Bayes to the rescue of Markowitz

*In this research letter, we turn our attention back to the **Finaltis EfficientBeta**TM strategy and in particular to our proprietary method for estimating the volatility and correlation.*

AUTHORS:

RÉMY CROISILLE
Head of Research
Senior Fund Manager

CHRISTOPHE OLIVIER
CIO

NICOLAS RENAUD
Head of Risk

LEWIS MEREDITH SMITH, CFA
Senior Fund Manager

IN BRIEF

1. Modern Portfolio Theory, as proposed by Harry Markowitz, is mathematically crystal clear. However, putting theory into practice presents many pitfalls and often ends in disappointing results.

2. One of the problems we face is estimating the parameters (returns, volatility and correlation) for the future, while only observations of the past are available. Practitioners resort to replacing the parameters by their estimators, which deteriorates the quality of the results, especially when the estimators are in fact highly unpredictable.

3. Bayes' theory allows for revising estimations of parameters based on observations and the probability laws of these observations: it therefore seems well suited to overcoming the pitfalls when putting Modern Portfolio Theory into practice.

4. The *Finaltis EfficientBeta*TM method uses Bayes' theorem to obtain an estimation of volatility which is different to the ubiquitous standard deviation. It incorporates not only the individual information of each stock but also collective information of observations of all volatilities. Additionally, it includes, through its probabilistic approach, the uncertainty inherent in the estimator.

SAVING PRIVATE MARKOWITZ

A large number of finance practitioners use a form of Modern Portfolio Theory proposed by Harry Markowitz [1952] especially in the form of Sharpe's Capital Asset Pricing Model [1964]. In the model, asset returns are modelled as random variables, with an expectation and variance. It follows that a portfolio is considered to be a linear combination of such assets. Therefore, the return of a portfolio is also a random variable with its own expectation and variance. We can deduce the portfolio parameters from those of its components:

$$\mathbb{E}[R_p] = \sum_i w_i * \mathbb{E}[R_i] \qquad \text{and} \qquad \sigma_p^2 = \sum_i \sum_j w_i * w_j * \sigma_i * \sigma_j * \rho_{ij}$$

Where:

- R_p is the portfolio return ;
- w_i the weight of asset *i* in the portfolio
- σ_p^2 its variance (and σ_p its volatility) ;
- R_i the return of asset i;
- σ_i^2 the variance of asset *i* (and σ_i its volatility);
- ρ_{ij} the correlation or the returns of assets *i* and *j*.

Under the joint hypothesis of market efficiency and investors rational risk aversion as estimated by the variance, one can calculate an optimal portfolio using quadratic optimisation which maximises the return/volatility ratio (the latter being the square root of the variance).

The mathematical theory is crystal clear and widely used in finance today. However, before receiving the Nobel Prize in 1990, Markowitz almost had his theory refused in 1955; Milton Friedman had judged that his work did not classify as economics…

In fact, putting the theory into practice is riddled with pitfalls and often produces disappointing results. One of the most notable of these is the uncertainty in the estimation of the parameters (return, volatility and correlation). In order to determine a portfolio of *n assets*, we would need to know the exact n returns and variances of the underlying assets and also the $^{n(n+1)}/_2$ correlations. While we need to estimate future parameters, we only have access to past observations. ***The practitioner is forced to replace the parameters by their estimations and that is by no means benign regarding the quality of the results, all the more so when the estimator is very uncertain.***

Using the analogy of tossing of a fair coin to produce a *heads* or a *tails*, to suggest that a stock's volatility parameter is 15% because its 65 day annualised historical volatility is 15% is like saying that the expectation of a fair coin producing *heads* is 55% because in 100 trial coin tosses, heads came up 55 times. It goes without saying that to use 55% instead of the true parameter 50% to predict future coin tosses would ruin the gambler; if we were to propose, win 90€ for every heads and lose 100€ for every tails :

- He would believe his chances of winning to be *55%*90€* + *45%*(-100€)* = *4.5€* per coin toss…
- … whereas he would in fact lose *50%*90€* + *50%*(-100€)* = *-5€* per coin toss.

Pursuing the same analogy… Suppose we wish to select the coin or coins which have the greatest probability of producing a heads in a future coin toss chosen from 3 coins:

- Two are fair coins (real parameter 50%) ;
- One is biased, which causes it to produce *heads* 45% of the time.

If during a test phase, the biased coin produced a score of 52% while the two fair coins produced 51% and 47% we would select the biased coin if we were to put all our faith in this one single experiment.

The beauty of the coin toss is our ability to repeat the experiment. As long as we can repeat he experiment, the estimator will converge towards its true parameter. On the other hand, equity investors do not have the luxury of being able to repeat experiments. They are limited to one single historical observation. As such, they are forced to accept the uncertainty in estimating the parameters.

The most visible consequence of the ***pitfalls of putting Markowitz' Modern Portfolio Theory into practice is the extreme concentration of the resulting portfolios*** obtained by quadratic optimisation and the variance/covariance matrices estimated using the classic estimators. The theoretical problem is more complex than our coin toss, because there are, if we focus on the denominator, two parameters - the variance and the correlation. That said, in practice the problem is the same: we tend to select a small number of stocks for which the recent observations of the parameters are better than the rest, which in no way guarantees that their real parameters are different or better than the others.

When the practitioner is confronted with the observed concentration problem, a quick remedy can be found by artificially increasing the diversification by adding constraints such as:

- A maximum exposure per stock, per sector or per country ;
- A rule which decreases the weight of each stock.

We have noticed that the result of such methods, which treat the symptoms rather than the cause, are determined by the constraints as opposed to the optimisation.

An alternative solution is to make additional hypotheses such as:

- All correlations are equal to zero; all the returns are equal, which leads to the *Risk Parity* approach;
- The returns are proportional to their volatilities, which leads to the maximum diversification approach.

The question that we must therefore ask is: are these hypotheses coherent with the reality of our observations? This we do <u>not</u> believe to be the case.

The Finaltis EfficientBeta™ method offers a third solution for saving Private Markowitz.

REVEREND BAYES' RECIPE

Not much is known about the Reverend Thomas Bayes (1702-1761), apart from the fact that he was made a vicar near London in 1742 and he published just two papers in his lifetime, one of these was theological and deals with the need to know one's weaknesses in order to advance on the journey to salvation. We are not even sure that the only remaining portrait of him is authentic. His second work on conditional probability, which earnt him posterity, was published after his death by one of his friends and later validated by Laplace. However, he did leave his name to one of the fundamental branches of probability theory.

In its simple form, Bayes' Theorem states:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

Where, for two events *A* and *B*:

- *P(A)* is the probability that event *A* occurs;
- *P(B)* is the probability that event *B* occurs;
- *P(B|A)* is the probability that event *B* occurs given that A has occurred and *P(A|B)* is the probability that event *A* occurs given that B has occurred.

By giving a formula to simply calculate conditional probabilities, Bayes' theorem allows us to revise our estimation of a probability based on new information. This opportunity to reduce uncertainty from our learning experience can by assimilated with his theological approach to the journey to salvation by knowing one's weaknesses.
We now look at Bayes' theorem when applied to the paradox of Monty Hall.

HOW BAYES HELPS YOU TO WIN A TV GAMESHOW

Monty Hall was the host of an American TV gameshow *Let's Make a Deal*. The Monty Hall paradox, a sort of probability brain teaser, is loosely based on the strategy game opposing the gameshow host and a member of the public. The contestant is faced with three closed doors; two of them hide a goat and the third a car. The contestant starts by picking a door. The host, who knows where the car is, then opens another door, behind which there is always a goat. Assuming the contestant wishes to win the car (and not a goat), he asks his guest if he would like to change door or stick to the original one. What should the candidate do to optimise his/her chances of winning: switch or not?

When this problem is posed and a quick answer required, there are generally two groups of respondents :

- Those who say that two doors remain, each having the same probability. So, it is pointless to change.
- Those who say that at the outset there was a 1 in 3 chance of winning and by sticking with the first choice, one still has a 1 in 3 chance. So, by changing one must have a 2 in 3 chance; it is best to switch.

It is in fact a classic conditional probability problem, which is easy to solve with Bayes' theorem. Let's suppose, for this demonstration, that the contestant had chosen door number 1 and that the host had then opened door number 3.

By noting :

- *$P(C_2)$* the probability that the car is behind door 2 ;
- *$P(D_3)$* the probability that the host opens door 3 ;
- *$P(C_2 | D_3)$* the probability that the car is behind door 2 given that the host opened door 3 ;
- *$P(D_3 | C_2)$* the probability that the host opens door 3 given that the car is behind door 2.

Then Bayesians will assert:

$$P(C_2|D_3) = \frac{P(C_2)P(D_3|C_2)}{P(D_3)} = \frac{1/_3 * 1/_1}{1/_2} = 2/_3$$

Of course, the host cannot chose door 1 because the contestant has chosen it. If he didn't know that the car was behind door 2, he has two choices (probability = $^1/_2$). If he does know where the car is, then he has only one choice (probability = $^1/_1$).

By opening a second door, which does not contain the car, we obtain new information. By correctly taking this into account using conditional probability, we significantly improve our chances of winning. The contestant should change door in order to double his expected gain.

BAYES WORSHIPERS

Bayes' theorem is not only useful when playing televised gameshows. It is widely used in what is called Bayesian inference to revise estimations of a probability or a parameter based on observations and probability laws.

The statistician Nate Silver, famous for correctly predicting the winner of the 2008 US presidential election in 49 out of 50 states, has devoted an entire book *The Signal and the Noise: Why So Many Predictions Fail – but Some Don't* to the Bayesian probabilistic approach, the book has received much acclaim and was named Amazon's No. 1 best non-fiction book for 2012; it has been translated into 10 languages (but not French!). The book lists several applications of Bayes' theorem.

- *Poker*: the most popular variant, the *Texas Hold'em*, is the perfect playground for fans of Bayes. A first round of betting takes place after the first two cards are dealt, then again after turning over three more cards (*the flop*), and again after dealing the next (*the river*) and finally the last (*the turn*). At the outset, your opponent can have, just as you can, one of a possible 1326 hands. Each event (cards appearing, bets) reveals new information about possible hands of your opponent. The best players are capable of improving their predictions of opponents' hands based on the information, until the field of possible outcomes is reduced to a small possible number and allows them to rationalise their betting. With a good understanding of conditional probability, a well informed player can considerably improve his chances of winning. However, his available starting budget may not allow him to play enough hands for his probabilistic advantage to profit before being ruined. And this does not account for the fact that the opponent will do everything to hide his hand's strength by bluffing …
- *The US presidential elections*: A large part of Nate Silver's 2008 success and consequently his reputation relies on the updating process of his initial probabilities based on newly available results.

- <u>*Betting on Sport*</u>: Nate Silver built a program of statistical predictions relating to the results and performance of base-ball payers named *PECOTA* (*Player Empirical Comparison and Optimization Test Algorithm*). The author refers to Bayes not just for his probabilistic vision of each prediction, but also by his continuous updating of initial predictions based on new events. He also refers to the success of his colleague Bob Voulgaris and his successful betting on basketball. Bob Voulgaris would bet on the total number of points scored in the *Cleveland Cavaliers* matches based on whether or not Rick Davis was playing. It just so happens, that this player was more interested in scoring than defending.

Bayesian inference proves to be a powerful tool for revising predictions for its proponents.

BACK TO MARKOWITZ AND FINALTIS EFFICIENTBETA™

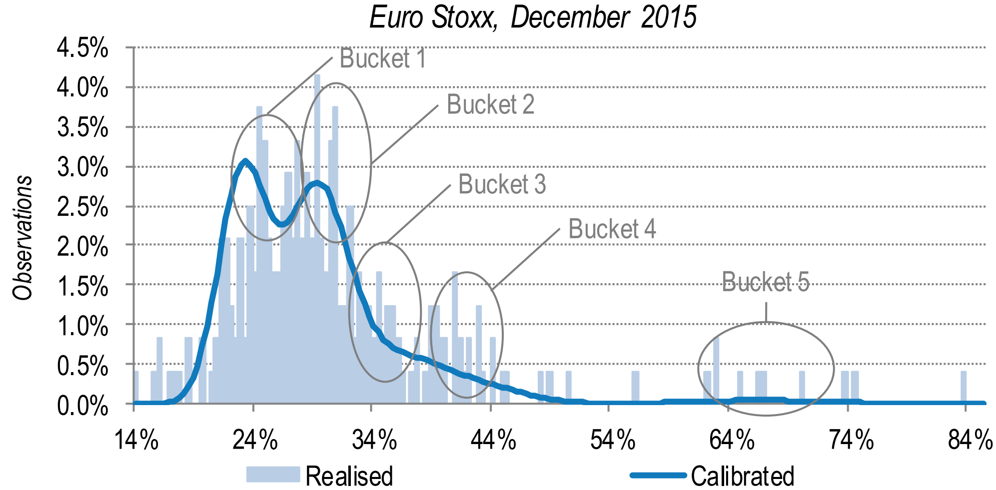
So, how does *Finaltis EfficientBeta™* use Bayesian probabilities to improve its estimations of the volatility and correlation parameters, and to obtain better optimal portfolios without adding artificial constraints? What is the additional information which allows us to revise our estimation?

Instead of starting from the theory, we proceeded by making a methodical analysis of the quarterly volatilities and correlations for all stocks across a period of 20 years. In the case of volatility for example, we observed, for each quarter in the study, a volatility histogram showing between 4 and 5 humps corresponding to volatility levels where there were a large number of stocks.

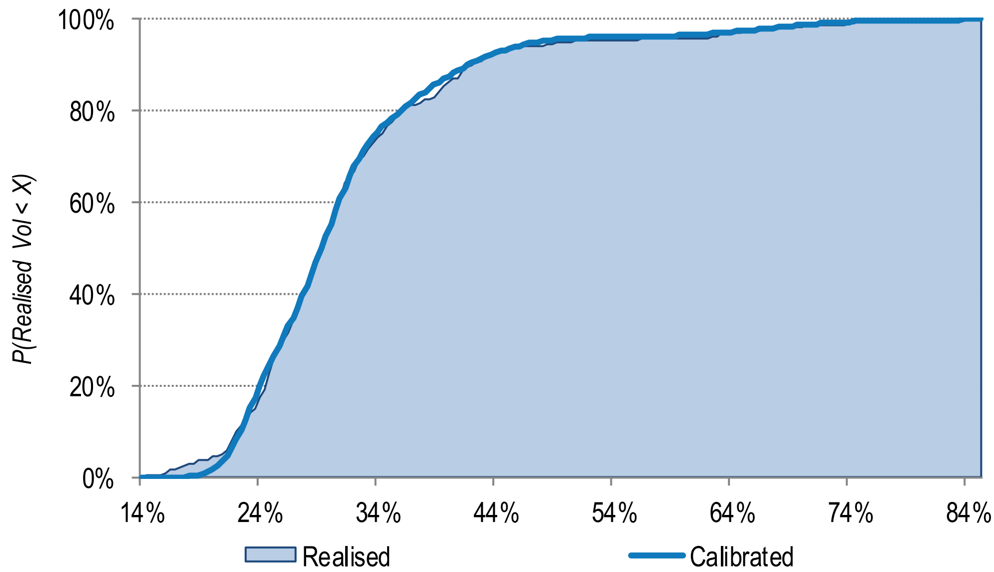
As such we have attempted to model the universe of stocks as having just five distinct levels of volatility; which translates, in mathematical terms, to the fact that the observed density function can be correctly approximated by a linear combination of 5 chi squared laws.

The following histogram illustrates the realised volatilities across the *Euro Stoxx* universe in December 2015 (date of the latest rebalancing of *Finaltis EfficientBeta™ Euro*). Although the histogram does not seem to perfectly represent the theoretical line (calibrated), examination of the cumulative probability in the second graph shows the model to be relevant.

Histogram of realised volatility
Euro Stoxx, December 2015

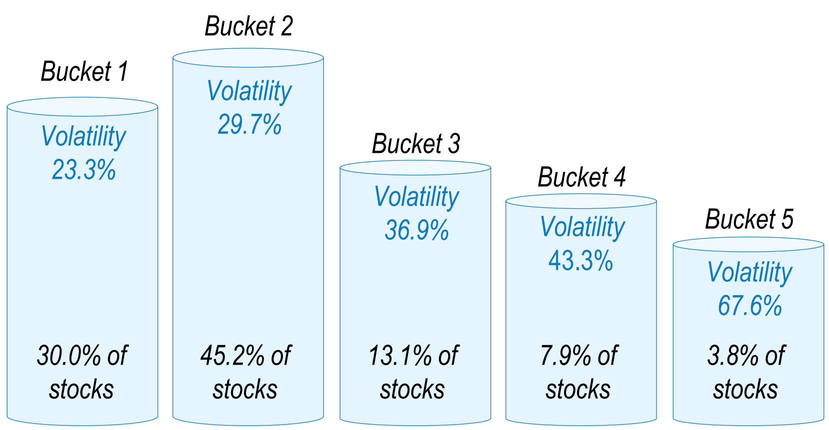


Distribution of realised volatility



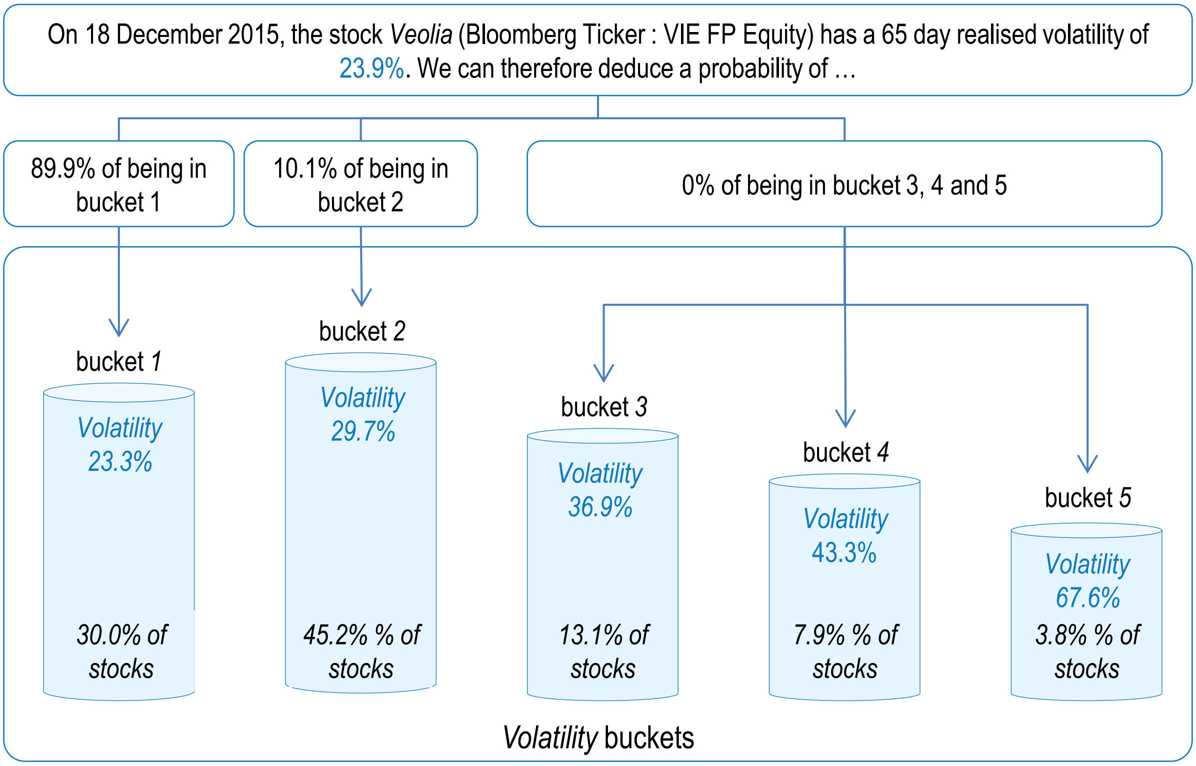
Source: Finaltis, data as of 18/12/2015

In December 2015, the interpretation of collective information, that is to say the observation of realised volatilities allows us to define volatility buckets (below) with their associated probabilities.



Source: Finaltis, data as of 18/12/2015

The volatility levels and corresponding probabilities of the five buckets change each quarter, but the underlying 5 bucket model remains valid. Returning to our analogy of heads and tails, it is as if we are making the hypothesis that there are two types of coins, those that are perfectly fair and those that are biased towards heads, and giving a probability to each group. One can now easily imagine conducting trials with the coins, to isolate the biased coins and eliminate them from our selection. We therefore find ourselves in the classic Bayesian inference framework where we can modify, based on our observations, the *a priori* probabilities of belonging to each group. The same goes for our stocks: the level of realised volatility will, by applying Bayesian probabilities, modify the probability of each stock belonging to each bucket. If a stock were to have a realised volatility of 23.9%, which is close to the bucket 1 volatility, the probabilities would adjust as follows:



Source : Finaltis, data as of 18/12/2015

From this we can deduce our volatility estimation, which differs from the standard deviation, which takes into account not just the individual information of each stock but also collective information of our observations of all stocks and integrates, through a probabilistic calculation, the uncertain nature of the estimator. Thanks, in part, to this revised estimator of volatility inspired by Bayesian probabilities, and a similar estimator of correlation, and a method inspired from Markov chains to model migration between buckets and an optimisation algorithm, the *Finaltis EfficientBeta™* method is able, without adding artificial constraints, to generate optimal portfolios, which offer return/risk properties which are far superior to those of the benchmark index.

Thanks Reverend Bayes !

BIBLIOGRAPHY

1. Bayes Thomas, Price Richard (1763) "*An Essay towards solving a Problem in the Doctrine of Chance*", Letter to John Canton

2. Croisille Rémy, Renaud Nicolas, Meredith Smith Lewis, Olivier Christophe (2013), "*Minimum Variance Portfolio in a Bayesian Context*", Finaltis

3. Markowitz Harry (1952) "*Portfolio Selection*", Journal of Finance, 7 (1), 77-91

4. Sharpe William F. (1964) "*Capital asset prices: A theory of market equilibrium under conditions of risk*", Journal of Finance, 19 (3), 425-442

5. Silver Nate (2012) "*The Signal and the Noise: Why so many Predictions Fail – but Some Don't*", Penguin Press

CONTACTS

63, Avenue des Champs Elysées
75008 Paris – France
www.finaltis.com

Denis Beaudoin	**Thierry Rigoulet**	**Yannick Robert**	**Mark Grobien**
Tel: +33 (0)1 55 27 27 01	Tel: +33 (0)1 55 27 27 07	Tel: +33 (0)1 55 27 27 99	Tel: +33 (0)1 55 27 27 08
dbeaudoin@finaltis.com	trigoulet@finaltis.com	yrobert@finaltis.com	mgrobien@finaltis.com

DISCLAIMER