Universal Portfolios and the Hindsight Bias Issue: A New Perspective?

David Edelman

University College, Dublin

1 Abstract

Recent results in dynamic portfolio allocation methods are used to develop a new paradigm for the assessment of Hindsight Bias in Model Selection, Forecasting, and Market Efficiency studies, calling into question previous accepted wisdom on the issue.

2 Introduction

As this paper involves the application of the Universal Portfolio concepts of Cover ([1][2]) to the Hindsight Bias issue, we begin by introducing each of these individually, beginning with the latter.

The Hindsight Bias problem, often also referred to as the 'Data Mining' problem, occurs when first, many possible models to describe an aspect of the past are posited, and then the best-fitting of these is selected *a posteriori* (i.e., after the fact). When a model thus selected is then applied as a model for the future, the results tend to be disappointing, as the following example illustrates.

> 8 coins are selected at random, and a gambling game is to be based on Heads outcomes being favourable and Tails results being unfavourable. An experiment is run, involving a single flip of each coin, with the result that 4 Heads are obtained and 4

> Tails. The 4 Head-falling coins are identified as 'Blue Chip' coins, and the others as 'Dogs'. An investment scheme is devised, involving investment solely in the 'Blue Chip' coins, but with the result that on the next round of tosses, the 'success' ratio of the 'Blue Chip' coins drops from 100% to 50%. Investors flee.

The term *Data Mining*, or looking for patterns in the past, some of which could be largely coincidental, has (somewhat confusingly) been recently rehabilitated to refer to the process of engaging in such an exercise while attempting to compensate for the potential coincidental component. To avoid confusion, we instead refer to the *Hindsight Bias* issue.

The term Universal Portfolio was first coined by Cover ([1][2]) to refer to a dynamic portfolio rebalancing scheme which asymptotically outperforms the best (*a posteriori*) single asset in the portfolio. It does this by tracking the best Constant Rebalanced Portfolio, which is a portfolio that maintains a fixed proportion of each asset in the portfolio, regardless of price changes.

In his seminal paper ([1]), Cover began with the following arguably counterintuitive statement:

> "We exhibit an algorithm for portfolio selection that asymptotically outperforms the best stock in the market"

As a simple example (see [2]), consider the two-stock problem in which the value of one stock is constant and the other alternately halves and doubles. While 100% investment in either stock can yield at most a doubling of one's wealth, a 50-50 Constant Rebalanced Portfolio will grow exponentially, increasing by a factor of 9/8 every two days.

On the other hand, in this simple case, if at each step, we were to determine which Constant Rebalanced Portfolio would have been optimal up to the present and then apply it to the following day, repeating this process day after day, we would indeed achieve exponential growth (thereby outperforming the best stock in the long-run), and in fact would soon be converging (in terms of relative difference) to a rule which was very similar to the optimal one. This is illustrated in

Figure 1., where the constant and 'sawtooth' raw return series for the individual stocks is graphed along with the optimal and the Universal Portfolio approaching it from below.

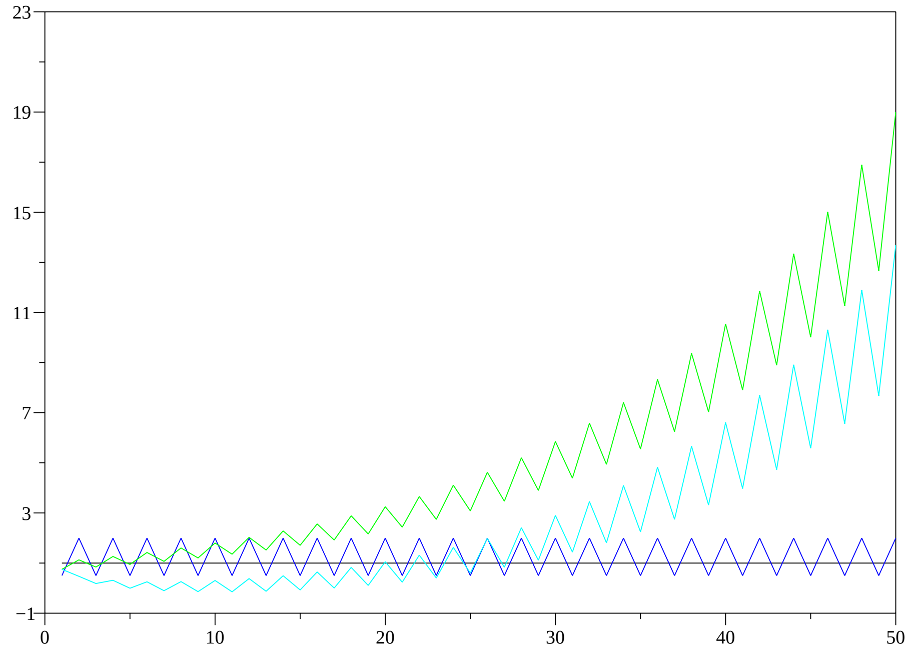

Figure 1.

Clearly, the Universal Portfolio outperforms any hindsight choice of the two individual stocks in the long run.

3 Universal Portfolios and the Hindsight Bias Issue

The relevance of Universal Portfolios to the Hindsight Bias Issue would seem rather natural, though surprisingly does not appear to have been noted previously in the literature. Imagine, hypothetically, for example, that two attempts using Neural Network forecasting methods on a Financial series had resulted in models yielding returns resembling the two patterns (constant, sawtooth) of the example of the previous section.

According to previous thinking [3] citing the return of the better of the two models as evidence of success would not only have been regarded as being flawed methodologically, likely resulting in the results being discarded as virtually worthless, but further in this case, the citing of the better of the two as evidence of success would have vastly *understated* the true worth of the research results.

Generally, then, it would appear that in building trading models, the methodology of running a Universal Portfolio algorithm on several modeling attempts to fit past data would be not only *desirable*, but might even be argued to be *essential* for any complete study.

Further, it would appear likely that at least some previous Market Efficiency studies of various methods for achieving ’excess returns’ might very well have been judged too harshly (or needlessly discontinued altogether), and perhaps might merit fresh perusal.

Indeed, it is hoped that this paper might result in just such types of outcomes, in addition to proposing a whole new methodological basis for model selection and evaluation in future studies.

References

[1] Cover, T. (1991) “Universal Portfolios”. *Mathematical Finance*, **1**(1):1-29, 1991.

[2] Cover,T. and Ordentlich,E. (1996). “Universal Portfolios with Side Information”. *IEEE Transactions on Information Theory*, **42**(2):348-363, March 1996.

[3] White, H., et al (1999), “Data-Snooping, Technical Trading Rule Performance, and the Bootstrap”, *Journal of Finance* **54**(5):1647-1691