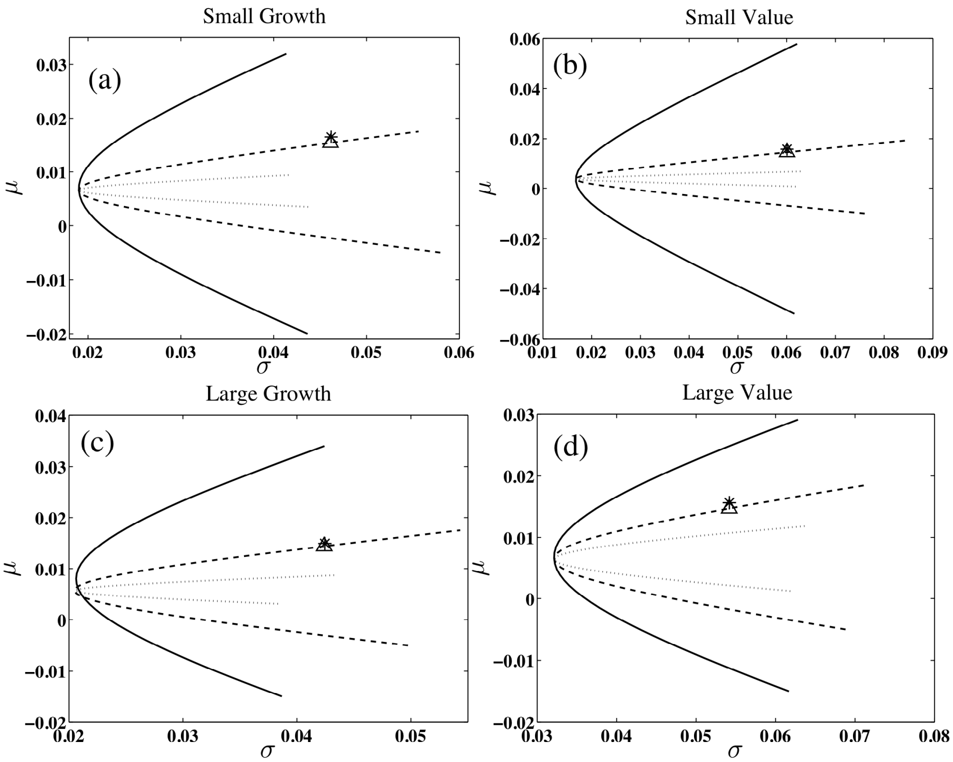
EXHIBIT 2

Number of Stocks in Each of the 25 Fama-French Portfolio [1]

Size quintile	Book-to-Market Equity (BE/ME)quintile				
	Low	2	3	4	High
Small	54	33	55	58	204
2	26	31	32	36	25
3	26	33	28	34	20
4	37	25	30	20	18
Big	52	33	21	16	2

EXHIBIT 3

The Efficient Frontiers of the four Special Asset Sets in the 25 Fama-French Portfolios



Note: In each of the four panels, the solid line represents the mean-variance frontier of the set of stocks described in the text with sample parameters, the dash line is the mean-variance frontier of the same asset set with parameters adjusted according to the Levy-Roll procedure, and the dotted-line corresponds to the hybrid function obtained by putting the adjusted parameters in the equation governing the solid line. The star denotes the market proxy with the sample parameters, and the triangle denotes the market proxy with the adjusted parameters. In panel (a), the Small Growth asset set is used, which contains in the 144 stocks with the two smallest sizes and the two smallest book-to-market values among the 25 Fama-French portfolios shown in Exhibit 2. In panel (b), the Small Value asset set is used, which contains the 323 stocks with the two smallest sizes and the two largest book-to-market values among the 25 Fama-French portfolios shown in Exhibit 2. In panel (c), the Large Growth asset set is used, which contains in the 147 stocks with the two largest sizes and the two smallest book-to-market values among the 25 Fama-French portfolios shown in Exhibit 2. In panel (d), the Large Value asset set is used, which contains the 56 stocks with the two largest size and two largest book-to-market values among the 25 Fama-French portfolios shown in Exhibit 2.

EXHIBIT 3 shows the mean-variance frontiers obtained for four asset sets among the 25 Fama-French portfolios, which are typical of the obtained results. One can observe a very large change in the shape and values of the mean-variance frontiers when going from the sample versions to the ones with parameters adjusted with the Levy-Roll procedure. In particular, we find a systematic large reduction of the expected returns for given risk levels, making the adjusted values look more conservative and reasonable. A good feature of these results is that the expected return and standard deviation of the market proxy remains practically unchanged when using the adjusted compared with the sample parameters. The rejection of the CAPM when using the sample parameters is seen by the fact that the market portfolio is far from the efficient frontier. In contrast, by construction, it lies on the mean-variance efficient frontier obtained with the adjusted parameters. Nonetheless, we have to observe that the CAPM alone is not enough to explain the risk-return tradeoff of the entire set of assets since the estimated risk-return couple of the mean-variance efficient portfolio that proxies the market portfolio significantly varies from one set of assets to the other. Therefore, even after adjustment, one single mean-variance efficient portfolio does not make vanish the size and book-to-market effect. Besides, based on the sample risk-return estimates for the set of small cap-value assets, investments in this style appear to be much more profitable than investments in the large cap-growth stocks at any risk level (see figures 3-b and 3c, solid line). However, using the adjusted risk-return parameters, the situation changes and the small-value stocks does not seem to outperform the large-growth stocks anymore (see figures 3-b and 3c, dashed line). This result challenges the conventional wisdom.

CONSISTENCY OF THE LEVY-ROLL METHOD

A natural concern is that the adjusted parameters found from the optimization problem 1 of Exhibit 1 may be specific to the selected set of stocks. When using a second set of stocks, it may be surmised that the adjusted parameters found for a given stock that is common to both sets and that solve the optimization problem 1 for this new set will be in general different from those obtained with the first set. If this would be the case, the Levy-Roll procedure would not reveal any insight on the cross-sectional properties of the universe of stocks. It would be just a mathematical exercise, where the optimization program of problem 1 leads to a kind of "over-fitting".

Levy and Roll [2010] partially addressed this issue by showing that the domains of compatibility found for different sets of stocks do not change too much when varying the number of assets in the optimization procedure. By domain of compatibility is meant the domain in the space (μ^*, σ^*) of solutions of problem 1 in Exhibit 1 such that 95% or more of the stocks have their parameters located in the 95% confidence intervals of their sample counterparts. Levy and Roll [2010] tested the stability of the domains of compatibility for the more constrained Problem 2 described in Exhibit 4, which adds to Problem 1 the conditions that the return and variance of the market portfolio are fixed to the pre-specified values (μ_0, σ_0) .

EXHIBIT 4

Optimization Problem 2:

Minimize with respect to (μ, σ):

$$\mathbf{D}((\mu,\sigma),(\mu,\sigma)^{sam}) = \sqrt{\alpha \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mu_i - \mu_i^{sam}}{\sigma_i^{sam}}\right)^2 + (1-\alpha)\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\sigma_i - \sigma_i^{sam}}{\sigma_i^{sam}}\right)^2}$$

Subject to:

(i) $$\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_N \end{bmatrix} \left[\quad \rho^{sam} \quad \right] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_N \end{bmatrix} \begin{bmatrix} x_{m1} \\ x_{m2} \\ \vdots \\ x_{mN} \end{bmatrix} = q \cdot \begin{bmatrix} \mu_1 - r_f \\ \mu_2 - r_f \\ \vdots \\ \mu_N - r_f \end{bmatrix}$$

(ii) $x'_m \mu = m\mu_0$

(iii) $$x'_m \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_N \end{bmatrix} \left[\quad \rho^{sam} \quad \right] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_N \end{bmatrix} x_m = \sigma_0^2 .$$

While encouraging, Levy and Roll's compatibility tests do not address the crucial question of the stability of the adjusted parameters for individual stocks, when they are included in different proxies of the market portfolios, with different other stock constituents and with different numbers of stocks. Indeed, it could be possible that two sets with different numbers of stocks have good overlapping compatible domain in the set of solutions (μ^*, σ^*), but may contain common stocks whose corresponding adjusted return parameters strongly differ. If this is the case, the Levy-Roll procedure would not reveal any insight on the cross-sectional properties of the universe of stocks and would just be a mathematical exercise, as already mentioned. It is thus essential to check that the adjusted parameters for the individual stocks are robust and genuinely uniquely associated with each stock.

This is done using the methodology explained in Appendix B. It consists in calculating the adjusted parameters of the common stocks in the two different asset sets. Then the two market portfolio proxies constructed on the common stocks with their estimators derived from the different asset sets are compared to decide whether the two estimators are compatible. Specifically, we compare the compatible domains in the parameter space (return, standard deviation) of the market portfolio proxies and define their intersection as the common domain. The results presented in Appendix B show that the common domains remain large, even when the number of stocks is varied a lot. This supports the concept that the Levy-Roll procedure is able to determine genuine hidden properties of the underlying assets.

DYNAMIC STABILITY

The previous series of tests have demonstrated the robustness and consistency of the Levy-Roll procedure with respect to variations of stock basket compositions. What about time consistency? This question is all the more acute, given the backward looking and unstable nature of the mean-variance efficient frontier as a function of time.

We use the 949 stocks from our cleaned Bloomberg database data from January 1992 to December 2009, organized in 16 rolling windows of three-year duration with yearly time step. The Levy-Roll procedure is applied to each of the sixteen 3-year-long windows to determine their corresponding adjusted parameters (μ and σ). EXHIBIT 5 summarizes the goodness of fit for each of the 16 time windows, using the two statistical tests on the vectors μ and σ described previously. Specifically, the column for μ reports the percentage of the assets whose estimators are located in the 95% confidence interval of the empirical μ . The column σ gives the percentage of assets whose estimators are located in the 95% confidence interval of the empirical σ . Twelve out of the sixteen three-year period confirm that the adjusted cross-sectional returns are fully compatible with the hypothesis that the market portfolio is mean-variant efficient.

EXHIBIT 5

The Goodness of the Parameters for the Portfolio in Different time periods

period	μ	σ	period	μ	Σ
1992-1994	93.89%	100%	2000-2002	98.84%	100%
1993-1995	94.84%	100%	2001-2003	98.10%	100%
1994-1996	94.63%	100%	2002-2004	22.87%	1.05%
1995-1997	16.75%	0%	2003-2005	100%	0%
1996-1998	96.94%	100%	2004-2006	98.63%	0%
1997-1999	96.52%	100%	2005-2007	93.47%	100%
1998-2000	95.26%	99.58%	2006-2008	97.79%	99.79%
1999-2001	98.84%	99.89%	2007-2009	99.79%	100%

For the four windows 1995-1997, 2002-2004, 2003-2005 and 2004-2006, the Levy-Roll procedure is rejected. Interestingly, they correspond to two periods (1995-1999 and 2003-2006) that have been documented to be primary examples of financial bubbles. Diagnosing bubbles in real time remains the ultimate challenge of the profession. However, following the logic that the Levy-Roll procedure indeed provides real hidden properties of the underlying assets, we propose that its rejection provides a novel tool to detect anomalous market regimes, such as bubbles.

The Levy-Roll procedure also allows us to construct ex-ante portfolios and study their performance. Specifically, we consider the minimum standard deviation portfolio and the tangent portfolio (i.e., whose point within the mean-variance plot is on the line tangent to the efficient frontier that intersects the ordinate at the risk free rate value). Their vectors of asset weights are given respectively by

$$\vec{W}_{min} = \frac{\Omega^{-1}\vec{1}}{\vec{1}^T\Omega^{-1}\vec{1}}, \vec{W}_{tgt} = \frac{\Omega^{-1}\left(\vec{\mu}-r_j\vec{1}\right)}{\vec{1}^T\Omega^{-1}\left(\vec{\mu}-r_j\vec{1}\right)}, \qquad (1)$$

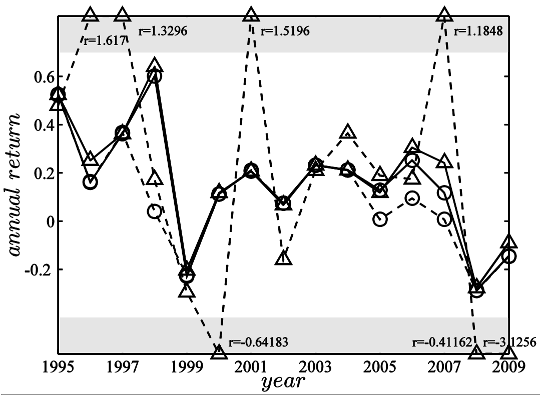
where Ω is the adjusted covariance matrix of the 949 stocks of our Bloomberg database. Their ex-ante annual returns R_{min} and R_{tgt} are given by

$$\begin{aligned} R_{tgt}(T) &= \vec{W}_{tgt}(T-1)\cdot\vec{R}(T) \\ R_{min}(T) &= \vec{W}_{min}(T-1)\cdot\vec{R}(T) \end{aligned}, \qquad (2)$$

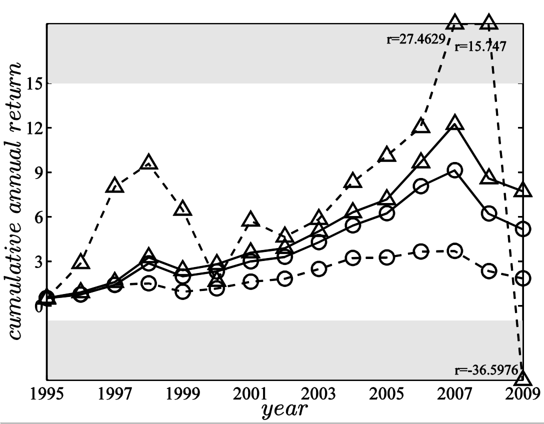
where $\vec{W}_{min}(T-1)$ and $\vec{W}_{tgt}(T-1)$ are the two row vectors of the weights calculated at the end of year $T-1$, and $\vec{R}(T)$ is the column vector of annual returns of all the stocks in the portfolio realized at the end of year T .

EXHIBIT 6

Annual Returns and Cumulative Returns of the Ex-Ante Portfolios



Panel A: Annual Returns of the Ex-Ante minimum standard deviation portfolio and of the tangent portfolio with both sample parameters and adjusted parameters. The dashed line with triangles plots the yearly returns of the tangent portfolio with sample parameters, while the solid line with triangles plots the yearly return of the tangent portfolio with adjusted parameters. The dashed line with circles plots the yearly returns the minimum standard deviation portfolio with sample parameters. The solid line with circles plots the returns of the minimum standard deviation portfolio with adjusted parameters. The crosses show the returns of the buy-and-hold strategy, i.e., to the portfolio of the 949 stocks.



Panel B: Cumulative Returns of the Ex-Ante minimum standard deviation and tangent portfolios, with the symbols as in panel A. The crosses show the cumulative returns of the buy-and-hold strategy, i.e., to the portfolio of the 949 stocks.

EXHIBIT 6 compares the yearly returns (panel A) and cumulative returns (panel B) of the minimum standard deviation portfolio and of the tangent portfolio, with stock weights and returns determined either from the sample parameters or from their Levy-Roll adjusted values. **EXHIBIT 7** lists the corresponding yearly returns for these four portfolios. One can see that the two portfolios with Levy-Roll adjusted values have the nice quality of being significantly less volatile and with more consistent positive returns than the optimal portfolios determined with the sample parameters. This supports further the notion that the Levy-Roll procedure extracts genuine and usable cross-sectional properties of the stock returns.

EXHIBIT 7

Table of the Annual Returns of the Ex-Ante Portfolios shown in Exhibit 6

year	$R_{tgt,sam}$	$R_{tgt,adj}$	$R_{min,sam}$	$R_{min,adj}$
1995	0.479	0.525	0.522	0.53
1996	1.617	0.252	0.164	0.16
1997	1.33	0.362	0.368	0.36
1998	0.172	0.641	0.04	0.6
1999	-0.295	-0.203	-0.224	-0.2
2000	-0.642	0.118	0.112	0.11
2001	1.52	0.207	0.211	0.21
2002	-0.16	0.067	0.074	0.07
2003	0.21	0.232	0.232	0.23
2004	0.364	0.211	0.213	0.21
2005	0.189	0.119	0.008	0.13

2007	1.185	0.242	0.008	0.12
2008	-0.412	-0.277	-0.288	-0.3
2009	-3.126	-0.09	-0.146	-0.1

Note: $R_{tgt,sam}$ denotes the tangent portfolio return calculated with the sample parameters, $R_{tgt,adj}$ denotes the tangent portfolio return calculated with the adjusted parameters, $R_{min,sam}$ denotes the minimum variance portfolio return calculated with the sample parameters, and $R_{min,adj}$ denotes the minimum variance portfolio return calculated with the adjusted parameters.

DYNAMICAL LEVY-ROLL OPTIMIZATION FOR SMALL PORTFOLIOS

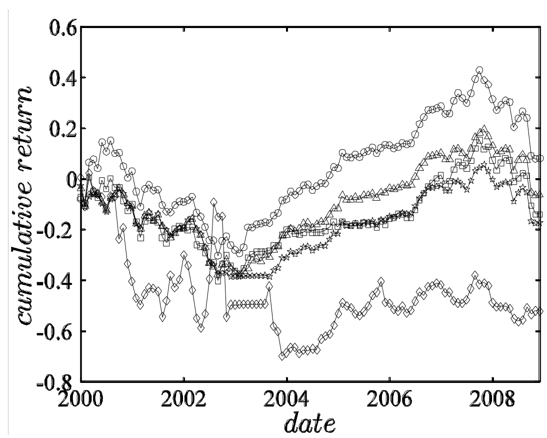
If the Levy-Roll approach truly reveals useful information on the cross-sectional properties of stocks, this should translate into improved portfolio selections. Here, we test this proposition on the basket formed by the 20 largest stocks by capitalization at the beginning of each period among all stocks in the CRSP database. We assume a monthly reallocation, based on the Levy-Roll procedure applied on the realized stock returns of the preceding five years. The data cover the period from January 1995 to December 2008, which correspond to one hundred and five monthly reallocations. The adjusted parameters (μ and σ) are estimated for each asset in each of the 105 5-year period and two portfolios are formed at the beginning of each month, with weights given by expression (1) now applied to the 20 largest stocks. The ex-ante annual returns R_{min} of the minimum variance portfolio and R_{tgt} of the tangent portfolio are determined with expression (2) with both sample and adjusted parameters for comparison. Their values known at the beginning of each month determine the trading strategy. If the predicted return is smaller than the risk free rate, the capital is invested on the risk free asset. If the predicted return is larger than the risk free rate, the capital is allocated to the portfolio with weights on each stock according to expression (1).

EXHIBIT 8

Average realized annual returns and cumulative returns of the tangential and minimum variance portfolio with both sample parameters and adjusted parameters.

	R_{market}	$R_{tgt,sam}$	$R_{tgt,adj}$	$R_{min,sam}$	$R_{min,adj}$
μ	-0,06	-0,06	0,06	-0,13	-0,03
σ	3,98	10,91	4,20	3,15	3,27
min	-11,68	-46,46	-12,05	-12,04	-12,07
max	11,31	49,97	17,76	6,45	6,44

Panel A: Average realized Annual Returns in % of the 5 Different Portfolios



Panel B: Realized Cumulative Returns of the 5 Different Portfolios

Note: The diamonds plots the return of the tangent portfolio with sample parameters, the circles plots the return of the tangent portfolio with adjusted parameters, the pentagram plot the minimum standard deviation portfolio with the sample parameters, the triangles plots the minimum standard deviation portfolio returns with adjusted parameters, and the squares plots the return of the value weighted portfolio as the bench mark .

EXHIBIT 8 shows the realized annual return of the tangential and minimum variance portfolio with both sample parameters and adjusted parameters, compared with the performance of the value weighted portfolio. Even for such a small basket of assets, one can observe that the portfolios with the Levy-Roll adjusted parameters are by far the least volatile and exhibit the best overall performance over the whole period.

CONCLUSION

We have presented a series of evidences suggesting that it is possible to infer more robust, consistent and useful cross-sectional properties of arbitrary baskets of stocks from a simple procedure proposed recently by Levy and Roll [2010]. Indeed, insisting that the market portfolio is mean-variance efficient is found to (i) explain the cross-sectional returns of the 25 Fama-French portfolios sorted by size and book-to-market values without the need for the two additional Fama-French factors, (ii) be consistent in extracting genuine, robust and idiosyncratic hidden properties of each stock return, independently of the basket of stocks, (iii) be dynamically robust over most studied time periods, (iv) and constitute a new method for optimal asset allocation with less volatile and larger realized returns. It thus seems that, not only the CAPM is not dead. On the contrary, insisting on its existence reveals novel cross-sectional properties of stocks that can be exploited operationally.

ENDNOTES

[1]The sum of the number of assets in each row (column) is different from one row (column) to the other due to the method retained to calculate the breakpoints of the size distribution of stocks. The quintiles are estimated on the basis of the stocks listed on the NYSE only, as in Fama and French approach. The breakpoints are then applied to the stocks listed on the AMEX, the NASDAQ and the NYSE.

[2]There are many problems and faults in the Bloomberg database. All stocks with missing data are removed from our database. About 3171 stocks remain after data cleaning, among which 949 stocks cover the period from July 1991 to June 2009.

[3]In our database, about 2335 out of 6546 stocks cover the period from January 1995 to December 2008.

[4]Since the 1-month Treasury Bill rate is not available on the website, we have to use 3-month Treasury Bill rate instead.

[5]The sample parameters in the whole paper are referred to the naïve average returns and standard deviations.

[6]The adjusted parameters in the whole paper are referred to the returns and standard deviations that are estimated from the Levy-Roll procedure.

[7]The market proxy in the whole paper is referred to value weighted market portfolio.

REFERENCES

Black, F. “Capital Market Equilibrium with Restricted Borrowing.” The Journal of Business, 45 (1972), pp. 444-455.

Carhart, M. “On persistence in mutual fund performance”, Journal of Finance 52 (1997), 57–82.

Fama, Eugene F., and French, Kenneth R. “Common Risk Factors in the Returns on Stocks and Bonds.” Journal of Financial Economics, 33 (1993), pp. 3-56.

Fama, Eugene F., and Richard Roll. “Some Properties of Symmetric Stable Distributions.” Journal of the American Statistical Association, 63(1968), pp. 817-36.

French, Kenneth R., Current Benchmark Returns,mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (2010)

Ledoit, O., and Wolf, M. “Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection.” Journal of Empirical Finance, 10 (2003), pp. 603-21.

Ledoit, O., and Wolf, M. “Honey, I Shrunk the Sample Covariance Matrix.” Journal of Portfolio Management, 30 (2004), pp. 110-119.

Levy, M., and Roll, R. “The Market Portfolio may be Mean-Variance Efficient After All.” Review of Financial Studies, 23 (2010), pp. 2464-2461.

Lintner, John. “Security Prices, Risk, and Maximal Gains from Diversification.” Journal of Finance, 20 (1965), pp 587-616.

Lintner, John. “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets.” Review of Economics and Statistics, 47 (1965), pp. 13-37.

Malevergne, Y., Santa-Clara, P., and Sornette, D. “Professor Zipf Goes to Wall Street.” NBER Working Paper #15295, 2009.

Markowitz, H. “Portfolio Selection”, Journal of Finance, 7 (1952), pp. 77-91.

Mossin, Jan. “Equilibrium in a Capital Asset Market.” Econometrica, 34 (1966), pp. 768-83.

Roll, R. “A Critique of the Asset Pricing Theory's Tests”, Journal of Financial Economics, 4 (March 1977), pp. 129-176

Ross, Stephen A. “The Arbitrage Theory of Capital Asset Pricing.” Journal of Economic Theory, 13 (1976), pp. 341-360.

Sharpe, William F. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." The Journal of Finance, 19(1964), pp 425-442.

Siegel, Jeremy, “Stocks for the Long Run”, 4th Edition, McGraw-Hill (2007).

Treynor, Jack L. “Toward a Theory of Market Value of Risky Assets.” Unpublished manuscript, 1961.

APPENDIX A: DATA INFORMATION

In this paper, two different databases are used, the Bloomberg database, and the CRSP database.

Bloomberg database. In order to construct the 25 Fama-French portfolios, the book value is needed for each firm. Since the book value cannot be found in the CRSP database, we use the data from Bloomberg instead. The data includes the stocks obtained from the NYSE, Amex and NASDAQ. Selecting the stocks that have a complete monthly return time series from July 1991 to June 2009, and after data cleaning, we are left with the 949 stocks used in the analysis presented in this paper.[2]

CRSP database. In the forward looking testing part of the paper, the book value is no longer needed. This allows us to use the more reliable data base – the Center for Research in Security Prices (CRSP) data base. The stocks of our sample data are still chosen as part of the NYSE, Amex and NASDAQ over the period from January 1995 to December 2008.[3]

Risk free data. Though the whole paper, the 3-month Treasury Bill rate is chosen as the risk free rate.[4] The data is released by the Board of Governors of the Federal Reserve System, which can be retrieved from http://www.federalreserve.gov/.

APPENDIX B: TECHNICAL DETAILS FOR THE COMPATIBILITY ANALYSIS

Here we will give the technical details of the compatibility test of the method mentioned in the former text. Before that, we first give some definitions for the asset that we will use in the following test. We define $\Phi(N)$ as a stock set, which contains N stocks with the highest market value (price times shares outstanding) from the observing sample. By definition, $\Phi(N+k) \supset \Phi(N)$, when $k \geq 1$. Considering $\Phi(N)$, we define $D\{N\}$ as its compatibility domain, which can be calculated by using Optimization Problem 2 given in Exhibit 4, and $\vec{\mu}_N^*$ and $\vec{\sigma}_N^*$ are the associated estimated parameter vector for the market portfolio proxy in the compatibility field of set $\Phi(N)$.

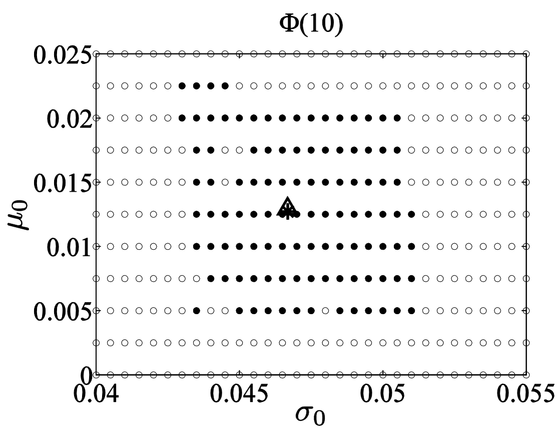
Given the set $\Phi(N+n)$, we define $D\{N+n \Rightarrow N\}$ as the compatibility domain of the set that are composed of the first N largest assets in set $\Phi(N+n)$. $D\{N+n \Rightarrow N\}$ is calculated by using $\vec{\mu}_{N+n\Rightarrow N}^*$ and $\vec{\sigma}_{N+n\Rightarrow N}^*$, where $\vec{\mu}_{N+n\Rightarrow N}^*$ and $\vec{\sigma}_{N+n\Rightarrow N}^*$ are the subspace of $\vec{\mu}_{N+n}^*$ and $\vec{\sigma}_{N+n}^*$ associated with the first N assets. For $n=0$, we define $D\{N \Rightarrow N\} = D\{N\}$.

To test the compatibility of the adjusted estimators, two different sets are built, using the CRSP database. First, we determine the intersection of $D\{N\}$ and of $D\{N+n \Rightarrow N\}$ for a fixed *n*. Second, we calculated the intersection of $D\{N\}$ with all the sets $D\{N+n \Rightarrow N\}(n \geq 1)$. This is performed for $N = 10, 20, \cdots, 100$, and n from 0 to $110 - N$ for each given N. Panel A of EXHIBIT 9 illustrates the compatibility field of the asset set $\Phi(10)$. The filled circles indicate the proxy (μ_0, σ_0) that is consistent with the mean-variance efficiency of the proxy portfolio and with the sample parameters. Panel 9 of EXHIBIT 9 gives the dependences of the domain areas, such as the one represented in panel A, corresponding to the two types of intersections as a function of *n*. The circles show the dependence of the domain area $A_N(n) = D\{N\} \bigcap D\{N+n \Rightarrow N\}$ as a function of *n*. The black triangles show the dependence of the domain area $I_N(n) = \bigcap_{k=0}^{n} D\{N+k \Rightarrow N\}$, for $N = 10$, and $0 \leq n \leq 110 - N$.

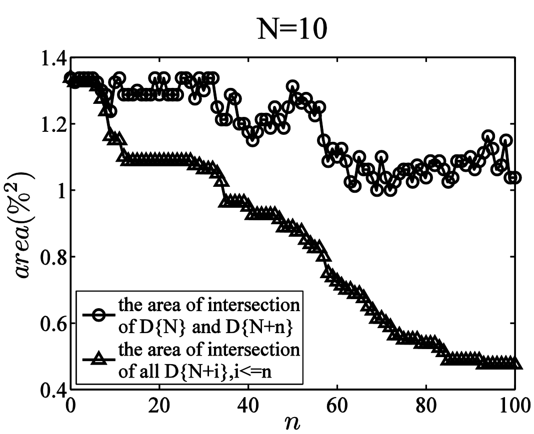
Panel B of EXHIBIT 9 shows that $A_{10}(n)$ remains remarkably stable with very little decrease for a large range of *n*. The more stringent test is of course provided by the dependence of $I_N(n)$, which without surprise decays with increasing n, but does not seem to shrink to the null set. Surprisingly, $I_N(n)$ remains quite large, even plateauing for n>>N, supporting the concept that the Levy-Roll procedure is able to determine genuine hidden properties of the underlying assets. We have carried these tests for sets $\Phi(N)$ with $N = 20, 30, \cdots, 100$ and find the same positive consistent results.

EXHIBIT 9

Compatibility Test for the Levy-Roll Procedure



Panel A: The Compatible Domain of $\Phi(10)$. The triangle corresponds to the parameters of the market portfolio proxy.



Panel B: The Intersection of the Compatible domains with different size