

Avery Bicks and Jay Weiss

Professor Smith

6 February 2020

ECON 190

Revisiting "The Next Best Thing to Knowing Someone Who is Usually Right"

Introduction

Mean-Variance Analysis is a common method used for making portfolio allocation decisions. The analysis uses the expected value, standard deviation, and covariance of several risky assets to produce a space which shows the infinite possible combination of these assets. When combined with a riskless asset, there is an “optimal risky portfolio” on the frontier which describes a portfolio that, at any level of risk, dominates all other portfolio allocations (Figure 1).

In order to create an optimal risky portfolio, we must estimate the portfolio’s assets’ expected values, standard deviations, and correlation with one another. For a portfolio to be considered optimal it should maximize expected return for a given level of risk. According to the Markowitz-Frontier, there exists a distribution of several portfolios that strikes an efficient balance between risk and return.

This paper seeks to update Smith, Steinberg, and Wertheimer (2008) which explores three different methods of estimating mean-variance analysis parameters, and their effect on portfolio performance. Using more recent data, this paper will examine the success of these three asset allocation strategies.

The first approach to estimating these inputs solely uses observed historical returns; the typical method for this analysis. However, another approach that will be tested involves informed predictions from experts. Like Smith, Steinberg, and Wertheimer’s paper, this study will use expert opinions from the Philadelphia Federal Reserve’s Semi-annual Livingston survey.

The Livingston Survey is given to expert economic forecasters who are asked to provide their best predictions of the S&P 500 index level along with long-term Treasury Bond yields in the following month, 6-months, and 12-months following the survey. Finally, this paper will update the findings of using regression towards the mean for asset risk and reward estimation by combining these expert predictions of these parameters with observed historical data. These three estimations will be implemented to create and test three portfolios whose average return, standard deviation, and Sharpe ratio will be measured over 6-month time horizons.

Smith, Steinberg, and Wertheimer's paper ultimately found that optimal risky portfolios that were selected on the basis of observed historical data performed the worst while the regression-to-the-mean portfolios not only had a higher average return but also a higher Sharpe ratio. Their paper also found that the predictions of S&P 500 returns in the Livingston Survey to be both substantially and significantly negatively correlated with actual stock returns.

Optimal Portfolios

As mentioned in the original paper, portfolio allocation involves a model of utility maximization by the investor given the distribution of asset returns. By using mean-variance analysis, the investor uses the expected value and standard deviation of portfolio returns to maximize this expected utility. Ultimately, the weighting of a portfolio, α_i , is chosen where the

portfolio mean, $\mu = \sum_{i=1}^n \alpha_i \mu_i$, for a given level of variance, $\sigma^2 = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \sigma_i \sigma_j \rho_{ij}$, constrained

by, $\sum_{i=1}^n \alpha_i = 1$, is maximized.

When looking at the Markowitz frontier, the upper section of the curve is deemed the “efficient frontier” where standard deviation is the smallest at the apex. This frontier identifies all potential portfolio allocations that provide the highest expected returns for each given standard deviation, and, when combined with a risk-free asset, such as a short-term treasury bill, there is an “optimal risky portfolio. According to Tobin’s Separation Theorem, this optimal risky portfolio is located where the risk-free asset is tangent to the frontier.

The portfolio allocation can only truly be optimal if the probability distribution is known, but, during estimation, there is often error with asset mean, standard deviation, and covariance that leads to suboptimal allocation, and, ultimately lower expected value than that of the optimal portfolio maximum value.

This loss in expected utility and value is the result of the incorrectly estimated frontier. There are several reasons why this frontier might have been incorrectly estimated, but it is often due to the estimates being made on the basis of historical distributions of returns. However, investors might be able to avoid this loss of utility if they are able to more correctly estimate these probability distributions from these observed historical returns.

Historical Estimates

In contradiction to the efficient market hypothesis, there is evidence that using observed historical data to make estimations in mean-variance analysis can still outperform the market. Cohen and Pogue (1967) showed that Markowitz-frontier portfolios that used historical data for estimation performed better than randomly selected portfolios.

While historical data can be directly inputted as means for estimating future Markowitz frontiers, one can use these data to instead inform their beliefs about future asset mean, variance, and correlation. If an investor sees that the S&P500 has returned more than 20% the previous 2 years but currently seems at an unreasonably high level, it may not be accurate to assume another 20% return, but rather weight the expected return with a smaller probability of a similar year.

There are rational reasons why significant portions of pre-existing literature argue that predicting financial market performance is “perhaps the greatest challenge” (Didier Sornette, 320). However, historical estimates can offer tremendous insight into investor sentiment during a given period, which could be useful in the construction of a strategy with a current, forward-looking scope. Since people’s expectations, greediness, and fear can “intertwine” and drive markets towards irrational highs and lows, awareness (Sornette, 321). An appreciation for this phenomenon and the tremendous uncertainty associated with it can be garnered from the examination of historical performance. In turn, an investor could use this information to adjust his forecasts and hopefully reduce the uncertainty surrounding an investment. Understanding the pitfalls of historical forecasts can allow an investor to contextualize historical data.

Likewise, the intuition holds for asset volatility. Observed historical volatility of asset prices do not need to be directly inputted as estimation but rather can aid in forming beliefs about volatility going forward. Utilizing this approach makes sense when making estimates regarding assets like Treasury Bill rates. In the past 90 years, historical data would suggest that the standard deviation of Treasury Bill rates around the mean is about -3% to 5% but, when allocating a portfolio using mean-variance analysis, an investor would never want to use this

variance because there is no uncertainty about returns on Treasury Bills, and, thus, will always provide a standard deviation of 0.

While the same logic regarding the volatility of Treasury Bills does not directly apply to the S&P 500 and long-term Treasury bonds, these assets should also reflect estimates of future uncertainty, rather than volatility during dissimilar economic periods. Smith, Steinberg, and Wertheimer's paper provides the example of the Federal Reserve announcing plans to hold Treasury rates steady which still holds. While the historical standard deviation of Treasury rates might suggest more volatility, an investor knows from this announcement that there is little uncertainty regarding volatility, and, thus, a zero standard deviation can be used as an estimate.

Lastly, estimations of correlations between stock and bond prices are rarely exactly the same as historical averages. Different economic periods in the past have produced both positive and negative correlation coefficients between stocks and bonds, so, ultimately, this correlation coefficient should depend on the uncertainty of interest rates and the strength of the economy for bond and stock return correlation.

Changing interest rates can signal varying scenarios of economic conditions, with an increase being a result of a quickly expanding economy or a Federal Reserve contractionary policy, either of which would result in differing correlations of stock and bond returns.

Regression Towards The Mean

Gary Smith and Teddy Schall provide a convincing framework to analyze regression to the mean using examples in Major League Baseball. In athletics, particularly baseball, observed performance is not always a reliable measurement for a player's true skill level (Smith & Schall, 231). For instance, a hitter who receives the Triple Crown, which occurs when a single player

leads the Major League in home runs, batting average, and runs batted in (RBIs), is highly unlikely to repeat this performance. Frequently, a player with a Triple Crown performance will have worse batting statistics the following year and the preceding year. The regression to the mean explanation for this phenomenon would be that the batter's skills did not increase dramatically and then "deteriorate, but rather that their unusually good performance [during the prior season] exaggerated their skills" (Smith & Schall, 231).

This analysis can be useful in explaining abnormal performance in economic and financial variables. For example, consider a country with a relatively consistent level of GDP growth. If this country's GDP growth jumps dramatically during one quarter, and there are no significant changes to macroeconomic health remains and the global economy, it is unlikely to witness this level of growth in the following quarter. Although this statement may seem pessimistic, the reverse is also true: a serious decline in a country's GDP growth during one quarter will likely not be as bad in the proceeding one, barring any significant macroeconomic disturbance.

Regression to the mean analysis is highly valuable for interpreting expert financial forecasters' predictions. Generally, analysts with extremely optimistic and pessimistic forecasts are likely too excessive in their respective directions. Keil, Smith, and Smith [2004] confirmed that the companies with overly optimistic (pessimistic) earnings predictions tended to do better (or worse) than average, but closer to the average than predicted.

Modeling and Implementing Regression Towards the Mean

Given an actual return, Y , which as has a probability distribution with expected value m at a point in time,

$$Y = \mu + \omega, \quad E[\mu] = 0 \quad (1)$$

We then assume, X , or the expert forecast return is equal to that of the actual return's expected value with an error score, ε , which has a matching distribution and is independent:

$$X = \mu + \varepsilon, \quad E[\varepsilon] = 0 \quad (2)$$

This paper's model of regression towards the means seeks to infer the expected value of the return from the forecasted return, rather than make predictions of forecasts given known value of μ . With data regarding μ and X , we would typically estimate:

$$\mu = \alpha + \beta X + v$$

With an OLS, or ordinary least squares, slope of:

$$\hat{\beta} = \frac{\text{cov}[X, \mu]}{s_x^2}$$

We then use this estimate to shrink the forecasts toward the mean:

$$\hat{\mu} - \bar{\mu} = \hat{\beta} (X - \bar{X})$$

Where we would expect the forecasts and actual values to be roughly the same as:

$$\hat{\mu} = \hat{\beta} X + (1 - \hat{\beta}) \bar{X} \quad (3)$$

Therefore, $\hat{\beta}$ is used to shrink the period's expert forecast toward the average forecast over time in order to predict expected return.

In this case, however, we are unable to use data for X and μ to estimate β , and we do not observe μ directly, but rather we have an observed Y , which is given by the actual return. We can then estimate this equation using OLS:

$$Y = \gamma + \delta X + \psi \quad (4)$$

And this will give an estimated slope of:

$$\begin{aligned}\widehat{\delta} &= \frac{\text{cov}[X, Y]}{s_X^2} \\ &= \frac{\text{cov}[X, \] + \text{cov}[X, \omega]}{s_X^2}\end{aligned}$$

If $\text{cov}[X, \omega]=0$, then $\widehat{\delta} = \widehat{\beta}$ and our estimate of δ can be used as an estimate of β . Thus, Equation (3) can be implemented. Ultimately, each period's prediction will be regressed toward the mean using the relationship between the predicted and actual returns.

Data

Similar to the previous paper on the topic written by Smith, Steinberg, and Wertheimer, this paper will use the Livingston Survey as the basis for the expert predictions. The Livingston Survey, initiated in 1946 by Joseph Livingston, consists of a semiannual survey that included predictions of macroeconomic variables by business economists. After Joseph Livingston passed away in 1978, the Philadelphia Federal Reserve Bank began to maintain and conduct the survey. The Federal Reserve of Philadelphia mails the survey to a wide variety of economic forecasters every May and November and the results are published each June and December.

In accordance with the previous paper, we will use the level of the S&P 500 index and the secondary market interest rates on 10-Year Treasury Bonds. The survey asks the forecasters to predict what the values of each of these assets will be on the last day of the survey release month (June and December) and both six and twelve months after that date.

The Livingston Survey started to report the yields of long-term treasury bonds and the level of the S&P 500 index in 1992, and up until December 2002, the survey asked to predict the yield on 30-year Treasury Bonds, and, in subsequent years, switched to forecasting the 10-year

yield. For the sake of consistency, this paper will also use the 6-month forecasts, and only uses the forecasts on 10-year Treasury Bonds for historical and forecasted returns.

Mean-Variance analysis requires the calculation of predicted returns from the predicted levels of both Treasury Yields and S&P 500. In order to do this, we use the predicted values on the last day of the survey release month and compare that with the predicted values six months thereafter. Thus, if the survey was mailed in November 2014 and released on the last day of December 2014, to calculate the implied predicted return from purchasing bonds on the last day of December 2014 and selling on the last day of June 2015, we use the difference between the predicted level of the 10-Year Treasury on December 2014 and June 2015.

To calculate the predicted return, \widehat{R}_S , for the S&P 500 index, the dividend yield is added to the forecasted capital gain:

$$\widehat{R}_S = \frac{D_1 + \widehat{P}_1 - \widehat{P}_0}{\widehat{P}_0}$$

Where \widehat{P}_0 is the forecasted level of the S&P 500 at the end of the survey release month and \widehat{P}_1 is the level six months after that. D_1 is the dividend paid during the six-month interval which is assumed to be known and paid upon the period ending. The same method is used for calculating the actual total return, \widehat{R}_S , by using the actual values of the S&P 500 index at the survey release date and six months after:

$$R_S = \frac{D_1 + P_1 - P_0}{P_0}$$

For the predicted return on the 10-Year Treasury bond calculation, it is assumed to have initially sold at par with a coupon to its yield to maturity. The predicted yield at the end of the

survey month is represented by \hat{y}_0 , the predicted yield six months thereafter is represented by \hat{y}_1 , the maturation value is represented by M where M = 100, and the number of semi-annual coupons is represented by n where n = 20. Thus, given $C = \frac{\hat{y}_0}{2} 100$, the bond's initial price is equal to:

$$\begin{aligned} P &= \sum_{t=1}^n \frac{\hat{C}}{(1+\hat{y}_0)^t} + \frac{M}{(1+\hat{y}_0)^n} \\ &= \sum_{t=1}^n \frac{(\hat{y}_0/2)100}{(1+\hat{y}_0)^t} + \frac{100}{(1+\hat{y}_0)^n} \\ &= 100 \end{aligned}$$

And the market price of the bond after six months is equal to:

$$\hat{P}_1 = \sum_{t=1}^{n-1} \frac{(\hat{y}_0/2)100}{(1+\hat{y}_0)^t} + \frac{100}{(1+\hat{y}_0)^{n-1}}$$

Where y_1 is equal to the forecasted yield to maturity when the bond is sold. The implied predicted return is then calculated as:

$$\hat{R}_B = \frac{\hat{C} + \hat{P}_1 + \hat{P}_0}{\hat{P}_0}$$

The same method of calculation was used for the actual returns, \hat{R}_B , using the actual yields to maturity.

Portfolio Construction

This paper, like Smith, Steinberg, and Wertheimer's, treats all semiannual data as independent, and, thus, we assume the investor knows about all other observations, regardless of

sequence. For example, if an observation were to be randomly selected, we could assume that all other observations in the data set have already occurred.

Like Smith, Steinberg, and Wertheimer's paper, we test the performance of three different portfolios. First, our benchmark portfolio is estimated using the historical return distribution, so the expected values, standard deviations, and correlations of assets are assumed to be equal to the averages of the data set. Moreover, our second portfolio estimation is provided by the Livingston Survey's forecasts of the S&P 500 and 10-Year Treasury Bond yields in each six-month interval. These forecasts give us our expected values in our portfolio composition. Lastly, our regression-to-the-mean portfolio uses the relationship between the forecasted and actual returns and implements Equation (3) to create expected value estimates. This is done by estimating the slope of the relationship between the Livingston Survey values and actual values for the periods in our data set and using this regression to shrink these values toward the average return for other periods.

Unlike Smith, Steinberg, and Wertheimer's previous paper, to construct these three portfolios we chose to estimate correlations rather than covariance. However, we first need to calculate the covariance matrix of bond and stock returns by calculating the predicted and actual return's squared difference for the six-month horizons, omitting the current period. The covariance matrix of squared errors for m observations serves as our posterior covariance matrix estimate of stock and bond returns:

$$\hat{\sigma}_{ij}^2 = \frac{\sum_{t=1}^m (R_t - \hat{R}_t)(R_j - \hat{R}_j)}{m}$$

Given $i = j = S$ for stock variance, $i = j = B$ for bond variance, and $i = S$ and $j = B$ for covariance. Similarly, historical forecasts use the historical covariance matrix derived from historical means \overline{R}_S and \overline{R}_B as estimates of expected returns:

$$S_{ij}^2 = \frac{\sum_{t=1}^m (R_t - \overline{R}_i)(R_t - \overline{R}_j)}{m-1}$$

In these periods, the observed differences between predicted and actual returns are used to estimate the covariance matrix in each period. We then divide by the product of the stock and bond standard deviations to determine estimates of correlation:

$$C_{ij} = \frac{cov[i,j]}{\sigma_i \sigma_j}$$

These expected returns, standard deviations, and correlation matrices are then used to create a Markowitz frontier from which we select an optimal risky portfolio. This paper also focuses on the optimal risky portfolio to combine with a riskless asset. The six-month Treasury bill is used as the riskless asset with return calculated from the quoted secondary market rates at the beginning of the six-month time horizons.

Results

Table 1 offers a comprehensive look at the predicted and actual returns for both stocks and bonds during each six-month period. The table also includes the quoted six-month Treasury bill returns that served as the risk-free asset. Similar to the results of the previous paper, the predicted stock returns varied greatly in magnitude, ranging anywhere from 2.14% to 47.78%, however, during no six-month period in the fourteen years did experts forecast a negative return on stocks. This is contrasted with predicted bond returns which not only had a small range, from -2.27% to 1.78% but also had only six periods in which experts predicted a positive return in

Treasury bonds. Moreover, it is clear that experts continue to take economic conditions into account, such as in December 2008, when forecasters predicted a recovery in the market after the crash. Unsurprisingly, survey predictions continue to be inaccurate with a correlation between predicted and actual returns of -0.07 for stocks and 0.384 for bonds.

Unlike the previous paper, the correlations between predicted and actual stock returns are no longer overwhelmingly negative. However, while they are still negative, they now seem to be mostly uncorrelated, causing the opportunity to profit off knowing the forecaster who is usually wrong to disappear. Now, actual and predicted bond returns are positively correlated by a similar coefficient as the negatively correlated actual and predicted stock returns in Smith, Steinberg, and Wertheimer's research. It has been consistently profitable to know the bond forecasters who are usually right.

Table 2 shows the returns of the three portfolio strategies implemented. The Sharpe ratio in this table is calculated by taking the average difference between the portfolio returns and the T-bill returns and dividing the portfolio standard deviation. While similar to the previous paper, the portfolio had higher average returns than portfolios based on historical estimates, it also performed better on average than the Regression-to-the-mean portfolios. Moreover, the Livingston portfolios had the highest Sharpe ratio of the three portfolios as well (0.71).

The adjusted forecasts of the Livingston predicted return estimates that were regressed to the mean no longer produced portfolios that had superior asset allocations to that of just expert predictions. While the regression to the mean portfolios still averaged higher returns than those based on historical estimates, the Sharpe ratio was substantially lower than that of the historical portfolios. A \$1 investment in December 2005 would have yielded \$5.01 by June 2018 in the

Livingston Survey Portfolios, \$3.83 in the Regression-to-the-mean portfolios, and \$2.56 in the historical estimate portfolios.

Conclusion

Historical return data is still used often as the basis of estimation when using mean-variance analysis. Moreover, a substitute for this method is to view these historical data and incorporate one's beliefs about asset expected return, standard deviation, and variance given the current economic environment. When allocating a portfolio, one might have information about the market that might be useful in selecting returns that are beyond what is observed in the historical data set. While this may be occasionally useful, even those deemed professional forecasters have difficulty understanding what future asset returns might look like. This finding is reasonable considering the tremendous macroeconomic factors, varying levels of idiosyncratic risk, and human spirits that impact individual stocks and the markets.

This paper ultimately found surprisingly conflicting results with that of Smith, Steinberg, and Wertheimer's previous paper on the topic and this might be due to several reasons. We used the predicted returns from the Livingston Survey and adjusted these predictions by regressing them to the historical average predicted returns. Similarly, we regressed the variance-covariance matrices to the mean as well and used these estimates to perform mean-variance analysis and create optimal risky portfolios.

In this case, however, these regression-to-the-mean portfolios failed to out-perform the Livingston Survey portfolios and failed to have a better risk-reward profile than portfolios based on historical data. The biggest driver of this seems to be the expert predictions of consistently negative bond returns. The Livingston Survey provided forecasts which made it very unlikely for

bonds to be incorporated into any optimal risky portfolio. Thus, while actual bond returns mostly outperformed the predicted estimates, most optimal expert portfolios were comprised of solely stocks, which during this 13-year period provided returns that were substantial enough to outweigh more diversified portfolios. Regression-adjusted portfolios still provided higher average returns than historical data portfolios but neither had a higher return or Sharpe ratio than that of the Livingston Survey portfolios.

References

- Cohen, Kalman J., and Pogue, Jerry A. (1967). An Empirical Evaluation of Alternative Portfolio Selection Models, *Journal of Business*, pp. 166–193.
- Markowitz, H. (1959). Portfolio selection: Efficient diversification of investments,”New York.
- Smith, Gary, Joseph Steinberg, and Robert Wertheimer. “The Next Best Thing to Knowing Someone Who Is Usually Right.” *The Journal of Wealth Management* 9, no. 3 (2006): 51–60.
- Smith, G. N. (2015). *The Art and Science of Investing*. CreateSpace Independent Publishing Platform.
- Sornette, D. (2017). PREDICTION OF BUBBLES, CRASHES, AND ANTIBUBBLES. In *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (pp. 320-354). Princeton; Oxford: Princeton University Press. doi:10.2307/j.ctt1h1htkg.13
- Tobin, James (1958). Liquidity Preference as Behaviour Towards Risk, *Review of Economic Studies*, 25, No. 67, pp. 65-86.
- Schall, Teddy, and Gary Smith. “Do Baseball Players Regress Toward the Mean?” *The American Statistician*, vol. 54, no. 4, 2000, pp. 231–235. *JSTOR*, www.jstor.org/stable/2685772. Accessed 17 Feb. 2020.
- Wikimedia Foundation. (2019, November 27). Efficient frontier. Retrieved March 7, 2020, from https://en.wikipedia.org/wiki/Efficient_frontier

Table 1 Predicted and Actual Returns

Forecast Date	Predicted Stock Return	Actual Stock Return	Predicted Bond Return	Actual Bond Return	Actual T-Bill Return
Dec 2005	3.92%	3.17%	-0.37%	-3.46%	2.11%
Jun 2006	3.36%	12.07%	1.78%	5.91%	2.52%
Dec 2006	3.46%	3.57%	0.76%	-0.04%	2.45%
Jun 2007	3.66%	1.84%	1.41%	10.26%	2.37%
Dec 2007	3.49%	-12.29%	0.40%	2.41%	1.69%
Jun 2008	5.48%	-27.58%	1.16%	16.80%	1.06%
Dec 2008	14.22%	10.43%	-1.34%	-9.13%	0.14%
Jun 2009	8.61%	14.06%	-0.78%	-0.76%	0.18%
Dec 2009	47.78%	-0.31%	-1.02%	9.16%	0.10%
Jun 2010	7.29%	15.21%	-1.46%	-1.19%	0.11%
Dec 2010	4.68%	3.85%	0.24%	2.63%	0.10%
Jun 2011	4.91%	-1.65%	-2.10%	12.76%	0.05%
Dec 2011	4.22%	10.89%	-1.05%	2.87%	0.03%
Jun 2012	4.90%	4.54%	-2.20%	-0.12%	0.08%
Dec 2012	3.24%	19.34%	-1.63%	-5.33%	0.06%
Jun 2013	3.05%	10.71%	-0.54%	-3.00%	0.05%
Dec 2013	2.91%	5.46%	-0.65%	5.80%	0.05%
Jun 2014	3.44%	7.67%	-2.01%	4.34%	0.04%
Dec 2014	2.66%	3.26%	-2.27%	-0.44%	0.06%
Jun 2015	3.18%	-1.82%	-1.92%	1.86%	0.06%
Dec 2015	2.99%	7.42%	-1.28%	8.02%	0.24%
Jun 2016	2.14%	4.05%	-1.84%	-7.35%	0.18%
Dec 2016	3.51%	11.35%	-0.82%	2.41%	0.31%
Jun 2017	3.20%	9.23%	-1.86%	0.39%	0.56%
Dec 2017	4.89%	6.44%	-1.51%	-2.52%	0.75%
Jun 2018	2.58%	-10.06%	-0.68%	2.76%	1.03%
Dec 2018	3.00%	19.98%	-0.43%	7.29%	1.25%
Jun 2019	4.21%	9.39%	-0.43%	1.69%	1.02%
Average	5.89%	5.01%	-0.80%	2.29%	0.78%

Table 2 Portfolio Returns

	Historical Data	Livingston Survey	Regress to Mean
Average Return	3.37%	6.03%	5.14%
Standard Deviation	4.01%	7.46%	9.13%
Sharpe Ratio	0.67	0.71	0.49

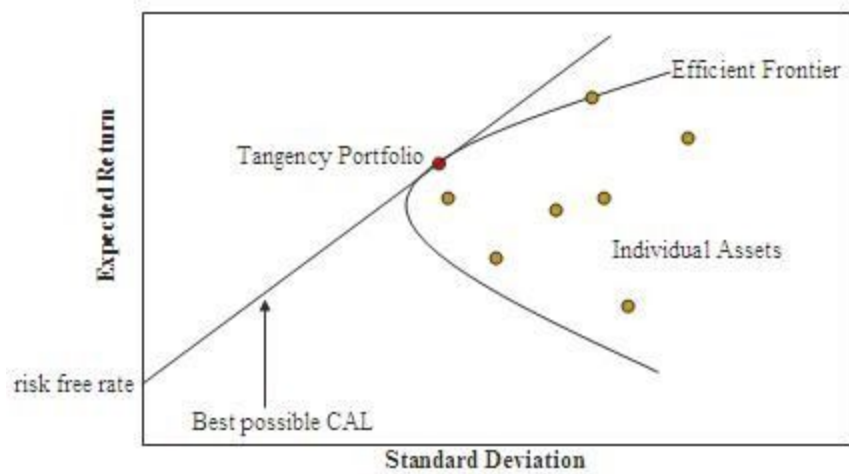


Figure 1 Optimal and Suboptimal Portfolios