OCTOBER 2008

Efficient Frontiers Constructed with Historical Data Can Be Misleading

JAMES L. DAVIS

When financial economists use terms like *ex ante* and *ex post*, the rest of the world usually tunes them out fairly quickly. This is too bad, because these are important concepts for investors.

Ex ante refers to the time before an event takes place, while *ex post* refers to the time after the event. Applied to investment returns, *ex ante* implies that investors have formed some expectation of what the returns will be. This expectation is typically expressed using the parameters of a probability distribution. The mean of the distribution represents the expected return, and the variance (or standard deviation, which is the square root of the variance) reflects the uncertainty associated with the ultimate outcome. The bigger the variance, the farther the actual outcome is likely to be from the expected value. In contrast, *ex post* refers to the case where the investment return has occurred and the investor knows the outcome.

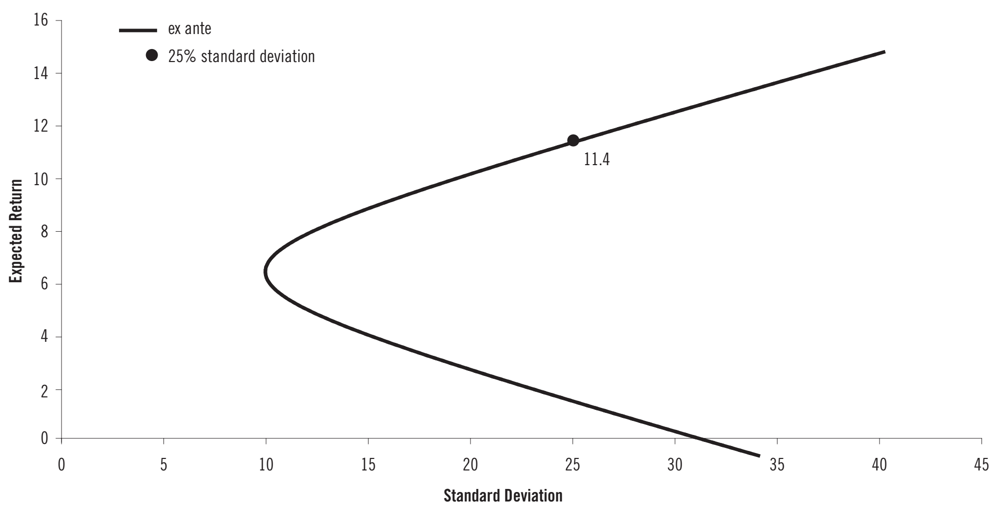
Drawing on the important work of Markowitz (1952), many investors are familiar with the concept of mean/variance efficiency. This concept is based on the assumption that investors would like to achieve a given expected return with as little uncertainty about the outcome as possible. Viewed another way, investors would like the highest

The helpful comments of Ken French, Inmoo Lee, and Weston Wellington are gratefully acknowledged.

possible expected return for a given level of return variation. Portfolios that achieve this objective are said to be mean/variance efficient. This idea can be expressed using a two-dimensional graph like Exhibit 1. Since investors prefer high expected returns and low standard deviations (or variances), they would like to go as far northwest as possible in this two-dimensional space. The available investments constrain how far they can go. The upward sloping part of the curve in Exhibit 1 shows a hypothetical *efficient frontier*. It shows the portfolios that are mean/variance efficient, since it sets the boundary for how far northwest investors can go. In this hypothetical example, Exhibit 1 shows that 11.4% is the highest expected return that can be achieved for a portfolio with a standard deviation of 25%.

Exhibit 1

Hypothetical Ex Ante Frontier



It is important to note that Exhibit 1 shows *expected* returns on the vertical axis. In other words, the efficient frontier is constructed *ex ante*. Investors want to form efficient portfolios before the fact, based on the probability distribution of returns. In spite of the clear *ex ante* nature of the efficient frontier, many investors like to construct charts like Exhibit 1 using historical returns. They turn an *ex ante* theoretical concept into an *ex post* tool for data analysis.

The main objective of this paper is to show that this exercise can be misleading. Specifically, constructing an efficient frontier using historical returns and then treating this frontier as the mean/variance tradeoff that is currently available to investors is likely to make investors feel a lot better off than they really are.

To see why, suppose there are six asset classes that are available for investment. These asset classes have the expected returns and standard deviations shown in the *ex ante* section of Table 1. Based on these parameters, and assuming the model for expected returns described in the appendix, the expected return/standard deviation frontier in Exhibit 1 is available to investors.

Now roll the clock forward 25 years and look at the returns for these six asset classes. Consider three different possibilities—market returns that were better than expected, close to expectations, and worse than expected. These three scenarios are labeled "good," "normal," and "bad" in Table 1. The details of these simulations are in the appendix. A good market is characterized by average returns that are generally higher than expectations, coupled with low standard deviations. A bad market displays the opposite combination.

Table 1

Summary Statistics for Hypothetical, Randomly Generated Asset Classes

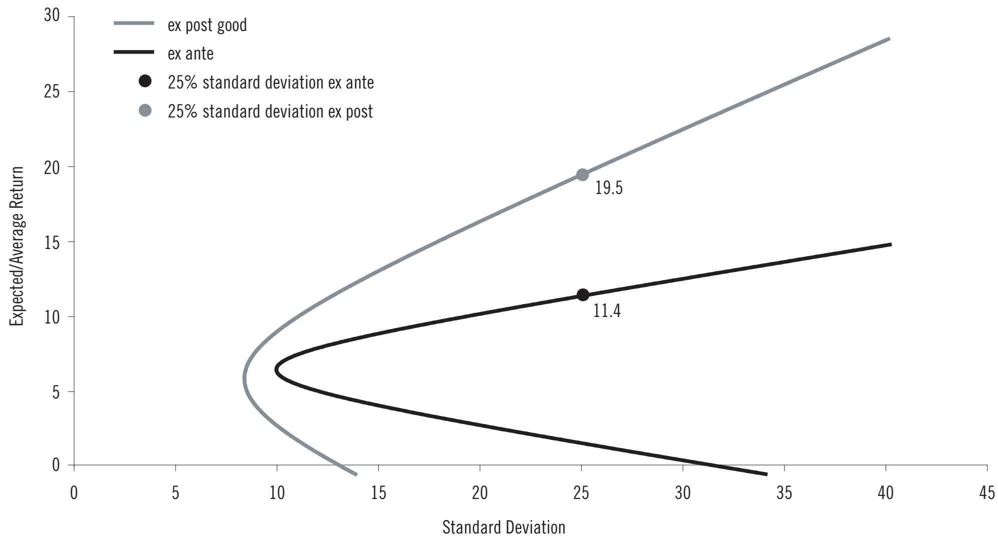
	Asset Class					
	1	2	3	4	5	6
Ex Ante						
Expected return	7.2	7.6	8.0	8.4	8.8	9.2
Standard deviation	11.3	12.4	13.4	14.5	15.6	16.7
Simulated Ex Post Returns						
Average Return						
Good	7.7	8.8	7.8	8.8	10.2	11.6
Normal	6.9	7.9	6.8	7.7	9.0	10.3
Bad	6.1	7.0	5.8	6.7	7.8	9.0
Standard deviation						
Good	10.8	11.4	11.8	12.3	13.1	16.6
Normal	11.5	12.2	12.7	13.4	14.2	17.8
Bad	12.1	12.9	13.6	14.3	15.3	18.9

How would frontiers constructed with these simulated historical returns compare to the *ex ante* frontier that was available to investors 25 years earlier? Exhibit 2 shows that the frontier constructed using good returns lies to the northwest of the *ex ante* frontier. This is probably not a surprise; higher averages combined with lower standard deviations should shift the frontier to the northwest. Exhibit 2 shows that the best portfolio with a standard deviation of 25% had a substantially higher average return than what was expected beforehand (19.5% vs. 11.4%).

Exhibit 2

Hypothetical Ex Ante and Ex Post Frontiers

Good Market Returns



What should we expect to see when frontiers are constructed using the normal and bad historical returns? Simple intuition might suggest that the normal frontier should lie close to the *ex ante* frontier, while the bad frontier might lie further to the southeast. In this case, simple intuition would be misleading. Exhibit 3 shows that the frontier constructed using normal returns lies to the northwest of the *ex ante* frontier. They are close near the bottom (i.e., near the minimum-variance portfolio), but the difference between the two grows as the standard deviation increases. Exhibit 4 shows a similar result even when historical returns are bad. Except for portfolios with relatively low standard deviations, the frontier constructed from simulated poor historical returns lies northwest of the frontier that investors expected at the outset. All three historical simulations suggest that an average return substantially higher than 11.4% was available for a standard deviation of 25%.

Why do we see this result? Levy and Roll (2008) provide the answer. The efficient frontier is the result of an optimization problem. It shows the set of portfolios that provide the best return/variability tradeoff, given the expected returns and covariances of the asset classes. A historical average return includes both the expected return and an average error—the part of the historical return that could not have been anticipated beforehand. While the *ex ante* efficient frontier optimizes with respect to expected returns, the *ex post* frontier optimizes with respect to both expected returns and errors. In other words, the *ex post* frontier treats the average errors as if they were expected. Compared to the *ex ante* frontier, portfolios on the *ex post* frontier tend to over-weight assets with positive errors and under-weight assets with negative errors.

Exhibit 3

Hypothetical Ex Ante and Ex Post Frontiers
Normal Market Returns

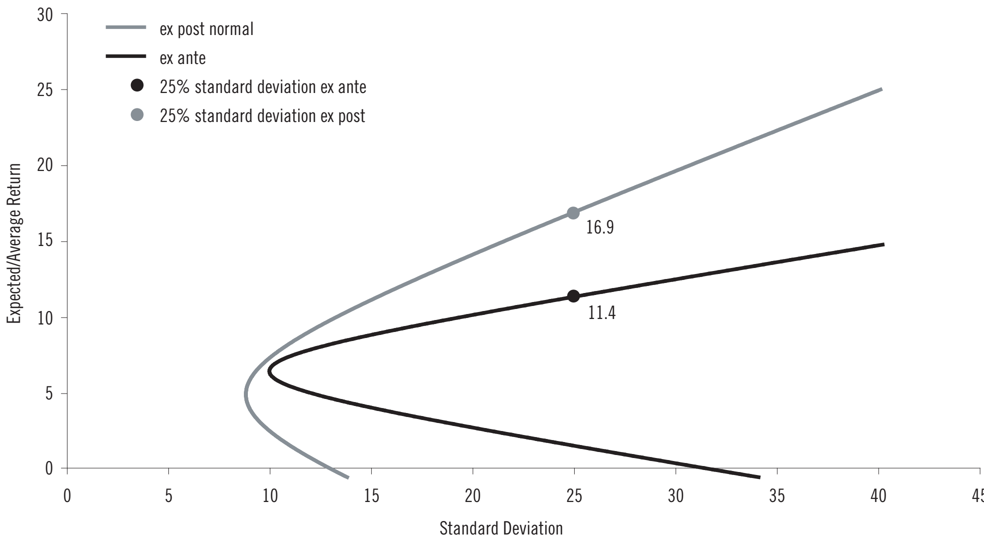
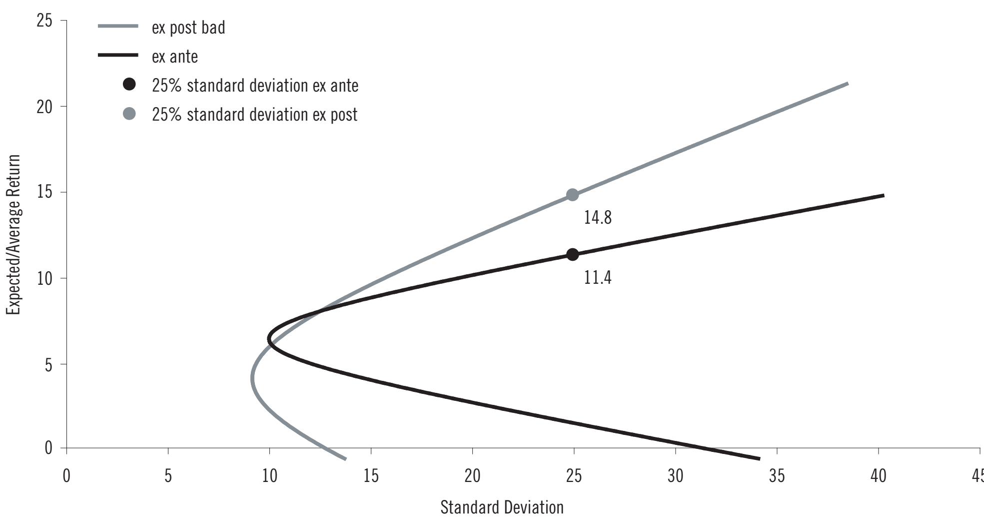


Exhibit 4

Hypothetical Ex Ante and Ex Post Frontiers
Bad Market Returns



This is what statisticians call a small sample problem. Because the sample has only 25 annual observations, some of the asset classes have average errors that are not close to zero, due to a few large (positive or negative) observations. These non-zero average errors give the optimizer more opportunity to push the boundary northwest. If the frontiers in Exhibits 2-4 had been constructed using 1,000 years instead of 25, this problem of an unrealistic *ex post* frontier would be much less severe, since the average errors would cluster around zero. In the real world, there aren't any 1,000-year data samples to work with, and there is no reason to believe that the distribution of returns would remain stable for that long anyway.[1]

There is nothing wrong with analyzing historical returns to see how assets would have combined into portfolios. There is something wrong, however, with treating the results of this exercise as the return/variability tradeoff that is available to investors on a forward-looking basis. To achieve the *ex post* return/variability tradeoff, investors would have to know the errors in advance. Exhibit 4 shows that even a bad market can produce an *ex post* frontier that gives investors an overly optimistic impression of the available investment opportunities.

What are the implications of these results? First, as emphasized by Clark (2004), investors should not treat the output from an optimizer as a reliable guide for portfolio allocations. When the inputs to the optimization problem are based on historical returns, the results are largely functions of random errors.

A second implication is that frontiers created using historical data should be treated as an upper bound on investment opportunities, not an unbiased estimate. Suppose a certain investor is about to retire, and she wants to run a Monte Carlo simulation to see how much of her savings can be spent each year. She assumes that the expected return/standard deviation combination from Exhibit 4 (14.8%/25%) is available, when the 11.4%/25% combination is what is really available. If she is willing to accept a 5% probability of running out of money within 30 years, the 14.8%/25% tradeoff would permit her to withdraw just over 4% of her savings next year, and a similar (nominal) dollar amount each year thereafter.[2] But since the tradeoff that is really available is 11.4%/25%, her probability of failure is actually close to 14%. She is not as well off as she believes, in spite of basing her projections on market returns that were below expectations.[3] Had she treated the 14.8%/25% tradeoff as an upper bound, she would have run additional simulations with different assumptions and would therefore have a better sense of her probability of failure. She would also have a better understanding of how sensitive these types of simulations can be to the underlying assumptions.

---

1. Using monthly returns instead of annual returns will not solve the problem. The positive and negative average errors for the sample will still be present.

2. This Monte Carlo result is a rough approximation, because it assumes returns are independently normally distributed. It also ignores inflation, taxes, other transfer payments, and other real-world complications. It further assumes only one deduction per year from savings at the end of the year, and it assumes that the portfolio allocation remains constant.

3. The situation is even worse if she bases her assumptions on good market returns. If her Monte Carlo simulations use the 19.5%/25% tradeoff from Exhibit 2, she will withdraw nearly 7% of her savings at the end of the first year. If she continues to withdraw this dollar amount every year, her probability of failure is in excess of 35%.

One way to attenuate the small sample problem is to use a larger sample. When the data are available, analysts should use longer samples, unless there is evidence that the distribution of returns has changed. Another way to attenuate the problem is to prohibit simulated portfolios from taking short positions. The frontiers in Exhibits 2-4 are pushed further northwest because portfolios are allowed to take short positions in asset classes with negative average errors. If short positions are not allowed, the optimizer's ability to take advantage of the sampling errors is reduced (but not eliminated).

If expected returns were observable, the easy solution to all this would be to use them in the construction of the mean/variance frontier, as prescribed by theory. Unfortunately, expected returns are not observable. While there are a number of models available for estimating expected returns (e.g., Fama and French 1993), the estimates will contain errors, because they are estimates. Optimizers will optimize with respect to these errors, just as they do with the errors in historical returns. Models of expected returns therefore do not provide the antidote for optimizers' lack of reliability. Nevertheless, using the expected returns from a well-specified model is probably a better approach than simply plugging in small-sample historical averages.

One other shortcoming of mean/variance optimization deserves mention. Two-dimensional charts like Exhibits 1-4 assume that expected returns and return variability are all that matter to investors. This implies that portfolio standard deviation is a sufficient measure of risk. In fact, there are good reasons to believe that many investors care about other sources of risk in addition to standard deviation. This belief is supported by both theoretical (e.g., Merton 1973) and empirical (e.g., Fama and French 1993) research in capital markets. As a result, optimizers may suffer from an even more basic problem than what has been discussed so far. The simulations in this study show that an optimizer may not provide a precise solution to the problem that it is given to solve. The additional complication is that a mean/variance optimizer may not be solving the right problem. It may be providing an imprecise answer to the wrong question.

Appendix

To keep the simulations as simple as possible, the expected returns on the six asset classes are assumed to be a function of a single pricing factor F. The return for asset class i in year t can be expressed as:

$$R_{it} = R_f + \beta_i F_t + e_{it}, \quad i = 1, 2, \ldots, 6$$

The risk-free rate R_f is fixed at 4%, and the pricing factor F has an expected value of 4% and a standard deviation of 12%. The sensitivity of asset class i to the pricing factor is measured by β_i. The sensitivities are assumed to have values $[\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6]$ = [0.8, 0.9, 1.0, 1.1, 1.2, 1.3]. The errors e_i each have a mean of zero, a standard deviation of 6%, and are mutually independent. The expected returns and standard deviations in the *ex ante* section of Table 1 are based on these assumptions.

For the simulated return histories, six series of 25 annual asset class errors are generated at random assuming a mean of zero and a standard deviation of 6%. The good, normal, and bad markets are differentiated by the assumptions used to generate random values for F_t. In the good market simulation, 25 observations of F are generated assuming a mean of 5% and a standard deviation of 11%. The corresponding (mean/standard deviation) assumptions for the normal and bad markets are (4%/12%) and (3%/13%), respectively. While both average returns and standard deviations are assumed to be affected by market conditions, the results shown in the Exhibits still hold if the standard deviations are held constant across the three scenarios. It is the presence of non-zero average errors that has the greatest impact on the location of the *ex post* frontiers.

References

Clark, Truman A. 2004. Stop playing with your optimizer. Working paper, Dimensional Fund Advisors.

Fama, Eugene F., and Kenneth R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3-56.

Levy, Moshe, and Richard Roll. 2008. The market portfolio may be mean-variance efficient after all. Working paper, Hebrew University and UCLA.

Markowitz, Harry. 1952. Portfolio selection. *Journal of Finance* 7:77-91.

Merton, Robert C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41:867-887.