TOWERS WATSON

The irreversibility of time

Or why you should not listen to financial economists

In this paper we continue to explore the subject of risk management. Our recent publication, 'The wrong type of snow – risk revisited'[1], took a high level overview of the subject. It defined risk as permanent impairment to an investor's mission and suggested that improved risk management required (1) a better risk framework, (2) better risk governance and (3) better tools. This paper has a much narrower scope and should be thought of as a contribution to the 'better tools' category. However, we believe it also sheds a useful light on the issue of 'permanent impairment to mission'.

We only live once, or at least we only live one life at a time[2], so why does finance and economics assume we have infinite lives all running in parallel? In this paper we draw on the recent work of Ole Peters and attempt to explain the difference between an 'ensemble average' and a 'time average'. These can be thought of as alternative terms for 'arithmetic mean' and 'geometric mean'[3] and so some readers may already be braced for a complicated statistical debate. However we believe the debate can be considerably simplified by invoking a rock-solid physical law. Our hope is that this paper brings clarity to a potentially nerdy subject and, at the margin, provides a positive contribution to our understanding and management of risk.

Take a gamble

To illustrate the point, consider the following gamble. You will roll a fair die, and if you roll a six I will pay you 10 times your current wealth. This is a thought experiment so we will gloss over my ability to pay – assume my credit is pristine. Imagine how much better your life would be if you were 11 times richer in the time it took to roll a die. The downside, paltry in comparison, is that if you roll any other number you will pay me your entire wealth – house, pension pot, pot plants, the lot. Will you take the gamble?

The way we have been trained to analyse the gamble means that we will consider all the possible future outcomes and then weight them in accordance with their probability. In effect we freeze time and take multiple copies of the world and then run the six versions forward as 'parallel universes'. In one of those worlds a one is rolled and we lose all our wealth. In the second a two is rolled with the same result. In the sixth world a six is rolled and we hit the jackpot and are paid 10 times our wealth. Having exhausted all the possibilities we travel back in time to the present and do our sums. The expected return of the gamble is the ensemble average – the average of all the possible independent outcomes. In this case the expected return is 83%[4] and so we would be 'crazy' not to take it.

So would you take the gamble? If you are like every other human we have met, the answer is "No". Instinctively, something does not feel right. Either you do not trust my credit, or the ensemble average (expected return or expected value) is misleading in some way. We know we should be attracted by such a high positive return, but risking our entire wealth puts us off.

1 Towers Watson, February 2012.

2 The author does not believe in reincarnation but respects the views of those who do.

3 Strictly, the time average will only be the geometric mean where the 'dynamic' between successive periods of time involves multiplication. If, instead, the dynamic involved addition then the time average would equal the arithmetic mean, and hence the ensemble average. In this latter case the system would be 'ergodic' – where the ensemble and time averages are the same.

4 The return is our final wealth ((5/6 x \$0)+(1/6 x \$11W) = \$1.83W) less the cost of the wager (\$1W), divided by our initial wealth (\$1W) which equals 0.83, or 83%.

We have built ourselves a paradox

In the thought experiment above, we have built our own version of the St Petersburg paradox which was put forward by Nicolaus Bernoulli in 1713. The paradox describes a lottery where a fair coin is tossed. If a head occurs on the first toss, the lottery pays out $1 and the game ends. If a tail occurs the coin is tossed again. If a head occurs on the second toss the lottery pays out $2 and the game ends, but continues to the third toss if a tail occurs. The payout of the lottery continues to double with each round until a head occurs which triggers the payout. The question is, how much should someone be willing to pay for a ticket which allows them to take part in this lottery?

To work out the expected value (ensemble average) of this lottery we again need to freeze time and create our parallel universes only, this time, we will need considerably more of them as there is a chance we will toss 100 tails in a row, and an even smaller chance we will toss 1,000 in a row (and so on). It turns out there is a chance the payout could be as large (or larger) than the entire wealth in the universe – it is just very unlikely. In non-technical terms the expected value of the payout from this lottery is very large indeed[5]. So N Bernoulli argued that a rational person should be willing to pay any price for a ticket in this lottery. The paradox arises because, in practice, people are rarely willing to pay more than $10 to take part.

Figure 01. Comparison of ensemble average and time average

Ensemble average	Time average
Arithmetic mean	Geometric mean*
Simple average of possible outcomes	Average rate of growth per time period
Outcomes considered in parallel	Outcomes considered in series
Implicitly or explicitly assumes parallel universes and/or time travel	Explicitly excludes parallel universes and time travel
For example, ask 100 people to roll their own die once	For example, ask one person to roll their die 100 times in sequence**

Note: We are comparing the two averages assuming a 'closed fund', for example a person's entire wealth which cannot be topped-up. Where new cash flows can be made into the 'fund' the conclusions noted in this paper will cease to hold.

* under a dynamic of multiplication between periods.

** the dynamic here would be additive so we would generate the same result as the ensemble.

5 In technical terms the expected payout is a 'diverging sum' which means it gets bigger the more rounds you average over, eventually getting to infinity.

How to resolve the paradox

1. Criticise the lottery

One way to get rid of a problem is to pretend it does not exist. As the expected payout is infinite (a) there are not enough goods and services in the world to pay it and (b) no one would offer the lottery because they would carry an infinite expected loss. The lottery is therefore unrealistic and irrelevant. Fair enough, but that doesn't exactly resolve anything.

2. Use the concept of utility

Another Bernoulli (Daniel, writing in 1738) was able to resolve the paradox by applying the concept of 'utility'. Under this framework there is no longer an objective test of the desirability of the lottery (the expected extra wealth) but instead there is a relative assessment given personal circumstances. Now the desirability of the lottery is judged by the usefulness of the wealth it produces. The more wealth you start with, the less utility there is in gaining more. D Bernoulli proposed that utility was logarithmic, and in so doing resolved the paradox[6]. As infinite wealth has little utility when you are already wealthy, there is no need to pay excessively for the chance to obtain it.

3. Abandon ensemble averages in favour of time averages

The above resolution of the paradox relies on the invention of a function, utility, that cannot be derived from fundamental considerations. Let us go back to our opening thought experiment which had an expected return of 83%. We noted that one reason for not taking the gamble might be because the ensemble average was misleading in some way. So let us consider the time average instead.

Instead of rolling the die once in each of six parallel universes, we will stay in our familiar universe and roll the same die six times in succession. Given that we have to risk all of our wealth for each roll of the die, it doesn't take too long to work out that the only way of walking away with any money is to roll six sixes in a row. Any other number at any point causes us to lose our entire wealth. But we are running ahead of ourselves. We compute the time average by taking each of the six possible outcomes and making them occur one after the other in our single, real, universe. We now compound our returns over the six periods and take the sixth-root to calculate our per-period expected (time average) return. It does not matter what order we roll each of the numbers one to six, we will lose all our wealth and so the time average is negative, and in a big way.

So the ensemble average **is** misleading. The 83% expected return unhelpfully disguises the large likelihood that we lose everything because of the small probability that we become fantastically wealthy.

In using the time average to solve the St Petersburg paradox, Ole Peters has (a) rejected the notion of parallel universes and (b) introduced the notion that we cannot go backwards in time (once we have lost everything we cannot go back and try again). This is more realistic. It also turns out that when you calculate the time average for the St Petersburg lottery it is mathematically identical to the logarithmic utility. So the paradox is resolved – but without the need to assume parallel universes, or to postulate a utility function. In this case the time average is (a lot) lower than the ensemble average. In fact, as a general rule, the time average will always be equal to, or lower than, the ensemble average – never greater.

So what?

We believe that a subtle but important shift would be beneficial. The way we are trained to think about investment and risk management is largely based on the ensemble average framework, with all that that implies – parallel universes, time travel and arbitrary utility functions. Where the ensemble average and time average are the same, or very close, then arguably this does not matter. But where they are different, this framework is wrong, and potentially dangerously so. The time average is better, because it shows an investor the return they are likely to get absent an ability to jump into a different parallel universe, or to rewind time and try again. In reality we only get to walk (invest/manage our wealth) down one path, and so the only thing that matters is what happens on that path – not what could have happened in some parallel universe.

So how different are ensemble averages and time averages in reality? In practice, in investment, the difference between the averages is usually small, but the difference grows and becomes important as volatility (the size of fluctuations) increases and as non-linearities enter the picture (such as through options). In other words, the time average becomes the better guide as risk increases. For example, when leverage within the system rises, when interest rates are driven lower, when correlations rise inhibiting diversification, and when incentive structures promote excessive risk taking (for example, annual bonuses with no claw back). If utility functions are not sufficiently restrictive regarding risk taking then ensemble averages will lead to excessive risk taking and eventual collapse – as in the global financial crisis.

6 A good proposal. Logarithmic utility is what is required for multiplicative dynamics.

Better risk management

1. Risk-return measure

As a consequence, in more extreme-risk cases, using time averages will allow us to manage risk better. To illustrate this consider a very simple coin-tossing example. It is a fair coin so there is a 50% chance of heads which will pay us a return of +10% on our stake (S), and a 50% chance of tails that will pay us a return of -10%. The ensemble average, or expected return, assumes these are independent events and averages across them, to give an expected value of 0%.

$$E(r) = \frac{(50\% \times 1.1S) + (50\% \times 0.9S) - S}{S} = 0\%$$

As the expected outcome is neutral we would need some other guidance to decide whether to take this bet, and this is where risk aversion or utility functions are brought in. If you are a risk seeker you may well take the gamble; if risk averse, not.

As discussed, the time average does not consider the two outcomes as independent (they are two sides of the same coin). So the time average assumes they occur one after the other over two periods, multiplies the returns and then takes the square root to give the expected per-period return:

$$T(r) = \sqrt{((1 + 0.1) \times (1 + (-0.1))} - 1 = -1\%$$

The time average shows that if we were to play this game repeatedly we would see our wealth erode gradually through time [7] – the outcome is not neutral. Notice that we do not need to know anything about our risk aversion/utility function. The time average gives sufficient guidance on its own.

We can see this even more clearly if we change the size of the payout for the coin toss. **Figure 02** shows that the ensemble average does not change as the size of the payout changes, so we have to include the standard deviation to show that the riskiness of the bet is increasing. Conversely the time average does change with the size of the payout. It consistently warns us against taking this bet, and gets progressively more 'vocal' in its warning as the riskiness of the bet increases. Note how the time average is particularly sensitive to an entire loss of wealth – the point of no return where the ultimate outcome cannot be rescued. The time average is acting as a risk-return measure in a single number.

2. Position sizing (optimisation)

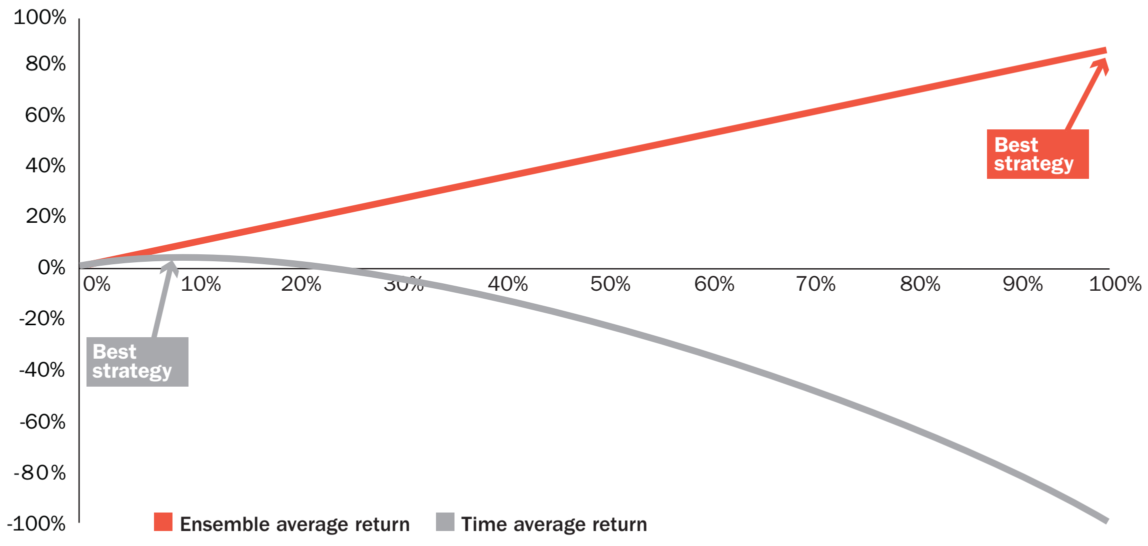
We will now show how the time average can be used to manage risk by adjusting how much we should wager ('position sizing' in the jargon). In our opening thought experiment we insisted that the wager had to be the player's entire wealth whereas in **Figure 03** we show the different returns for different wager sizes.

Figure 02. The time average return is a risk-return measure

Payout – heads	Payout – tails	Ensemble average	Standard deviation	Time average
10%	-10%	0%	10%	-1%
20%	-20%	0%	20%	-2%
30%	-30%	0%	30%	-5%
40%	-40%	0%	40%	-8%
50%	-50%	0%	50%	-13%
60%	-60%	0%	60%	-20%
70%	-70%	0%	70%	-29%
80%	-80%	0%	80%	-40%
90%	-90%	0%	90%	-56%
100%	-100%	0%	100%	-100%

7 We have been advised that some people, anchored in ensemble averages, will find this result hard to accept and so warrants further explanation. Consider betting \$1 on the coin toss and playing for two rounds. Because it is a fair coin, one round will be heads and the other tails. If we get heads first, our \$1 becomes \$1.10 (a 10% gain). In the second round we get tails, a 10% loss. 10% of \$1.10 is \$0.11, so we end up with \$0.99. Alternatively, if we get tails first, our \$1 falls to \$0.90. The subsequent heads wins us \$0.09 and we, again, end up with \$0.99. In the real world we can get lots of heads in a row, but we can also get lots of tails in a row. If the coin is fair we will see half of each side and an expected loss of 1% per round. This illustrates the fundamental difference between an additive dynamic and multiplicative dynamic.

Figure 03. The relationship between returns and position size



The right hand side of the lines on the chart shows our earlier results for the 'entire wealth' bet: an ensemble average return of 83% and a time average return of -100%. We have already established that the ensemble average is blind to risk, and so the ensemble average return increases linearly with position size. If we were to use this measure for 'optimisation', it would push us to wager 100% of wealth (and almost certain financial death). Of course, no investor looks at return in isolation but at this point we are simply comparing the characteristics of the two averages. The time average return does incorporate risk, and is not linear, so we can use it to optimise[8] our wager size. In this case the optimal position size is a stake of around 10% of wealth, and we can see that wagers much above 20% of wealth will tip us from making money into losing money. If we were to zoom in on the chart we would see that the actual optimal position size, in this case, is 9.1% of wealth which would give us a per-period rate of return of 3.7%. Wagers smaller than this do not pay off enough when we roll a six, while wagers bigger than this cost us too much of our wealth when we roll one to five and hence the per-period rate of return is lower.

It is a bit of a leap from rolling a die or tossing a coin to real-world investing, but we believe the analogy is sound. For any investment, considered in isolation, the time average return will show an upper limit to the amount of risk we should take in that investment. For low volatility, unleveraged investments (for example, cash) the time average will be the same as the ensemble average, and so the upper limit is likely to be 100% of the portfolio but for other, riskier, investments it is likely to give us a lower upper limit. Note that this stage of the analysis does not require us to make any assumptions about future correlations.

Practical implications

It should be clear from the above that we believe the time average has an important role to play in the realm of risk management. However we are not claiming that it is 'the answer' to all risk management needs.[9] It is also our belief that the time average is a relatively unfamiliar concept in finance. Consequently we give our initial thoughts as to what the practical implications could be. Greater familiarity with and experience of using time averages should increase this list.

- **Improved risk management:** we have made the case above for the time average being a better risk-return measure than currently calculated expected returns, and for how it can be used to inform position sizes. We do not want to over-state the case, and have already noted that in many cases the difference between the ensemble and time averages will not be material. For example our investment strategy team has tended to use the ensemble average for risk budgeting exercises (typically over a one-year horizon) and the median result from the parallel universes for long-term (typically 10-year) projections.[10] The mathematical relationship between the various statistics shows that most of the time this is acceptable practice. However, in cases where leverage is higher or option strategies are used, we suggest that time average returns should be calculated and included on an asset owner's risk dashboard.

8 This is an artificial example where the payoffs are known precisely, so optimisation is possible. We are not claiming that time average returns can be used to optimise in real world situations that have uncertain outcomes, but we believe they should still be a useful guide.

9 For a discussion regarding what we think is necessary for an improved risk management framework please see our recent publication *The wrong type of snow*, Towers Watson, February 2012.

10 This statement is more applicable to our UK team, with the US team historically making more use of geometric means/time averages.

- **Development of stochastic modelling:** much of current risk modelling relies in some way on stochastic projections – the random projection of many future outcomes so as to better understand the likely probability distribution. As this is a version of invoking parallel universes, care should be taken that the modelling results respect the irreversibility of time. In a sense, current practice does this implicitly. By projecting over 10-year horizons there is a significantly increased chance that a rare, negative one-year event (the impact of which is under-represented in the one-year arithmetic average) will surface over the 10-year term. We are not modelling experts but a potential enhancement, in our minds, would be to run the generated paths consecutively through time and calculate back the time average return. If any single path entails a complete loss of wealth, then the time average will suggest that the strategy under consideration is too risky, irrespective of how attractive the median path appears. Our current strategic investment advice already recognises that certain paths represent no-return outcomes but the time average brings the benefit of replacing intuition with hard, physical reality.
- **Limit on leverage:** having lived through the global financial crisis, we already know intuitively that there is a point at which leverage becomes toxic. We believe that the use of time averages will help us better understand where that point may lie. It would not surprise us to discover that time averages showed some highly-leveraged investment strategies to be unattractive no matter how small the proposed position size.
- **Risk-based asset allocation:** on the one hand we expect that time averages will show the greater spread of capital in risk-based asset allocation to be attractive. However, linked to the previous point and leading to the following point, risk-based allocation tends to make extensive use of leverage to maintain overall portfolio returns. This suggests to us that further work is required to better understand how much leverage remains prudent when adopting risk-based allocation.
- **Limit on target returns:** in our thought experiments we saw that the time average was sensitive to large potential drawdowns in wealth, and particularly when they were close to 100% of wealth. Again, we believe further work is required here, but just as the tortoise beat the hare in Aesop's fable we have a hunch that a low-volatility, robust portfolio built on time averages will compound through time at a faster rate than a 'racy' portfolio built on ensemble averages – albeit that it will look recklessly conservative at times.[11] In any event, we believe that time averages show that a 100% equity strategy (as found in some defined contribution pension plans) is not appropriate. To explain, a 100% equity strategy implies that wealth will suffer considerable volatility, say 20% in round numbers. So a 50% drop in wealth would be a rare-but-possible event. If immediately followed by a 100% gain, then the ride was bumpy but of no lasting consequence. However, if the 50% drop was due in part to over-valuation and not followed by a rebound then there will have been a loss in wealth – and a permanent loss in wealth if this occurs just before the end of the investment horizon. If moderate levels of leverage are permitted, it is almost always possible to build a superior portfolio by adding a diversifying asset. The moderate leverage allows the asset owner to target the same expected return while the diversification reduces the volatility and potential hit to wealth.
- **Cash flow is king:** we will argue slightly against our last point – that a 100% equity strategy is inappropriate for defined contribution plans. Our paper has only considered 'closed-wealth' systems whereas in the real world cash flows, both in and out, are an integral part of investing, insurance and risk management systems. Hence our conclusions that time averages should be favoured for higher volatility, higher leverage and non-linear returns require qualification. If the inflows of cash are large relative to the initial assets then we would favour using the ensemble average. Cash flows make the time average harder to define, and in effect we would need to calculate internal rates of return for each of our 10,000 future projections, which under current technology would likely be sufficient to kill the modelling system. Conversely, it is clear that cash outflows make risk management considerably more important. Not only can a risk event cause a drop in wealth, but the cash outflow reduces the amount of assets that can be 'sweated' to recoup the lost wealth. The transition from positive to negative cash flow is a distinctly non-linear event.

11 If you have the ability to dynamically adjust your asset allocation successfully this will leave the static, low-volatility portfolio in a cloud of dust. In aggregate, however there is no evidence that the majority of asset owners can pull this off, making the robust portfolio a sensible choice for many, if not most, investors.

12 More granularity could be added by including additional categories such as 'middle-aged' and 'very mature'.

- **Investor type matters:** we typically categorise investors according to the purpose for which the asset pool is established – pension fund, insurance company, endowment, sovereign wealth fund and the like. The above consideration of cash flows suggests it may be useful to use a different taxonomy such as ‘young’ and ‘mature’.[12] Young investors would have positive cash flows and higher tolerance for volatility and leverage. They may feel that the easier-to-work-with ensemble averages are fit for purpose in terms of setting and managing their risk profile. At the other extreme, the most mature investors with large outflows would use time averages in their risk management, particularly if still wishing to run a relatively aggressive strategy. As a generalisation, pension funds would start young and grow inexorably more mature. Other investor types are arguably less tied to the passage of time; for example, a large rise in the price of oil for sovereign wealth funds, a jump in alumni donations for endowments, and an improvement in business mix for insurance companies could see their investor type become less mature.
- **Catastrophe bonds:** financial economists wedded to modern portfolio theory (and ensemble averages) have repeatedly argued that since catastrophe risk is not systematic, it can be fully diversified, and therefore catastrophe bonds deserve no expected return above the risk-free rate. So far, the markets (practitioners) disagree with this theoretical reasoning as the expected returns on catastrophe bonds remain high (typically around 15%). On a time average basis the offered return has to be high enough to cover the more-than-immaterial chance of 100% loss, and we do not expect this to be diversified away.
- **Improved regulation:** the above implications apply to individual asset owners or asset classes, but time averages also speak to macro issues such as controlling the aggregate amount of leverage within the financial system, what margin requirements or minimum capital levels should be, and maximum permissible loan-to-value ratios for mortgages.
- **Casinos:** as a final aside, it is typically assumed that with respect to casinos, the ‘house’ makes its money by having a slender edge in the (ensemble) odds over multiple independent bets. Perhaps time average thinking suggests a subtle difference – that the house makes money because most people keep playing repeatedly until they eventually lose (the bets are not independent, and you can’t travel back in time to replay your stake money).

Conclusion

The fields of finance and economics have focused on trying to rectify the shortcomings of ensemble averages with ever more sophisticated considerations of utility. While academics may not have spent sufficient time grappling with time averages, investment and risk practitioners have definitely ‘dirtied their hands’. However, to our knowledge, the debate has typically been statistical and technical and consequently has not broken into the mainstream.

The reality of the irreversibility of time quickly gets us to a better starting point and, we believe, in a manner that is accessible to a wide audience. From that point we can then layer on considerations of risk aversion and utility if we so desire. We hope we have made the case that adding time averages to our existing use of ensemble averages, is a relatively simple step and one that would be positive, at the margins, for risk management.

References

Mauboussin, Michael (March 23, 2012), *Shaking the foundation – an interview with Ole Peters*, Legg Mason Capital Management.

Peters, Ole (2011), *Optimal leverage from non-ergodicity*, Quantative Finance, 11:11, 1593–1602.

Peters, Ole (2011), *The time resolution of the St Petersburg paradox*, Philosophical Transactions of the Royal Society (2011) 369, 4913-4931.

Peters, Ole (2012), *Menger 1934 revisited*, in review.

Thinking Ahead

This publication is written by members of our Thinking Ahead Group (TAG) who are part of the investment business at Towers Watson. Their role is to identify and develop new investment thinking and opportunities not naturally covered under mainstream research. They seek to encourage new ways of seeing the investment environment in ways that add value to our clients. The contents of individual articles are therefore more likely to be the opinions of the respective author(s) rather than necessarily representing the formal view of the firm. No action should be taken on the basis of any article without seeking specific advice. If you would like to discuss any of the areas covered in more detail, please get in touch with the consultant who normally advises you at Towers Watson, or:

Carole Judd
+ 44 1737 274329
carole.judd@towerswatson.com

Appendix: the maths behind ensemble and time averages

For simplicity we consider an investment opportunity with a starting value of $P_{t=0}$, which has equal probability of reaching N possible outcomes of $P_{t=1}$ $(i=1,2,3\ldots N)$ at the end of one time period. The rate of change (r) associated with each outcome is therefore calculated by

$$r(i) = \frac{[P_{t=1}(i) - P_{t=0}]}{P_{t=0}}$$

Ensemble average – $E(r)$ – is an answer to the question "what is the rate of change on the investment, computed from an average over all possible outcomes (universes)?" and defined as

$$E(r) = \sum_{i=1}^{N} \frac{r(i)}{N}$$

Time average – $T(r)$ – is an answer to the question "what is the rate of return on this investment averaged over time" and defined as

$$T(r) = \prod_{i=1}^{N} [1 + r(i)]^{1/N} - 1$$

For the technically minded who prefer to think of investment in continuous time, the stock price in standard finance literature normally follows a geometric Brownian motion (GBM). In other words the price is log-normally distributed with a constant percentage rate of change (μ) and constant percentage volatility (σ):

$$\frac{dP_t}{P_t} = \mu dt + \sigma dW_t$$

Where:

dP_t is the infinitesimal change of P in continuous time

dt is the infinitesimal increment of time

W_t is a continuous-time stochastic process known as a Wiener process or Brownian motion. Intuitively, it is a process that jiggles up and down in such a random way that its expected change over any time interval is 0 with a variance term T over time T.

Skipping a few steps of derivations that involve stochastic calculus, the ensemble-average growth rate is given by $E(r)=\mu$, and the time-average growth rate by $T(r)= \mu - \sigma^2/2$. The formula clearly demonstrates the effect of volatility "drag" from risk taking not captured by the ensemble average.

About Towers Watson

Towers Watson is a leading global professional services company that helps organisations improve performance through effective people, risk and financial management. With 14,000 associates around the world, we offer solutions in the areas of employee benefits, talent management, rewards, and risk and capital management.

TW-EU-2012-28847. October 2012.

towerswatson.com

TOWERS WATSON