Momentum, Markowitz, and Smart Beta
A Tactical, Analytical and Practical Look at Modern Portfolio Theory

Wouter J. Keller (Flex Capital)[1]
June 13, 2014, v0.9

Abstract

In this paper we will try to improve on the Modern Portfolio Theory (MPT) as developed by Markowitz (1952). As a first step, we combine the MPT model with generalized momentum (see Keller 2012) in order to arrive at a "tactical" MPT. In our second step, we will use the single index model (Elton, 1976) to arrive at an analytical solution for a long-only maximum Sharpe allocation. We will call this the MAA model, for Modern Asset Allocation.

In our third step, we use shrinkage estimators in our formula for asset returns, volatilities and correlations to arrive at practical allocations. In addition, as a special cases, we arrive at EW (Equal Weight), Minimum Variance (MV), Maximum Diversification (MD) and (naïve) Risk Parity (RP) submodels of MAA. These EW, MV, MD and RP models are sometimes called "smart-beta" models.

We illustrate all these different models on three universes consisting of respectively 10 and 35 global ETFs, and 104 US stocks/bonds, with daily data from Jan. 1998 – Dec. 2013 (16 years), monthly rebalanced. We show that all these models beat the simple EW model consistently on various return /risk criteria, with the general MAA model (with return momentum) also beats nearly all of the "smart beta" models.

Keywords: Tactical Asset Allocation, Momentum, Markowitz, Elton, MPT, mean variance, minimum variance, Sharpe, EW, SIM, smart beta

JEL Classification: C00, C10, G00, G11

Affiliations and contact information: Wouter J. Keller, Flex Capital BV, Rotterdam and Vrije University, Amsterdam, The Netherlands, email wkeller@flexcapital.nl

---

1. Introduction

The Modern Portfolio Theory (MPT) goes back to a seminal article from Markowitz (1952). It's also known as the "mean-variance" or "tangency" solution for the optimal portfolio allocation. In principle, it's a strategic approach, or, in other words it aims at a long run (multi-year) allocation.

The core of the MPT is the classical mean-variance solution for portfolio selection, also known as the maximization of the Sharpe ratio. To compute the corresponding optimal portfolio, one needs to estimate the expected mean and covariance of assets returns, e.g. by their sample estimates from historical return data. These estimates often contain substantial estimation errors, especially for the mean return.

In the strategic (multi-year) MPT framework, the expected mean and covariance of assets returns are often estimated over a multi-year historical window, say five to ten years (60-120 months). In a seminal article, DeMiguel (2007) have shown that the strategic sample-based MPT allocation is nearly always outperformed by a simple equal weight (EW) allocation[2]. They show that this also holds true for most of its extensions designed to reduce estimation error, e.g. when shrinkage estimators are used. A similar conclusion was recently arrived at by Ang (2012) and Jacobs (2013). All authors used multi-year windows (60-120 months) and therefore a strategic approach.

In this paper, we will take a more tactical approach to MPT and limit the estimation (lookback) window to 12 months or less. The approach relies heavily on the return momentum anomaly, see eg. Jegadeesh (1993) and Faber (2007 and 2010). We also apply the same short-turn momentum ("persistence", see also Keller, 2012) approach to the estimation of the future volatilities and correlations. Together with the long-only maximization of the Sharpe ratio as optimization criterion we arrive at our "Tactical MPT" model.

In order to get more grips on the optimal allocation in analytical terms, we also used the assumption of a single-index model (see Elton, 1976) where the returns are related to a market (index) return, like EW. Now, we can arrive at an elegant analytical formula for the (long-only, single-index) maximum Sharpe ratio allocation. We will refer to the resulting model (with the single-index, long-only and generalized momentum assumptions) as the "MAA model", for *Modern Asset Allocation*. The optimal MAA allocation can be expressed analytical as a near closed- form solution with only one non-linear parameter, which can be solved very easily and rapidly computationally.

The MAA model shows the effects of *four components*: return (R), volatility (V), market (M) and correlation (C). Later on we will connect these four components with the shrinkage of asset returns, asset volatilities, market variance and the cross-correlations of assets.

The first component R represents the return effect: the optimal (long-only) MAA allocation for an asset is directly proportional with its momentum returns and zero for assets with nega-

---

[2] Kritzman (2010) found that EW can be beaten by "hand picking" the expected returns from even longer lookback periods than the 60 or 120 months of DeMiguel.

tive momentum returns. This is similar to the well-known relative and absolute momentum effect, but now resulting from the MPT theory. The second component V, represents the volatility effect: the optimal long-only MAA allocation for an asset is inversely proportional to the square of its idiosyncratic volatility. This effect is similar to the well-known low-volatility anomaly, but now resulting from the MPT theory.

Components M and C do work together and reflect the so-called "systematic" effect of the market and different correlations. If the market effect (component M) is assumed to be zero, all assets (with positive returns) are included in the optimal long-only MAA allocation. In this case, the effect of correlations on the optimal long-only allocation completely disappears. If the market effect is positive, there is a threshold for assets to enter the long-only allocation, depending on the correlations. We can therefore increase the number of selected assets by simply shrinking the market effect towards zero, giving more diversification. This avoids a well-known practical problem with MPT where the number of selected asset is often very limited. The larger the "systematic" or market effect is (measured by the market variance), the larger the effect of correlations. When in addition the component C is non-zero, the stronger is the effect of different correlations, while a zero component C corresponds to the MAA model with a constant-correlation model.

We also cope with practical estimation errors in this model and use shrinkage estimators for momentum returns, volatilities, etc. to improve these estimators. In addition, by complete shrinkage of each asset return to a constant or to its volatility, we also arrive at special cases of this practical MAA model: Minimum Variance (MAA-MV), Maximum Diversification (MAA-MD) and Risk Parity (RP) models. These models are sometimes referred to as "smart-beta" models (see eg. Maillard (2009), Scherer (2010) , Clark (2011 and 2012), Choueifaty (2011), Schoen (2012), Jurczenko (2013), and Roncalli (2013)). We will compare all our models with the Equal Weight (EW) model, which can be shown to be also a special case of our MAA model.

Notice that for all these smart beta models, we assume a restricted covariance matrix, arising from the single-index model. As a result of this single-index assumption, all the optimal smart-beta allocations can be expressed as near closed-form solutions with only one non-linear parameter, just like the general MAA model. This is in contrast to the complex non-linear solution of the general minimum variance and maximum diversification models (without the single-index assumption)[3].

We will demonstrate that all these "smart beta" models (and in particular the general MAA model with momentum) easily outperform the equal weight (EW) allocation (taken as benchmark) for different universes, showing the usefulness of the Modern Portfolio Theory. Before that, however, we will first look at our analytical model in more detail.

---

[3] Although we will subsequently refer to the MAA-MV and MAA-MD model as MV and MD, the reader should remember that these models are submodels of the MAA model where we assumed a restricted covariance model according to single-index model. The EW and RP models are not restricted in this way.

2. The single-index model

Our basic MPT assumption is the maximization of the Sharpe ratio (the so-called tangency or mean-variance solution) in the single-index model of Elton (1976). The core of this model is the distinction between the systematic effect, which relates the return of an asset to the return of a single market index (like the EW index) through the so-called "beta" coefficient on the one side and the residual (or idiosyncratic or non-systematic) effect on the other. By using this simple model we are able to reduce the number of parameter estimates from the NxN covariance matrix of the returns to the more manageable N beta's, where N is the number of assets in the universe.

We will assume no short sales (long only), no leverage and a risk-free rate of zero for simplicity[4]. Then the optimal long-only asset allocation which maximizes the Sharpe ratio can be expressed as an elegant analytical formula. We derived this MAA formula as a generalization of the formula for a long-only MV or MD portfolio by Clark (2012). Later we learned that a similar formula as ours appeared long ago in a classical paper from Elton (1976), who called it the single-index model (SIM).

Our single-index maximum Sharpe formula expresses the long-only optimal asset allocation shares w_i as a function of the expected returns r_i, the expected idiosyncratic variances s_i and the expected beta's b_i of the assets (i=1..N) for a given universe. This is the main formula of our paper. The **MAA formula** is (see also Appendix A for proofs):

$$(1) \quad w_i \sim (1 - t/t_i) \ r_i \ / \ s_i \qquad \text{for } t_i > t, \text{ else } w_i = 0, \text{ for } i=1..N$$

where "~" stands for "proportional to", and

r_i is the return of asset i,
s_i is the idiosyncratic variance of the returns of asset i,
t_i is the Treynor ratio of asset i (with $t_i = r_i/b_i$),
b_i is the beta of asset i wrt. the market return,
t is the long-only Treynor *threshold*, so that $w_i=0$ for all assets with $t_i<t$

The MAA formula eq. (1) gives us the optimal portfolio allocation as an elegant analytical formula, given the long-only and single-index assumptions. Besides the returns r_i and the idiosyncratic variances s_i, an important role is played by the Treynor ratio t_i (see Treynor, 1996). Selected (long-only) assets should have a Treynor ratio t_i above the Treynor threshold t. This threshold t is a function of the optimal weights w_i, and therefore endogenous. The optimal MAA allocation can be expressed analytical as a "near closed-form" solution with only one non-linear parameter, which can be solved very easily and rapidly computationally.[5]

When the market variance s (*component M*) is zero, the systematic part t/t_i of eq. (1) becomes zero since the threshold t is zero (see eq. (A.6)) and there is no beta or correlation effect. In

---

[4] We use this assumption mainly for notational simplicity. Mathematically, one can easily include a non-zero risk-free rate by replacing returns by excess-returns (above the risk-free rate) in most formulas. Practically, we found only small effects when using excess-returns, also because we had short-term bonds in most of our universes.

[5] We can numerically solve this problem by iterating between w_i and t until convergence, which is very fast in practice.

that case, the optimal long-only MAA allocation shares for an asset are zero if its return r_i is negative, or proportional to its returns if positive. We call this the effect of *component R*. The share of an asset is also inversely proportional to its idiosyncratic variance s_i, which equals the square of their residual volatilities (*component V*). When component M is not zero, only assets with a *Treynor ratio above the threshold* t are included[6] in the portfolio, with larger shares for assets with larger Treynor ratios. Notice that larger Treynor ratios correspond to smaller betas and therefore to smaller market correlations, given certain returns.

3. Special Cases of the MAA formula

The biggest practical problem with MPT is the determination of the expected returns. This also holds for the MAA formula (1) above. Traditionally, with "strategic" MPT, lookback periods of 60 or 120 months (5-10 years) of data are used to estimate expected returns (and volatilities and correlations). In the next section we will consider more "tactical" (short-term) lookback periods. Alternatively, one might use restricted models *without* the need to estimate the expected returns. Examples are Minimum Variance (MV), Maximum Diversification (MD) and naïve Risk Parity (RP), often called "smart beta" models. These can be shown to be special cases of the MAA model given some simplifying assumptions on the expected returns (see also eg. Hallerbach, 2013).

Minimum Variance (MV). An alternative for maximization of the Sharp ratio is minimization of the portfolio variance. See eg. Scherer (2010) and Clark (2011). It can easily be shown that the corresponding optimal allocations are a special case of the maximum Sharpe allocations when the expected return for each asset is assumed to be constant over all assets. After substitution of $r_i=r$ in eq. (1) we arrive at the **MV formula** for the optimal long-only singe-index allocation:

(2) $w_i \sim (1- b_i/b) / s_i$ for $b_i < b$, else $w_i = 0$, for i=1..N

where
   b   is the long-only beta *threshold,* so that $w_i = 0$ for all assets with beta $b_i > b$.

When the market variance s (component M) is zero, the idiosyncratic variance s_i equals the asset variance ($s_i = v_i^2$, with v_i the volatility of asset i) and the systematic part b_i/b of eq. (2) becomes zero (since 1/b=0, see the Appendix). In that case, the optimal long-only allocation share for an asset is only inversely proportional to its *variance* (equal to the square of its volatility v_i). So all assets are now included. When component M is not zero, only assets with a *beta below the threshold* b are included in the portfolio, with larger shares for assets with smaller beta's.

Maximum Diversification (MD). An alternative for maximization of the Sharp ratio is maximization of the *diversification* of the portfolio. See Maillard (2009), Choueifaty (2011), Jurczenko (2013) and Roncalli (2013). It can easily be shown that the corresponding optimal allocations are a special case of the maximum Sharpe allocations when the expected return for

---

[6] This is under the assumption that all beta b_i are non-negative (which holds for nearly all universes). The special case of assets with negative beta's is (although rare) most interesting since these assets can act as "hedge". See also the term (1-t/ti) in eq. 1 which becomes >1 when $t_i<0$ and t>0. For more details see Elton (1976) and the Appendix.

each asset is assumed to be proportional to its volatility. This implies that the Sharpe ratios r_i/v_i are the same for all assets. After substitution of $r_i = v_i$ in eq. (1), we arrive at the **MD formula** for the optimal long-only single-index allocation:

(3) $w_i \sim (1 - c_i/c) \, v_i / s_i$ for $c_i < c$, else $w_i = 0$, $i = 1..N$

where

c_i	is the correlation of asset i with the market,
v_i	is the volatility of asset i,
c	is the long-only correlation *threshold,* so that $w_i = 0$ for all assets with $c_i > c$.

When the market variance s (component M) is zero, the idiosyncratic variance again equals the asset variance and the systematic part c_i/c of eq. (3) becomes zero (since $1/c = 0$, see the Appendix). In that case, the optimal long-only allocation share for an asset is only inversely proportional to its *volatility*. This is the "**naïve" Risk Parity (RP)** solution as a special case of the MD allocation, where all allocation weights w_i are proportional to the inverse of the volatility v_i. All assets are now included. When component M is not zero, only assets with market correlation *below the threshold c* are included in the portfolio, with larger shares for assets with smaller correlations.

Note that when we refer to MAA, MV and MD, we will always assume the single-index model and therefore a restricted covariance matrix. A special case of the MAA, MV and MD model is the **Constant-Correlation** (CC) version (Elton, 1976) where we assume constant cross-correlations among assets, which is the case when the market correlation (c_i) is constant over assets. Now we say that *component C is zero*. For the MAA-CC model we arrive at formula similar to eq. (1) but with the Treynor ratio (r_i/b_i) replaced by the Sharpe ratio (r_i/v_i). Now only assets with a Sharpe ratio (instead of the Treynor ratio) above a certain long-only threshold will be included, with larger shares for assets with larger Sharpe ratios. Finally, if we assume that all returns, volatilities and all market-correlations are the same for all assets, we arrive at the **Equal Weight (EW)** allocation as a very special case of the MAA allocation where all weights are 1/N, identical to the market (index) allocation.

4. Estimation of the general MAA model: momentum and shrinkage

In the above MAA model, future parameters like expected returns r_i etc. are assumed to be known, which is not realistic. In practice we have to use sample estimates of these parameters based on the past. So the quality of our optimal portfolio allocation model depends primary on the quality of these estimates.

Momentum. We will estimate all expected returns by the rate of change (ROC) of the asset price over a certain lookback period, assuming some persistence over time. In the traditional "strategic" MPT model, the lookback period is often several years (60-120 months). We will take a more tactical approach by focusing on lookback period of maximum twelve months, in order to make use of the well-known *momentum* anomaly.

Looking at eq. (1), we see that the optimal long-only shares w_i are only positive if the return r_i is positive[7]. This corresponds to *absolute momentum* or *trend-following*, see eg. Faber (2007), Hurst (2012), Antonacci (2013). When these shares are positive, the optimal long-only shares w_i are proportional with r_i. This corresponds to *relative momentum* anomaly, see eg. Jegadeesh (1993), Faber (2010) and Butler (2013). So both effects results directly from our MAA model.

Besides returns, using daily historical data we can also arrive at sample estimates for expected asset volatilities v_i, for expected market volatilities v (=sqrt(s)) and market correlations c_i using historical estimates over similar lookback periods. With e.g. 4 months lookback we can use around 84 days of total return data for computing historical volatilities and correlations, next to returns. Just like return momentum (for component R), we use the assumption of persistence ("generalized momentum", see Keller, 2012) to arrive at estimates for expected asset volatilities (component V), market volatilities (component M) and correlations (component C) in the future. And remember that with tactical (ie. monthly) rebalances, the future is also just months away, instead of years as in case of "strategic" MPT.

So "momentum" refers here to not only returns. Therefore we can take advantage of the flexibility of short-term momentum also for volatilities and correlations. This is relevant since, even when volatilities and correlations are more stable over time than returns, they do change, in particular in times of crisis like 2008 (see Chin 2013, Butler 2012, Schoen 2010 and New-found, 2013). That this flexibility might be relevant for asset allocation is also shown in the present discussion of low-volatility and low-beta indices and ETFs (see e.g. Blitz, 2012).

Notice that, as a consequence of the single-index model, only the N market correlations c_i (and thus N beta's) are to be estimated. This is in contrast to the full NxN covariance matrix which might be singular with limited (short-term) data. Notice also that in principle the lookback period can be different for all these components R, V, M and C. As said, for our "tactical MPT" allocation, we assume all lookback periods to have a maximum length of 12 months. As default in our empirical test we will actually use a lookback period of 4 months for all components, since this turns out to give good results for many different universes. In the next section we will look at different lookback periods for a particular universe (N=10) to check the robustness of this default.

Shrinkage. The disadvantage of using short lookback periods is that estimated expected returns, volatilities and beta's can change rapidly over time, leading to large errors and possibly to even more extreme weights in our optimal portfolio allocation than in a strategic MPT allocation. Therefore we will use simple shrinkage methods to reduce these errors[8]. See also Ledoit (2004) and DeMiguel (2009 and 2013). We will simply shrink all returns r_i towards the average return, all volatilities v_i towards the average volatility, the market variance s towards zero and all market correlations c_i towards the average market correlation by introducing "shrinkage weights" W_R, W_V, W_M and W_C (see below) for the components R, V, M, and C.

---

[7] Since the idiosyncratic variance s_i and the systematic part $(1-t/t_i)$ are always positive if $t_i>t$. We discard for the moment the rare "hedge" case where $b_i<r_i<0$ which also results in a positive share w_i. See also Elton (1976) and note 4.

[8] As an example, the shrinkage formula for return r_i equals $w*r_i+(1-w)*r_m$, where w is the shrinkage weight (w=0 and 100% corresponds to full and no shrinkage, respectively) and where r_m is the (average) market return. Shrinkage will reduce the mean-square error by decreasing the variance, while increasing the bias of the estimator.

All averages are also short-term (eg. over 4 months) and cross-sectional, so eg. the average return equals the EW return ($r_m = \sum r_i/N$).

By shrinking the market variance s, we decrease the Treynor threshold t and therefore allow for more assets in the portfolio (see A.6 in the Appendix). This puts more emphasize on diversification and less on the systematic component. When the shrunk variance s becomes zero, the threshold t is zero, and all assets (with positive returns) are included in the optimal allocation. This is the case where we say that component M is zero.

We use "weights" W_R for the component R such that a weight $W_R=100\%$ implies no shrinkage of return r_i and $W_R=0\%$ implies full shrinkage towards the (market) mean, effectively disabling the variations in returns r_i. The same applies to components V, M and C, with weights W_V, W_M and W_C between 100% (no shrinkage) and 0% (full shrinkage). Therefore the shrinkage weight W reflects the importance of a component, with a maximum effect of the component when W=100%.

Besides shrinking a component to 50% to improve the (robustness of the) estimates of the expected value (eg. for return, volatility, variance or correlation) we can also put these weights "on" (100%) or "off" (0%) to arrive at some special cases. And instead of shrinking the return r_i to the average (market) return, we can also shrink it to (or substitute it by) the volatility v_i of asset i, indicated by $W_P=100\%$ (with P for Parity)[9]. Below we have summarized all models used in the next sections in terms of shrinkage weights W, where **bold** weights are model defaults (W=0 or 100%) and the rest our (arbitrary) own defaults of W=50% (or $W_V=0\%$ for MAA-Offensive, see the next sections).

Model[10]	W_R	W_V	W_M	W_C	W_P
MAA	50%	50%	50%	50%	**0%**
MAA-Off	50%	0%	50%	50%	**0%**
MV	**0%**	50%	50%	50%	**0%**
MD	**0%**	50%	50%	50%	**100%**
RP	**0%**	**100%**	**0%**	**0%**	**100%**
EW	**0%**	**0%**	**0%**	**0%**	**0%**

Table 1. Shrinkage weights for various models

---

[9] It is even possible to make a mix of MS and MD by shrinking r_i both ways (eg. $W_R=50\%$ and $W_P=50\%$).

[10] For the last universe (N=104) we will use $W_R= 10\%$ (instead of 50%) as default choices for MS and MS-Off, and $W_V=10\%$ (instead of 50%) for MS, MV and MD, in view of the high volatility of the individual (Nasdaq100) stocks. See also section 8.

5. Tactical MPT in practice: data and methodology

In the following sections we will apply our models to three universes from 1997 to 2013 (16 year), to demonstrate the superiority of "tactical MPT" over EW and other smart-beta models.

We will present backtest results for three universes of increasing size (N=10 and N=35 both for global ETFs and N=104 for Nasdaq100 stocks), using mostly the same default shrinkages (50%) and the same default lookback periods (4 months) for all universes as parameters in order to limit the risk of datasnooping. In the next section, we will also explore our models for the first universe (N=10) in more detail for other parameter values, to test for robustness. As assets we will use global ETF's for the first two universes (N=10 and 35) for equities, bonds, alternatives, etc., for both US and abroad (IM and EM) and mainly Nasdaq stocks for the third universe. We assume monthly rebalancing.

The daily total-return data for the three universes is from Bloomberg and Yahoo and the timeframe for all backtest is December 31, 1997 to December 31, 2013 (16 year). When historical data is unavailable from the start of 1997 (for the maximum lookback period of twelve months) for certain ETFs and stocks, we will extend them to the past by using (the returns of) similar (highly correlated) index funds.

Rebalancing is done on the first close of the new month, based on the (adjusted) data for the last close of the old month. For transaction costs we will use 10 bps[11]. We assume there is no leverage possible and all trades are long only. For simplicity, we will assume that the risk-free rate is zero in all our models. [12]

The legend for the various backtest statistics is:

R = CAGR, so annual Return (in %)
V = annual Volatility (in %)
D = maximum Drawdown over the full backtest 1997-2013 (in %)
T = annual Turnover
SR = Sharpe ratio (with an annual 2.5% risk-free rate)
OR = Omega ratio (with an annual 0% target return)
CR = Calmar ratio (with an annual 5% target return)

The Omega ratio OR reflects the "gains to losses" ratio (around a target return of 0%). The Sharpe ratio SR gives the annual return R above the average historical T-Bill rate[13] divided by the volatility V. The Calmar ratio CR gives the ratio of the return above an annual 5% target return and the max drawdown D. By using a 5% return target (instead of the T-Bill rate of 2.5%) this ratio is more sensitive for higher returns (than SR and OR). By using the maximum drawdown D (instead of the volatility), the Calmar ratio CR is more sensitive for negative

---

[11] We did some sensitivity tests for higher (one-way) transaction costs than 10bps, but most results for our models in comparison to EW in terms of return/risk stay valid up to a maximum of 50-100 bps.

[12] See note 4. The risk-free rate is also relatively low over the 16 years. For the Sharpe Ratio (SR) we will use the average T-Bill rate (ca. 2.5%) as risk-free rate, for the Calmar Ratio (CR) double this as target rate (5%).

[13] The 3-month T-Bill has an average annual return (CAGR) of approximately 2.5% over the period considered (1997 – 2013).

deviations than S. Notice that all backtest statistics (including the max drawdowns) are based on monthly (instead of daily) measurements.

We prefer the Calmar ratio CR over SR (and OR) as the best metric to judge the return/risk performance of a backtest. This can be confirmed, in our opinion, by visual inspection of the equity graphs for different values of CR for different universes.

In the following we will often refer to the MAA and the MV, MD and RP submodels. Please remember that for all these models not only the short-term (momentum) shrinkage estimates are used but also the restricted covariance matrix from the single-index model. For the EW model no shrinkage or covariance restrictions are relevant. The same holds for the naive RP model, where only the (short-term) non-shrunk volatility estimates are used in order to arrive at the traditional volatility weighted allocation.

6. The small global universe (N=10)

Before we arrive at the empirical validation of our models for the two other (larger) universes, in this paragraph we explore the various corners of our models applied to the first multi-asset universe (N=10), in order to assess its robustness. The N=10 universe consists of 10 global ETFs, representing US, international and EM stocks (VTI, VGK, EWJ, EEM), two government bonds (IEF and TLT), two REITs (IYR, RWX) and two commodities (DBC, GLD). Data (daily adjusted close) is from Bloomberg (Jan. 1997- Dec. 2013), the backtests start in Jan. 1998. All lookback periods are set to 4 months, and all shrinkage weights W_R, W_V, W_M, W_C are initially (rather arbitrarily) set to 50%, except for MV ($W_R=W_P=0$), for MD and RP ($W_R=0$ and $W_P=1$), and for EW ($W_R=W_V=W_M=0$). Besides the default MAA variant there is also an "offensive" variant (MAA-Off) where W_V is set to zero. By changing W_V we can actually control V nearly linear.

In fig. 1 we present the statistics and the equity graph (log scale) for all our six models including EW for this small global universe (N=10). As can be seen, the MAA models are clearly superior not only to EW but also to the MV, MD and RP submodels. These latter models are very similar in return R as the EW benchmark, but with smaller max drawdown D (and volatility V) and therefore with better (higher) return/risk figures as expressed by the Omega, Sharpe, and Calmar ratios. In terms of the CR ratio both MAA models wins hands-down, while the MAA-Off variant delivers slightly higher returns R without much decrease in return/risk. We include this variant because it shows how our models can show some "pseudo-leverage" (without lending) by fully shrinking the V component ($W_V=0$). Notice that turnover T is near zero (minimal) for the EW model, with the MD and RP models second and the MV model third, and maximal for both MAA models. In the rest of this section we will examine this global multi-asset universe (N=10) in somewhat more detail to get a feeling of the robustness of our models.
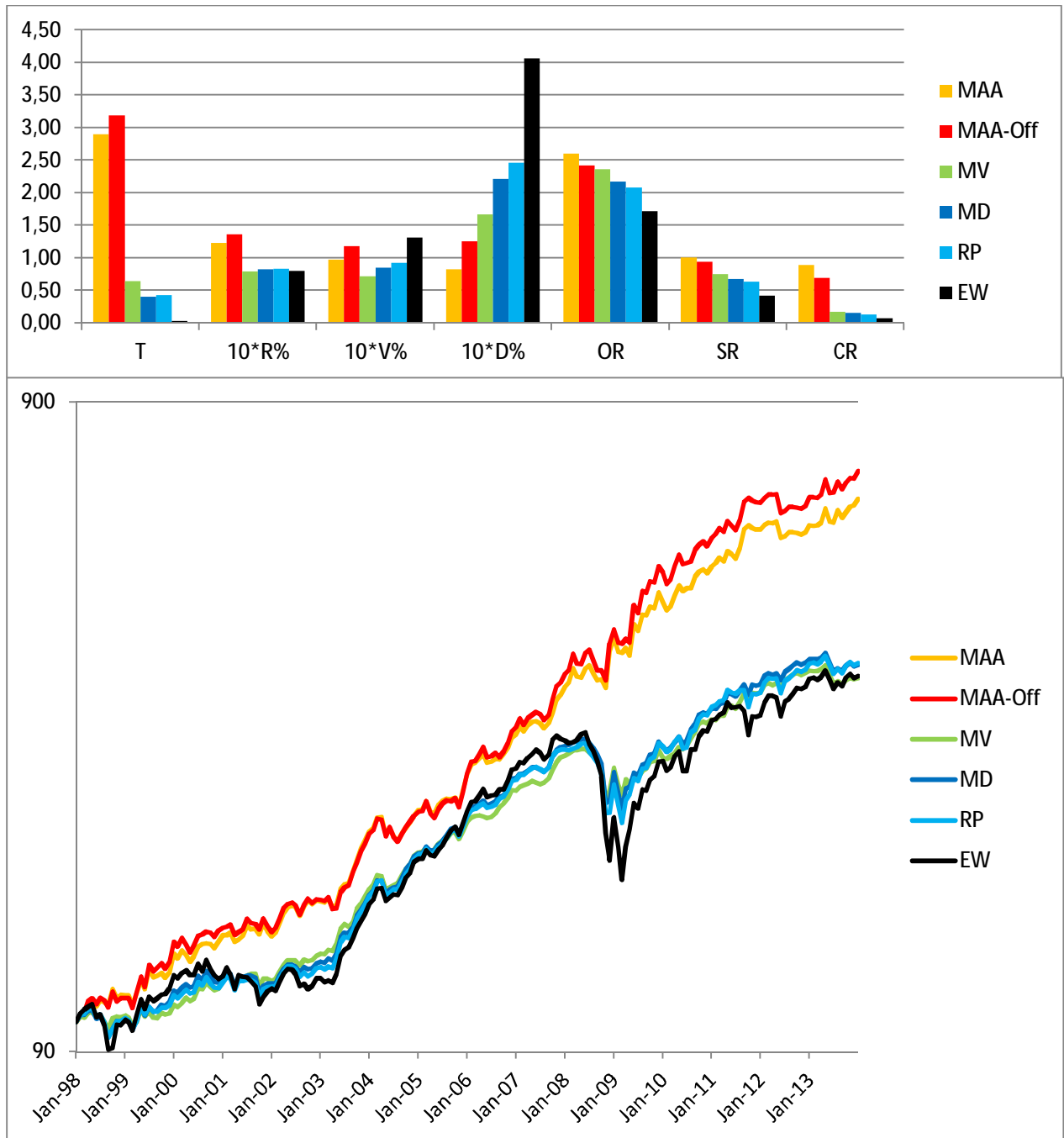
Fig. 1. Statistics and equity line (log scale) for the small global universe (N=10)

Decomposing the MAA model into components . In this section we will examine the impact of the different components R, V, M and C on the various return/risk ratios of the MAA model, by rebuilding the model step-by-step, starting at the EW benchmark and adding these components one by one, using shrunk and non-shrunk estimates.

When we assume that components R, V, and C are all irrelevant, we can shrink the corresponding estimates fully to their means by putting the shrinkage weights W_R, W_V and W_C equal to zero. In this case, we arrive at the **Equal Weight** (EW) model as the market index

(or the benchmark) for which $W_R=W_V=W_C=0$. This is our starting point of fig. 2 (left bar), where we show the Sharp and Calmar Ratio (CR left, SR right) for EW. In the graphs we also show these return/risk ratios for an increasing number of components (R, RV, RVM, and RVMC), including the absolute momentum (AM) case where only the sign instead of the return size (as in R) is used in eq. (1). We give these ratios for both the default shrunk model (with all W=50%) as the non-shrunk (NS) model (with all W=100%). As an example, for the RVM case we have $W_R=W_V=W_M=50\%$ (or =100% for NS) and $W_C=0\%$. This is the Constant Correlation model (here for MAA). Notice that the RV case has no systematic (market) effect, while the RVMC case reflects the systematic (market) effect including different correlations.
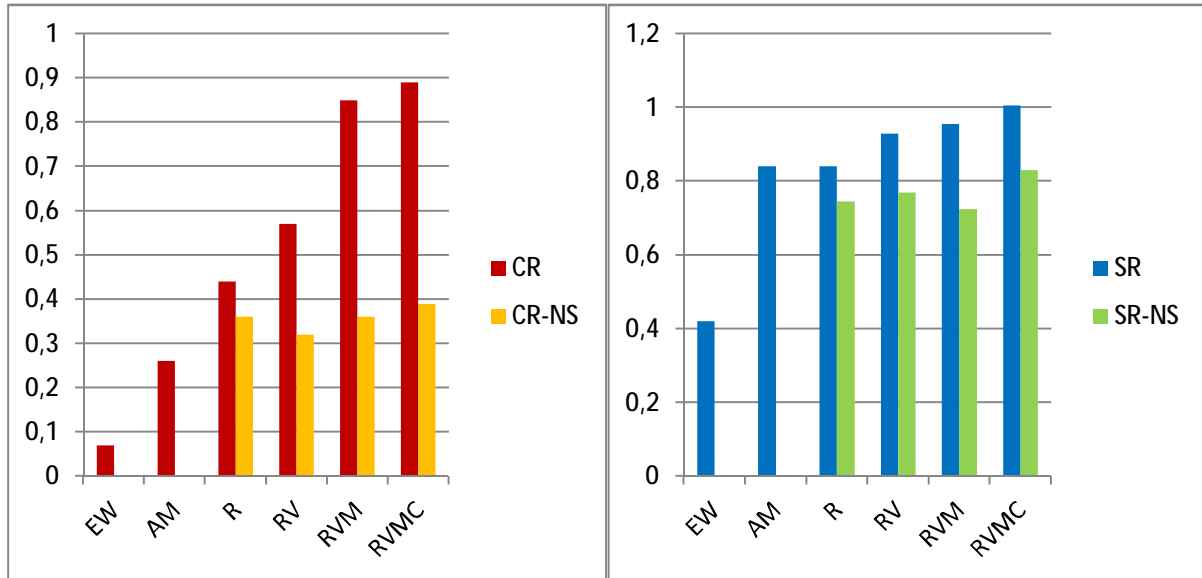

Fig. 2. Effect of factor combinations / EW on the Sharp (SR) and Calmar (CR) ratios, shrunk and non-shrunk (NS) for MS (N=10)

Both the Sharpe and the Calmar ratio improves monotonic for the shrunk model when the number of components increases from EW and AM towards to full RVMC model. This is not the case for the unshrunk components, where also the return/risk ratios are lower than in the shrunk case. Notice the improvement of the Calmar ratio CR from EW towards RVCM for the shrunk case. The same holds for the Sharp ratio (SR) from EW/AM/R. There is (in terms of SR) not much difference between AM and R, but adding V, M and C improves the return/risk performance, as do shrinkage.

Different shrinkage weights. In fig. 3 we present the effect on the Calmar ratio (CR) for the MAA model (N=10) as a function of different shrinkage weights (W_R, W_V, W_M, W_C) plus the scores for EW (right bar) for comparison. For example, in the left group we see the effect on CR of changing W_R (=0,10, .., 100%), holding all the other weights (W_V, W_M, W_C) at the default value (W=50%) for the MAA model. The effect on the Sharpe ratio (not displayed) is very similar to that on the Calmar ratio but less pronounced.

From the graph it is clear that the effects of the shrinkage weights W_R on the return/risk ratio CR are the most substantial of all four components, with the best CR around $W_R=50\%$ (red bar). The next components which show some sensitivity are W_V and W_M, with the same pattern. The effect of W_C is negligible in this case (MAA, N=10), making clear that the effect of different correlations is very limited. All CR ratios are much better than those for EW (black bar), including those for the non-shrunk case (W=100%, orange bar). This also shows that

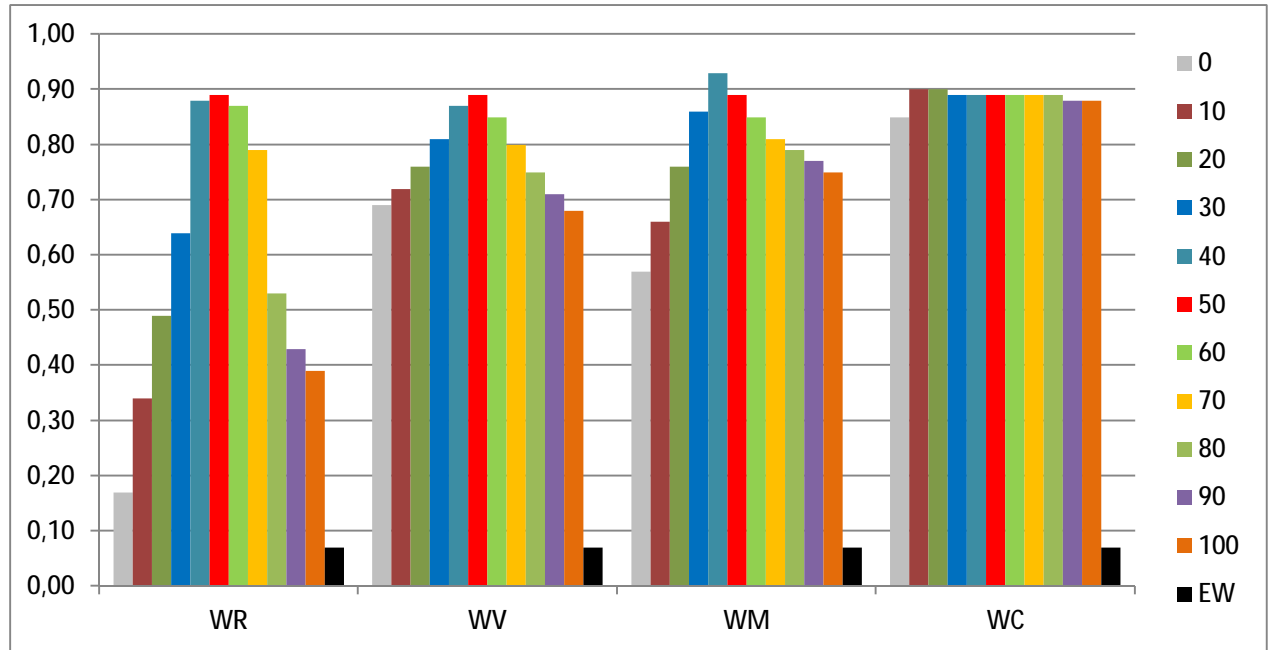(for N=10) our MAA results relative to the EW benchmark are rather robust for different degree of shrinkage.

Fig. 3. Effect of shrinkage weights and EW on Calmar ratio (CR) for MS (N=10)

Different lengths of the lookback period. Now we will check the robustness of the MAA model (N=10) for a different lookback period for R, V, M and C (returns, volatilities, variances and correlations) for the N=10 universe, both for 1998-2005 (8 years) and the default 1998-2013 (16 years). See fig. 4 where we show the effect on the Calmar ratio (CR) for lookback lengths of 1- 6, 9 and 12 months in months plus EW. For all our models in this paper, we used a default lookback period of length 4 months for the four components R, V, M, C. All lengths in fig. 4 are also the same for all components. So when we use eg. a lookback period of 12 months, the returns, volatilities/variances and correlations are all estimated on a historical lookback period of 12 months.
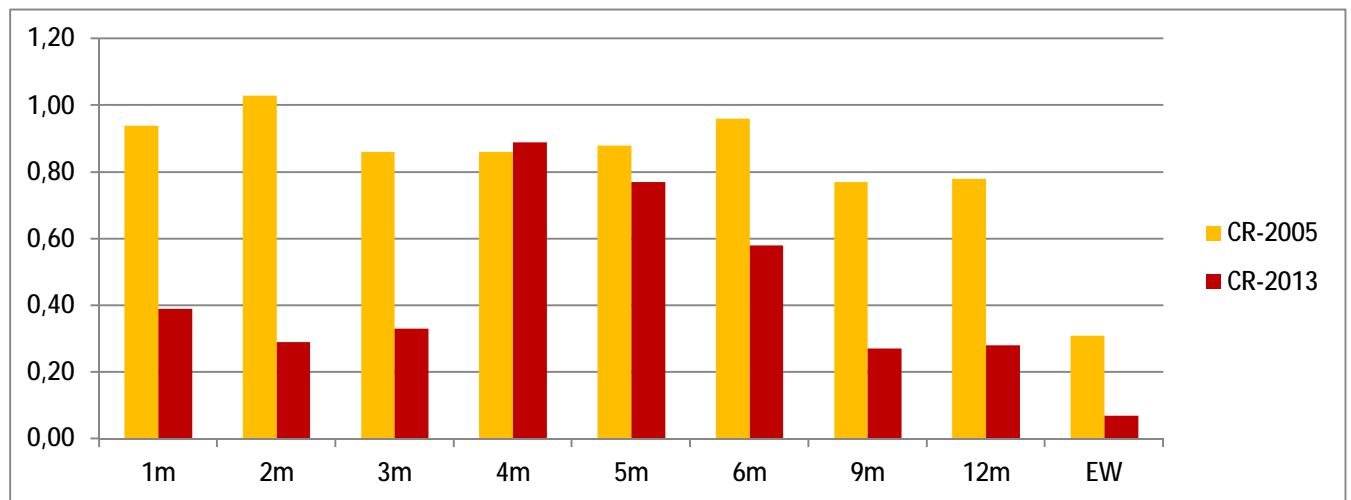

Fig. 4. Effect of lookback length (months) and EW on the Calmar ratio (CR), for 1998-2005 and 1998-2013 for MS (N=10)

We assumed again a default shrinkage of 50% for all our four components. Notice that for the longest lookback period of 12 months we needed data from begin 1997.

From fig. 4 we conclude that for the full 16 year period (1998-2013), a lookback length of four months is clearly optimal. For the first 8 year period (1998-2005, excluding the financial crisis in 2008/9), all short lengths (1-6 months) are good, with a slight optimum at two months. So four months seems a good compromise. Per period, all CR scores are always better than the EW benchmark.

7. The large global universe (N=35)

In this paragraph we examine our models applied to a larger global universe (N=35). The N=35 universe consists of 35 global ETFs (VTI, IWM, VIG, QQQ, XLF, XLY, XLP, XLU, XLV, XLB, PFF, VGK, EWJ, EPP, SCZ, FXI, ILF, EWX, SHY, IEI, IEF, TLT, TIP, MUB, MBB, CIU, LQD, HYG, BWX, EMB, VNQ, RWX, DBE, DBC, DBP), extended to 1997 by corresponding index funds if necessarily. Data (daily adjusted close) is from Yahoo (Jan. 1997- Dec. 2013), the backtests start in Jan 1998. We use the same models, weights and lookbacks as for the N=10 universe, incl. W_V=0 for MAA-Offensive.

In fig. 5 we present the statistics and the equity graph (log scale) for all the models including EW for this universe (N=35). Disregarding the MAA-Off model (red bar), nearly all conclusions for the N=10 also holds for this universe. In particular, the risk/return statistics OR, SR and CR are best for MAA and subsequently lower for MV, MD, RP and EW. In particular the Calmar ratio CR is dramatic low for MV, MD, RP and EW because of low return R and high max drawdown D, while the Sharpe (SR) and Omega (OR) ratio are lowest for EW.

The most interesting model is the offensive MAA-Off (with W_V=0%) which has a much better return (R=14.3%) than MAA (10.7%) but with nearly double the volatility (V=14% vs. 8%) and more than double the max drawdown (D=17% vs. 7%). Still, this form of "pseudo-leverage" easily wins over EW (R=9.5%, V=14%, D=37%) and is in return/risk scores SR and CR only bypassed by MAA.
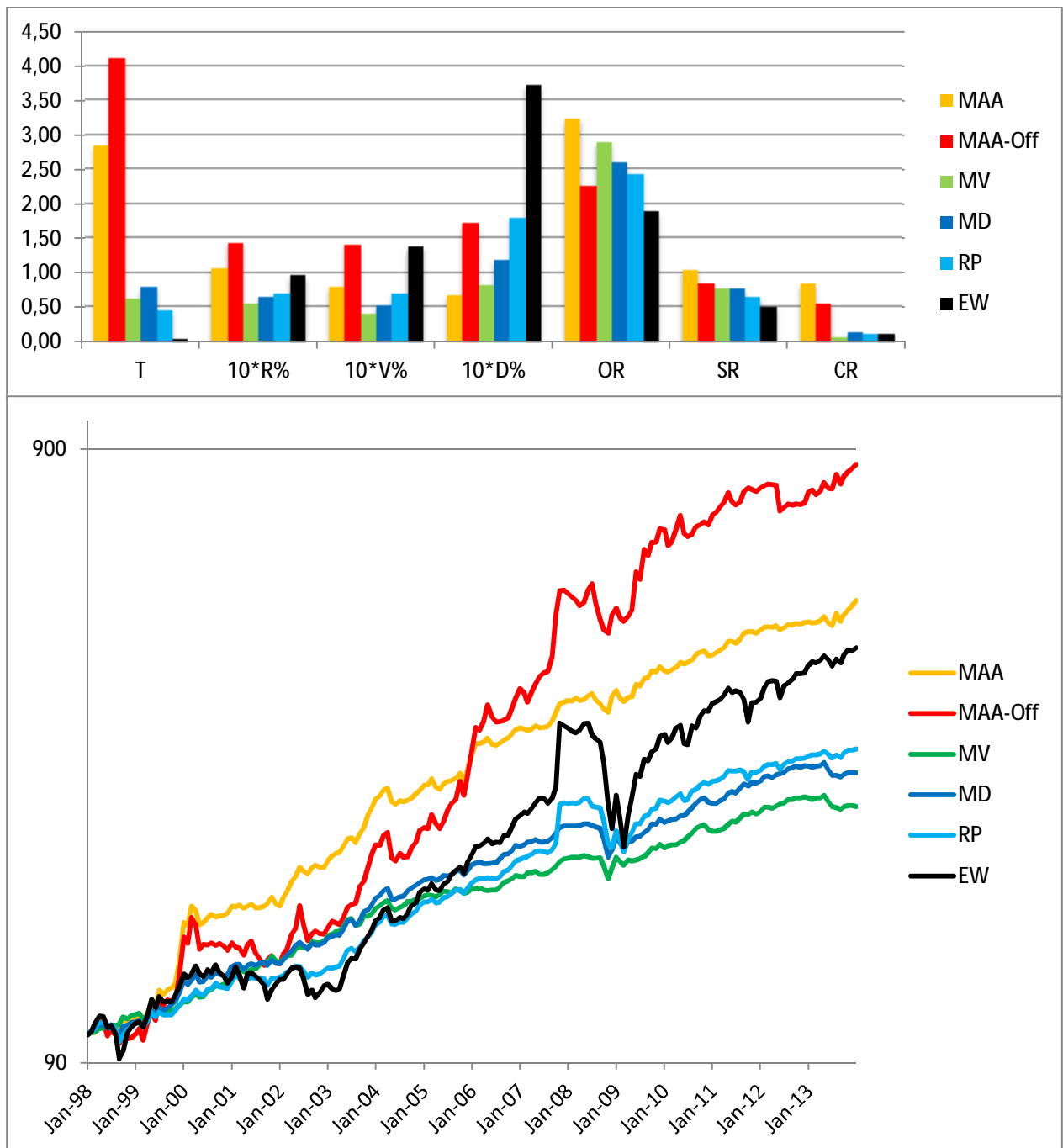
Fig. 5. Statistics and equity line (log) for the large global universe (N=35)

8. A large US stock universe (Nasdaq100)

In this paragraph we examine our models applied to the third and final universe (N=104), which consists of all the recent 100 Nasdaq100 stocks together with 4 US government bond ETFs (IEI, IEF, TLT and EDV). All stocks are extended to 1997 by the Nasdaq-100 index (^NDX) [14] and by corresponding index funds for the bonds. Data (daily adjusted close) is from Yahoo (Jan. 1997- Dec. 2013), the backtests start in Jan 1998. Because of the much higher volatility of individual stocks we decided to shrink W_V and W_R further down to 10% (instead of the default 50%), but else we have used the same models, weights and lookbacks as for the N=10 and 35 universe, incl. W_V=0 for MAA-Offensive.

In fig. 6 we present the statistics and the equity graph (log scale) for all our models including EW for this universe (N=104). Nearly all conclusions for the N=10 and N=35 also holds for this universe. In particular, the Calmar ratio CR is best for MAA and MAA-Off (around CR=1.5), and subsequently lower for MAA-MV, MAA-MD (around CR=1.1), and dramatic lower for RP and EW (around CR=0.4) because of high maxdrawdowns (D=38% and 47%, respectively).

Omega (OR) and Sharpe (SR) ratios are similar for MAA, MAA-Off, MAA-MV and MAA-MD (around OR=4 and SR=1.6, resp.), while the Calmar ratio is dramatically low (around CR=0.4) for RP and EW.

Notice that MAA-Off scored an incredible return (R=36% annual) without any leverage, but also the other models fared well with the lowest return (R=20%) for RP and the best for MD and MAA (29%), with even 24% for EW (including the 4 bonds). Volatility V is between 13 and 21% for all non-EW models and 27% for EW. Besides both MAA models also the performance of MAA-MD and MAA-MV is impressive.

---

[14] Some stocks available only recently (like Facebook), are extended by us by ^NDX (the Nasdaq100 index) back to 1997. By doing so, we reduce the survivorship bias somewhat since ^NDX also includes non-survivors from the past. However, the combination of a high beta and average return of ^NDX often prohibits selection into the optimal allocation.
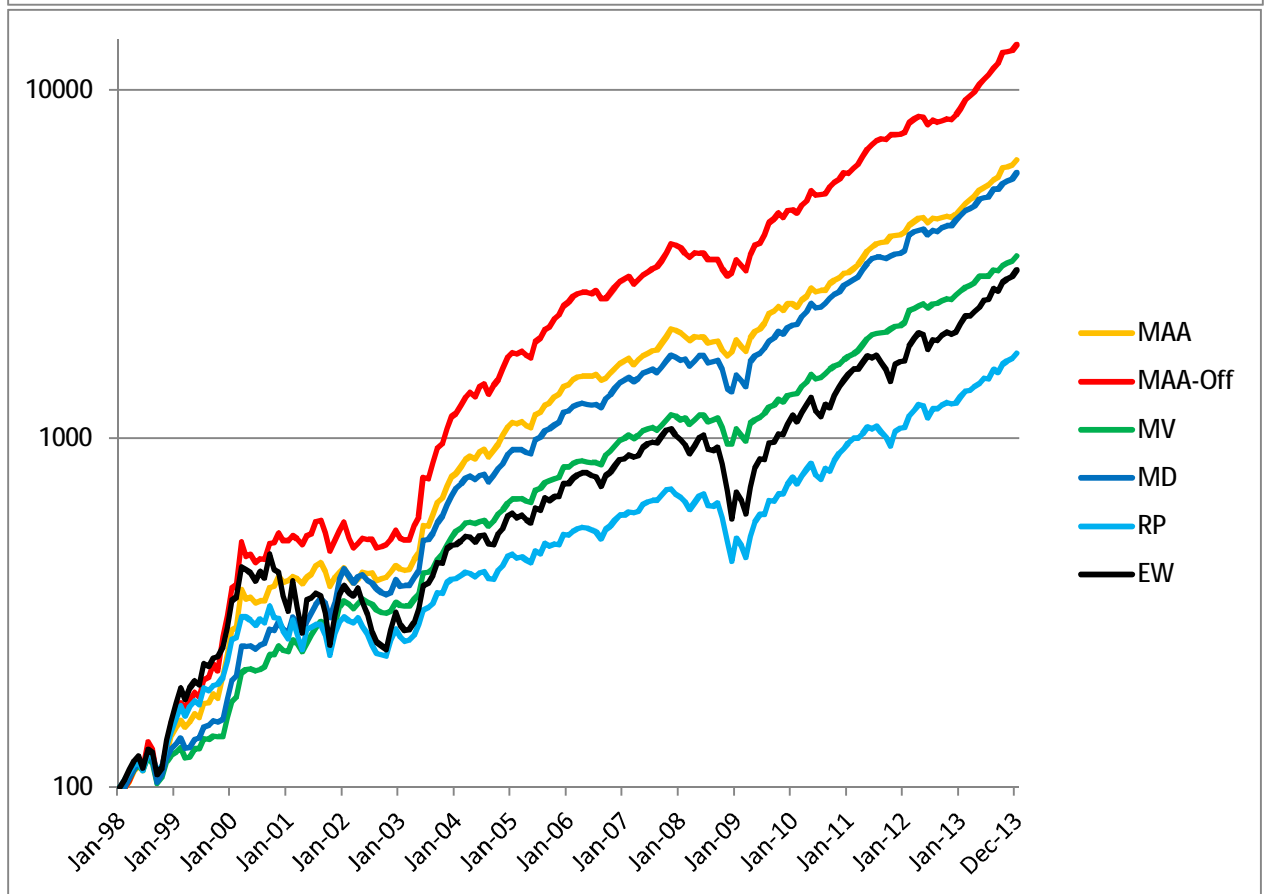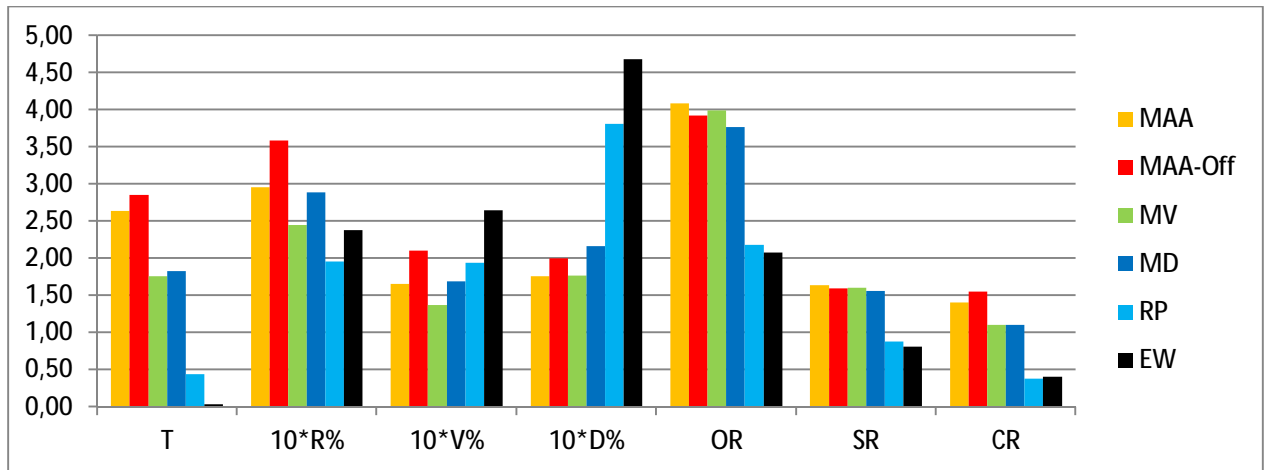
Fig. 6. Statistics and equity line (log) for the large stock universe (N=104)

9. Conclusions

We have tried to improve upon MPT by using a more tactical, analytical and practical approach. In our tactical MPT model, we use short-term lookback periods in order to have more flexibility and to benefit from the momentum effect. This small step has the biggest impact on the success of MPT, as we have proved empirically. In addition we used a simple four-component single-index model to enable an elegant analytical interpretation of the long-only allocation. Finally we add shrinkage of the estimators for returns, volatilities, and correlations to arrive at our practical solution. This includes maximum Sharpe (MAA) and "smart-beta" models, like minimum variance (MV), maximum diversification (MD), and naïve risk parity (RP) models. We run monthly backtests from 1998 to 2013 (16 years) for three universes of 10 and 35 global ETFs and 100 Nasdaq stocks. All our models beat EW in terms of various return/risk statistics and in nearly all cases our MAA models beat the smart-beta models.

For future research we would like to consider more advanced (eg. EMA and GARCH like) models for estimating the expected R, V, M and C components rather than the simple 4 month lookback. We also have done some preliminary tests with the unrestricted tactical MPT model, using the Critical Line Method (CLA) of Markowitz (see also Nawrocki, 1996) to numerically invert the covariance matrix. In addition we did some datasnoopings test on our MAA models based on Bailey (2013). The first results for both look promising.

In conclusion, we think there are enough topics for future research when we take a more 'tactical' approach to good old MPT. So yes, the reports of the death of Markowitz's MPT have been greatly exaggerated!

Appendix. Proof of eq. (1)

What follows is a matrix representation of the proof by Elton (1976). Let **S** be the expected (semi-positive definite and symmetrical) NxN covariance matrix, **w** the optimal weight vector and **r** the vector of expected (excess) returns, **l** a vector of one's, all of length N. Then the MAA solution maximizes

(A.1) **r'l** / sqrt(**w'Sw**),   **w'l** =1

The solution is given by the vector of optimal weights

(A.2)   **w** = s_p **S**$^{-1}$ **r**

where s_p is the normalization constant (and variance of the optimal portfolio), following from the constraint $\sum w_i = 1$. The single-index model (SIM) is, for assets i=1,..,N

(A.3) $r_i = a_i + b_i r_m + e_i$

where r_m is the market or index (excess) return, a_i is the "alfa" for asset i, b_i is the market (or index) "beta" for asset i, and e_i is the "idiosyncratic" residual for asset i, assumed to be independently and randomly distributed. When r_m equals the EW index, there is a small dependency which we will disregard for simplicity and since it is of the order 1/N and therefore small for large N (see Fama, 1968).

The NxN covariance matrix **S** equals, given the single-index assumption,

(A.4)   **S** = s **bb'** + Diag (**s^e**)

where **b** equals the N-vector of beta's b_i (=$v_i c_i/v$), s (=v^2) the market variance, **s^e** the N-vector of residual (or idiosyncratic) variances s_i (= $v_i^2 - s b_i^2$), and c_i the correlation of asset i with the market, i=1..N. The inverse of the matrix **S** equals (see eg. Clarke, 2012)

(A.5)   **S**$^{-1}$ = Diag (1/**s^e**) − (**b/s^e**)(**b/s^e**)' / (1/s + (**b/s^e**)'**b**)

Substitution of eq. (A.5) in eq. (A.2) gives eq. (1) with the "long-only" *Treynor threshold* t

(A.6)   t = (s $\sum_p r_j b_j/s_j$) / (1 + s $\sum_p b_j^2/s_j$)

where $\sum_p$ equals the summation over all assets j in the portfolio (ie. with w_j>0) and $t_i = r_i / b_i$ equals the Treynor ratio for asset i (i=1,..,N). This can easily be implemented when b_i>0 by sorting all assets on their Treynor ratio (highest t_i first) and computing the Treynor threshold t for those assets already included until an asset t_i exceeds the Treynor threshold t (see also Elton 1976). In practice one computes eq. (1) starting with t=0 and iterates then between t and w_i until convergence (which is often very fast). This also works when b_i<0 for some i.

In the MV (full shrinkage of all returns to r_m) and MD case (to v_i), we can compute the beta threshold b=1/t and the correlation threshold c=1/t assuming r_m=1, since r_m cancels in the quotient t_i/t in eq. (1) while the constant r_m is absorbed in the normalization constant. Then eq. 1 becomes wi ~ (1-b_i/b) / si for bi <b else w_i=0 for MV, and w_i ~ (1-c_i/c) / s_i for c_i <c else w_i=0 for MD. We can also proof that in case of MAA with constant correlation (c_i independent of i) there holds w_i ~ (1-h/h_i) / s_i for h_i>h else w_i=0 where h_i=r_i/v_i equals the Sharpe ratio and h=1/t is the Sharpe threshold, again assuming r_m=1.

References

Ang, A., 2012, Mean-Variance Investing, SSRN 2131932

Antonacci, G., 2013, Absolute Momentum: A Simple Rule-Based Strategy and Universal Trend-Following Overlay, SSRN 2244633

Bailey, D.H., Borwein, J.M., López de Prado, M., Zhu, J., 2013, The Probability of Backtest Overfitting, SSRN 2326253.

Blitz, D., vanVliet, P., 2012, Low-Volatility Investing: Collected Robeco Articles, Robeco, Rotterdam

Butler, A., Philbrick, M., Gordillo, R., 2013, Adaptive Asset Allocation: A Primer, SSRN 2328254

Chin, C.J., 2013, Correlations Have Personality, Too: An Analysis of Correlations between Assets, paper, Price Asset Management

Choueifaty, Y., Froidure, T., Reynier, J., 2011, Properties of the Most Diversified Portfolio, SSRN 1895459

Clarke R., De Silva, H., Thorley, S., 2011, Minimum Variance Portfolio Composition, The Journal of Portfolio Management 37/2 Winter, pp.31-45

Clarke R., De Silva, H., Thorley, S., 2012, Risk Parity, Maximum Diversification, and Minimum Variance: An Analytic Perspective, SSRN 1977577

DeMiguel, V., Garlappi, L., Uppal, R., 2009, Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?, The Review of Financial Studies 22/5, pp.1915-1953

DeMiguela, V. , Martin-Utrerab, A., Nogalesb, F., , 2013, Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection, SSRN 1891847

Elton, E.J. Gruber, M.J., Padberg, M.W., 1976, Simple Criteria for Optimal Portfolio Selection, Journal of Finance 31, 5, pp. 1341-1357

Faber, M. T. (2007), A Quantitative Approach to Tactical Asset Allocation. Journal of Wealth Management, Spring 2007. Update (2009) as SSRN 962461.

Faber, M. T. (2010), Relative Strength Strategies for Investing, SSRN 1585517

Fama, E.F, 1968, Risk, Return and Equilibrium: Some Clarifying Comments, Journal of Finance, 23, pp. 29-40,

Hallerbach, W.G., 2013, Advances in Portfolio Risk Control. Risk ! Parity ?, SSRN 2259041

Hurst, B., Ooi, Y.H., Pederson, L.H., 2012, A Century of Evidence on Trend-Following Investing, working paper, AQR Capital Management.

Jacobs H., Müller, S., Weber,M., , 2013, How should individual investors diversify? An empirical evaluation of alternative asset allocation policies, SSRN 1471955

Jegadeesh, N., Titman,S.,1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, Journal of Finance XLVIII, 65/91.

Jurczenko, E., Michel, T., Teiletche, J., 2013, Generalized Risk-Based Investing, SSRN 2205979

Keller, W.J. and Van Putten, H., 2012, Generalized Momentum and Flexible Asset Allocation (FAA): An Heuristic Approach, SSRN 2193735

Keller, W.J. and Van Putten, H., 2013, Tactical MPT and Momentum: The Modern Asset Allocation (MAA), SSRN 2373086

Kritzman, M., Page, S., Turkington, D., 2010, In Defense of Optimization: The Fallacy of 1/N, Financial Analysts Journal 66-2, pp. 31-39

Ledoit, O., and Wolf, M., 2004. Honey, I Shrunk the Sample Covariance Matrix: Problems in Mean-Variance Optimization. Journal of Portfolio Management 30:110–19

Nawrocki, D., 1996, Portfolio analysis with a large universe of assets, Applied Economics 28, pp. 1191-1198.

Newfound, 2013, Allocating Under Uncertainty: Simple Heuristics & Complex Models, paper, Newfound Research.

Niedermayer, A., Niedermayer, D., 2006, Applying Markowitz's Critical Line Algorithm, Discussion paper 06-02, University of Bern, Dept. of Economics.

Maillard, S., Roncalli, T., Teiletche, J., 2009, On the properties of equally-weighted risk contributions portfolios, SSRN 1271972.

Markowitz, H. M., 1952, Mean-variance analysis in portfolio choice and capital markets, Journal of Finance 7, 77–91

Roncalli, T., 2013, Introduction to Risk Parity and Budgeting, Chapman & Hall/CRC Financial Mathematics Series.

Sharpe, W.F., 1963, A Simplified Model for Portfolio Analysis, Management Science, Vol. 9, No. 2, pp. 277-293

Scherer, B., 2010, A New Look At Minimum Variance Investing, SSRN 1681306

Schoen, R.J., 2012, Parity Strategies and Maximum Diversification, Putnam Investments, paper, Putnam Investments

Treynor, J.L.,1966, "How to Rate Management of Investment Funds", Harvard Business Review 43, pp.63-75