Abstract

Modern Portfolio Theory as introduced by Markowitz (1952) frames the time dimension of investing as a single period over which the parameters of the probability distribution of asset returns are both known with certainty and are unchanging. We know that neither assumption is true in the real world. The second assumption has received attention in the theoretical literature but there has been little progress in terms of practical implementations available to financial practitioners. In this paper, we will discuss simple approximations that can improve the ability of investment professionals to rebalance portfolios in ways that efficiently control portfolio turnover and the related costs of rebalancing.

Introduction

Modern Portfolio Theory as introduced by Markowitz (1952) frames the time dimension of investing as a single period over which the parameters of the probability distribution of asset returns are both known with certainty and are unchanging. In essence, we treat the future as a single period which begins at the present but ends only at some unknown moment in the future. This second assumption has received attention in the theoretical literature but there has been little progress in terms of practical advances available to financial practitioners.

The underlying assumption that legitimizes the “single-period” framework for portfolio optimization is that all frictions that impact portfolio formation and rebalancing are inconsequentially small. If the cost of rebalancing is zero, then the single period assumption does no harm. If conditions change such that our portfolio should be changed, we can simply declare that the current single period has ended and a new one has begun. We can then rebalance our portfolio accordingly at no cost. By doing so, we wish away the relevance of routine trading costs, market impact, and taxes on capital gains. While these costs may be very small for some investment assets held by some investors (e.g. liquid futures contracts held by tax-exempt entities) the necessary conditions are not true in most cases.
Smarter Rebalancing: Using Single Period Optimization In a Multi-period World

For almost all real world investors, portfolio rebalancing is not free, and often is very costly. For taxable investors holding illiquid assets such as real estate, private equity or investment collectibles (e.g. art), the cost of transactions often dominant risk and return considerations unless holding periods exceed multiple decades.

Almost all widely available portfolio optimization systems are based on mean-variance optimization methods which maximize the expectation of the risk adjusted return on a portfolio over a single period, as described in Markowitz and Levy (1979).

Multi-Period Optimization and Approximations Thereof

Much research has been done on the formulation of full multi-period optimization. Mossin (1968) suggested an explicit multi-period formulation for portfolio optimization. A paper by Cargill and Meyer (1987) focused on the risk side of the multi-period problem. This was followed by Merton (1990) who introduced a continuous time analog to mean-variance optimization, and by Pliska (1997) who provided a discrete time analog to the single period approach.

Other papers such as Li and Ng (2000) provide a framework for multi-period mean variance optimization (MVO) using dynamic stochastic programming. Such stochastic programming methods investigate a range of potential paths of future events and outcomes, and select the set of asset weights that best fulfills a stated objective and set of constraints. Such techniques are often used for asset allocation for complex taxable investors such as insurance companies and high net worth individuals. The computational intensity of the method often limits the number of assets that can be handled. Even with today’s improvements in computational efficiency, the methodology is still only viable for portfolios with a small number of assets. Using a set of simplifying assumptions, Sneddon (2005) provides a closed form solution to the multi-period optimization including optimal turnover that could be applied to problems with large numbers of assets.

Explicit multi-period optimization and related processes have had very limited adoption among practitioners because of these constraints. The most significant impediment is parameter estimation error. Investment practitioners often have great difficulty accurately estimating the future distribution and correlations among the assets in their universe of choices. To fully exploit the benefits of multi-period optimization we must be able to not only forecast these distributions, but also forecast how these future distributions will change from one future moment of time to the next and so on. Obviously, being able to accurately forecast such changes appears to be beyond the predictive power of real world investors.

Another line of research in this area focuses on the idea of creating rules that inform investors when it is really necessary to rebalance their portfolio weights. In the absence of statistically significant and economically material advantage such rules simply tell the investor to do nothing, hence avoiding rebalancing costs altogether. Early research in this area includes Rubenstein (1991) who examines the efficiency of continuous rebalancing and proposes a rule for avoiding unnecessary turnover. Kroner and Sultan (1993) propose a “hurdle” rule for rebalancing currency hedges when return distributions are time varying, while Engle, Mezrich and Yu (1998) propose a similar hurdle on alpha improvement as the trigger for rebalancing actively managed asset allocations.
Smarter Rebalancing: Using Single Period Optimization In a Multi-period World

Bey, Burgess, Cook (1990) use bootstrap resampling to identify “indifference” regions along the efficient frontier. In Gold (1995), a similar technique is to define indifference to rebalancing portfolios for illiquid asset classes such as real estate. Michaud (1998) uses a parametric form of resampling to measure the confidence interval on portfolio return and risk to form a “when to trade rule.” The rule was elaborated upon in Michaud and Michaud (2002) and patented.

Markowitz and Van Dijk (2003) propose a rebalancing rule (“MvD”) based on game theory to approximate multi-period optimization, but argue it is mathematically intractable (at least in closed form) for large problems. In Kritzman, Mygren and Paige (2007) the authors are able to test the efficiency of MvD against full dynamic programming for cases up to a maximum of five assets, as dynamic programming becomes computationally too burdensome for larger numbers of assets. The authors are able to extend the MvD method to one hundred assets.

Practitioner adoption of some form of “all or nothing” rebalancing rule is fairly common, but at least two negative arguments have come to light. The first is that when active managers are “inactive” because the potential benefits of rebalancing are too small, this lack of trading is perceived by clients as the manager being neglectful rather than as an analytically-driven decision to reduce trading costs. Presumably this objection could be overcome by appropriate communication between the asset manager and their investors.

The second argument is that after a period of inactivity, the eventual rebalancing concentrates the required trading into a particular moment in time. For large investors the market impact arising from doing trades that are a larger fraction of available trading volume per unit time will create higher transaction costs than if the trading had been done gradually between the previous portfolio rebalancing and the current one.

For portfolios that are composed of assets with homogeneous transaction costs it is common to simply place heuristic limits on the amount of turnover allowable in a given rebalancing. Grinold and Stuckelman (1993) consider a value added/turnover efficient frontier. They derive that under certain common assumptions, value added (improvement in utility) is approximately a square root function of turnover (“half the turnover gets you three quarters of the available improvement”). As such, investors can optimize their portfolios without consideration of transaction costs, and then simply choose the intermediate point (e.g. the amount of turnover) between the initial portfolio and the optimal portfolio that gives the most favorable balance of utility improvement relative to the expense of trading costs.

A Single Period Model with Trading Costs

A typical way of accounting for transaction costs in an optimization is to deduct these costs from the mean-variance objective function, amortizing the expected costs over the expected holding period of the investment. Put differently, the expected holding period is just 100% / M%. 
Smarter Rebalancing: Using Single Period Optimization In a Multi-period World

\[ U_1 = \alpha - \lambda \sigma^2 - (\Sigma_{j=1}^{N} \text{abs}(W_{ij} - W_{fj}) * K_j) * M \]

Where

- \( N \) = the number of assets in the portfolio
- \( W_i \) = the initial weight of asset \( J \)
- \( W_{fj} \) = the optimal weight of asset \( J \)
- \( K_j \) = the unit cost of trading asset \( J \)
- \( M \) = the expected % one way turnover per annum
- \( \alpha \) = the expected annual % return on the portfolio
- \( \sigma \) = % expected volatility of annual portfolio returns
- \( \lambda \) = investor risk aversion scalar

One obvious improvement to the simple linear amortization of costs is to realize that expending transaction costs today reduces the value of the portfolio on which returns can be earned. For portfolios with small costs and low returns, the linear approximation is sufficient, but if costs or returns are large we need to consider the effects of return compounding. For example, assume a trade with 20% trading costs and an expected holding period of one year that provides a return improvement of 20% per year. Despite the apparent equivalence of costs and benefits, if we give up 20% of our portfolio value to costs now we would end the year with only 96% of our initial wealth. The solution is to adjust the amortization rate to reflect the opportunity costs of the lost returns.

\[ M_{\alpha} = M * (1 + \alpha/100) \]

It should be noted that multiple distinct amortization rates may be required in dealing with taxable portfolios at the individual tax lot level. For example, we might have a portfolio that normally has 20% per annum turnover. However, within the portfolio there is a single security that has a negative expected return, but a substantial embedded short-term capital gain. We would like to sell the security because of the negative return expectation, but we would also not like to not sell the security because of the substantial tax consequences. If the point in time at which our short term gain would revert to a long term capital gain that is less taxed is near, we would hold. If the point of conversion to long term capital gain status is far away, we would sell. To optimize this portfolio correctly we must use our regular amortization rate for the long term portion of cost of the capital gain tax, while the incremental portion of the tax cost of selling a short term capital gain should be amortized over the time interval between today and the conversion to long term status.

The Probability of Realization

Imagine we have a portfolio, \( P_1 \) with return \( \alpha \) (net of fees and expenses) and standard deviation \( \sigma \). Our usual utility function would say:
Smarter Rebalancing: Using Single Period Optimization In a Multi-period World

\[ U_1 = \alpha - \lambda \sigma^2 \]

Now let’s imagine that another portfolio, \( P_2 \), has a higher utility, because either the return is higher or the standard deviation is lower. This portfolio has completely different positions than the initial portfolio. Let’s assume that this portfolio has a higher return by positive increment \( \Delta \), so:

\[ U_2 = (\alpha + \Delta) - \lambda \sigma^2 \]

Since \( U_2 \) is greater than \( U_1 \), we should be willing to pay some transaction costs to switch from \( P_1 \) to \( P_2 \). Now let’s consider a different way to improve our returns. We go back to the manager of Portfolio 1 and ask them to reduce their fees by \( \Delta \), so now our revised utility on \( P_1 \) is \( U_{1L} \) for “lowered fees.” Notice that \( U_{1L} \) and \( U_2 \) are equal. So if we invest our money in either \( P_2 \) or \( P_{1L} \) (after lowering the fees), the expected value of wealth at the end of time is the same. This suggests that we should be willing to pay the manager an upfront fee to lower ongoing management costs equal to the trading costs we would be willing to pay to switch from the initial portfolio \( P_1 \) to \( P_2 \). **As long as conditions never change, this is valid.**

However, since \( P_2 \) and \( P_{1L} \) have different securities, the performance will be different from day to day and month to month, even if the long term average return and volatility are identical. So over any finite time horizon, we cannot be sure which of the portfolios will perform better. On the other hand, \( P_{1L} \) will always perform better than \( P_1 \) over all time horizons, because it is just the same portfolio with lower fees. For \( P_{1L} \), the probability of outperforming \( P_1 \) is always one. \( P_2 \) is guaranteed to be better than \( P_1 \) in the long run if conditions don’t change, but the probability that \( P_2 \) will actually outperform \( P_1 \) over any finite horizon is between one half and one. We call this probability value the “probability of realization” of the utility increase.

Amortizing transaction costs over a single period is equivalent to assuming that the probability of realization is always one. Let’s assume a single period optimization for a strategy with expected turnover of \( M\% \) per annum. In our multi-period world, we want to amortize by \( M \) divided by a scalar that reflects the probability that the revised portfolio will actually realize a better risk adjusted return over the finite holding period. That is:

\[ M^* = \frac{M \alpha}{K} \]

Where

\[
\begin{align*}
M^* & \text{ is the adjusted amortization rate} \\
K & \text{ is a chosen scalar in the range } [Z:1] \text{ where } Z \text{ is the probability of realization}
\end{align*}
\]

To calculate \( Z \) we can use the tracking error (expected volatility of the return differences) between any two portfolios as a standard error on the expected differences in utility. We obtain the tracking error value from whatever model we are using to estimate \( \sigma \) in our utility function. We can then calculate a T-statistic on the
expected difference in utility, calculating Z under a one-tailed test and our choice of probability distribution. In our example above, the value of Z is one for the incremental utility of P1L relative to P1. This is because the tracking error between P1 and P1L is zero, as they are the same portfolio with different fee structures. Highly risk-averse investors will chose values for K close to Z (the expected benefit in utility over a finite horizon is uncertain while the costs of trading are certain so we want to reduce turnover). More risk-tolerant investors will chose K close to one.

The Next Rebalancing

The process above to amortize rebalancing costs reflects only the trading costs that are going to be experienced as a result of the current portfolio revisions. Another way to get closer to multi-period optimization is to give some forethought to the expectation of ongoing costs that are likely to be experienced in future rebalancing events. In high turnover strategies, the changes in the asset weights will be dominated by the volatility of the asset values and the lack of correlation across the asset returns. The expected amount of trading required at the next rebalancing will be greater for asset universes with higher variety (i.e. cross-sectional dispersion of returns). For low turnover strategies, the differences in asset weights at the next rebalancing will be dominated by differences in the mean returns achieved by the assets. The larger the difference in expected returns across the asset universe, the larger the required turnover is likely to be at the next rebalancing. If some assets have materially higher trading costs than others, we can decrement the expected returns of the individual assets to reflect the expectation of higher ongoing trading costs that will be experienced in future rebalancing events.

Many portfolio strategies also have implicit beliefs about auto-correlation in asset returns. For example in a stock portfolio with a momentum strategy, we are implicitly expecting positive serial correlation in returns. Stocks that realize a high return in one period will have an increased expected return for the next period. In that an above average return relative to other assets implies an increase in portfolio weight, the portfolio is already positioned to reflect the higher expected return for the next period that would justify an increase in the asset’s weight in the new optimal portfolio. This congruence reduces turnover.

Conversely, a value strategy implicitly expects negative serial correlation in returns. Stocks that realize negative returns in one period will have a higher expected return for the next period. The negative return will reduce the weight in these assets, while the higher expected return is apt to call for a higher optimal weight for the next period. This discrepancy of direction implies greater portfolio turnover and costs over future rebalancing cycles. To the extent that a portfolio strategy leads to a larger or smaller expectation of turnover we may choose to change the magnitude of the suggested decrement to expected returns.

Conclusion

While full multi-period optimization seems obviously preferable to the single period representation there are many impediments which prevent practical implementation. “When to rebalance” rules are often helpful in reducing the
economic cost of portfolio turnover, but may be detrimental to large investors where the market impact of trading is a material issue.

The MvD solution has been demonstrated to be useful for small cases, but implementation of the technique for portfolios of large numbers of assets positions or large numbers of active positions (i.e. a small portfolio measured against a broad benchmark) is problematic.

We suggest that much of the potential benefit of multi-period optimization can be obtained by incorporating trading costs of the current rebalancing into the optimization process in a thoughtful way. Additional benefits can be obtained by considering the long-run rebalancing costs of different investment assets, and adjusting return expectation to be net of these long-term costs.

References


Smarter Rebalancing: Using Single Period Optimization In a Multi-period World


