Spreadsheets in Education (eJSiE)

Volume 10 | Issue 3 Article 5

5-16-2018

An active learning exercise showing some fundamentals of financial portfolio construction

Graham D. Barr
University of Cape Town, gdi@iafrica.com

Leanne Scott
leanne.scott@uct.ac.za

An active learning exercise showing some fundamentals of financial portfolio construction

Abstract
This paper focuses on a key statistical insight, namely, that the statistical variability of a sum of variables may be substantially reduced by combining individual components that have low correlation. We demonstrate this idea on a spreadsheet through the use of simulated data with associated statistics and scatter plots but note that, in addition, these ideas have particular relevance in financial applications. We apply these variability reducing principles in the field of finance to demonstrate, on a spreadsheet, the circumstances under which we may get significant risk reduction in a portfolio of 2 shares, compared to holding only one share. However, rather than simply coming up with a single "true" efficient frontier based on assumed "true" parameters, we suggest using repeated sampling with realistic underlying parameters to simulate sets of share returns and risk and then construct sets of simulated efficient frontiers. These sets of efficient frontiers reflect the underlying uncertainty of future share returns and the variability of the returns. We can then embed the idea in students that far from being an exact optimization problem, portfolio construction requires circumspection and a subtle appreciation of statistical variability. This Excel-based didactic approach has been used to introduce students to the principles of variance reduction and the construction of efficient frontiers in the portfolio paradigm, as a component of Honours level courses in the department of Statistical Sciences at the University of Cape Town. Students in these courses consistently found this spreadsheet-centred approach to be a very useful active learning tool for understanding these principles.

Keywords
Statistical Variability, Return, Risk, Portfolio, Efficient Frontier

1 Introduction

There are many didactical exercises which employ spreadsheets to good effect to handle the computational side of statistical analysis and to demonstrate results graphically; see, for example, [10] and the history of the use of spreadsheets [1]. However there are not many published exercises which showcase the extraordinary ability of spreadsheets to demonstrate the effects of the elusive statistical concept of random variation. Spreadsheets, through the medium of simulation, provide an environment for students to develop an appreciation of the effects of random variation that is simply not accessible through theory or through analysis of 'static' data sets, see, for example, [2] and [3]. This paper uses the simulation approach to first demonstrate the variability reduction when random variables are combined and then applies this simulation approach to constructing a portfolio of shares to demonstrate risk (proxied as standard deviation) reduction. Critically, in providing a 'real world' example of statistical theory in action, it showcases how important it is to look at portfolio construction in a simulated empirical paradigm rather than making the unrealistic assumption that financial parameters are known with certainty. It should be emphasized that, while students may well be familiar with the theory of variance/standard-deviation reduction (either through their Statistics studies or, as Economics students, through exposure to Markowitz's approach to portfolio risk reduction through diversification), simulation is used here as a didactic tool to effect a deeper and more practical understanding of these theoretical concepts.

This paper and active learning exercise is aimed particularly at students who are beginning a course in Portfolio theory and Econometrics within an honours degree in Applied Statistics in the Science Faculty at the University of Cape Town (UCT). The exercise is intended to show statistical theory in action, and critically, to indicate that there is a practical point to a piece of theory which we can use to advantage in a real life exercise in share portfolio risk reduction. A spreadsheet approach has been used to demonstrate financial concepts in this journal before. In two recent papers [6], [7] Kwan has considered the effect of shrinkage estimators for the correlation matrix of share returns and demonstrated how this approach can be used to strike a balance between the reduction of forecast errors and the retention of existing idiosyncratic correlation features. This paper considers the problem from a somewhat different perspective, focusing on the distributional characteristics of the problem by simulating different return realisations using a spreadsheet. It demonstrates to the student how a fixed set of underlying statistical parameters, including the correlation matrix, can be realized as a varying range of return outcomes pointing to a range of different investment decisions.

Moreover, the approach also demonstrates the fact that one can take advantage of knowledge of the correlation between individual share returns when constructing a portfolio of shares with the intention of reducing the risk (represented by the standard deviation of return) of the resultant portfolio. For example, returns from gold shares tend to correlate closely to changes in the gold price which in turn often moves counter cyclically with respect to the share returns of large American banks. Hence gold shares are commonly seen as a hedge against poor overall performance of the world financial sector and banking shares. The actual details of algorithmic programming behind portfolio construction via spreadsheet methodology have been handled in the spreadsheet didactic literature (see for example [5]) but are not the focus of this paper. This paper focuses on the statistical principle that low correlation between variables reduces variability of the sum total of variables and applies this to a simple 2-share portfolio construction, allowing for a key statistical nuance. Namely, that in practice, when constructing portfolios, we never know the realised values of share price parameters and therefore cannot come to deterministic solutions regarding the optimal portfolio composition. A better approach is to consider simulated samples which reflect sampling variation in order to obtain a more realistic sense of the uncertainty surrounding investment decisions for the future. The example in this paper is deliberately kept at a simple 2-asset level. More complicated examples (3 or 4 or n-asset cases) have been considered, but the associated increase in the number of required parameter estimates, and the difficulty with the graphic representation of these more complicated cases tends to confound the didactic advantage of a visual spreadsheet approach. Restricting our analysis to the 2-asset case yields an effective graphical depiction of risk reduction on the spreadsheet, and allows the pertinent statistical principles to be clearly demonstrated and easily interrogated by students.

1.1 The Paper Structure

The paper is structured as follows. As a first didactic step we depict on the spreadsheet 2 simulated and correlated variables X and Y drawn from underlying exhibiting distributions with specified expected values, standard deviations and correlations. We consider specifically how the variance (and standard deviation) of the sum of the two variables compares with the sum of the variances (and standard deviations) of the individual components. We then apply these ideas to a simple portfolio problem with two shares and no risk-free rate by applying portfolio weights to the two shares which constitute the portfolio. We consider the risks and returns of this portfolio and how we may construct an efficient frontier to examine which share proportions could give an efficient frontier and how the shape of this efficient portfolio changes as we vary the correlation between the share returns. We then pursue the fact that in practice we may be more interested in the set of *empirical* frontiers and simulate, under repeated sampling, empirical efficient frontiers which capture the type of overall variation of the realized efficient frontier that we may expect to encounter in practice.

2 Demonstrating how the variability of the sum of 2 variables may be reduced by low correlation between them

The first step in the teaching process is to demonstrate that the correlation between X and Y has another important statistical effect, namely it affects the extent to which the variability (either variance or standard deviation) of the sum of the two variables X and Y is different from the sum of the individual variabilities of X and Y considered separately. In many practical problems we are interested in how we may reduce variability. In a financial context where there is an incentive to reduce risk, which is measured by the standard deviation of return, we do this by combining variables with low correlation or combining uncorrelated scenarios. In particular, we will look below at how this affects portfolio construction (investing in different shares or in different countries).

We first look at the statistical theory underpinning these important ideas.

2.1 The Statistical theory behind variation reduction

When random variables are added or combined together, the resulting variable has a variation which, *inter alia*, reflects the correlation between the variables.

For random variables X and Y, we have

$$Var\,(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y).$$

Equivalently, if we let ρ be the correlation coefficient between X and Y,

$$Var\,(X+Y) = Var(X) + Var(Y) + 2\,\rho\,\sigma_{(X)}\,\sigma_{(Y)}\,. \qquad (1)$$

We may therefore surmise that when $\rho \leq 0$,

$$\boxed{Var(X+Y) \leq Var(X) + Var(Y).}$$

Also, we may write from (1) that

$$\sigma_{(X+Y)} = \sqrt{\sigma_{(X)}^{\ 2} + \sigma_{(Y)}^{\ 2} + 2\ \rho\,\sigma_{(X)}\,\sigma_{(Y)}} \quad . \tag{2}$$

and, we may therefore surmise that, since $|\rho| \leq 1$,

$$\boxed{\sigma_{(X+Y)} \leq \sigma_{(X)} + \sigma_{(Y)}}$$

whatever the value of ρ.

We have thus shown that for two random variables X and Y with standard deviations $\sigma_{(X)}$ and $\sigma_{(Y)}$, the standard deviation of their sum, denoted by $\sigma_{(X+Y)}$ can be written in terms of the component standard deviations, $\sigma_{(X)}$ and $\sigma_{(Y)}$ as well as the correlation between X and Y, denoted by ρ.We have also shown that the standard deviation of the sum of X and Y is, in fact, smaller than the sum of the individual standard deviations of X and Y, whatever their correlation. Using these formulae we may demonstrate how the variability of a combination of variables may be significantly less than the variability of the component parts.

In the case when X and Y are (continuously compounded) share returns, and weighted w_1 and w_2 in a portfolio of two shares (such that the weights add up to 1.0), we may write:

$$\sigma_{(w_1X+w_2Y)} = \sqrt{w_1^{\ 2}\sigma^2_{(X)} + w_2^{\ 2}\sigma^2_{(Y)} + 2w_1w_2\rho\ \sigma_{(X)}\sigma_{(Y)}} \tag{3}$$

We will then take this argument one step further in section 3 and demonstrate how the risk of a portfolio of two shares may be reduced through combining the two shares.

2.2 Demonstrating the variable reduction idea on the spreadsheet

We showed above that the variance of a sum of two variables will be less than the variances of the components when $\rho \leq 0$ but that the standard deviation of a sum

of two variables is always less than or equal to the standard deviation of the components. In order to demonstrate these variance reductions in an empirical framework, we now give some theoretical backdrop as to how one might generate bivariate data with any given mean, variance and correlation.
We assume that *X* and *Y* are normally distributed, and, in order to sample from this bivariate distribution *(X, Y)* we use the Cholesky decomposition method (see, for example, [4]).

For the simple 2-variable case where the correlation matrix $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ we will have $L = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{(1-\rho^2} \end{pmatrix}$

Then $\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} X \\ X^* \end{pmatrix} = \begin{pmatrix} X \\ \rho X + \sqrt{1-\rho^2}\, X^* \end{pmatrix}$ where, say, *X* is drawn from some distribution and X^* is independently drawn from the same distribution. Then the generated *Y* will be from the same distribution and have an expected correlation of ρ with *X*. This Cholesky decomposition allows us to generate a bivariate distribution $\begin{pmatrix} X \\ Y \end{pmatrix}$ with correlation matrix Σ. Note that *Y* and *X* can then be suitably scaled and mean shifted so as to have individually any mean and variability (variance or standard deviation) required, allowing us to generate a bivariate distribution $\begin{pmatrix} X \\ Y \end{pmatrix}$ with any given parameters μ_X , μ_Y , σ_X , σ_Y and ρ.

In our spreadsheet exposition, we could let:
cell C2 = μ_X , cell C3 = σ_X ; E2 = μ_Y , cell E3 = σ_Y and cell G2 = ρ
Then the scaled, and mean shifted, Cholesky decomposition for a normally distributed bivariate distribution translates on the spreadsheet to:

X [cell C7] = $C$2+NORMINV(RAND(),0,1)*$C$3 , and
Y[cell D7] = $E$2+$E$3*($G$2*(C7-$C$2)/$C$3+SQRT((1-$G$2^2))*(NORMINV(RAND(),0,1)))

where NORMINV(RAND(), 0, 1) is the Excel formula which generates independent drawings from a N(0, 1) distribution.

We can now generate a range of examples, but the one generated here is instructive. It assumes the means and variances (and standard deviations) of both X and Y separately are equal at 0 and 1 respectively and considers 3 different randomly sampled cases where the true value of rho equals, -0.5, 0.0, and 0.5 respectively. We also list the implied values for the standard deviation of $X+Y$ and the regression coefficients alpha and beta. We can then simulate sets of estimates for the case n equal 50, for these three different values of rho. . We can then capture the stochastic nature of the simulated data through tabulating the empirical mean, variance and correlation of sets of 50 points of the generated data. In the tables below we first give the true parameters assumed (and an example of the true implied parameters for the case rho = -0.5) as well as a particular set of empirical estimates for the 3 cases rho equals, -0.5, 0.0, and 0.5

Table 1: Assumed Parameters (in red) and implied parameters in blue
Case of rho = -0.5

True Values Assumed (and implied true Std.Dev.(X+Y), Cov(X, Y), α, β)					
E (X)	0.0	E (Y)	0.0	Corr (X,Y)	-0.50
Std. Dev (X)	1.0	Std. Dev (Y)	1.0	Std.Dev(X+Y)	1.00
Alpha	0.00	Beta	-0.50	Cov(X,Y)	-0.50

Table 2: Set of *Empirical Estimates* for Case of rho = -0.5

Estimates (50 pts)	X	Y	X+Y
Mean	-0.126	0.138	0.013
Variance	0.730	0.822	0.603
Standard Deviation	0.854	0.907	0.776
Cov(X,Y)	-0.475	Corr(X,Y)	-0.613
Alpha	0.057	Beta	-0.650

Table 3: Set of *Empirical Estimates* for Case of rho = 0.0

Estimates (50 pts)	X	Y	X+Y
Mean	0.211	-0.242	-0.031
Variance	0.750	0.877	1.573
Standard Deviation	0.866	0.937	1.254
Cov(X,Y)	-0.027	Corr(X,Y)	-0.034
Alpha	-0.234	Beta	-0.036

Table 4: Set of *Empirical Estimates* for Case of rho =+0.5

Estimates (50 pts)	X	Y	X+Y
Mean	0.107	-0.055	0.052
Variance	1.378	1.354	4.607
Standard Deviation	1.174	1.164	2.146
Cov(X,Y)	0.937	Corr(X,Y)	0.686
Alpha	-0.128	Beta	0.680

There are several important conclusions we can now reach:
We have shown that the theoretical variance of the sum of X and Y is less than the sum of variances when rho is less than 0. Repeated sampling over sets of 50 points will show this to be the case in practice as well. The student will also observe how empirical estimates of the correlation estimate for 50 data points under repeated sampling, as well as estimates for the mean, standard deviations and variances, will vary with repeated sampling.

We may show, in a similar vein, that the standard deviation of the sum of X and Y is always less than the sum of the standard deviations (whatever the correlation between X and Y) and we may demonstrate this fact empirically.

3 The Portfolio problem in Finance

We now turn our attention to a practical example from the field of finance; the construction of a simple portfolio of two shares with some expected set of returns, standard deviation of returns and correlation between the shares. The simulation has particular relevance in this application as the variability of the outcome can be demonstrated with respect to the original assumptions. By starting off with, for example, share A with expected return of 1% (per month) and standard deviation of return 1%, combined with share B which has an expected return of 1.5% (per month) with standard deviation of return 1.5%, we can demonstrate the observed return and risk of a portfolio (of the two shares) as we vary the weights of the individual shares in the portfolio.

3.1 The Efficient Frontier of the Portfolio

A plot of the returns and risks together for portfolios resulting from different weights (or proportions) of share A and share B allows us to demonstrate the so-called efficient frontier ([8], [9]). If we were able to ascertain the true expected return and risk, this would give us the theoretical efficient frontier.

From the theory above (see equation (3)) we may infer that as long as the share returns are not perfectly correlated, then the standard deviation of the portfolio return of a combination of A and B (and hence the associated risk) will be reduced. In this case, we assume the correlation between the share returns (labelled X and Y) is zero. Working with the true expected values, we first explore the combination of return and risk (standard deviation of return) associated with different portfolios comprising a% of A and (100-a)% of B, where $0 \le a \le 100$ to construct Table 5 below.

Table 5: Table of values for constructing an Efficient Frontier

True Values Port. Components		Corr(X,Y)	0.0
E(X)	1.0	Std. Dev (X)	1.0
E(Y)	1.5	Std. Dev (Y)	1.5
Wt(A)	Wt(B)	True Port. Ret	True Port. Risk
0%	100%	1.50	1.50
10%	90%	1.45	1.35
20%	80%	1.40	1.22
30%	70%	1.35	1.09
40%	60%	1.30	0.98
50%	50%	1.25	0.90
60%	40%	1.20	0.85
70%	30%	1.15	0.83
80%	20%	1.10	0.85
90%	10%	1.05	0.91
100%	0%	1.00	1.00

From Table 5 we see that for the range of weights considered, the standard deviation (Risk) of portfolio return is minimized when Wt(A) is 70% and Wt(B) is thus 30%.

The values in Table 5 are plotted in Figure 1 and Figure 2 below (blue lines)

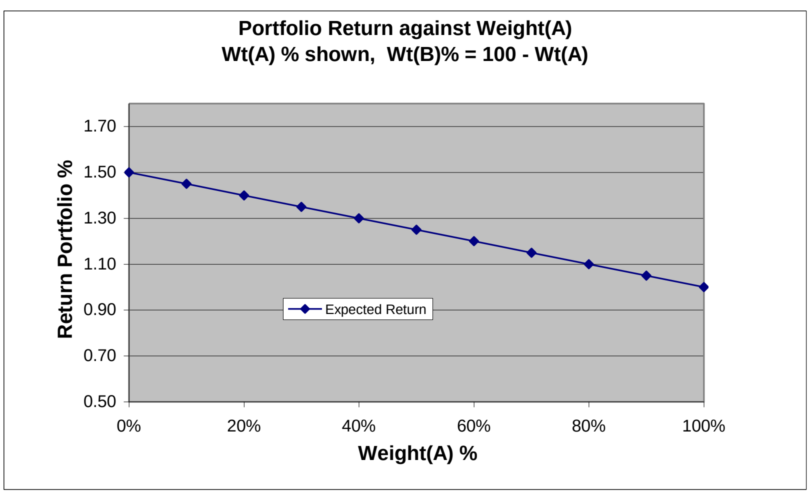

Figure 1: "true" Portfolio returns against weight of share A

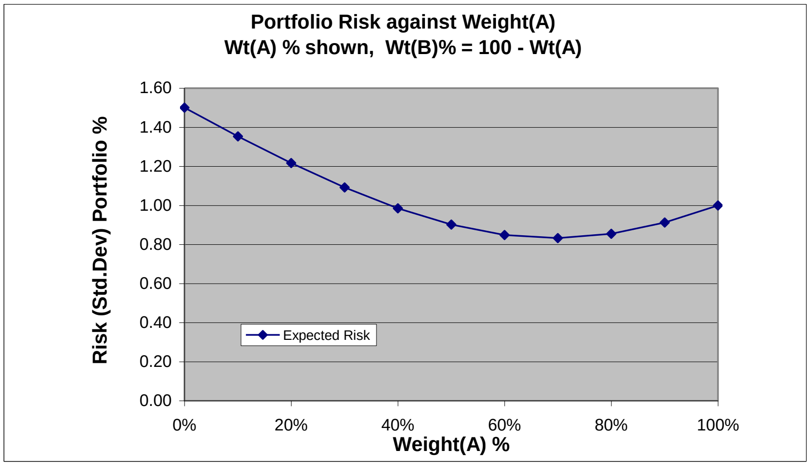

Figure 2: "true" Portfolio Risk against weight of share A

If we keep the expected returns and standard deviations the same but vary the correlation (say to -0.5) we would observe that the reduction of risk effect becomes more marked. We now combine Figures 1 and 2 onto a single diagram (Figure 3) which displays the return and risk associated with differently weighted portfolios into a single diagram. Note that the lower section of this frontier is dominated by the upper section and thus , the lower section is not efficient. For ease of graphical exposition on the spreadsheet we will plot these frontiers with both lower and upper parts.

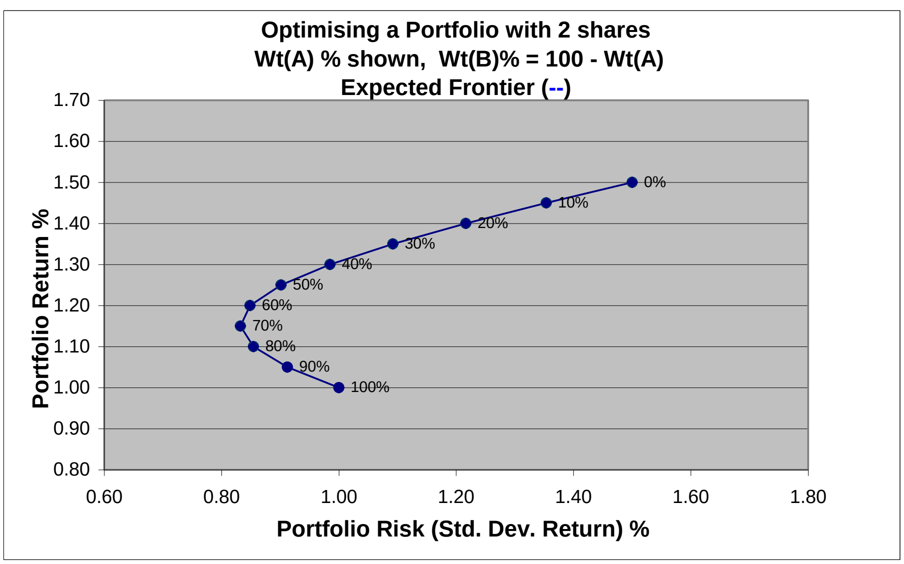

Figure 3: The "true" Efficient Frontier (showing also dominated portfolios below the efficient frontier).

An investor would be seeking to maximise return and minimise risk. In terms of return, depending on the expected return of share A and B, as the portfolio contains more of the share with the higher expected return, so the return of the portfolio will increase and vice versa as one adds less of that share. So the direction of return is monotonic (increasing in the direction of increased proportions of the share with the higher expected value). However, the story with the risk associated with the portfolio is different. This is generally not monotonic in changes in portfolio composition. As noted above, diversification often can be seen to bring risk down so that by combining two shares in the portfolio one can increase return and reduce risk, up to a certain point. Beyond this point, the addition of increased proportions of a second share requires an increase in risk for increased return.
The result is that a section of the consolidated risk return curve is "dominated"; in other words, for (certain) given levels of risk there will be two different combinations of X and Y (portfolios) that yield this level of risk: one will give a higher return than the other. The one with the higher return is clearly the preferred portfolio; the other is "inefficient". The efficient frontier can thus be used to form optimal portfolios under various criteria such as providing the weights which would give the minimum total risk portfolio.

Students may experiment with different correlations to examine how the effect becomes more marked as the correlation reduces to -1. In Table 6 below we keep the expected monthly returns and risks the same but change the correlation to -0.5.

Table 6: Table of values for true Efficient Frontier

True Values Port. Components		Corr(X,Y)	-0.5
E(X)	**1.0**	**Std. Dev (X)**	**1.0**
E(Y)	**1.5**	**Std. Dev (Y)**	**1.5**
Wt(A)	Wt(B)	**True Port. Ret**	**True Port. Risk**
0%	100%	1.50	1.50
10%	90%	1.45	1.30
20%	80%	1.40	1.11
30%	70%	1.35	0.94
40%	60%	1.30	0.78
50%	50%	1.25	0.66
60%	40%	1.20	0.60
70%	30%	1.15	0.61
80%	20%	1.10	0.70
90%	10%	1.05	0.84
100%	0%	1.00	1.00

The more negatively correlated the component shares are, the better the reduction in risk. However, even when there is positive correlation between the shares or when there is zero correlation, there is often some value to diversification across two shares, provided the one share is not very much more risky than the other.

We will now show that repeated sampling drawn from the (assumed normal) distribution will give us multiple empirical sets and give us some idea of how variable the empirical efficient frontiers will be for some given underlying expected return and risk characteristics. Moreover, we may show firstly which weighted combination of A and B would be expected to be appropriate for producing a portfolio with optimal characteristics (for example minimizing total risk of the portfolio) for the true expected case but secondly how variable these weights may be when we consider a number of empirical efficient frontiers.

3.2 Sampling from the true distribution to form empirical estimates

Of course, in practice, we never know what the exact movement of a share will be at any point in time (i.e. the movements are subject to random variation and are thus stochastic) so although we can mathematically calculate the theoretical returns and risk for a portfolio, the empirical (observed on the spreadsheet) share returns and risks will be subject to random variation. One approach to assessing the effect of random variation on an efficient frontier is to perform a standard sensitivity analysis, varying the input parameters and then considering the impact on the optimal weights and hence the frontier. In this paper, we adopt the alternative approach of simulating different underlying data sets sampled from distributions which are characterized by *a given* set of parameters. Under the assumption of normality, we sample returns from a normal distribution with the given expected values, to form empirical sets of share return data and demonstrate to students the range of possibilities for return and risk which could occur in practice. The number of data points sampled may be set to any value but for demonstration purposes here we sample 50 data points and compute the empirical average return and risk from these sampled 50 data points.

We can then superimpose various *sampled* empirical efficient frontiers on the *true* efficient frontier plotted assuming perfect foresight (that is, computed using the true expected returns and risks). In Figure 4 the sampled empirical frontier is plotted in red. The red frontier constitutes the summary results from a random sample of share returns and their associated risk.

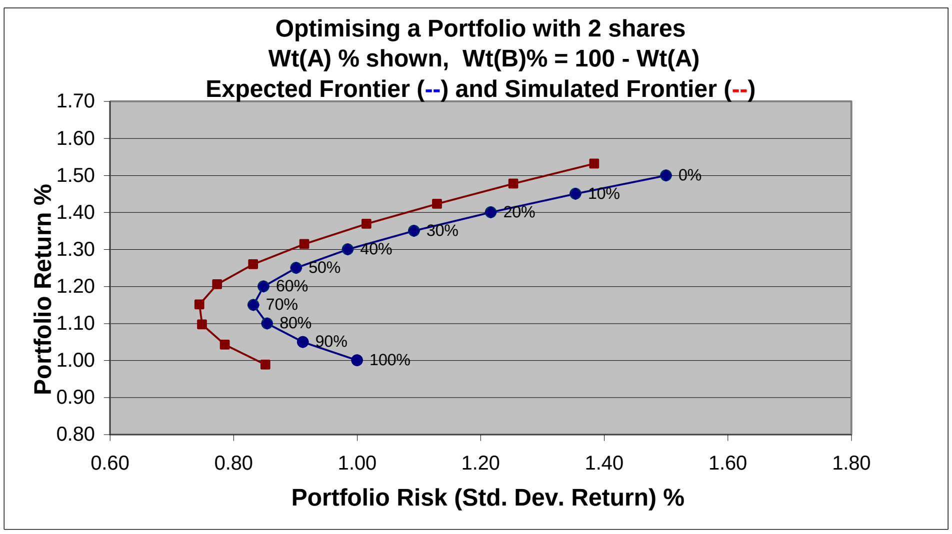

Figure 4: "true" Efficient Frontier (blue) with a simulated Efficient Frontier imposed (red) (showing also dominated portfolios below the efficient frontier)

We now show how repeated sampling may change the empirical frontiers (red) relative to the fixed true frontier (blue) and show 3 further cases (Figures 5, 6 and 7) where in each case the expected frontier remains fixed, but the empirical frontier varies randomly as it would do in practice.

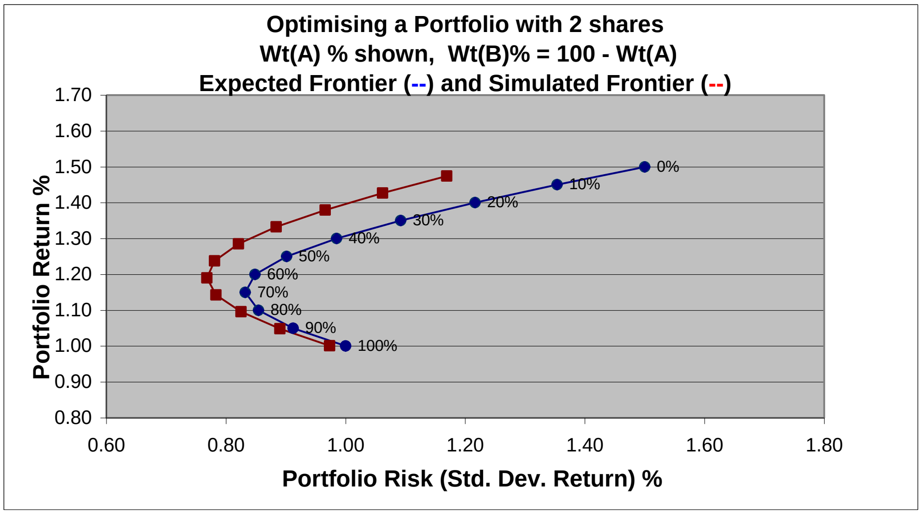

Figure 5: "true" Efficient Frontier (blue) with a (further) simulated Efficient Frontier imposed (red) (showing also dominated portfolios below the efficient frontier)

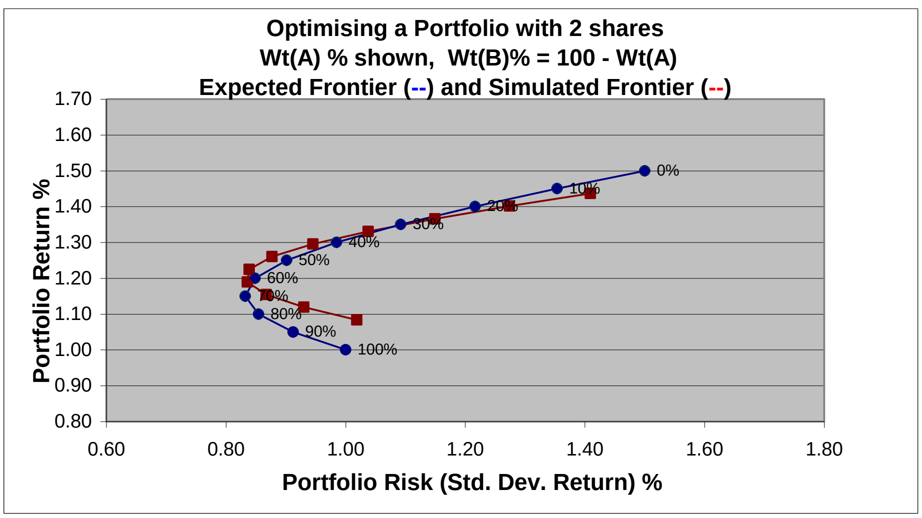

Figure 6: "true" Efficient Frontier (blue) with a (further) simulated Efficient Frontier imposed (red) (showing also dominated portfolios below the efficient frontier)

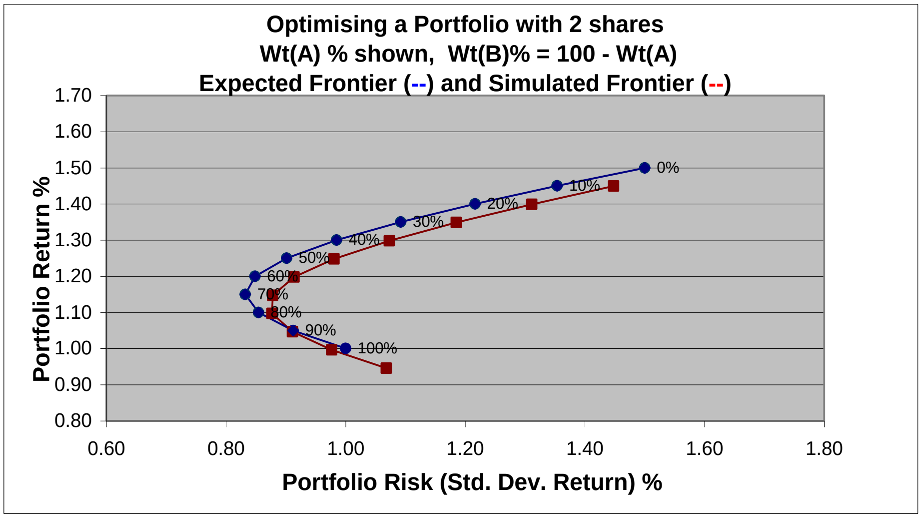

Figure 7: "true" Efficient Frontier with a (further) simulated Efficient Frontier imposed (showing also dominated portfolios below the efficient frontier)

3.3 Repeating the Simulation Process to capture the Distribution of Empirical Portfolio Weights

One can then repeat the portfolio estimation process for the same set of fixed underlying parameters, but with different simulated data. To demonstrate this, 20 sets of data were generated with 50 (bivariate) point sets and the 50 associated efficient frontiers generated. These can be accessed on the associated spreadsheet, with each generated frontier on a separate page. A useful didactic feature of this analysis should be a discussion around the distribution of the portfolio weights, and one of particular interest is the distribution of weights associated with the minimum risk portfolio. One can then generate different empirical renditions of the distribution by simply pressing the F9 (recalculate) key, and the one depicted below represents one particular rendition. In addition, we show a distribution for the empirical minimum risk itself. Again, of course, this constitutes a random rendition of the empirical distribution and different renditions can be seen by pressing the F9 key. The distribution of weights (for A) shown in Figure 8 is intuitive, with most (optimal) weight measures falling close to a 60% weight for share A (and hence a 40% weight for share B).

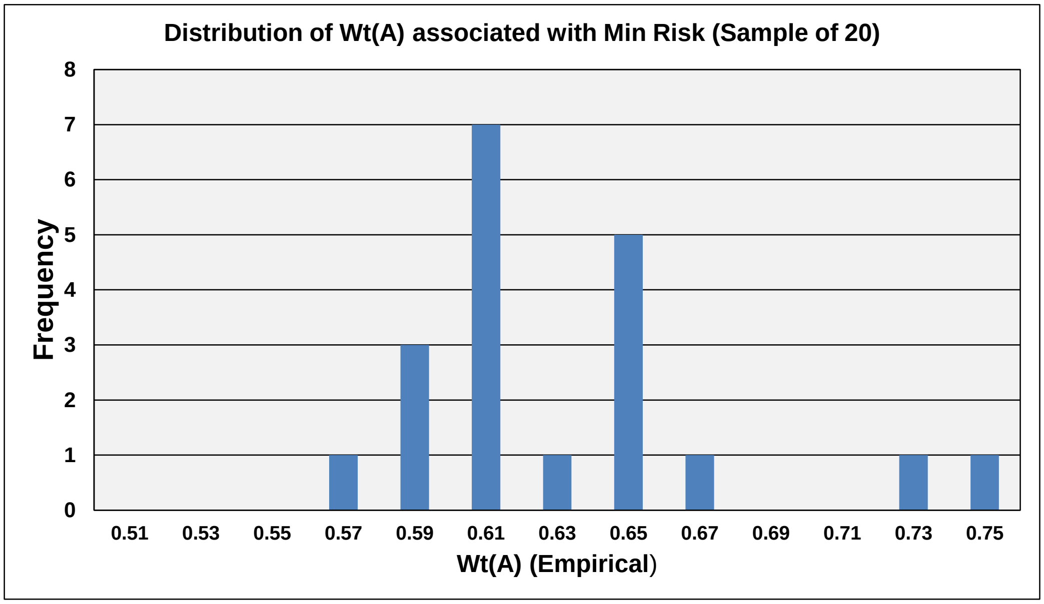

Figure 8: A Rendition of the distribution of Empirical Weights for Share(A) associated with the minimum risk portfolio.

As mentioned above, we also include one of the generated distributions for the empirical value of the actual minimum risk over the sample of 20 simulations.

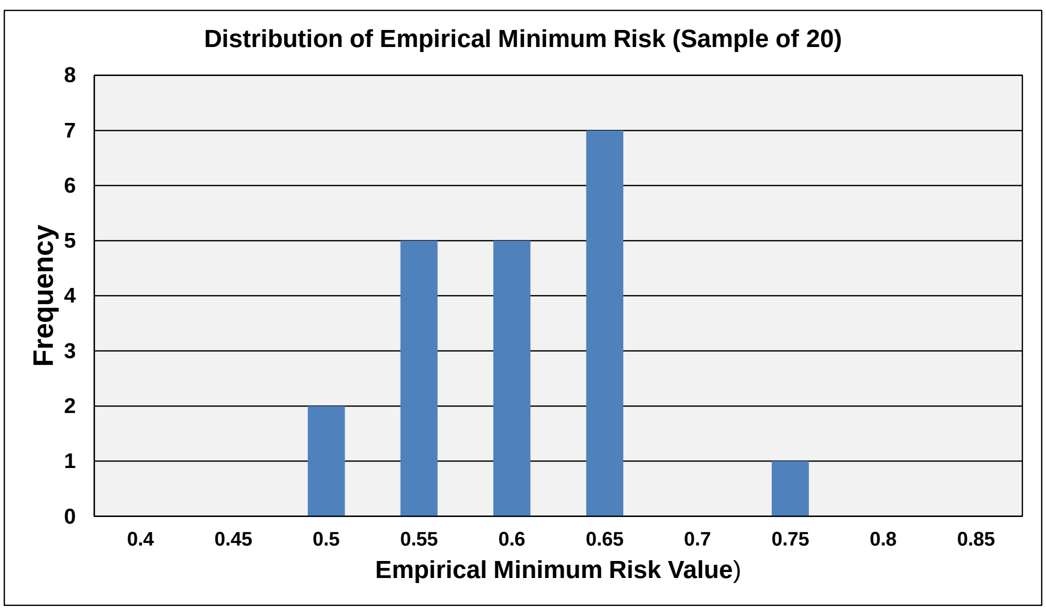

Figure 9: The distribution of the Empirical Minimum Risk values.

Again, the distribution of the values themselves are intuitive, and as expected the empirical values in Figure 9, are centred between 0.55 and 0.65.

4. Student reaction to the exercise as a means for demonstrating variability reduction and the application to share portfolio outcomes

This spreadsheet based, active learning approach has been demonstrated to Honours (post-graduate fourth year) students in the courses Econometrics and Portfolio Theory in the department of Statistical Sciences at UCT over the three year period from 2014 to 2016. It generally requires three teaching sessions to adequately cover these concepts; at least one lecture for the correlation concepts and two for the portfolio concepts. The students found the visual demonstration of variance reduction when random variables were combined enlightening – it was a concept that they had been taught from a mathematical perspective in their undergraduate curriculum but felt that the simulation approach was key to establishing a deeper and more solid understanding of the concept. The approach consolidated the fact that knowledge of the true correlation, mean and variance of two random does not give a deterministic estimation outcome but a whole range of possible estimation outcomes. This leads to an understanding that the estimated correlation coefficient between two random variables is a random variable, itself. However, the portfolio example had particular didactic traction in the repeated sampling context. Students were familiar with the idea of risk reduction through portfolio construction but had focused on the idea of deterministic portfolio optimization with a *particular* data set yielding a *particular* optimization result. That is they had not absorbed the key idea that the efficient frontier *itself* is subject to random variation and that any optimization process has to be seen in the context that the optimization result itself is simply one result drawn from a distribution. Some specific comments from students were collected over a three-year period from 2014 to 2016 and are included below.

4.1 Specific comments from students for the courses Econometrics and Portfolio Theory in Applied Statistics Honours programme at UCT

Econometrics 2014

Think the interactive mode very helpful. The effect of changing correlation on the portfolio efficient frontier very nice to see visually.

Energising to see these things visually. Tired of notes and boards. Nice to play with it myself.

Liked the mix of seeing the Maths on the one-side and the visual demonstration on the other.

Econometrics 2015

Chalk and cheese difference in teaching. Really helped understanding.

Never properly understood where that frontier came from and the fact that it moves for a given correlation is interesting.
Love visual approaches. Really understand how the efficient frontier works now.
Tried it out myself at home. Actually really good.
Econometrics 2016
Think we should do more visual demonstrations. Really gets ones attention.
Portfolio simulation something I had never thought of. Interesting!
Maths approach very dull compared to this visual approach.
Portfolio Theory 2015
Definitely helpful to see where the efficient frontier comes from.
The Maths behind the efficient frontier impossible to understand but this really helps.
Have never seem portfolio frontier dome like this – excellent teaching tool.
Portfolio Theory 2016
The frontier demo was a hit. Much better than doing via a text book.
Interesting the sensitivity of portfolio selection to underlying parameters. Thought that very good.
Really nice that we were given the program to use ourselves.

5 Conclusions

We have lead students through a set of useful statistical paths and associated insights. Firstly, we demonstrate the fact that statistical variability may be reduced by combining entities that have low correlation. In particular, when 2 variables are combined which are less than perfectly positively correlated the standard deviation of the sum will always be less than the sum of the standard deviations of the 2 variables by themselves. This key idea is particularly relevant to the field of Finance, where risk is often proxied by standard deviation of return. We consider 2 sets of share returns and their associated risk and apply these ideas to the construction of a 2 share portfolio and demonstrate how we may construct an efficient frontier; that is a set of weighted combinations of the 2 shares such that the combination has maximum possible return for each risk. We enhance this idea with an important fact, namely that, in practice, we never know what the realized share returns and risks will be, even though we may be able to form estimates of their expected values. The share returns and risks will always be subject to random and unpredictable variation and thus the efficient frontier looking forward is not a fixed deterministic quantity but a whole range of possibilities.
We show that by simulating efficient frontiers through repeated sampling from some assumed set of underlying share return parameters the efficient frontier can display considerable variation itself. This is important for embedding the idea that the process of share selection for a portfolio using Markowitz principles is not a precise exercise but one which requires circumspection as share returns and share return variability are not

predictable with any precision but subject to the vagaries of random variation. The spreadsheet teaching approach to the issue of risk reduction, as well as efficient frontier simulation, has been well received by students in senior courses and seen as a refreshing and energizing change from the standard text-book-based presentation.

References

[1] Baker, J. and Sugden, S. (2003). Spreadsheets in education. The first 25 years. *eJournal of Spreadsheets in Education*, 1(1), Article 2. Available at https://epublications.bond.edu.au/ejsie/vol1/iss1/2/ [Accessed 3 May 2018].

[2] Barr, G. and Scott, L. (2011). Teaching statistics in a spreadsheet environment using simulation. *eJournal of Spreadsheets in Education,* 4(3), Article 2. Available at: http://epublications.bond.edu.au/ejsie/vol4/iss3/2. [Accessed 3 May 2018].

[3] Barr, G. and Scott, L. (2015). Spreadsheets and simulation for teaching a range of statistical concepts. In: H. MacGillivray, B. Phillips, M. Martin, ed., *Topics from Australian conferences on teaching statistics: OZCOTS 2008-2012*, 1st edition, New York: Springer, pp 99 - 117.

[4] Horn, R. and Johnson, C. (1991). *Topics in matrix analysis*. Cambridge: Cambridge University Press, pp 242 - 268.

[5] Kwan, C. (2007). A simple spreadsheet-based exposition of the Markowitz critical line method for portfolio selection. *eJournal of Spreadsheets in Education,* 2(3), Article 2. Available at: http://epublications.bond.edu.au/ejsie/vol2/iss3/2 [Accessed 8 May 2018].

[6] Kwan, C. (2011). An introduction to shrinkage estimation of the covariance matrix: A pedagogic illustration. *eJournal of Spreadsheets in Education*, 4(3), Article 6. Available at https://epublications.bond.edu.au/ejsie/vol4/iss3/6/ [Accessed 8 May 2018].

[7] Kwan, C. (2017) Shrinkage of the sample correlation matrix of returns towards a constant correlation target: A pedagogic illustration based on Dow Jones stock returns, *eJournal of Spreadsheets in Education,* 10(1), Article 3. Available at http://epublications.bond.edu.au/ejsie/vol10/iss1/3 [Accessed 8 May 2018]

[8] Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 5, pp 77-91.

[9] Markowitz, H. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, 3, pp 111-133.

[10] Mills, J. (2002). Using computer simulation methods to teach statistics: A review of the literature. *Journal of Statistics Education*, vol. 10 (1).