2013

Portfolio Optimization with Correlation Matrices: How, Why, and Why Not

Manuel Tarrazo
University of San Francisco, tarrazom@usfca.edu

Recommended Citation
Manuel Tarrazo. Portfolio Optimization with Correlation Matrices: How, Why, and Why Not. Journal of Finance and Investment Analysis, vol. 2, no.3, 2013, 17-34.

Journal of Finance and Investment Analysis, vol. 2, no.3, 2013, 17-34
ISSN: 2241-0998 (print version), 2241-0996(online)
Scienpress Ltd, 2013

Portfolio Optimization with Correlation Matrices: How, Why, and Why Not

Manuel Tarrazo[1]

Abstract

Correlation is used frequently both in the classroom and in professional environments to illustrate and summarize investment know-how, especially with regard to diversification. Pedagogically, the initial build-up on correlation, which reaches its climax while describing a hypothetical two-variable optimization case, abruptly disappears when the discussion reaches optimizations of several securities, thereby stopping short of running a full-fledged, correlation-based optimization. Why is that so? We offer some explanations. First, correlations initially seem to provide clarification of the workings of the optimization, specifically with respect to how security risk-relations affect optimal weights. However, the variable transformation required changes coordinates, thus making correlation-based optimal weights and the desired information hard to understand. Second, correlation-based optimizations may be counterproductive. Nobody with a minimum of financial sophistication would try to make up covariance estimates; correlations, however, are easy to make up, which may make one overstate their practical value. Third, while mean-variance optimal weights can be easily constructed from correlation-based optimal numbers, not transforming the optimal numbers back to the mean-variance values deforms the information processed. We do not expect correlation-based optimizations to replace mean-variance ones except in specialized cases (e.g., small portfolios where investors may have extra-knowledge of security relationships).

JEL classification numbers: C61, G11
Keywords: Portfolio optimization, mean-variance, correlations, regressions.

---

[1]School of Management, University of San Francisco. 2130 Fulton St., San Francisco, CA 94117-1045, Tel: (415) 422-2583.

Article Info: *Received :* June 17, 2012. *Revised :* July 14, 2013.
Published online : August 15, 2013

1 Introduction

Diversification is the elusive reward that portfolio theory offers to investors. To study such potential diversification one can go directly to its source: the variance-covariance matrix quantifying relationships among securities. In the case of stocks and many other investments, this matrix is hard to understand because of the units in which the magnitudes are expressed (roughly corresponding to squared returns). A more attractive prospect is to look at the matrix of correlation coefficients, which is easy to understand and, perhaps because of that, popular enough to occasionally even appear in newspapers, e.g., [1]. A correlation coefficient quantifies how the two variables move together and is usually expressed in percentage form. Presumably, optimizing a portfolio using a correlation matrix would link optimal weights to correlations and enhance our understanding of the whole process. Unfortunately, such optimizations are nowhere to be seen (textbooks, literature), a rather peculiar absence that motivated our investigation. We provide correlation-based portfolio optimizations and study their practical and pedagogical value.

The correlation coefficient indicates how one variable moves with another. It is computed to range between +1 and -1. These extremes indicate full synchronization: positive, when the variables move up/down together; negative, when one variable moves up and the other moves down. Moving away from ±1 indicates same "qualitative" movement but with less strength. The intermediate point, a correlation of zero, indicates that, based on the data being used, the movements of each variable do not seem to be co-related to the movements of the other variable. A correlation of zero is associated with "insurance" --if a life insurance company selling policies to two individuals; it is best for them not to be related at all. It is appropriate to note that the concept "relationship" is wider and more complex than that of co-relation, which is restricted to a paired, observation-by-observation association of numerical variables. The study of co-relationships using leads and lags is one of the more complex areas of econometric analysis.

The co-movement described in the previous paragraph applies to investments as well. Clearly, if one investment goes down, it would be good if the other investment doesn't follow. But if one investment goes down, a negative correlation would offer support to the idea that at least one of the assets will go up –this is the core of the "diversification" concept. A third major risk management principle, "hedging", is also best explained with the help of correlation. Establishing positions on perfectly negatively correlated assets –the perfect hedge, as in "hedging your bets"-- would offer the highest likelihood to having some up position at the end of the trading horizon. Derivative securities were custom-made for hedging: instead of buying and selling (or selling short) the same asset to get the perfectly negative correlation, one would take a given position in a given asset and the contrary position on its derivative (forward, futures, or options). Further, to cap it all, the correlation coefficient can easily be expressed as a percentage, which facilitates taking advantage of the information it conveys.

Again, why do we not run portfolio optimizations with correlation matrices? In order to answer, we must first examine the effects of certain transformations of variables. Next, we must evaluate the benefits and the limitations of using correlation matrices in portfolio optimization. What we discover in this analysis may not appear favorable to correlation-based optimizations. However, this does not imply correlation analysis has no usefulness. For example, correlation-based analysis may be helpful when considering portfolios of a few securities (limited diversification) where investors may have some extra-knowledge of variable relationships.

2 Variable Transformations, Regression Analysis, and Portfolio Optimization

Optimal portfolio weights can be obtained by maximizing the following unrestricted function:

$$F(x) = - \tfrac{1}{2} x' A x + x' b \tag{1}$$

This is a quadratic equation composed of a quadratic form (x' A x) and a vector (x' b), where A and b represent the covariance matrix and the vector of average individual stock returns, respectively. The first order conditions provide a set of simultaneous equations, A x = b. Optimal weights are calculated by normalizing the x auxiliary variables; that is, $w_i^* = x_i / \Sigma x_i^*$. These are the expressions for the portfolio return and its variance, respectively: $r_p = w' b$, and $var_p = w' A w$.

The optimal portfolio thus calculated is the so-called "tangent" portfolio, which is the one that maximizes the return-to-risk ratio. This is also the portfolio that includes the usual (linear) arbitrage relationships ($\Sigma w_i = 1 = w_p$, $\Sigma w_i r_i = w_p r_p = r_p$), which implies that the portfolio cannot have more value than any of its parts. The algebraic formulation of portfolio optimization in (1) above keeps the analysis in the well-known area of simultaneous equations systems (SES), which is also shared by regression analysis.

We could think of applying linear algebra techniques to the first order conditions of the mean-variance model to study what types of equivalent transformations would change the usual mean-variance optimization into a standardized mean-correlation specification (C x = d; where C and d would now represent correlation matrix, and the vector of standardized means, respectively). As per common usage in linear algebra, equivalent transformations are those that do not alter the set of optimal solutions in a system of simultaneous equations. We started to pursue this course of action and, right away, we noticed that the potential changes would not only change the coordinate system of reference, but would alter the right-hand side by changing the units of the average returns as well. Unfortunately, equivalent transformations seem to cloud the optimization with seemingly arbitrary changes, and they cannot produce a straightforward way to re-state mean-variance results.

A more advantageous tactic is to exploit the relationship between regression analysis and portfolio optimization to study the effects of certain variable transformations which, as it happens, have been well-known to statisticians since the dawn of econometric analysis. The regression approach to portfolio optimization was first developed by Jobson and Korkie [2], and further studied by Britten-Jones [3] and Tarrazo [4].

Tables 1 and 2 study the effects of some well-known transformations in the context of ordinary, multivariate regression (y = X b + u). Some results could be obtained using probability distributions and mathematical statistics but it is more advantageous to keep to straightforward linear algebra. Table 1 shows the matrix approach to ordinary least squares for both the original and the mean-adjusted variables. Note that mean-adjusting the regressors, but not the regressand, would produce the same slope estimates but higher fitting errors, which means we need to adjust means of all variables, or none. When we do so, the two regressions in Table 1 are exactly equivalent (same R-squared and associated goodness of fit indicators). Note the role of the interplay between the intercept, calculated with the usual vector of ones, and alternative mean specifications. We are using "population" formulas for clarity.

Table 2 is more interesting for our purposes. The top shows the multivariate regression for both mean-adjusted and standardized variables, which is often used to obtain variables thought to be normally-distributed, and with expected value and variance of 0 and 1, respectively. This regression is helpful because it shows very clearly how the transformed regression coefficients (bx_i*) are related to the original coefficients (bx_i). For reasons that will become clear later on, we would like to retain the mean vector in the optimization; therefore, the bottom regression is the one of highest interest at this point in the analysis.

3 Portfolio Optimizations Using Correlations: How

Tables 3 and 4 carry the analysis over to the portfolio optimization arena. The top of Table 3 shows the atypical regression that yields optimal portfolio weights. The data matrix X, which includes vectors x1, x2, and x3, represents security returns. We regress these returns on the y variable, which is simply a vector of ones (actually any constant would do) and calculate OLS estimates (bx). Then, we compute optimal portfolio weights ($w^* = [w_1^*\ w_2^*\ w_3^*]'$) by normalizing betas: $w_i^* = bx_i / \Sigma\, bx_i^*$. The conventional mean-variance optimization appears at the bottom of Table 3 as well. It boils down to solving a simultaneous equation system (Ax = b) which, through a normalization, produces the optimal portfolio weights.

Table 1: Initial data and deviations from the mean

Initial data

y	X intercept	x1	x2	
3	1	3	5	
1	1	1	4	
8	1	5	6	
3	1	2	4	
5	1	4	6	
4		3	5	mean
5.6		2	0.8	varp
2.366432		1.414214	0.894427	stdp

X'X				X'y
	5	15	25	20
	15	55	81	76
	25	81	129	109

intercept	4
bx1	2.5
bx2	-1.5
	5

Deviations from mean

y-meany	X	x1-mean1	x2-mean2	
-1		0	0	
-3		-2	-1	
4		2	1	
-1		-1	-1	
1		1	1	
0		0	0	mean
5.6		2	0.8	varp
2.366432		1.414214	0.894427	stdp

X'X			X'y
	10	6	16
	6	4	9

intercept	0
bx1-mean1	2.5
bx2-mean2	-1.5
	1

Table 2: Mean-adjusting and standardizing

Mean-adjusted, standardized data = (var-mean)/std

y*		X x1*	x2*	
-0.42258		0	0	
-1.26773		-1.41421	-1.11803	
1.690309		1.414214	1.118034	
-0.42258		-0.70711	-1.11803	
0.422577		0.707107	1.118034	
0		0	0	mean
1		1	1	varp
1		1	1	stdp

X'X			X'y
	5	4.743416	4.780914
	4.743416	5	4.2521

(intercept	0)			
bx1*	1.494036	=	2.5	= bx1* σy* / σx1*
bx2*	-0.56695	=	-1.5	= bx2* σy* / σx1*
	0.927089		1	

Divided by their std, but retaining means

ys	X x0s	x1s	x2s	
1.267731	1	2.12132	5.59017	
0.422577	1	0.707107	4.472136	
3.380617	1	3.535534	6.708204	
1.267731	1	1.414214	4.472136	
2.112886	1	2.828427	6.708204	
1.690309	1	2.12132	5.59017	mean
1	0	1	1	varp
1	0	1	1	stdp

X'X			X'y
5	10.6066	27.95085	8.451543
10.6066	27.5	64.03612	22.70934
27.95085	64.03612	161.25	51.49766

bx0s	1.690309	4	= bx0s *σy = ysmean * σy
bx1s	1.494036	2.5	= bx1s* σy / σx1
bx2s	-0.56695	-1.5	= bx2s* σy / σx2
	2.617398	5	

Let us observe carefully the atypical regression. Note that a) the values of the "y" in the regressions of Table 1 now appear as those of another "exogenous variable"; b) the usual intercept of ones has been moved to the left-hand-side; and c) the regression is run without an intercept. From a financial point to view, whether the vector [3 1 8 3 5] is a security or a portfolio does not matter. What matters is that the optimization will insure that each variable is properly valued and arbitraged in reference to the other securities. In passing, this table also makes obvious that the "tangent" portfolio, in addition to implementing arbitrage conditions, is also an optimal predictor. This means it performs best within the class of linear estimators, under some conditions, and also the best in a larger class of estimators that do not include linearity restrictions. D1, D2, and D3 refer to the determinants of order (k) in the matrix A. Their values (positive for all ranks) indicate that A is positive definite, as any variance-covariance matrix must be.

Table 4 presents the results we are after. A simple standardization, dividing every observation by its standard deviation provides the sought-after standardized mean-correlation system. The correlation matrix can be calculated using matrix functions in EXCEL with the variance covariance-formula when the variables have been standardized: target cells = mmult(transpose(data range – mean vector), data – mean vector).

Note that the values in the right-hand-side vector, [1.690309, 2.12132, 5.59017]', are the values of the means of the original variables divided by their standard deviation: that is, 1.690309 = 4/2.366432. This means the correlation optimization has the correlation table in the left-hand-side and r_i/std_i in the right-hand-side. Noticing this is critical to establish the analytical equivalency of mean-variance and correlation-based optimizations and is something that remains implicit in the numerical examples.

The investor performing the optimizations should have two objectives. The first one is to calculate the optimization numbers, which are found as the solution to the simultaneous equation system A x = b, where A and b are now the correlation matrix and the mean for the standardized data, respectively. The solution vector is bxs = [31.55243 -78.9603 53.66563]', which must be normalized to function as optimal portfolio weights (w_is^*) for the standardized variables. The second objective is to trace these weights back to the original variables, which requires dividing the solution coefficients by the standard deviation of the corresponding original variable (e.g., 13.33333333 = bx_is / $stdx_i$ = 31.55242551 / 2.366431913); this operation returns the non-normalized solutions to the mean-variance optimization (bottom of Table 3, bx and w_i^*). The values we found last must be normalized to provide the original portfolio weights –that is, w_1^* = ($bx_1s/stdx_1$) / sum($bx_is/stdx_i$) = 13.33333333 / 17.5 = 0.761904762). See appendix for further detail and analytical proof.

Table 3: From regressions to optimal mean-variance portfolios

"Portfolio" data: Obtaining optimal portfolio weights with regressions

X

ones	x1	x2	x3	
1	3	3	5	
1	1	1	4	
1	8	5	6	
1	3	2	4	
1	5	4	6	
1	4	3	5	mean
0	5.6	2	0.8	varp
0	2.366432	1.414214	0.894427	stdp

X'X				X'y
108	76	109		20
76	55	81		15
109	81	129		25

		wi*
bx1	0.071365	0.761905
bx2	-0.29884	-3.19048
bx3	0.321142	3.428571
	0.093666	1

Conventional mean-variance portfolio optimization

X

x1	x2	x3	
3	3	5	
1	1	4	
8	5	6	
3	2	4	
5	4	6	
4	3	5	mean
5.6	2	0.8	varp
2.366432	1.414214	0.894427	stdp

variance-covariance matrix				means
5.6	3.2	1.8		4
3.2	2	1.2		3
1.8	1.2	0.8		5

		wi*		
bx1	13.33333	0.761905	D1 =	5.6
bx2	-55.8333	-3.19048	D2 =	0.96
bx3	60	3.428571	D3 =	0.048
Sum bxi	17.5	1		

Table 4: Standardized variables and mean-correlation optimization

Portfolios with regressions: standardized variables

ones	X x1s	x2s	x3s	
1	1.267731	2.12132034	5.59017	
1	0.422577	0.70710678	4.472136	
1	3.380617	3.53553391	6.708204	
1	1.267731	1.41421356	4.472136	
1	2.112886	2.82842712	6.708204	
1	1.690309	2.12132034	5.59017	mean
0	1	1	1	varp
0	1	1	1	stdp

X'X			X'y
19.28571	22.70934	51.4976594	8.451543
22.70934	27.5	64.0361226	10.6066
51.49766	64.03612	161.25	27.95085

		wis*		wi*	
bx1s	0.16888	0.071365	= bx1s/stdx1	0.761905	= wi*s/sumwi*s
bx2s	-0.42262	-0.29884	= bx1s/stdx2	-3.19048	
bx3s	0.287238	0.321142	= bx1s/stdx3	3.428571	
	0.033494	0.093666		1	

Correlation-table portfolio optimization

X x1s	x2s	x3s	
1.267731	2.12132034	5.59017	
0.422577	0.70710678	4.472136	
3.380617	3.53553391	6.708204	
1.267731	1.41421356	4.472136	
2.112886	2.82842712	6.708204	
1.690309	2.12132034	5.59017	mean
1	1	1	varp
1	1	1	stdp

correlation matrix			means
1	0.956183	0.85042006	1.690309
0.956183	1	0.9486833	2.12132
0.85042	0.948683	1	5.59017

		wis*	bxis/stdxi	wi*	
bx1s	31.55243	5.042096	13.3333333	0.761905	(1)
bx2s	-78.9603	-12.6179	-55.833333	-3.19048	
bx3s	53.66563	8.575799	60	3.428571	
	6.2578	1	17.5	1	

(1) = (bxis/stdxi)/sum(bxis/stdxi)

4 Portfolio Optimizations Using Correlations: Why And Why Not

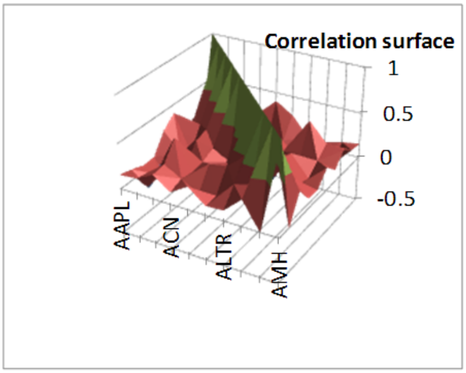
In the introduction, we indicated that the concept of correlation is very helpful summarize a great deal of investing know-how related to diversification. In the previous section, we have shown that using correlation matrices to optimize a portfolio is rather straightforward. Why not, then, use the standardized mean-correlation analysis? There are at least five reasons that seem to push in that direction. However, upon close examination, they actually strengthen the case for the traditional or conventional mean-variance optimizations.

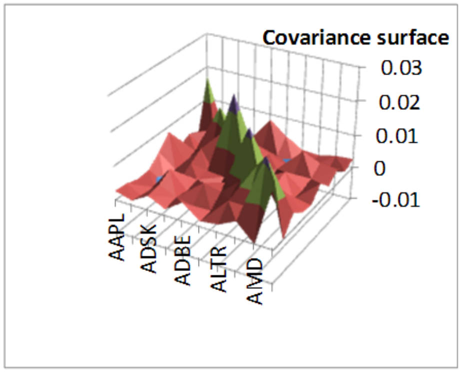
First. Clarifying the link between optimal portfolio weights and correlations was one of our hopes at the outset of our analysis. Unfortunately, the transformation required changing the mean reference indicator (vector b of the original means), and effecting a change of coordinates by altering the A matrix. Within the realm of the optimization, using correlations does not add any clarity, it is rather the opposite.

Second. The beta coefficients with the original variables measure how much the dependent variable will change when the independent variable changes, keeping everything else constant. The betas for the standardized variables measure how the ratio (y-mean/σy) changes when (x-mean/σx) changes. The ratio mean-to-risk (μ/σ) is extremely important in portfolio optimization. It effectively ranks securities to the point of anticipating the ranking of optimal portfolio weights, in some cases with the same ordinal rankings. It is easy to show that when the covariances (or correlations) are equal to zero, the individual return-to-risk ratios determine optimal portfolio weights. At the portfolio level, it can work as the objective variable of the optimization, as shown by the early work of Elton and Gruber and Padberg on simple criteria for portfolio optimization, [5]. Furthermore, Tarrazo [6] shows how to optimize a portfolio by ranking the securities by their $\mu i/\sigma i$ ratios first and adding one security at a time --if the weight is positive, keep the security; if negative, discard the security. The optimization is terminated, and the optimal portfolio achieved, when there are no more securities with positive average returns ($\mu i > 0$) remaining. The ranking provided by individual return-to-risk ratios would anticipate optimal portfolio weights perfectly if not for covariance effects, which in some cases are strong enough to alter the individual return-to-risk ranking of the securities. This lessens the importance of having the return-to-risk proportions as the right-hand-side variable in the simultaneous equations system of the optimization. In addition, investors may find it easier to keep track of the relationship between optimal weights and individual mean returns.

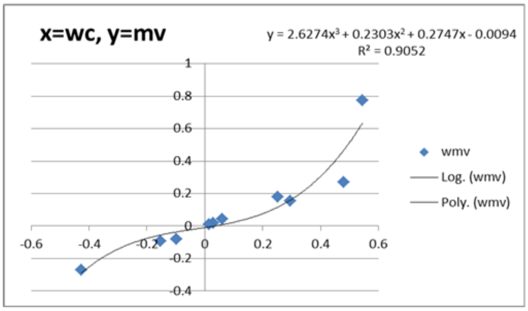
Third. We had expected that correlation-based optimization should help visually, at least, given that that the correlation coefficients can easily be read as percentages. It turns out that visualizations with correlation matrices, in general, may convey less about the risk structure than mean-variance ones do.

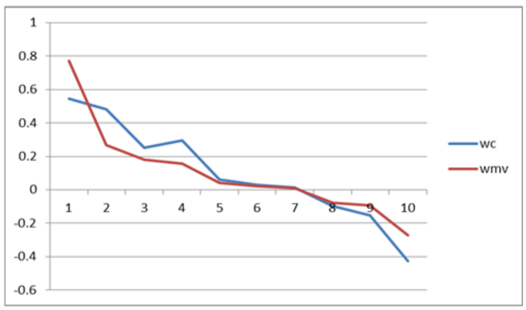
Table 5: An example with stock market data





	wc	wmv
AAPL	-0.42593	-0.27247
ABM	0.545004	0.771621
ADSK	0.028833	0.020297
ACN	-0.09867	-0.079
ADBE	0.294749	0.15625
ADPT	0.251965	0.178349
ALTR	0.480862	0.268352
AMAT	0.061139	0.041127
AMD	0.014719	0.008286
AMH	-0.15267	-0.09281
rp	0.05576	0.040849
σp	0.094868	0.063014
rp/σp	0.587759	0.648258
yearly rp	66.91%	49.02%
yearly σp	32.86%	21.83%





The top of Table 5 shows two surfaces that resemble those used in orography when studying the formation and relief of mountains and other geographical objects. In the case of portfolio optimization these surfaces show the volatility "map", where a mountain with a higher represents a riskier security. The risk surface is dominated by the main diagonal of the risk matrix, which divides the terrain diagonally as if it were a powerful mountain range. The correlation matrix makes all these diagonal peaks equal to one, which makes us lose sight of which securities are riskier. It also does not enhance the perception of off-diagonal correspondences. (The charts shown are non-auto scaled. Auto-scaling them makes the off-diagonal areas nearly undistinguishable.)

Fourth. The equivalence of optimization for the tangent portfolio is straightforward, but calculating optimal portfolios under other restrictions (e.g., required return) may appear to be complicated to students, as it cannot be done with the usual textbook procedure (i.e., Lagrangian method), with the same straightforwardness. Adding additional variables (e.g., a risk-free rate) in the correlation-based framework is also problematic for efficient, not tangent, portfolios.

Fifth. The most important reason not to use standardized mean-correlation optimization without transforming the weights to their mean-variance values is a very serious one. At first, it seems it should be easy to pass from the mean-variance system to the standardized mean-correlation. But that is not the case –dividing the right hand side of the mean variance system does yield standardized means but it does not produce the correlation matrix. The required changes cause non-linear changes that alter the data in hard to assess ways. This is why using equivalent transformations, which are linear, was doomed to fail. Further, the transformation produces a tangent portfolio that is not mean-variance efficient (unless, of course, weights are transformed back). The lower part in Table 5 shows how the correlation-based weights (wc) compare to the mean-variance (wmv) weights. We are using an old data set that has good pedagogical characteristics and with which we are very familiar. The correlation portfolio is not efficient, not because its return-to-risk ratio is lower than the mean-variance ratio, but because a mean-variance portfolio offering the returns of the wc portfolio has a lower variance. The nonlinear deformations of the weights are observable by looking at the charts relating both sets of weights; one of them is an "x-y" chart, and the other a regular "line" chart. The profile for wc weights is flatter than that for the wmv weights, no doubt due to the standardization of the right-hand side variables.

The last point deserves further explanation. Standardization is sometimes used to better understand the numbers we have. Some variables (e.g., income, age, gender, civil status, years of schooling, room temperature) lend themselves better to measurement than others (feeling cold, education, satisfaction with products or services, capability of solving certain complex problems, and so on). In some cases, standardizing helps to analyze the variables and assess behavior over the whole observed ranges. Such standardization can be implemented by 1) using percentile ranking, 2) z-scores, or 3) partitioning the data as in the Goldfeld-Quandt procedure, to detect heteroscedasticity. Standardization may also help when processing variables with different units of measurements that, nonetheless, need to all be included in a common model (e.g., panel data). Z-scaling is a linear transformation, and therefore the effects of relative differences are respected, however, but it does not produce normality, unless the actual data is normally distributed. Percentile-scaling rearranges the variables into neater subdivisions based on the number of percentiles but something is lost in the process. Turning cardinal variables into ordinal ones disregards the effects of relative distances (i.e., a non-linear transformation). Standardizations generate a modified sample that changes if we take out observations, or add observations to the sample. Note that in the case of portfolio theory, standardizations directly impact the risk-structure of the problem and, therefore, may actually ruin the optimization.

Our findings do not recommend using correlation-based portfolio optimizations unless the optimal weights are transformed back to mean-variance ones. However, five positive observations can be made:

a) Correlation-based optimizations respect the ranking that security return-to-risk ratios impose on optimal weights, a fact which may have further theoretical and practical value.
b) Paradoxically, when we try to make the best of correlations, we find strong reasons to restrict their use, especially in textbooks, which often depict "dream cases" of extreme correlations (perfectly negative, positive, or zero) among securities, which are all unreal. Using the actual values of correlations in the actual optimizations would be one valuable step in the right direction because it may induce more practical varieties of classroom work. For example, low and negative correlations are not so abundant.

Theoretically, Samuelson (1968) shows that there are limits to how negative correlation can be under general statistical conditions: "Although there is no limit on the degree to which all investments can be positively intercorrelated, it is impossible for all to be strongly negatively correlated. If A and B are both strongly negatively correlated with C, how can A and B fail to be positively intercorrelated with each other? For three variables, the maximum common negative correlation coefficient is - 1/2; for four variables, - 1/3; for n variables, (–l / (n - l)),"[5]. In practice, Elton and Gruber found that an average correlation coefficient of +1/2 (50%, positive) was representative of usual market conditions, [6]. Piling up companies in the hope of gaining diversification is a risky activity, which Peter Lynch often referred to as "diworsification".

c) It is worth noting that interest in implied correlations among particular securities and other actual investment positions has increased recently, especially in derivative analysis.
d) There may be situations where investors may hold portfolios of a few securities. In this case, correlation coefficients may easily capture any extra knowledge concerning the relationships between the variables of interest.
e) Everything considered, it seems correlation concepts will continue being employed in research of both a theoretical and applied nature; see [8] and [9], respectively.

At a more general level, our study stresses the connections between correlation, arbitrage, and mean-variance analysis, and builds upon the relationship between regressions and portfolio optimizations, which highlights the conditional properties of portfolios. Ultimately, investors will appreciate portfolio theory because of its usefulness, which is based on the strength of the signals it provides and on its reliability, which may preclude employing certain transformations.

It is appropriate to close by providing further information and references to the material employed throughout the study.

The data used in the first two tables is from Johnston [10, p.178 and ss.]. Ezequiel's (1959) text, [11], whose first edition was written at the birth of modern econometrics (1930), shows a great deal of care in properly applying statistical techniques to data and interpreting the numbers correctly. Part of his effort focused on studying the logical analysis of relationships between variables as warranted by their statistical association. The author placed special emphasis on three measures –the standard error of the estimate, the coefficient of determination, and the correlation coefficient; and also on cases when variables are expressed in alternative measurements, and the purported relationship may not be necessarily linear. Because the size of the usual regression coefficients varies with the units in which each variable is expressed, Ezequiel investigated whether expressing each variable in terms of its own standard deviation would make the regression coefficients more comparable. He referred to these coefficients as "beta coefficients". Ezequiel [op. cit, p. 196] provided the formulae for comparing the original and the "beta" regressions, which we will reproduce in a more conventional notation:

$$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + u \tag{2}$$

$$y/\sigma y = \alpha + \beta_1 x_1/\sigma_1 + \beta_2 x_2/\sigma_2 + \beta_3 x_3/\sigma_3 + v \tag{3}$$

$$y = \alpha\, \sigma y + (\beta_1\, \sigma y/\sigma_1)\, x_1 + (\beta_2\, \sigma y/\sigma_2)\, x_2 + (\beta 3\, \sigma y/\sigma_3)\, x_3 + u \tag{4}$$

Equation (4) shows the relationship between the coefficients of the ordinary regressions, equation (2), and those of the standardized ones, equation (3): a = α σy, $b_1 = \beta_1$ σy ⁄ σ_1, $b_2 = \beta_2$ σy ⁄ σ_2, and b_3 = σy /σ_3 . In the ordinary or two-variable regression beta-hat = β σy ⁄ σx = ρ_{xy} σy ⁄ σx, which means $\beta = \rho_{xy}$. It is intriguing that the coefficient of the regression between a single security, or a single portfolio, stock returns and the returns on the market index is called "the beta coefficient" in the Capital Asset Pricing Model, but it is calculated with the same process as an ordinary regression. Ezequiel's betas and their potential uses have not been ignored by econometricians in recent times. Maddala summarizes the reasons for the apparent neglect: "In simple regression the beta coefficient is identical to the correlation coefficient... However, in multiple regression there is no relationship between the beta coefficients and the simple- or partial- correlation coefficients. Hence, often not much use is made of the beta coefficients," [12, p. 119]. Kmenta also notes, "Since the beta coefficients do not solve the main problem of separating the effects of each explanatory variable on the dependent variable any better than the usual regression coefficients, their use in econometrics has been rather rare," [13, pp. 422-3].

Sewal Wright studied transformations that would make regression coefficients and correlations identical. Like Ezequiel's betas, these correlations would also measure the fraction on standard deviation caused by some variables in others. In addition, they would be most easily understood and could represent causation pathways. Wright's efforts resulted in what is currently known as "path analysis," see [14, 15], and Duncan [16].

The regression approach to portfolio optimization was first developed by Jobson and Korkie [2, op. cit.], and further studied by Britten-Jones [3, op. cit.] and Tarrazo [4, op. cit.]. Tarrazo [6, op. cit.] employs individual return-to risk heuristic to fully optimize a portfolio, which is accomplished by sequentially forming a portfolio of the two highest return-to-risk securities and by successively trying securities with positive returns ranked according to their return-to-risk ratio. The role played in the previous author's analysis by graphical-oriented, information-organizing objects such as paths, graphs, and networks is interesting to note. The notion of a "balanced" graph –a collection of equally-signed (all positive, or all negative) connected nodes— was instrumental in understanding the link between positive returns and positive portfolio weights in Tarrazo [6, op. cit] and, therefore, finding the point where no more securities would be added to the optimal set. Tarrazo [17] focuses on calculating optimal portfolio weights without mathematical programming or Lagrangians. Graph theory is a powerful tool in discrete mathematics, with applications in many other areas such as combinatorial optimization and networks. Frank Harary [18] contributed heavily to graph theory and its related areas and developed the concept of signed graphs out of a particular problem in sociology. Ellerman [19, 20] employs graph theory in his specialized analysis of arbitrage, which appears as the force or medium (path) relating one security to another.

5 Concluding Comments

The apparent attractiveness of correlation concepts in portfolio optimization initially motivated this research. By studying the effects of variable transformations in regressions we quickly ascertained how to perform portfolio optimizations using mean-correlation, instead of mean-variance analysis. Of course, the two alternative set-ups produce equivalent optimal portfolio weights if correlation-based number are transformed back to mean-variance ones. Without such transformation, correlation-based optimizations present some a priori advantages –clarification of risk-relations among securities and direct use of return-to-risk measures which link individual return-to-risk and portfolio ratios through the optimal weights. The advantages, however, are outweighed by the inherent limitations –the optimizations effect a transformation that makes interpreting the resulting weights difficult and may deform the risk-structure of the data in a nonlinear and hard-to-assess manner. On a more positive note, the analysis presented strengthens the role of regression methods in portfolio analysis. Further, the difficulty in finding transformations of the data that would both clarify the relationship between individual security characteristics and portfolio weights in a practical manner. In sum, correlation concepts clearly have some pedagogical value, and correlations do provide easy-to-understand information on potential diversification (or lack thereof) that one can obtain before running the optimization.

References

[1] J. Marte, Why the Math of Correlation Matters, *The Wall Street Journal*, Monday, October 2010, R7.

[2] J. Jobson, and R. Korkie, Statistical Inference in a Two-Parameter Portfolio Theory with Multiple Regression Software, *Journal of Financial and Quantitative Analysis*, **18**(2), (1983), 189-197.

[3] M. Britten-Jones, The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights, *The Journal of Finance*, **44**(2), (1999), 655-671.

[4] M. Tarrazo, On Portfolios and Regressions, *The Journal of Financial Education*, **35**, (2009), 56-74.

[5] E. Elton, M. Gruber, S. Brown, and W. Goetzman, Modern Portfolio Theory and Investment Analysis, 8th. ed., Wiley, New York, 2010.

[6] M. Tarrazo, Identifying Securities to Buy: The Heuristic ri/stdi, *Research in Finance*, **25**, (2009), 229-268.

[7] P. Samuelson, A General Poof that Diversification Pays, *Journal of Financial and Quantitative Analysis*, **2**(1), (1967), 1-13.

[8] H. Levy, The Demand for Assets Under Conditions of Risk, *The Journal of Finance*, **28**(1), (1973), 79-96.

[9] A. Rubin, and J. Morgan, A Portfolio Diversification Index, *The Journal of Portfolio Management*, **32**(2), (2006), 81-89.

[10] J. Johnston, Econometric Methods, 3rd ed. McGraw-Hill, 1984.

[11] M. Ezequiel, Methods of Correlation and Regression Analysis, John Wiley & Sons, Inc., New York, 1959.

[12] G. Maddala, Econometrics, McGraw-Hill, 1977.

[13] J. Kmenta, Elements of Econometrics, 2nd ed., The University of Michigan Press, Ann Arbor, MI, 2000.

[14] S. Wright, Path Coefficients and Path Regressions: Alternative or Complementary Concepts? *Biometrics*, **16**(2), (1960), 189-202.

[15] S. Wright, The Method of Path Coefficients, *The Annals of Mathematical Statistics*, **5**(3), (1934), 161-215.

[16] O. Duncan, Path Analysis: Sociological Examples, *The American Journal of Sociology*, **72**(1), (1966), 1-16.

[17] M. Tarrazo, Portfolio Optimization without Programming or Lagrangians –Implications for Investing, Proceedings of the Decision Sciences Institute Conference, San Francisco, (2012), 6220-62233.

[18] F. Harary, On the Notion of Balance of a Signed Graph, *Michigan Mathematical Journal*, **2**(2), (1953), 143-46.

[19] D. Ellerman, An Arbitrage Interpretation of Classical Optimization, *Metronomica*, **41**(3), (1990), 259-276.

[20] D. Ellerman, Arbitrage Theory: A Mathematical Introduction, *SIAM Review*, **26**(2), (1984), 241-261.

Appendix

The first order conditions from the equation (1), for the mean-variance and the correlation-based portfolio optimizations models mode can be formulated as linear simultaneous equation systems (A x = b) which, for a two-variable case, are the following:

Mean-variance

$$\begin{vmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = \begin{vmatrix} r_1 \\ r_2 \end{vmatrix} \qquad \text{A(1)}$$

Correlation-based model

$$\begin{vmatrix} \rho_{11} & \rho_{12} \\ \rho_{12} & \rho_{22} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = \begin{vmatrix} rs_1 \\ rs_2 \end{vmatrix} \qquad \text{A(2)}$$

The equivalence between the two models is more easily seen when we notice that:

a) The correlation coefficient between a variable and itself is equal to one.
b) The average of the standardized variable is the mean of the original variable divided by its standard deviation.
c) The formula for the correlation coefficient is the following:

$\rho_{12} = \sigma_{12} /(\sigma_1\sigma_2)$.

Then we can express the correlation model in a manner close to the mean-variance specification:

$$\underset{R}{\begin{matrix} 1 & \sigma_{12} /(\sigma_1\sigma_2) \\ \sigma_{12} /(\sigma_1\sigma_2) & 1 \end{matrix}} \quad \underset{x}{\begin{vmatrix} x_1 \\ x_2 \end{vmatrix}} \underset{=}{=} \underset{c}{\begin{vmatrix} r_1/ \sigma_1 \\ r_2/ \sigma_2 \end{vmatrix}} \qquad \text{A(3)}$$

Cramer's rule conveniently separates denominator effects, which are directly associated with coordinates, the matrix A, from coordinate-independent effects.

Applying Cramer's rule to solve A(3), we find that

$$x_1 = \frac{\begin{vmatrix} r_1/ \sigma_1 & \sigma_{12} /(\sigma_1\sigma_2) \\ r_2/ \sigma_2 & 1 \end{vmatrix}}{|R|} = \frac{r_1/ \sigma_1 - (r_2/ \sigma_2 (\sigma_{12} /(\sigma_1\ \sigma_1)))}{|R|} \qquad \text{(A4)}$$

$$x_2 = \frac{\begin{vmatrix} 1 & r_1/ \sigma_1 \\ \sigma_{12} /(\sigma_1\sigma_2) & r_2/ \sigma_2 \end{vmatrix}}{|R|} = \frac{r_2/ \sigma_2 - (r_1/ \sigma_1 (\sigma_{12} /(\sigma_1\sigma_2)))}{|R|} \qquad \text{(A5)}$$

This can be expressed as:

$$x_1 = \frac{(r_1\ \sigma_{22} - r_2\ \sigma_{12}) / (\sigma_{22}\ \sigma_1)}{|R|} \qquad \text{(A6)}$$

$$x_2 = \frac{(r_2\ \sigma_{11} - r_1\sigma_{12}) / (\sigma_{11}\ \sigma_2)}{|R|} \qquad \text{(A7)}$$

Here, we can already visualize that dividing each x_i by its corresponding σ_i would yield a common denominator that would simplify away, as that corresponding to the |R| terms, giving the mean-variance optimal (non-normalized) weights: $xmv_1 = r_1\ \sigma_{22} - r_2\ \sigma_{12}$, $xmv_2 = r_2\ \sigma_{11} - r_1\sigma_{12}$. The normalized mean-variance optimal weights are $w_i = xmv_i/\mathrm{sum}(xmv_i)$.
In effect, normalizing the x_i's above by dividing each of them by their sum ($x_1 + x_2$), and calculating the ratio x_1/x_2, confirms why, in order to pass from the correlation-based weights to mean-variance ones we must divide each of the normalized correlation-based weights (wc_1, wc_2) by their corresponding standard deviation and normalize again. (It is helpful to track the value 13.3333 in Tables 3 and 4.)

$$\frac{wc_1}{wc_2} = \frac{(r_1\ \sigma_{22} - r_2\ \sigma_{22})}{\sigma_{22}\ \sigma_1} : \frac{(r_2\ \sigma_{11} - r_1\sigma_{12})}{\sigma_{11}\ \sigma_2} = \frac{\sigma_1}{\sigma_2}\ \frac{w_1}{w_2} \qquad \text{(A8)}$$

The equivalence in terms of relative weights shows that both systems A(1) and A(2) share the same homogeneous coordinates. One is obtained from the other as if modeling clay, without tears or discontinuities, and what one learns in one system can be applied to the other.