Breadth Momentum and Vigilant Asset Allocation (VAA);
Winning More by Losing Less

By Wouter J. Keller and Jan Willem Keuning[1]
July 14, 2017, v0.99

Summary

VAA (Vigilant Asset Allocation) is a dual-momentum based investment strategy with a vigorous crash protection and a fast momentum filter. Dual momentum combines absolute (trendfollowing) and relative (strength) momentum. Compared to the traditional dual momentum approaches, we have replaced the usual crash protection through trendfollowing on the asset level by our *breadth momentum* on the universe level instead. As a result, the VAA strategy is on average often more than 50% out of the market. We show, however, that the resulting momentum strategy is by no means sluggish. By using large and small universes with US and global ETF-like monthly data starting 1925 and 1969 respectively, we arrive out-of-sample at annual returns above 10% with max drawdowns below 15% for each of these four universes.

1. Introduction

The VAA (Vigilant Asset Allocation) strategy described in this paper is a follow-up on our PAA strategy (Protective Asset Allocation, see Keller, 2016), targeting higher annual returns while at the same time offering stricter crash protection than PAA. Our target for VAA is offensive returns with defensive risks: winning more by losing less. To be more precise, with VAA we aim at moderate but offensive returns above 10% but with defensive drawdowns of less than 20%, preferably less than 15%. We will use monthly data starting Dec 1925 and Dec 1969 respectively for various US and global asset classes (as proxies for present day ETFs). More on our data later (see section 7).

VAA is part of the class of momentum based tactical asset allocations. Momentum (or "price persistence") can be applied to stocks (see eg. Jegadeesh, 1993) or asset-classes (see eg. Faber, 2007). We focus here on asset-classes (or assets, for short). *Relative* momentum (also called cross-sectional or relative strength momentum, see eg. Faber 2010, Moskowitz 2011, Asness 2014 and Faber 2015) uses only the best (top T) performing assets within a universe (of size N>T) while *absolute* momentum (also called time-series momentum or trendfollowing, see eg. Moskowitz 2011, Antonacci 2013a, and Levine 2015) selects only the assets with positive momentum. The combination of absolute and relative momentum is often called *dual* momentum (Antonnacci, 2013b, 2014). For a historical overview, see also Faber (2013), Newfound (2015) and Antonacci (2014). As with most dual momentum models we will restrict ourselves to long-only trading strategies (no short-selling) with monthly portfolio reforms and rebalances.

[1] We thank Adam Butler, Walter Jones, Steve LeCompte, Bas Nagtzaam, Michael Roovers, and Valeriy Zakamouline for useful comments on earlier versions of this paper. All errors are ours.

Electronic copy available at: https://ssrn.com/abstract=3002624

To arrive at a vigilant model, we apply a "vigilant (fast and vigorous) crash protection (CP) strategy based on a "fast" filter for absolute momentum. We define a "bad" asset as an asset with non-positive momentum. Our crash protection is based on the number of bad assets in the universe, instead of the replacement of individual bad assets, as with traditional absolute (and dual) momentum approaches. So, we use a kind of market breadth (in terms of our fast momentum filter) as crash indicator, as we did with PAA. Therefore, our *breadth momentum* extends the traditional absolute momentum when it comes to crash protection.

Compared with PAA, however, we will use the number of bad assets relative to a protection threshold as a more granular crash indicator. As we will show, in-sample optimization of this threshold frequently results in an out-of-market allocation (i.e. cash), even when just one or only a limited number of the assets are bad. As a result, we will be in cash on average more than 50% over time in our four universes. When not fully in cash we will use the same (fast) filter for relative momentum to arrive at a more offensive strategy with only a limited number (top T) of best performing assets.

As we will see later, our "cash" concept is not limited to a risk-free asset but also includes various bonds in a separate "cash universe", besides the main universe of "risky" assets (which also might include bonds). Analogous to our relative momentum strategy for the risky universe, we will always choose the single best performing bond for cash, but without looking at the sign of the momentum of these bonds (ie. no absolute momentum for cash).

As return measure, we will focus on CAGR (Compound Annual Growth Rate, so geometrical returns) while for risk we mainly take maximum drawdown in consideration, since we believe this is often felt as the more important risk indicator for investors than the traditional volatility measure. As our (in-sample) optimizing target we will therefore use a new return/risk measure: returns adjusted for drawdowns (RAD), besides the more familiar Sharpe (excess return/volatility) and MAR ratio (return/max drawdown). More on RAD later (see section 6).

2. Fast momentum

For our momentum filter we will use a variant of the often used 13612 filter (see eg. Faber, 2007 and Keller, 2015), but now with an even faster response curve by using the average *annual* returns over the past 1, 3, 6 and 12 months. We will denote this filter as 13612W. The traditional 13612 filter uses the average *total* returns over the same four periods.

In Fig. 1 we give the monthly return weights[2] for various momentum filters (see also Beekhuizen, 2015, Zakamulin, 2015, and Keller, 2016, note 2), including our new 13612W filter. Notice that our faster 13612W filter gives a weight of 40% (19/48) to the last month return as compared to 8%, 15% and 18% for the simple 12-month return filter (RET12) as used eg. by Moskowitz (2011) and

[2] Monthly return weights are defined as the effect on the momentum filter of the return in a particular historical (lagged) month.

Antonnacci (2013a), the SMA12-based filter as used by Keller (2016) for PAA, and the 13612 filter as used by Faber (2010), Hurst (2012), and Keller (2015), respectively[3].

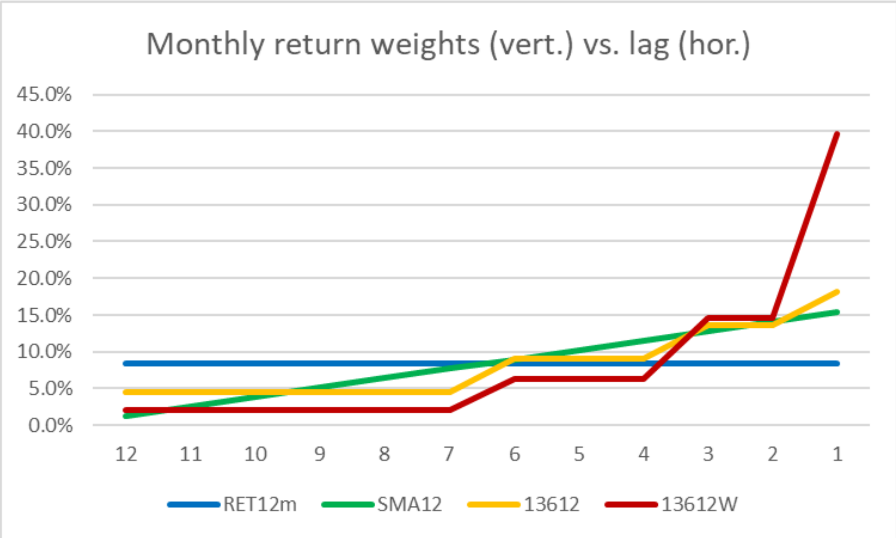

Fig.1. Monthly return weights for RET12, SMA12, 13612, and our 13612W filter

Because of the heavy weight to the last month in 13612W, we assume no 1-month reversal effects in our asset-class data. Although 1-month reversals are common for individual stocks, we found the opposite for most of our asset-class data. In-sample testing also shows better return/risk performance when using the 13612W filter instead of the 1-month filter or one of the other momentum filters (SMA12, RET12, and 13612).

Now we can define our momentum measure of asset (class) i=1..N for (the end of) each month as the average weighted lagged price according to the 13612W filter applied to the (dividend etc. adjusted ie. Total Return/TR) lagged monthly prices of the asset. Notice that we need 12 lagged prices for the 13612W filter, so our backtests always start one year after the start of the data. We will use the same 13612W filter for absolute and relative momentum in contrast to other authors (eg. Faber, 2007 using dual momentum). [4]

3. Fast Crash Protection

For our crash protection (CP) we go along the lines of our PAA model where we used the number of so-called *bad assets* b (with non-positive momentum) in our universe to define the degree of cash. We will call this universe based approach *breadth momentum*, in contrast to traditional absolute momentum, which operates at the individual asset level by trendfollowing. Breadth momentum was

[3] The RET12 filter equals p0/p12 -1, the SMA12 filter equals p0/SMA12 -1, the 13612 filter equals p0/p1+p0/p3+p0/p6+p0/p12 -4, and the new 13612W filter equals 12*p0/p1+4*p0/p3+2*p0/p6+1*p0/p12 -19, where pt equals price p with lag t.

[4] With different filters for traditional absolute and relative momentum, it is possible that eg. in the Top T, the second best asset is replaced by cash but not the third best, which seems less logical. Therefore, we will always use the same filter for both relative and absolute momentum.

also used with our PAA strategy (Keller, 2016). However, the crash protection algorithm for VAA allows for much more granularity and aggressiveness than with PAA.

We therefore define the *breadth protection threshold* B (or “breadth B”, for short) as the minimal number of bad assets b for which we go 100% to cash, while we use the fraction b/B (for b<=B<=N) as Cash Fraction, CF. In formula:

CF= b/B if b<B, and CF=1 when b>=B (with b=0,1,..,N, and B<=N)

Notice that this is strikingly different from the traditional dual momentum approach (see Faber 2007, and others[5]) where CP is based on the number of bad assets in the relative momentum based best (top T) assets. So VAA allows us to go to cash more vigorously (ie. faster). When B<N/2, VAA is also faster than the most protective PAA variant, PAA2. In fact, the crash protection of PAA2 is equivalent to VAA with B=N/2 except for the momentum filter: SMA for PAA, 13612W for VAA.

Fig. 2 illustrates the vigorous crash protection (CP) of VAA. Here we show the cash fraction CF as a function of the number b of bad assets for three strategies (all with universe size N=12 and Top T=3):
- VAA with B=4
- PAA2 (=VAA with B=N/2=6)
- Dual momentum, with T=3

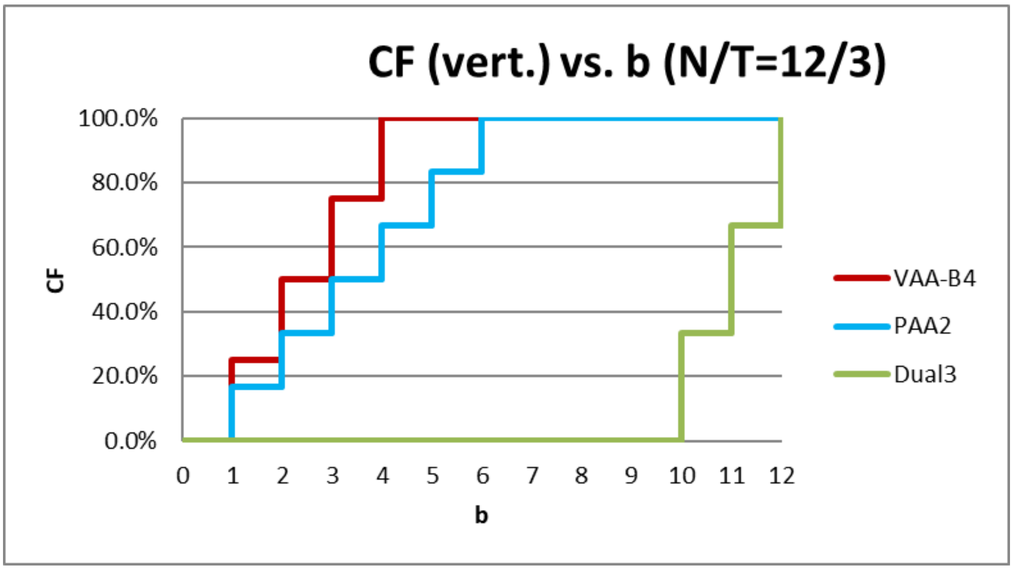

Fig. 2. CF(b) for b=0..12, N/T=12/3 for VAA (B=4), PAA2 (i.e. B=6) and Dual (T=3)

Dual (Dual momentum) is defined here as the traditional dual momentum strategy where you select the best performing top T assets with share w=1/T (equal weight) while replacing the bad assets (ie.

[5] Notice that Antonacci (2014) used a slightly different definition of Dual momentum for his specific GEM strategy, using only SPY for CP (see his p.98), in contrast to his informal (more usual dual) approach (in his flowchart on p. 101). Our definition of dual corresponds to Antonacci (2013b) and Faber‘s GTAA approach (2007, 2010).

with non-positive absolute momentum) in this top T by cash[6]. As can easily be seen from the figure, for a given b<N (and B<=N) the cash fraction CF for VAA will always be more than that for Dual, ie. $CF_{VAA}(b) > CF_{Dual}(b)$ for b=0,..,N-1. Only when b=N (all assets bad) there holds CF=100% for VAA, PAA and Dual.

Notice again that for Dual the chosen relative momentum filter (for the top T selection) is sometimes different from the absolute momentum filter (see eg. Faber 2007). In contrast, for VAA our fast 13612W filter is used for both types of momentum. Notice also that absolute momentum (the direction of the trend) in VAA works at the universe or market level while it works at the individual asset level in Dual approaches. In other words, in VAA absolute momentum only defines the number of bad assets b (and therefore CF) in the universe and not the individual bad assets to be replaced by cash, as in Dual.

In fact, one might recognize in the ratio (g-b)/N (with g=N-b being the number of *good* assets, ie. with positive momentum) the well-known *market breadth,* now based on our 13612W momentum filter. This market breadth ratio equals one when all assets are good, minus one when all assets are bad and zero when the market is fifty/fifty. So, you might say that our cash fraction is a function of this market breath and the relative protection threshold B/N. This also explains our term *breadth momentum* as the force behind our aggressive crash protection.

4. Easy Trading and the cash universe

As is shown in Fig. 2, the CF fractions differ between VAA, PAA and Dual for b<N. With dual momentum (Dual), the selected top T<N asset fractions equal the cash fractions. If eg. T=3 all three selected assets have an equal share of 33%, which is replaced by an equal share of cash in case of bad assets in the top T. This will make trading easy: you simply replace bad assets in the Top T by cash.

Trading is, however, less easy with VAA. Take, for example, the case where N=12, T=3 and B=4 as displayed in Fig. 2. Starting with no bad assets at all in the universe of N=12 assets, so b=0 and w=1/3 for each of the top3 assets, when next month b=1 (one bad asset in the universe) the cash fraction becomes CF= b/B= 1/4. So, w= (1-CF)/T= (1-1/4)/3= 1/4 for each of the top3 assets and for cash (PAA style). As a result, we have to sell part of all three assets and buy cash to replace the remainder, which results in much more trading (and therefore possibly more slippage) than in the “easy trading” case of dual momentum where assets in the Top are sold and replaced by cash.[7]

To force “Easy Trading” (ET) in the case of VAA we map the fractions b/B to a multiple of the Top asset fractions w=1/T, and remove the corresponding worst asset(s) from the Top T. This is simply achieved by rounding down the raw fractions b/B to multiples of 1/T. So, if eg. T=3 and b=1 with B=4, no cash replacement is required and we keep the Top3 allocation since b/B=1/4<1/3=1/T, for which rounding down results in CF=0. If eg. T=3 and b=2 with B=4, the CF=1/3 as the rounddown

[6] See also note 5.

[7] Apart from the required rebalancing of open positions to their prescribed weights w=1/T.

result of b/B=1/2, consequently we replace the worst asset from the Top3 by cash. Explanding the same example to b=3 or b=4 which gives b/B=3/4, or b/B=1, we arrive at cash fraction of 2/3 or 1, respectively. The worst assets are the assets with the lowest 13612W momentum in the top T. In general, the formula for CF with ET rounding becomes:

CF=(1/T)*rounddown(bT/B), with max(CF)=1.

By following this method, we always replace the worst asset(s) of the top T by cash instead of sizing down all top T assets. This VAA-ET mechanism is also an essential difference between VAA on one hand and PAA at the other, since no ET was defined in case of PAA (see Keller, 2016).

When B=T no Easy Trading rounding is necessary, as was the case with eg. the N12 universe in our PAA2 model (where B=T=6). When B=1 or T=1 the whole portfolio is fully invested in cash (when b>=1) or fully invested in the top T risky asset(s) (when b=0). Notice that when B>>T, the rounding of CF might give rise to less granular crash protection than with b/B (see T/B=3/8 in Fig. 3), while when T>>B the crash protection becomes less granular than all possible multiples of 1/T (see T/B=6/3 in Fig. 3). For example, if T=6 and B=2, only CF= 0, ½, 1 (for b=0, 1, 2+ resp.) will be used, so we replace groups of 3 assets at a time by cash (and vice versa). In Fig. 3 we have depicted CF for various combinations of B and T.

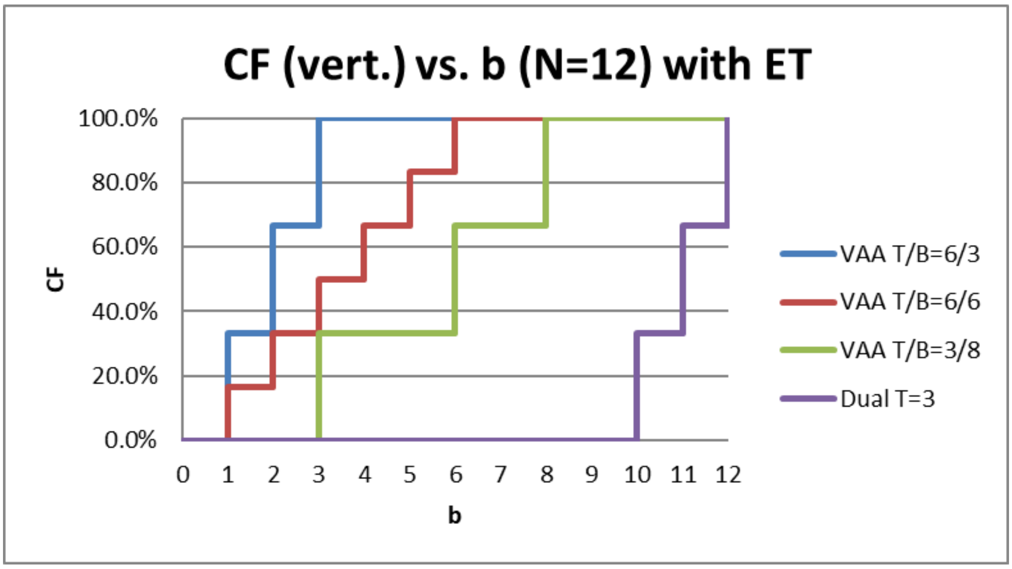

Fig. 3. CF(b) for VAA, PAA, Dual (b=1,..,12, for T/B=6/3, 6/6, 3/8 and N=12)

Finally, some words about “cash”. Traditionally, Faber (2007) uses (riskfree) 90-day TBill as cash by default, but also introduced alternative cash strategies with the 10y Government bonds as “cash”-proxy. We took cash one step further in our PAA paper by introducing a cash universe populated with SHY and IEF (short-term and intermediate-term US-treasury ETFs, respectively), picking each month the best of the two (in terms of momentum) as the cash asset. With VAA we extend the cash universe to three bond-like assets: SHY, IEF and LQD (or 30d TBill, IT Gov and LT Corp from Ibb for the longer backtests from 1927, see Section 7 below) and use our 13612W relative momentum filter to select the best bond. Notice that we use the same cash concept for the Dual strategy in our backtests too, where we replace the assets in the Top T with non-positive momentum by the best cash asset out of our cash universe. We also use the same 13612W momentum filter for Dual as with VAA.

5. Backtesting: In-sample optimization and out-of-sample validation

In each of our four VAA backtests below, we will split our backtest in two nearly equal parts: the first part being the in-sample "optimization" (or learning) period, followed by the second out-of-sample "test" (or validation) period. This way, "datasnooping" (a.k.a. overfitting or datamining, see e.g. Harvey 2013, 2014) for the optimal parameters is limited to the in-sample period. We also need the first year of the data for our momentum filter, hence each of the backtests requires a one year initialization period.

Therefore, in our short dataset from Dec 1969 – Dec 2016 (47 years) we will use Dec 1970 – Dec 1993 (23 years) as the In-Sample (IS) optimization period to determine the best parameters. Then, we will check these parameters in the Out-of-Sample (OS) test period of Dec 1993 – Dec 2016 (23 years). Besides IS and OS periods we will also define the Full Sample (FS) as the combination of IS+OS, so FS in our example equals Dec 1970 – Dec 2016 (46 years), and the Recent Sample (RS) as the last decade (Dec 2006 – Dec 2016, 10 years). Similarly, with the long dataset from 1925, IS and OS for the longer backtest equals Dec 1926 – Dec 1970 (44 years) and Dec 1970- Dec 2016 (46 years), respectively.

What parameters are optimized over IS in the VAA model and how? We will optimize both the breadth threshold B (eg. B=4 when CF=1 for b>=4) and the "top" T (eg. T=3 when we select the top3 of best performing assets). Since we can't select more assets than the full universe population, T is constrained by 0<T<=N. Similar, we have the feasible range 0<=B<=N (with B=0 we will always be in cash). As can be seen in Fig.2, in the case where b<N, the (rounded) cash fraction CF is always higher with VAA than in the traditional Dual case. So no bad assets will show up in the remaining Top assets (after removing the CF fraction of worst assets for VAA with ET).

To allow for vigorous crash protection for the three larger universes with N>=6, we will limit the breadth B<=6, while we also restrict diversification (the Top selection) by limiting T<=6. For the smallest universe with N=4 we simply take B,T<=4. So, the number of BxT values (scenarios) to consider in-sample is 6x6=36 for the larger universes while for the smallest universe the number is 4x4=16. Given the long in-sample periods (23 and 46 years, respectively) combined with only two parameters (T and B) and the limited number of scenarios (16-36), we expect limited datasnooping effects on eg. the estimated Sharpe ratios.[8]

[8] Harvey (2013) shows that to test the hypothesis Sharpe S=0 without datasnooping, one can use the t-test with t=S*sqrt(Y) where Y is the sample size in years and S is the annualized in-sample Sharpe ratio (over risk-free). When there is datasnooping with less than 100 scenarios (here 16-36), a haircut of 50% could be applied for S around 0.5, or 25% if the S is around 1 or higher. Recently, Paulsen (2016) proved that asymptotically (for large Y) the haircut equals k/Y for an in-sample Sharpe ratio S=1 and 4k/Y for S=0.5. For VAA, this results in a haircut of 9-35% (for S between 0.5-1) when k=2 is the number of optimized parameters (T, B) and Y=23 year

Notice that, although we limited the datasnooping bias by this in/out-of sample tests for both parameters T and B, there are always other choices which are not tested this way. In this respect, one might consider in particular our choice of the 13612W momentum filter. Although there were some theoretical reasons for selecting this filter (above the usual 13612 filter), and although we have mainly focused on its in-sample performance when selecting it, we might not have chosen it with a bad out-of-sample performance. However, the fact that it works so well for all our 4 different universes and two different backtest periods, both in- and out-of-sample, might indicate some robustness and therefore limited datasnooping bias. Also, our choice of the 4 universes might be a source for datasnooping, although this choice was mainly determined by the available data and previous studies (PAA and GEM), and partly (VAA-U6) by some in-sample testing only. And finally, we like to focus on the new "breadth B" parameter, so that is the main object in our in-sample optimizations. But still, datasnooping is a serious risk when backtesting, so the best test is a live test (see, e.g., Jones 2017).

The next question is: what performance indicator do we optimize in-sample?

6. Returns adjusted for drawdowns (RAD)

As said in the introduction, with our VAA strategy we aim at combining high (offensive) annual returns with low (defensive) drawdowns. Therefore, we will introduce a new return/risk measure of the resulting VAA equity line, called Returns Adjusted for Drawdowns (RAD). Its formula is

$$RAD = R*(1-D/(1-D)), \quad \text{if } R>=0\% \text{ and } D<=50\%, \text{ and } RAD=0\% \text{ otherwise,}$$

where R=CAGR and D=Max Drawdown (D>=0, measured EOM in our case) of the VAA equity line over the chosen backtest period. Since RAD is an adjusted return, its interpretation is similar to any return (a simple percentage).

This RAD measure is based on the observations that a max drawdown of 50% often leads to the liquidation of a hedge fund. In this case our RAD=0%, independent of R. We also recognize in RAD the term D/(1-D) which is the necessary increase in price to recover to the previous top portfolio capital level after a drawdown of D. When D=50%, this price increase equals 100%, so RAD=0%, reflecting the difficulty of getting back to the previous top portfolio capital level after a severe drawdown. [9]

Why do we opt for this RAD measure instead of the usual ones like the Sharpe and MAR ratio? The Sharpe ratio is defined as the return R (often in excess over a target return like the risk-free rate) divided by the (annual) volatility V of the returns. The MAR ratio (similar to the Calmar ratio) is

---

for IS, and 5-18% when Y=46 years. Notice that these haircuts hold for (in-sample) Sharpe ratios, but not for MAR and RAD.

[9] Using similar ideas, one might also construct a Return Adjusted for Volatility, e.g. RAV=R*(1-2V/(1-2V)) with a max Volatility of 25% for which RAV=0%.

simply return R divided by max drawdown D (with D>=0). Both measures assume that you can apply leverage to arrive at higher R, V and D combinations with the same Sharpe and MAR ratio. But, as we know from leveraged ETFs, this only holds for constant growth (and lending rate equal to the risk-free rate) , while in practice your Sharpe ratio will be much less after leverage. And not all investors have access to cheap leverage at the risk-free rate. So, when optimizing Sharpe or MAR ratios you might be stuck at relative low returns R with low risk, especially when we use a low or zero target (or risk-free) return for Sharpe.

This can be demonstrated eg. in Table 14, where EWC (the equal weight of our cash universe) has MAR=0.66 with R/D= 2.9/4.4% while VAA has MAR=0.29 with R/D=7.4/25.4%. Also, assuming a zero-target return, EWC's Sharpe=2.9/2.1= 1.4 is better than VAA's Sharpe=7.4/10.7= 0.7. Still, VAA trumps EWC on RAD (4.9% vs. 2.7% for EWC) which seems appropriate for most non-defensive investors. This is also the reason why we take the (often higher than the risk-free rate) EWC return as target return in our Sharpe formula in all tables below, since a higher target return in Sharpe will give higher (more offensive) optimal returns.

And while volatility V is statistically a much nicer risk measure than max drawdown D, because most stable stochastic processes have stable V but increasing D over time (see eg. Goldberg, 2016, for a review of the literature), most retail investors commonly identify true risk with D over V. For example, during the 2008 subprime crisis the SP500 (TR index) crashed over 50% in approximately 1.5 years from its late 2007 peak with 3 years to recover to break even, leaving B&H investors without any positive returns over nearly five years. The biggest drawback of the measure D is that it is a "tail risk" which can only be meaningfully assessed in large backtests (preferably covering multiple decennia) with many drawdowns, since any backtest limited to only the last decennium (Dec 2006 – Dec 2016) will be based on a single event (ie. the 2007-2009 crash) and therefore is not very meaningful as some kind of average (as R and V are). Notice, that in this paper D is measured monthly (EOM prices) so a flash crash might get unobserved.

To conclude, we will optimize RAD in-sample over many decades, but also give Sharpe and MAR ratios (besides R and V) for both the in- and out-of-sample backtests. Since we use monthly data, D and V are also only measured using End-Of-Month (EOM) prices, so eg. the daily max drawdown might be greater.

7. The Data

To test VAA, we will use two monthly datasets, one short global from 1969 and one long for the US from 1925.

The first (short, global) dataset is from ourselves (denoted KK, see also PAA, 2016, and TrendXplorer, 2017) and runs from (ult.) Dec 1969 to Dec 2016 (47 years). The second (long, US-only) dataset is combined from Ibbotson/Morningstar[10] (denoted Ibb) and Fama-French (FF, see French, 2017) from (ult.) Dec 1925 until Dec 2016 (91 years). Since we need the first 12 months for our momentum filter

[10] The data has been generously provided to us by Morningstar.

(see above) the actual data lengths used for backtesting are 90 years for Ibb/FF (from Dec 1926) and 46 years for KK (from Dec 1970). All sets contain monthly total return (TR) prices, so including dividends, etc.

The shorter KK dataset contains (among others) the following global asset-classes (ETF proxies, see also Keller, 2016), for a total of 17 global assets from Dec 1969– Dec 2016: SPY, IWM, QQQ, VGK, EWJ, EEM, EFA, ACWX, IYR, GSG, GLD, SHY, IEF, TLT, LQD, HYG, AGG. [11]

The larger Ibb/FF dataset contains the following 21 US-only asset-classes (index series) from Dec 1925 – Dec 2016: SP500, Small Caps, 30d T-Bill, IT Gov, LT Gov, LT Corp, and High Yield from Ibbotson/Morningstar (Ibb for short) and 10 US sectors and 4 US factors (Large/Small Cap x Growth/Value), from Fama French (FF). Here IT/LT stands for Intermediate/Long Term maturities, and Gov/Corp bonds for Government (Treasuries) and investment grade Corporate bonds.

All early ETF data from 1969 are based on ETF proxies constructed by us (see Keller, 2016, and TrendXplorer, 2017 for details). By calibration, all our ETF proxies include ETF fees, etc. Notice that all Ibb and FF data represent TR index prices, so no corrections for ETF fees etc. were made in this case. Only the recent years of the KK database has observed and tradable ETF prices, all the other historical prices are non-tradable.

The in-sample out-of-sample (IS/OS) split will be half ways (see Section 5). Notice that in both IS periods (in particular in the 1971-1993 period) there are rising (and decreasing) bond yields, so we optimize ("train") our VAA strategy in a similar rate-environment as we might experience in the (near) future.

8. Four VAA universes

Using this data, we will consider four "market" universes. For each dataset we will choose a large (with size N=12 or N=15 assets) and a small (with size N=4 or N=6) universe to test VAA. In all our shorter backtests (from Dec 1970) we will us SHY, IEF and LQD as "cash universe" for the "out-of-the-market" allocation. Compared to our choice of cash in our PAA paper we added LQD to get slightly more "pizazz" when out of the market. We will use a similar combination for the longer backtests from Dec 1926 with SHY replaced by 30day T-Bill, IEF by IT Gov, and LQD by LT Corp (all three from Ibb). Notice that the VAA model (like the PAA model) always chooses the single best cash asset in terms of our (13612W) momentum filter, irrespective of their sign (so no absolute momentum filter for the cash universe).

[11] We actually used (proxies for) Vanguard ETFs VEA, VWO, VNQ, and BND instead of the mentioned (and more common) iShares ETFs EFA, EEM, IYR, and AGG, respectively, in nearly all our backtest since these ETFs has lower fees and similar AUM's as the iShares ETFs. When we used Vanguard or iShares ETFs, our proxy ETFs from 1969 were also calibrated using data for recent Vanguard or iShares ETFs, respectively. See also note 16 for some difference in RAD etc. between using iShares and Vanguard ETFs.

The first two universes are based on the global KK dataset from 1969 with backtests from Dec 1970. The large universe (N12) is the same as we use in our PAA paper:

- VAA Global 12 (VAA-G12): SPY, IWM, QQQ, VGK, EWJ, EEM, IYR, GSG, GLD, TLT, LQD, HYG[12]

The small universe (N4) is inspired by Antonnacci 's GEM (see Antonacci, 2014), where we replaced his ACWX (World ex US) by EFA (developed markets) and EEM (emerging markets) and added AGG to arrive at a N4+3 instead of his N2+1 universe:

- VAA Global 4 (VAA-G4): SPY, EFA, EEM, AGG[13]

For cash we again use SHY, IEF and LQD (see above).

The last two universes are based on the Ibb/FF dataset for the US from Dec 1925 to Dec 2016 with backtests starting Dec 1926, where we also have chosen a large and a small universe. The large universe (N15) is based on FF's 10 US-sectors assets augmented with all the available five Ibb bonds (30d T-Bill, IT Gov, LT Gov, LT Corp, and High Yield):

- VAA US 15 (VAA-U15): FF 10 sectors, 30d T-Bill, IT Gov, LT Gov, LT Corp, High Yield

The small universe (N6) consists of the four US FF factors (Large/Small Cap x Growth/Value) with two Ibb. bonds added, so:

- VAA US 6 (VAA-U6): FF 4 factors, T-Bill, LT Corp

For cash we used 30-day T-Bills, IT Gov and LT Corp (all from Ibb, see above), equivalent to SHY, IEF, LQD in the shorter backtests.

The choice of Ibb's T-Bill (30d) and LT Corp for the smaller (risky) VAA-U6 universe was made by manual optimization on the in-sample period only. We opted for not too small (N4 global and N6 US) universes to have enough "breadth" for our crash-protection mechanism. Therefore, we preferred the 4-factors over Ibb's SP500/Small Cap combi for the small US universe.

For all universes, we used a 0.1% (one way) transaction fee to account for brokers commission, slippage, etc. This is certainly too low for 1926 and probably too high for 2016, but it seems reasonable for ETFs over the last decade with EOM/EOD trading. We also show the yearly Total Transaction Costs (TTC) as turnover indicator, which equals 2.4% per year (=2x12x0.1%) when we have full monthly turnover.

We also test our breadth momentum strategy for our each of the four VAA universes against the traditional Dual (or GTAA, see Faber 2007, 2013) momentum strategy. To make a fair comparison with VAA we will use the same data (incl. ETF fees) and cash universe, the same costs (TC=0.1%) and the same 13612W momentum filter for Dual, while optimizing in-sample the top parameter T. We also tested the (somewhat special) GEM strategy from Antonacci, 2014 using his N2+1 universe and (excess RET12m) momentum filter against our VAA-G4.

[12] We used VWO and VNQ instead of EEM and IYR in our actual backtest. See also note 11.

[13] We used VEA, VWO and BND instead of EFA, EEM and AGG in our actual backtest. See also note 11.

9. The VAA-G12 universe

The first universe is the same global universe as we used in our PAA paper, with 12 risky assets but now with three cash assets and some Vanguard data (SPY, IWM, QQQ, VGK, EWJ, VWO, VNQ, GSG, GLD, TLT, LQD, HYG with SHY, IEF, and LQD as cash). Our in-sample (IS) period is from (ult.) Dec 1970 – Dec 1993 (23 years), while our out-of-sample (OS) period is Dec 1993 - Dec 2016 (also 23 years). Data is from our own KK dataset, corrected for ETF fees etc. (see Keller 2016, and TrendXplorer 2017 for details).

In-sample (IS)

As said, we compute RAD over IS on top T=1..6 (selected number of Top assets) and breadth B=1..6 (protection threshold) for a total of 6x6=36 values, using our 13612W momentum filter. Dual is the dual momentum strategy (called GTAA by Meb Faber) computed for various T=1..6 using our own momentum filter for both absolute and relative momentum.

From Table 1, the best VAA score at IS is RAD=19.7% for T=2 and B=4. For Dual optimal it is RAD=15.7% at T=3. The best RAD values are clearly concentrated around this optimum and part of a "ridge" of high RAD values at B=4. [14] We will use the optimal T/B=2/4 for VAA-G12 and T=3 for Dual in our backtests.

RAD					T		
	19.7%	1	2	3	4	5	6
	0	8.7%	8.7%	8.7%	8.7%	8.7%	8.7%
	1	8.5%	11.8%	12.2%	12.3%	11.6%	11.6%
	2	12.1%	14.2%	14.5%	14.0%	13.8%	13.3%
	3	15.7%	17.6%	15.9%	16.4%	15.3%	14.7%
	4	18.7%	19.7%	18.8%	17.5%	17.4%	16.8%
	5	5.3%	16.1%	17.4%	18.2%	16.7%	15.8%
B	6	5.2%	15.9%	17.3%	16.2%	15.9%	15.5%
	Dual	4.2%	11.3%	15.7%	14.3%	13.2%	13.8%

Table 1. RAD for VAA-G12 on IS (dec1970-dec1993) for T/B=1..6

We can see more detail behind these T/B values when we draw the R/D points (like normally done for R/V) as function of T and B, as in Fig. 4. As with the R/V frontier curve, points in the upper/left corner (with high return R and low drawdown D) are preferable over points in the right/lower corner. Points in the left/upper corner will be called "efficient" when there are no other points with better R/D combinations (ie. no points in the upper/left quadrant), and "optimal" when they have the best R and best D values of all points.

[14] When using Sharpe or MAR instead of RAD we found a much more dispersed and irregular pattern, indicating less robust optimization results.

All VAA-T curves in Fig. 4 start with points with low B (high protection) at the left/under (so low R and D) and go to the right/upper (so high R and D) for higher B values (low protection). The VAA-B (and both Dual) curves in Fig. 5 (and Fig. 4, resp.) go from high T left/under (high diversification) to low T (less diversification) at the right/upper part. Notice that points left/under corresponds to more protection (lower B) and/or higher diversification (higher T), and vice versa for points right/upper.

From Fig. 4, it is clear that the point R/D=21.0/5.8% on the red VAA-T2 line (for B=4) is not only RAD optimal (in terms of return/risk) on IS (see Table 1) but also efficient. The point R/D=23/44% on the VAA-T1 (yellow line) wins in terms of only R but with a much larger D>20%.

Notice that from Table 1 we see that B=4 is also high on RAD for other values of T. This is even more clear when we draw in Fig. 5 the R/D frontier for various B, which shows that the red VAA-B4 line (for various T) is efficient to all other values of B. This also indicates some robustness (and possibly less datasnooping) for the B=4 choice. The first left/under points of all VAA-T lines in Fig. 4 corresponds to B=1, the next to B=2, etc. Similary, the right/upper points of all VAA-B in Fig. 5 (and Dual) lines corresponds to T=1, the next to T=2, etc. As we see from Fig. 5 (and Table 1), T=3 is RAD optimal for Dual (with R/D=22/21%), see the black line in Fig. 5. From Fig. 4 and 5 it is also clear how worse Dual (the black line) is wrt. VAA-T2 and VAA-B4, especially in terms of max drawdown D. In particular in Fig. 5, the VAA-B4 curve (for various T) dominates the Dual line (for various T) which resembles the VAA-B4 line shifted to the right side. Dual achieves nearly the same R but only at the cost of a higher D.

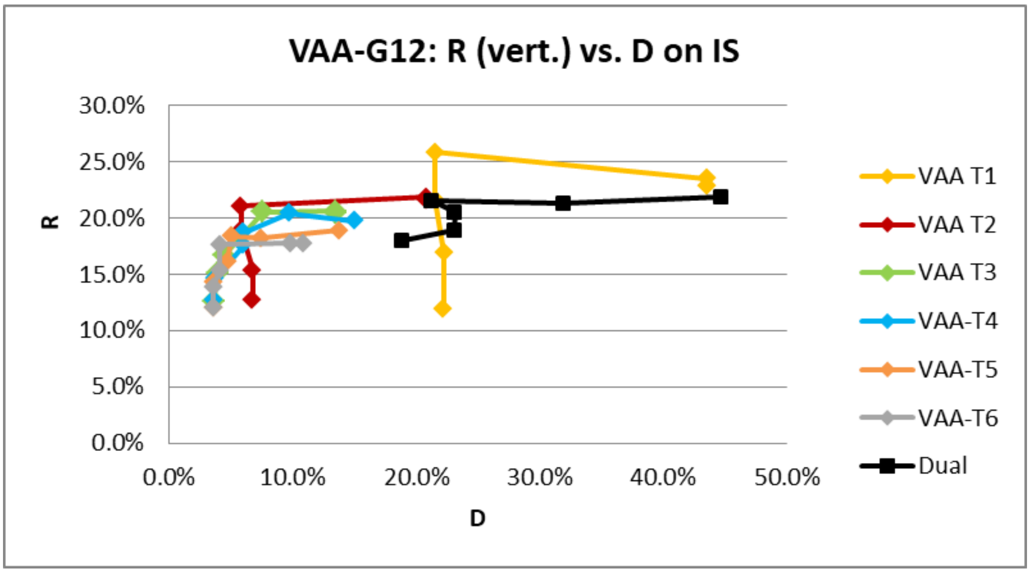

Fig. 4. VAA-G12: R/D frontier on IS (dec1970-dec1993) for VAA-T and Dual

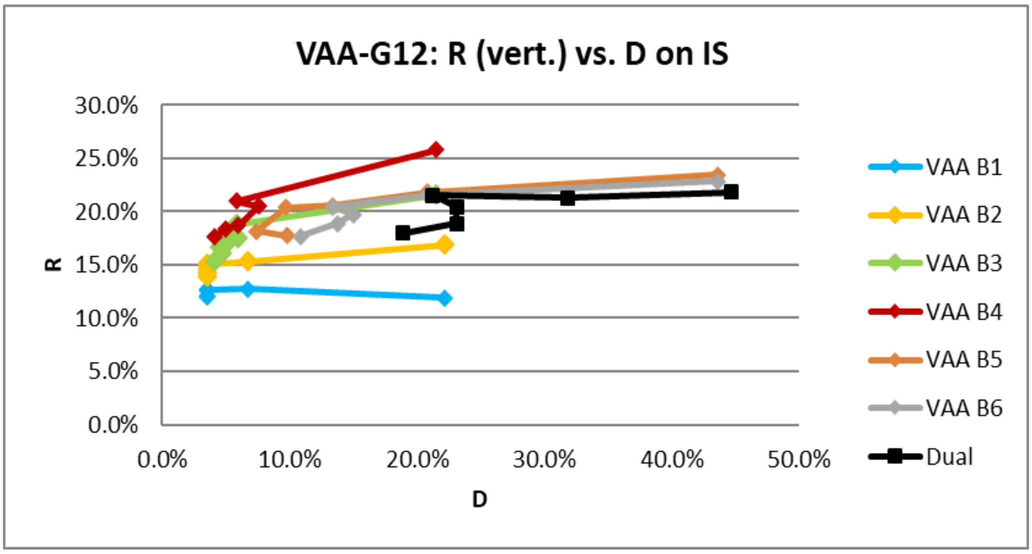

Fig. 5. VAA-G12: R/D frontier on IS (dec1970-dec1993) for various VAA-B and Dual.

Now we focus on the VAA-T2 curve, see Fig. 6 (red curve). This VAA-T2 curve manifests its typical "bent" appearance: starting at B=1 left under (with R/D=13/7%), we go upwards (higher R) through B=2, B=3, to the optimal point or the "VAA knee" at B=4 (with R/D=21/6%), all with nearly the same D but with increasing R for B=1..4. The upper leg of the curve after the "knee" stretches nearly horizontally to the right (higher D for nearly the same R) through B=5, and 6 (D=21%). We see the same characteristic bent curve (but shifted to the right) for VAA-T1 (yellow curve). Surprisingly, when we do the same analysis for OS, we found a similar curve with the same "knee" at B=4 for VAA-G12 (but at T=1 instead of T=2). As we will see with other universes later, this "bent" R/D line for increasing B is typical for VAA. This could indicate that there might be some interesting timing mechanism at work here.

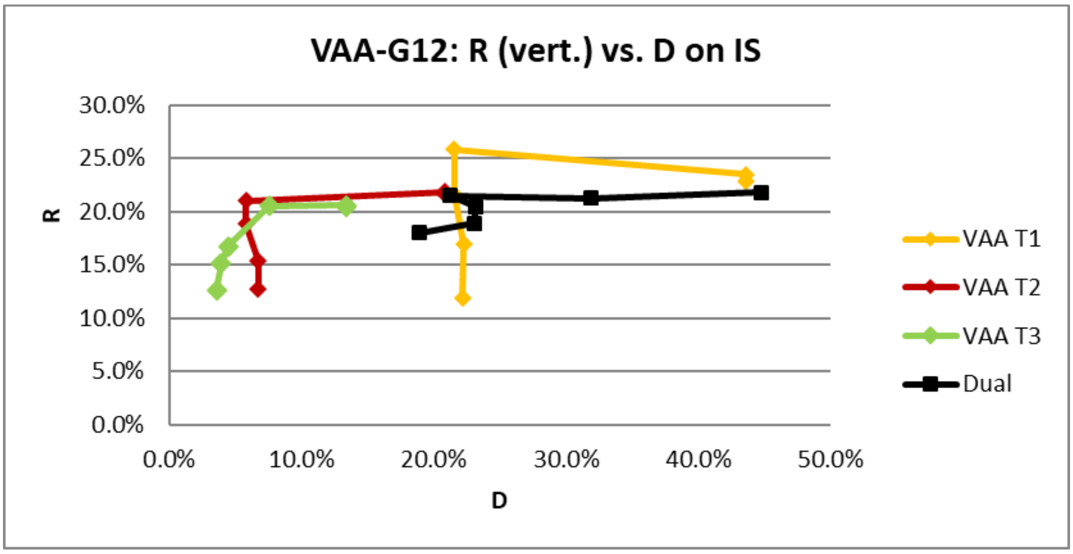

Fig. 6. VAA-G12: R/D frontier on IS (dec1970-dec1993) for VAA-T1, T2 and Dual.

From now one, we will refer to VAA-G12 as the optimal strategy (in terms of RAD) with T/B=2/4 for the N12+3 universe. For this optimal VAA-G12 strategy, we arrive at the following performance indicators on IS for VAA, EW (equal weight of the global risky universe, our benchmark), EWC (equal weight of the Cash universe of assets SHY, IEF and LQD), the well-known 60/40 (SPY/IEF) portfolio, SPY, and Dual. See Table 2, where R, V, D equals return (CAGR), (annual) Volatility and the max Drawdown (EOM), and where the Sharpe ratio is computed with cash rate equal to the return of EWC, the MAD ratio equals R/D and RAD is the Return Adjusted for Drawdown.

As can be seen from Table 2, the return of VAA on IS is a staggering R=21% with an unbelievable low drawdown D=5.8% (over 1970-1993) and a volatility V=10%, resulting in very high Sharpe=1.24, MAR=3.62 and RAD=19.7%. VAA's R is nearly twice that of EW, 60/40 and SPY while VAA's D is 5-7 times smaller. By definition, Sharpe EWC=0, while Sharpe VAA is more than 2, 5 and 6 times larger than EW, 60/40 (SPY/IEF) and SPY, while MAR is 7, 9, and 13 times larger, respectively. Dual (with optimal T=3, see Table 1) is slightly better for R (21.5% vs. 21.0% for VAA), but worse for V, D (21.3% vs. 5.8% for VAA) , Sharpe, MAR, and RAD (15.7% vs. 19.7% for VAA). But remember, these results are the outcome of in-sample (IS) optimization, so possibly the result of datasnooping for parameters T and B.

IS	R	V	D	Sharpe	MAR	RAD
VAA-G12	21.0%	10.0%	5.8%	1.24	3.62	19.7%
EW	13.2%	10.4%	24.6%	0.45	0.54	8.9%
EWC	8.6%	5.6%	9.1%	0.00	0.95	7.7%
60/40	11.1%	10.6%	27.4%	0.24	0.41	6.9%
SPY	11.6%	15.4%	42.5%	0.20	0.27	3.0%
Dual3	21.5%	15.4%	21.3%	0.84	1.01	15.7%

Table 2. Performance indicators for VAA-G12 at IS (dec1970-dec1993)

Out-of-sample

The out-of-sample (OS) results are displayed in Table 3, where we show all performance indicators for OS (Dec 1993 – Dec 2016) as well as some for the last 10 years, denoted Recent-Sample or RS (Dec 2006 – Dec 2016). As we see in Table 3, return R and max drawdown D over OS are "back to normal" (compared to IS), but are still within our preferred target range: R>10% and D<15%. This also holds for the last decade or RS (including the Global Financial Crisis), with an appealing D=8%, compared to 40% (EW), 29% (60/40), and 51% (SPY), while return R (11%) and RAD (10%) of VAA-G12 are also much better than the R and RAD readings for EW, 60/40, SPY, and Dual. This is all out-of-sample (OS), so without datasnooping for parameters T and B.

The resulting Full-Sample (FS: Dec 1970 – Dec 2016, so 46 years) outcomes are displayed in Table 4. Return R=16% and max drawdown D=13% are well within our preferred VAA target range (R>10%, D<15%) resulting in RAD=13.3%, with corresponding high Sharpe and MAR ratios. Dual is slightly better at R (17% vs. 16% for VAA) but worse at V, Sharpe, MAR and especially D (26% vs. 13% for VAA) and RAD (11% vs. 13% for VAR).

OS	R	V	D	Sharpe	MAR	RAD	RS	R	V	D	RAD
VAA-G12	10.5%	10.3%	13.0%	0.51	0.81	8.9%		10.6%	10.3%	7.7%	9.8%
EW	7.3%	11.1%	39.6%	0.19	0.19	2.5%		4.9%	12.7%	39.6%	1.7%
EWC	5.3%	4.4%	4.7%	0.00	1.11	5.0%		4.3%	4.5%	4.7%	4.1%
60/40	8.2%	8.7%	29.4%	0.34	0.28	4.8%		6.7%	8.7%	29.4%	3.9%
SPY	9.1%	14.6%	50.8%	0.26	0.18	0.0%		6.9%	15.2%	50.8%	0.0%
Dual3	11.9%	13.7%	25.6%	0.48	0.46	7.8%		6.3%	13.4%	25.6%	4.1%

Table 3. Performance indicators for VAA-G12 at OS (dec1993-dec2016) and RS (dec2006-dec2016)

Over FS the Total Transaction Costs per year are high (TTC=1.4%/y) indicating a monthly turnover of nearly 60% (1.4/2.4= 58%). The average Cash Fraction over FS equals CF=59%, reflecting the vigorousness of the Crash Protection. Together with the resulting R>15% and D<15% we think this underlines our VAA motto “Winning more by Losing Less”. Notice that Dual has a much lower CF (2.2% vs. 59% for VAA), but a similar TTC as VAA (1.2% vs. 1.4% for VAA).

FS	R	V	D	Sharpe	MAR	RAD	TTC	CF
VAA-G12	15.6%	10.2%	13.0%	0.85	1.20	13.3%	1.4%	58.9%
EW	10.3%	10.8%	39.6%	0.31	0.26	3.5%	0.0%	0.0%
EWC	6.9%	5.1%	9.1%	0.00	0.76	6.2%	0.0%	0.0%
60/40	9.6%	9.7%	29.4%	0.28	0.33	5.6%	0.0%	0.0%
SPY	10.3%	15.0%	50.8%	0.23	0.20	0.0%	0.0%	0.0%
Dual3	16.6%	14.6%	25.6%	0.66	0.65	10.9%	1.2%	2.2%

Table 4. Performance indicators for VAA-G12 at FS (dec1970-dec2016)

We also show in Fig. 7 the (log) equity line on FS, as well as the rolling 3 year annual returns (Fig. 8) for both VAA-G12 and its benchmark (EW). We also provide the relative price (green line) of VAA vs. EW in the equity line curve, which indicates for how far VAA is behind (sloping downwards) or ahead (upwards) of its benchmark. As Fig. 7 shows, from 2009-2016 the relative price is near horizontal, indicating similar price performance for the strategy as compared to the benchmark. It is also clear from Fig. 7 that recent returns of VAA above EW is made in the crash years like 2000/2002, 2008/9 (winning more by losing less) while in the years before 2000 VAA harvested risk premia during uptrending markets too.

The chart in Fig. 8 compares the rolling 3 year annual returns for VAA-G12 and its EW benchmark. The green curve depicts periods with relative outperformance of VAA over EW (curve above zero) and underperformance (curve below zero). Fig. 9 shows the drawdown profile of VAA against that of its EW benchmark. With only three outliers above the D=10% mark, VAA-G12 depicts a contained drawdown profile.

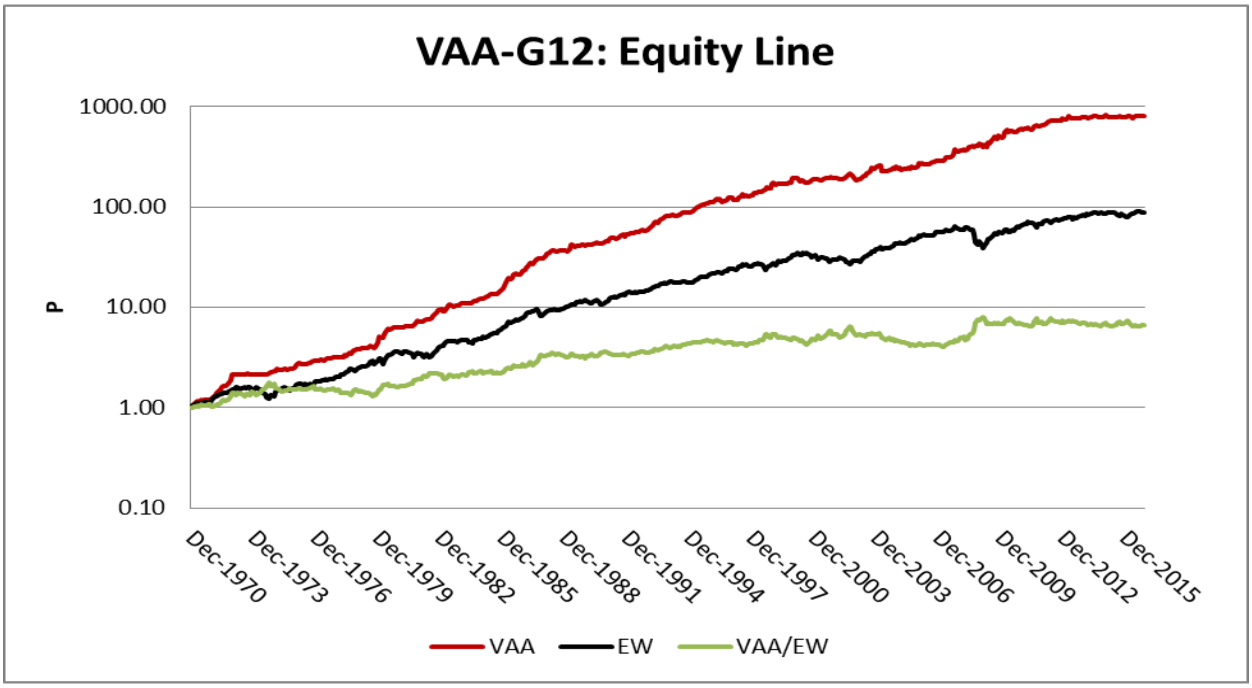

Fig. 7. Equity line VAA-G12, EW and their relative price (log scale)

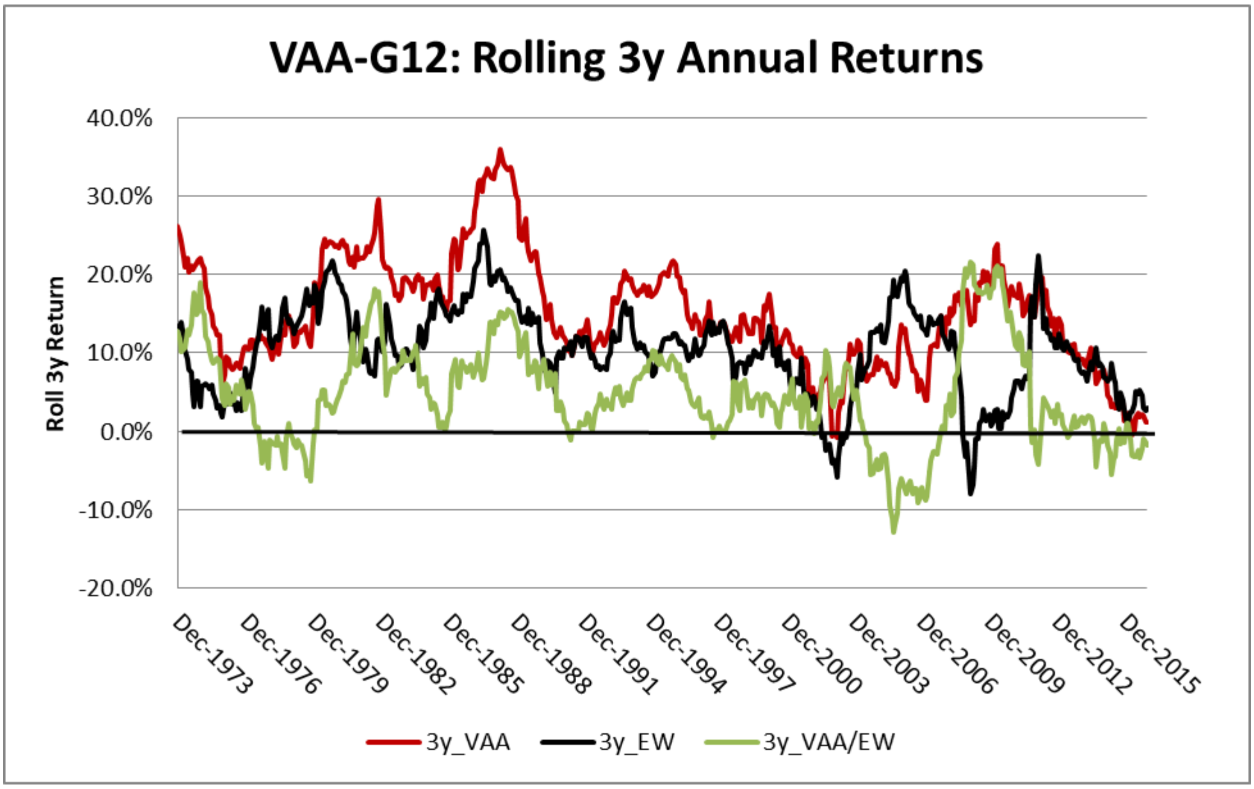

Fig. 8. Rolling 3-year returns of VAA-G12, EW and their relative price

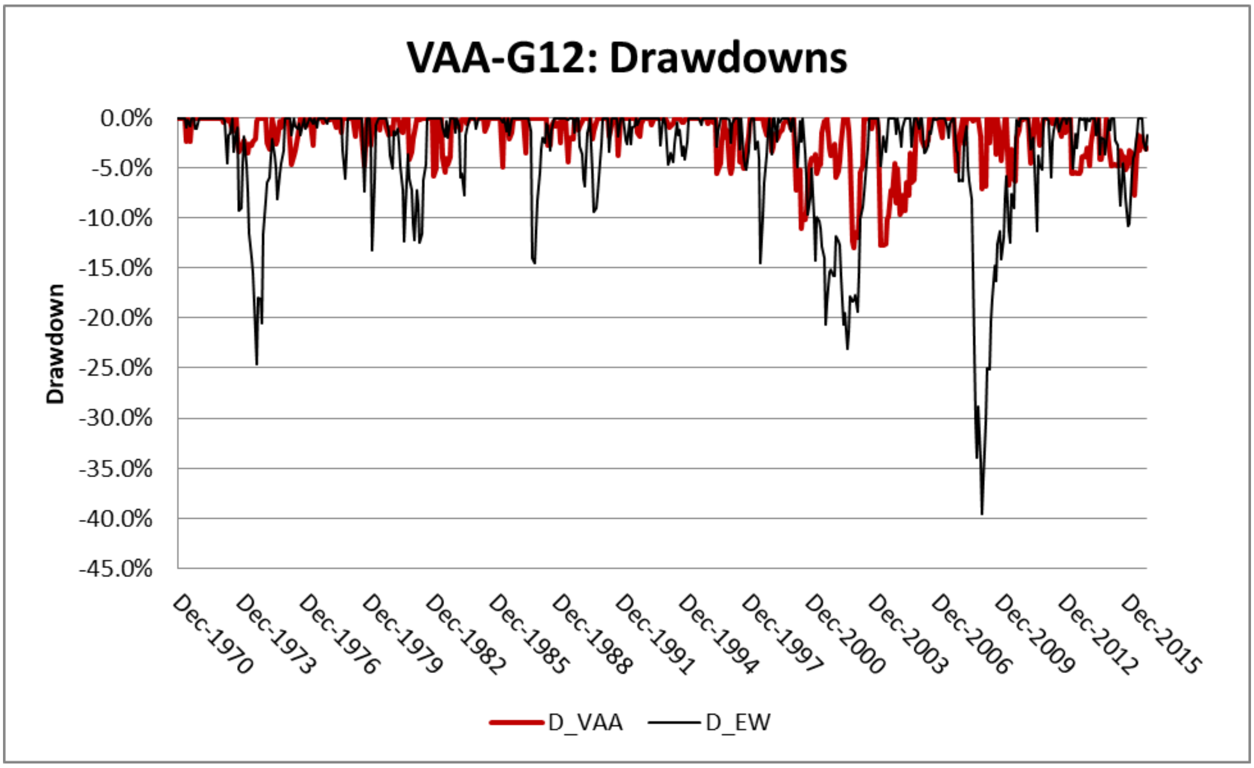

Fig. 9 Drawdown of VAA-G12 and EW

10. The VAA-G4 universe

The second universe is inspired by Antonacci's GEM (Antonnacci, 2014). He uses a risky universe of only two assets, SPY, and ACWX (Global Stocks ex US), and AGG (aggregate US bond) as cash. For VAA we will split ACWX into Vanguard's VEA (International Markets, including Europe, Japan and IM Pacific) and VWO (Emerging markets), and add one bond for some breadth to arrive at N=4. Surprisingly, the best in-sample bond was BND (Vanguard's AGG), so we arrived at N=4 for our risky universe with SPY, VEA, VWO, and BND. Notice that adding a bond like BND to the risky universe has a somewhat similar effect as taking excess returns (as Antonacci did with his RET12m momentum) over that bond for the other (more risky) ETFs, since the best (Top1) asset should always be better than BND (or BND itself).

As before, we will use SHY, IEF, and LQD as cash. We will denote our VAA strategy with this N=4 global universe as VAA-G4. Our in-sample (IS) period is (like VAA-G12) from (ult.) Dec 1970 – Dec 1993 (23 years), while our out-of-sample (OS) period is Dec 1993 - Dec 2016 (also 23 years). Data is from our own KK dataset, corrected for ETF fees etc.

In-sample (IS)

Since N=4, we compute RAD over IS on top T=1..4 and breadth B=1..4 for a total of 4x4=16 values (scenarios), using our 13612W momentum filter. See Table 5. Dual is the dual momentum strategy computed for various T=1..4 using our own momentum filter (13612W) for both absolute and relative momentum.

RAD		T			
18.7%		1	**2**	**3**	**4**
	0	8.7%	8.7%	8.7%	8.7%
	1	18.7%	17.6%	14.8%	13.9%
	2	0.4%	10.5%	10.3%	11.0%
	3	0.0%	9.2%	10.2%	10.2%
B	4	0.0%	9.0%	9.1%	10.2%
	Dual	0.0%	8.6%	9.1%	10.2%

Table 5. RAD for VAA-G4 on IS (dec1970-dec1993) for T/B=1..4

From Table 5, the best score for VAA is RAD=18.7% for T/B=1/1 and for Dual it is RAD=10.2% for T=4. So we use T/B=1/1 and Dual4 in our backtests here. We can see more detail behind these T/B values when we draw the R/D frontier as function of T and B, as in Fig. 10 and 11. As with the R/V frontier curve, points in the upper/left corner (with high return R and low drawdown D) are preferable over points in the right/lower corner. To show the “knee” in the VAA-T curves, we have added the B=0 (100% cash) strategy with R/D=9/9% (independent of T), so now the range in eg. the VAA-T1 line in Fig. 10 is B=0..4 (5 points).

From Fig. 10, it is clear that the point R/D=22/13% upper/left on the red VAA-T1 line (for B=1) is not only RAD optimal (in terms of return/risk, see Table 5) but is also optimal in terms of R. However, the point R/D= 19/8% upper left on the yellow VAA-T2 line is only slightly worse on RAD (17.6 vs. 18.7%, see Table 5) but has a substantial smaller D (8% instead of 13% for T1). [15]

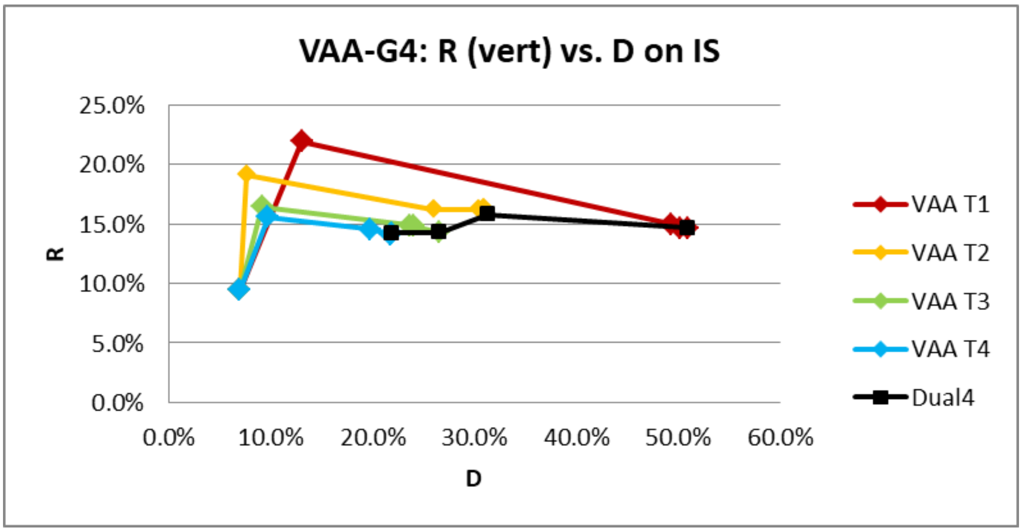

Fig. 10. VAA-G4: R/D frontier on IS (dec1970-dec1993) for VAA-T and Dual

[15] The T/B=2/1 setting resulted in an FS performance of R/V/D= 16.0/10.1/9.1% and therefore RAD= 14.4% over Dec 1970 – Dec 2016. We also used Meb Faber's Backtester (see Faber, 2017) with this T/B=2/1 setting for a preliminary test of this VAA variant from Dec 1927 - Dec 2016 and found an FS performance well within our target range of R>10% and D<15% over nearly 100 years.

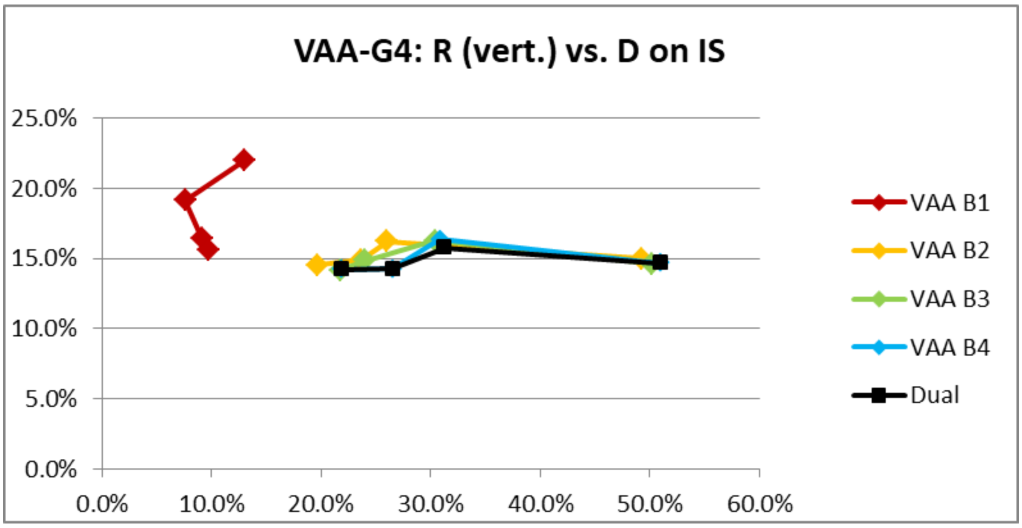

Fig. 11. VAA-G4: R/D frontier on IS (dec1970-dec1993) for VAA-B and Dual

Fig. 10 clearly shows the bend VAA curves as we had seen with the VAA-G12 strategy, starting at the B=0 point (giving a 100% cash strategy with R/D=9/9%) and bending at the "knee" at R/D=22/13% at B=1 and similar for all other VAA-T lines: they all have the knee at B=1. From Fig. 11 (where we display the lines for various B) it is clear that the line VAA-B1 (for all T=1..4) is optimal both in terms of R as in D over nearly all other B, including the Dual strategy, of which even the best T=4 solution at the left with R/D= 14/22% is not near our red VAA-B1 solutions. This demonstrates (as with our VAA-G12 model) clearly the power (and robustness) of our new breadth parameter B for vigilant asset allocation. The same B=1 (and T=1) turns out also to be RAD optimal in our OS period (Dec 1993 – Dec 2016), showing remarkable stability, as with VAA-G12.

From now on we will refer to VAA-G4 as the optimal strategy with T/B=1/1. In Table 6, we show the usual performance indicators for VAA-G4 (with optimal T/B=1/1) at IS (Dec 1970 - Dec 1993).

IS	R	V	D	Sharpe	MAR	RAD
VAA-G4	22.0%	13.7%	13.0%	0.98	1.69	18.7%
EW	13.8%	12.9%	42.4%	0.41	0.33	3.6%
EWC	8.6%	5.6%	9.1%	0.00	0.95	7.7%
60/40	11.1%	10.6%	27.4%	0.24	0.41	6.9%
SPY	11.6%	15.4%	42.5%	0.20	0.27	3.0%
Dual4	14.2%	10.3%	21.9%	0.55	0.65	10.2%

Table 6. Performance indicators for VAA-G4 at IS (Dec 1970-Dec 1993)

As we see in table 6, the return R is nearly double that of EW, 60/40 and SPY, while the max drawdown D is at less than half (and 1/3 of that of SPY). Sharpe and MAR for VAA-G4 are multiples of those of EW, 60/40 and SPY, as is RAD. Dual is worse for R (14.2% vs. 22.0% for VAA), as is D (22% vs. 13% for VAA), Sharpe, MAR and RAD (10.2% vs. 19% for VAA). This is all in-sample (IS), so possibly with datasnooping for parameters T and B.

Out-of-Sample (OS)

The performance indicators for OS (Dec 1993 – Dec 2016) are shown in table 7. The return R is again (as IS) roughly double that of EW, 60/40 and SPY, while the max drawdown D is less than 1/3 (and 1/5 of that of SPY). Sharpe and MAR for VAA-G4 are again multiples of those of EW, 60/40 and SPY, as is RAD. Dual is worse for R (9.6% vs. 16.0% for VAA), as is D, Sharpe, MAR and RAD (8.3% vs. 14.1% for VAA). Similar results (although with a smaller R and RAD) hold for the last decade (RS). This is all out-of-sample (OS), so without datasnooping for parameters T and B.

OS	R	V	D	Sharpe	MAR	RAD	RS	R	V	D	RAD
VAA-G4	16.0%	11.7%	10.4%	0.92	1.54	14.1%		11.4%	12.3%	10.4%	10.1%
EW	6.4%	12.7%	44.9%	0.09	0.14	1.2%		4.0%	14.0%	44.9%	0.7%
EWC	5.3%	4.4%	4.7%	0.00	1.11	5.0%		4.3%	4.5%	4.7%	4.1%
60/40	8.2%	8.7%	29.4%	0.34	0.28	4.8%		6.7%	8.7%	29.4%	3.9%
SPY	9.1%	14.6%	50.8%	0.26	0.18	0.0%		6.9%	15.2%	50.8%	0.0%
Dual4	9.6%	8.5%	11.8%	0.52	0.82	8.3%		7.4%	8.5%	11.8%	6.4%

Table 7. Indicators for VAA-G4 at OS (Dec 1993 - Dec 2016) and RS (Dec 2006 - Dec 1993)

In Table 8 we show the main statistics for the Full Sample (FS: Dec 1970 - Dec 2016, so 46 years), including the yearly Total Transaction costs (TTC) and the average Cash Fraction (CF), both over FS. All the performance stats are impressive (R/D=19/13% over nearly 50 years!) and similar to those at OS, showing limited datasnooping bias at IS. RAD (=16.1%) is much better than EW, EWC, 60/40, SPY, Dual and GEM (with RAD=13.0%[16]). The TTC shows a heavy monthly turnover of ca. 55% (=1.3/2.4) while average cash is CF=57%, both high and similar to VAA-G12. Dual is roughly half of VAA on TTC and CF (with a very low TTC=0.26% for GEM, see note 16).

FS	R	V	D	Sharpe	MAR	RAD	TTC	CF
VAA-G4	18.9%	12.7%	13.0%	0.94	1.46	16.1%	1.3%	57.1%
EW	10.1%	12.8%	44.9%	0.25	0.22	1.9%	0.0%	0.0%
EWC	6.9%	5.1%	9.1%	0.00	0.76	6.2%	0.0%	0.0%
60/40	9.6%	9.7%	29.4%	0.28	0.33	5.6%	0.0%	0.0%
SPY	10.3%	15.0%	50.8%	0.23	0.20	0.0%	0.0%	0.0%
Dual4	11.9%	9.42%	21.9%	0.53	0.54	8.6%	0.6%	27.1%

Table 8. Performance indicators for VAA-G4 at FS (Dec 1970 - Dec 2016)

[16] We have for the "official" GEM model (see Antonacci, 2014, and also TrendXplorer, 2016) but with our KK data (including ETF fees and a 0.1% transaction cost), and Vanguard instead of iShares ETFs (SPY, VEU, and BND), the following performance statistics over Dec 1970 – Dec 2016: R/V/D= 16.7/13.0/17.8%, Sharpe=0.70 (with EWC=AGG=7.6%), MAR= 0.93, and RAD=13.0%, with TTC=0.26% and Av CF= 28%. With Antonacci's original SPY, ACWX and AGG from iShare, we found R/V/D= 16.0/12.9/19.4%, Sharpe=0.65 (with EWC=AGG=7.6%), MAR= 0.83, and RAD=12.2%, with TTC=0.27% and Av CF= 28%.
When we used SPY, EFA, EEM, AGG for our VAA-G4 the FS results were R/V/D= 18.8/13.1/16.4%, Sharpe=0.90 (with EWC=6.9%), MAR= 1.15, and RAD=15.1%, with TTC=1.31% and Av CF= 57%. So, we conclude that both strategies (Antonacci's GEM and our VAA-G4) were dependent on the particular ETFs (iShares or Vanguard) chosen (thanks Walter). Notice that we also calibrated our proxy ETFs from 1969 using Vanguard or iShares ETFs, respectively; see also note 11.

Below (see Fig. 12, 13, and 14) we also give the (log) equity line, rolling 3y annual returns and drawdowns for VAA-G4, EW and the relative performance of VAA/EW for the full sample (FS). Notice that the relative price in Fig. 12 is flat from 2009, indicating that returns (not risks) are similar to the benchmark (EW). The rolling 3y returns (Fig. 13) of VAA/EW are negative from 1977 to 1981, and positive most of the other years. Drawdowns (Fig. 14) for VAA-G4 are maximal at 13% in July 1990, and less than 10% in other years, except in 1974.

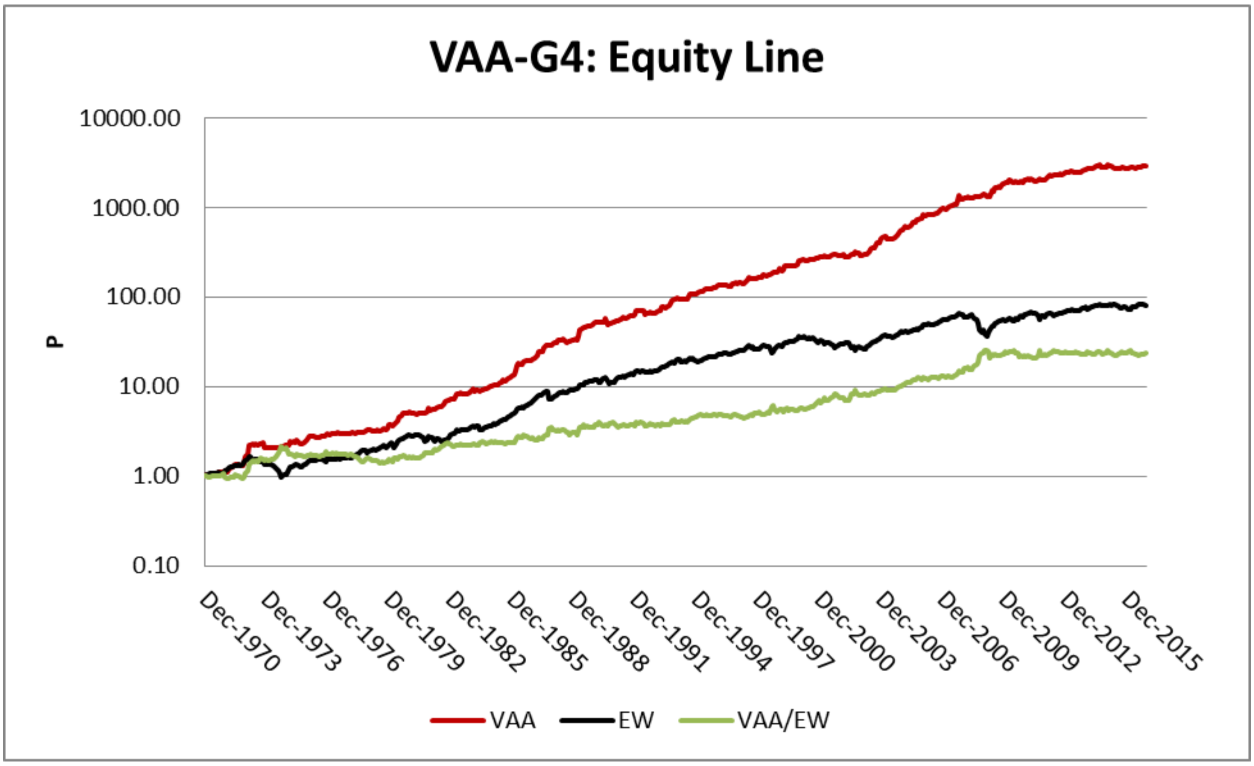

Fig. 12. Equity line for VAA-G4, EW and VAA/EW for FS (Dec 1970 – Dec 2016), log scale

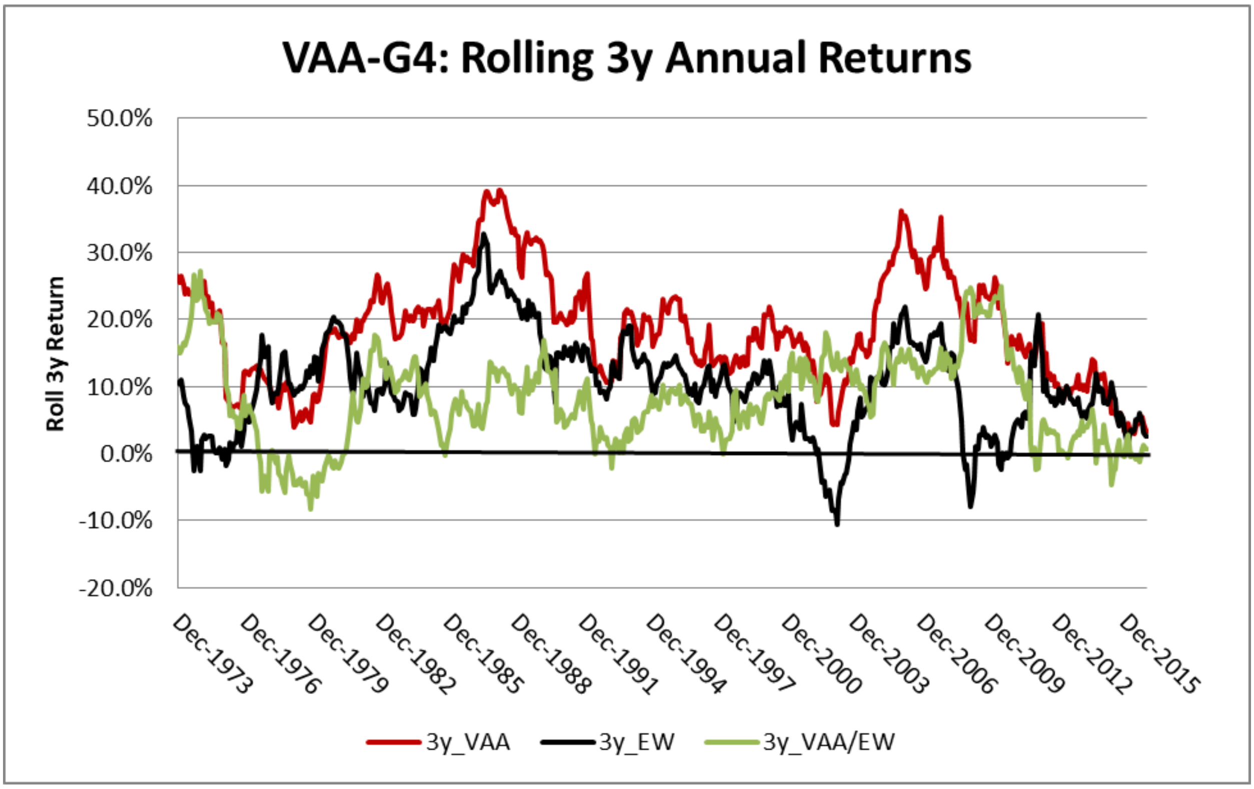

Fig. 13. Rolling 3y Returns for VAA-G4, EW and VAA/EW for FS (Dec 1970 – Dec 2016)

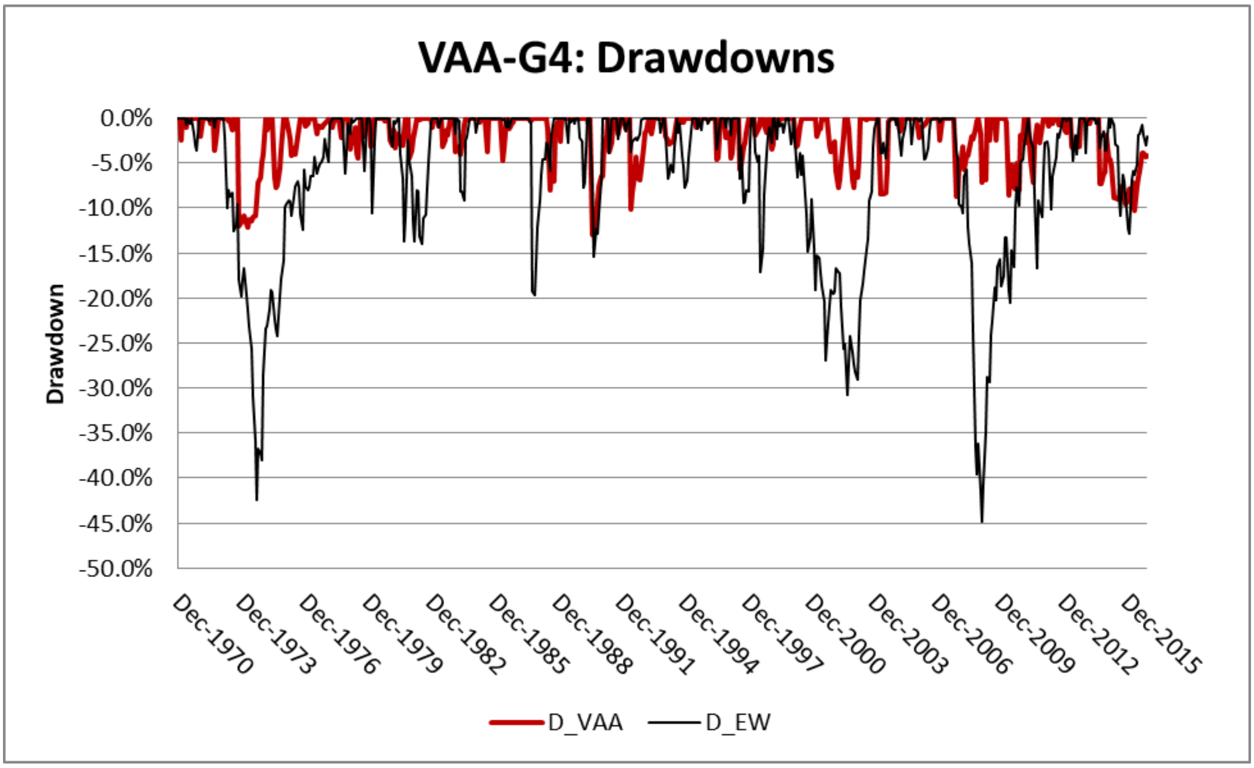

Fig. 14. Drawdowns (EOM) for VAA-G4, EW and VAA/EW for FS (Dec 1970 – Dec 2016)

11. The VAA-U15 universe

The third universe is a long backtest (from Dec 1926) with a large universe, using the Ibbotson (Ibb) and Fama French (FF) dataset. We will use a risky universe of 15 assets, consisting on the ten US sectors from FF and all the five US bonds (T-Bill, IT Gov, LT Gov, LT Corp, and HY) from Ibb, so N=15. We will use 30d T-Bill, IT Gov and LT corp from Ibb as Cash, similar to SHY, IEF and LQD for the shorter dataset (see Sections 9 and 10 above). So, all assets refer to the US. We will therefore denote our VAA strategy with this N15+3 US universe as VAA-U15.

Our in-sample (IS) period is from (ultimo) Dec 1926 – Dec 1970 (44 years), while our out-of-sample (OS) period is Dec 1970 - Dec 2016 (46 years). We use index data (EOM Total Returns) from Ibb/FF, so not corrected for ETF fees etc.

In-sample (IS)

Like with VAA-G12, we compute RAD on IS for this larger universe again over top T=1..6 and breadth B=1..6 for a total of 6x6=36 values (scenarios), using our 13612W momentum filter. The result is shown in Table 9.

RAD							**T**
	7.4%	**1**	**2**	**3**	**4**	**5**	**6**
	0	2.3%	2.3%	2.3%	2.3%	2.3%	2.3%
	1	3.6%	4.6%	4.8%	5.5%	5.1%	4.9%
	2	1.4%	3.8%	4.9%	5.6%	4.8%	4.6%
	3	4.6%	6.2%	7.0%	7.4%	7.1%	6.6%
	4	0.4%	4.6%	5.1%	6.6%	6.4%	6.5%
	5	0.4%	4.8%	3.5%	5.3%	6.4%	6.0%
B	6	0.4%	5.0%	3.6%	5.4%	5.2%	6.1%
	Dual	0.0%	0.0%	0.0%	0.0%	1.7%	2.5%

Table 9. RAD for VAA-U15 on IS (Dec 1926 – Dec 1970) for T/B=1..6

Dual is the dual momentum strategy computed for various T=1..6 using our own momentum filter (13612W) for both absolute and relative momentum. From Table 9, the best score is RAD=7.4% for T=4 and B=3, the best Dual strategy is Dual for T=6, called Dual6. We can see more detail behind these T/B values when we draw the R/D frontier as function of T and B, as in Fig. 15 and 16.

As we see from Fig. 15, 16 (red lines) and Table 9, the best point is T/B=4/3 with R/D=11/24% which is optimal for low D in Fig. 15. The lines for T= 5 and 6 are close to T=4 and therefore hidden to keep the picture clean. B=3 is also clearly optimal (and robust) for low D in Fig. 16 (where we hide VAA-B6: too close to VAA-B5). The VAA-T4 line in Fig. 15 has less clear "bend and knee" than in our global (G12/4) universes, but there is still a knee at B=3 (R/D= 11/24%) in the red VAA-T4 line. The Dual line is clearly inferior in D to nearly all VAA-T (Fig. 15) and VAA-B (Fig. 16) lines, while optimal in R.

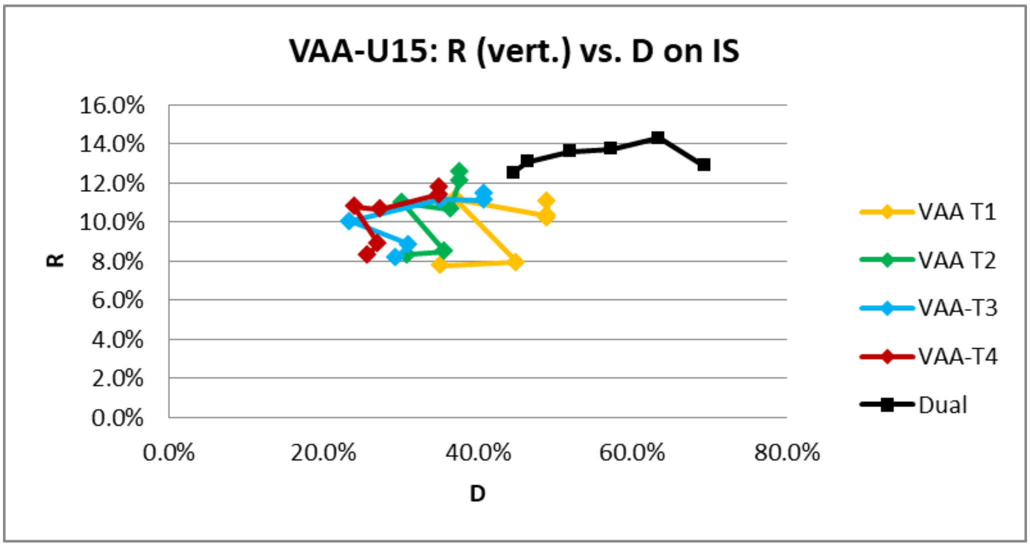

Fig. 15. VAA-U15: R/D frontier on IS (Dec 1926 – Dec 1970) for VAA-T and Dual

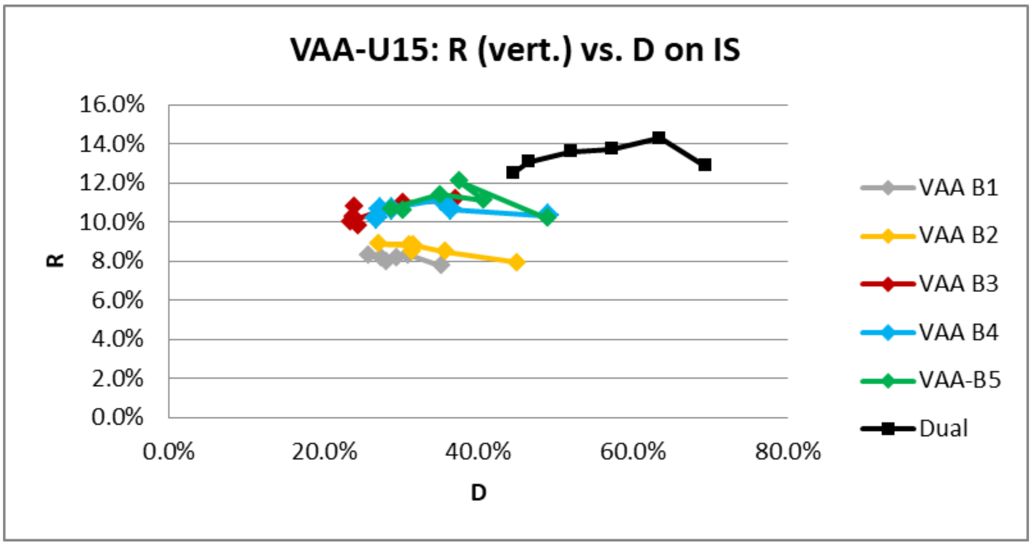

Fig. 16. VAA-U15: R/D frontier on IS (Dec 1926 – Dec 1970) for VAA-B and Dual

From now on we will refer to VAA-U15 as the optimal strategy with T/B=4/3. In Table 10, we show the usual performance indicators for VAA-U15 (with optimal T/B=4/3) on IS (Dec 1926 - Dec 1970).

As we see in Table 10, the return R=11% is larger than that of EW, 60/40 and SPY, while the max drawdown D=24% is less than half of EW (D=67%) and 60/40 (D=62%) and less than 1/3 of that of SPY (D=83%). Sharpe and in particular MAR for VAA-U15 are multiples of those of EW, 60/40 and SPY. Since max Drawdown D for EW, 60/40 and SPY are all larger than 50%, RAD=0%, in contrast to our VAA-U15 which is RAD=7.4%. Dual is better on R but worse on V, D (45% vs. 24% for VAA), Sharpe, MAR and RAD (2.5% vs. 7.4% on VAA). This is all in-sample (IS), so possibly with datasnooping for parameters T and B.

IS	R	V	D	Sharpe	MAR	RAD
VAA-U15	10.8%	12.0%	24.0%	0.66	0.45	7.4%
EW	8.2%	14.6%	67.1%	0.36	0.12	0.0%
EWC	2.9%	2.1%	4.4%	0.00	0.66	2.7%
60/40	7.6%	13.4%	62.1%	0.35	0.12	0.0%
SPY	9.6%	22.1%	83.4%	0.30	0.11	0.0%
Dual6	12.5%	15.5%	44.6%	0.62	0.28	2.5%

Table 10. Performance indicators for VAA-U15 at IS (Dec 1926-Dec 1970)

Out-of-Sample (OS)

The performance indicators for OS (Dec 1970 – Dec 2016) of VAA-U15 (with T/B=4/3) are shown in Table 11. As we can see in Table 11, the return R and max Drawdown D for VAA-U15 out-of-sample are both better than EW (the benchmark), 60/40 and SPY for R (12% vs. 10% for EW and SPY, less for 60/40) but much better for D (12% vs. 30%, 35%, and 51% for EW, 60/40, and SPY, resp.). Sharpe is roughly double that of EW, 60/40 and SPY, while MAR is 3-4 times as good. RAD for VAA-U15 is also positive (RAD=10%) while RAD of EW, 60/40, and SPY at OS are all around 5% or zero (SPY: D>50%). Dual is slightly better on R, but worse on V, D (24% vs. 12% for VAA), Sharpe, MAR and RAD (8.1% vs. 10.1% for VAA). This is all out-of-sample (OS), so without datasnooping for parameters T and B.

OS	R	V	D	Sharpe	MAR	RAD	RS	R	V	D	RAD
VAA-U15	11.7%	8.5%	12.1%	0.57	0.96	10.1%		8.6%	9.75%	10.46%	7.6%
EW	10.4%	10.6%	34.7%	0.34	0.30	4.9%		7.5%	10.7%	34.7%	3.5%
EWC	6.8%	4.7%	8.3%	0.00	0.83	6.2%		3.9%	4.5%	5.0%	3.7%
60/40	9.4%	9.5%	29.9%	0.27	0.31	5.4%		6.0%	8.8%	29.9%	3.5%
SPY	10.4%	15.1%	51.0%	0.24	0.20	0.0%		7.0%	15.3%	51.0%	0.0%
Dual6	11.8%	11.1%	23.8%	0.45	0.50	8.1%		10.5%	10.3%	17.2%	8.3%

Table 11. Indicators for VAA-U15 at OS (Dec 1970-Dec 2016) and RS (Dec 2006- Dec 2016)

In Table 12 we show the main statistics for the Full Sample (FS: Dec 1926 - Dec 2016, so 90 years), including the yearly Total Transaction costs (TTC) and the average Cash Fraction (CF), both over FS.

All the performance stats are better than EW, 60/40, and SPY (R/D=11/24% over nearly 100 years) and return is similar to those at OS, showing limited datasnooping bias at IS. Notice the D=24% (on FS) in the early years (see Fig. 19), which is large for VAA but small compared to D=62-83% for EW, 60/40, and SPY. Dual is better on R, but worse on V, D (45% vs. 24% for VAA), Sharpe, MAR and RAD (2.4% vs. 7.7% for VAA). The TTC shows a monthly turnover of 58% (=1.4/2.4) while average cash is CF=61%, both high and similar to other VAA models.

FS	R	V	D	Sharpe	MAR	RAD	TTC	CF
VAA-U15	11.3%	10.4%	24.0%	0.61	0.47	7.7%	1.4%	60.7%
EW	9.3%	12.7%	67.1%	0.35	0.14	0.0%	0.0%	0.0%
EWC	4.9%	3.7%	8.3%	0.00	0.59	4.4%	0.0%	0.0%
60/40	8.5%	11.6%	62.1%	0.31	0.14	0.0%	0.0%	0.0%
SPY	10.0%	18.8%	83.4%	0.27	0.12	0.0%	0.0%	0.0%
Dual6	12.2%	13.4%	44.6%	0.54	0.27	2.4%	1.0%	6.1%

Table 12. Performance indicators for VAA-U15 at FS (Dec 1926-Dec 2016)

Below (see Fig. 17, 18, and 19) we also give the (log) equity line, rolling 3y returns and drawdowns for VAA-U15, EW and the relative performance of VAA/EW for the full sample (FS). Notice that the relative price in Fig. 17 is roughly flat from 1926, 1939, 1974 and 2009, indicating that returns (not risks) in these three periods are similar to the benchmark (EW). The rolling 3y returns (Fig. 18) of VAA/EW are more than -10% in 1935, 1941 and from 2012, and around zero most of the other years. Max drawdowns (Fig. 19) for VAA-G4 are 24% before 1945, and 12% in all other years. So VAA-U15 shows outperformance in bear markets and similar performance in bull markets, as compared to EW.

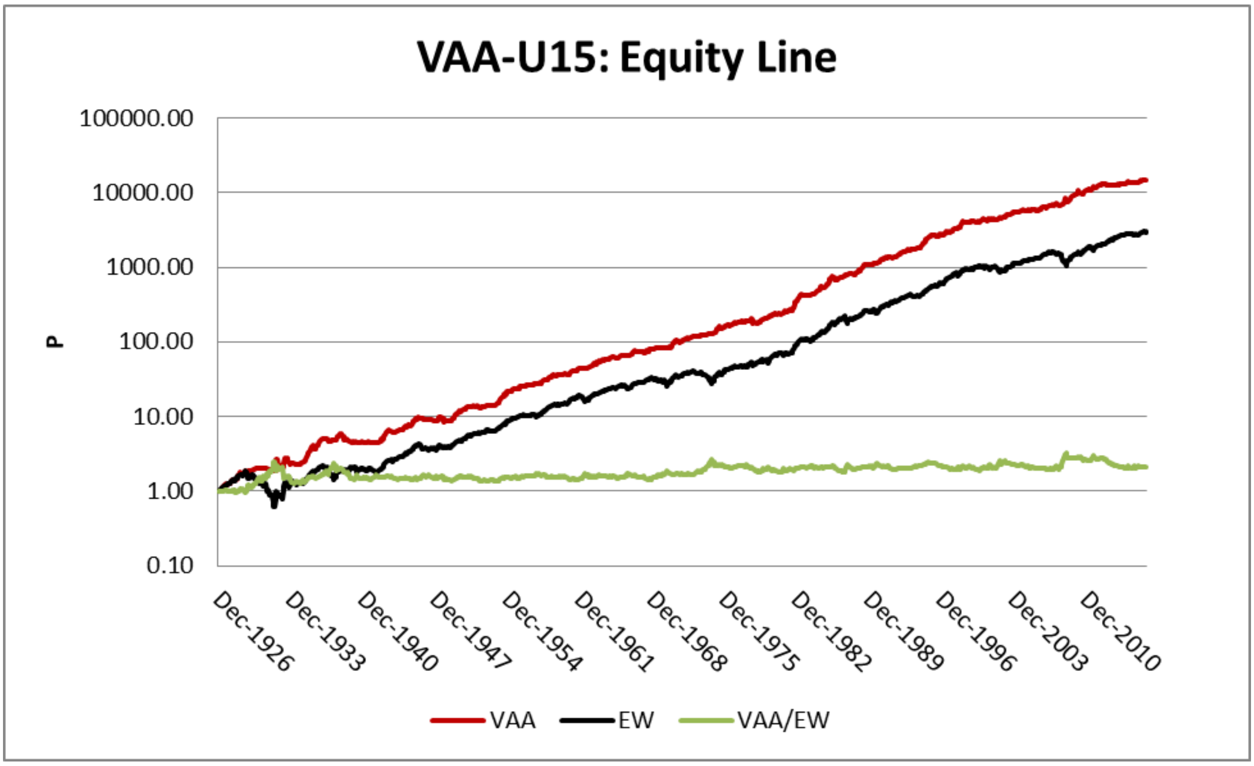

Fig. 17. Equity line for VAA-U15, EW and VAA/EW for FS (Dec 1926 – Dec 2016), log scale

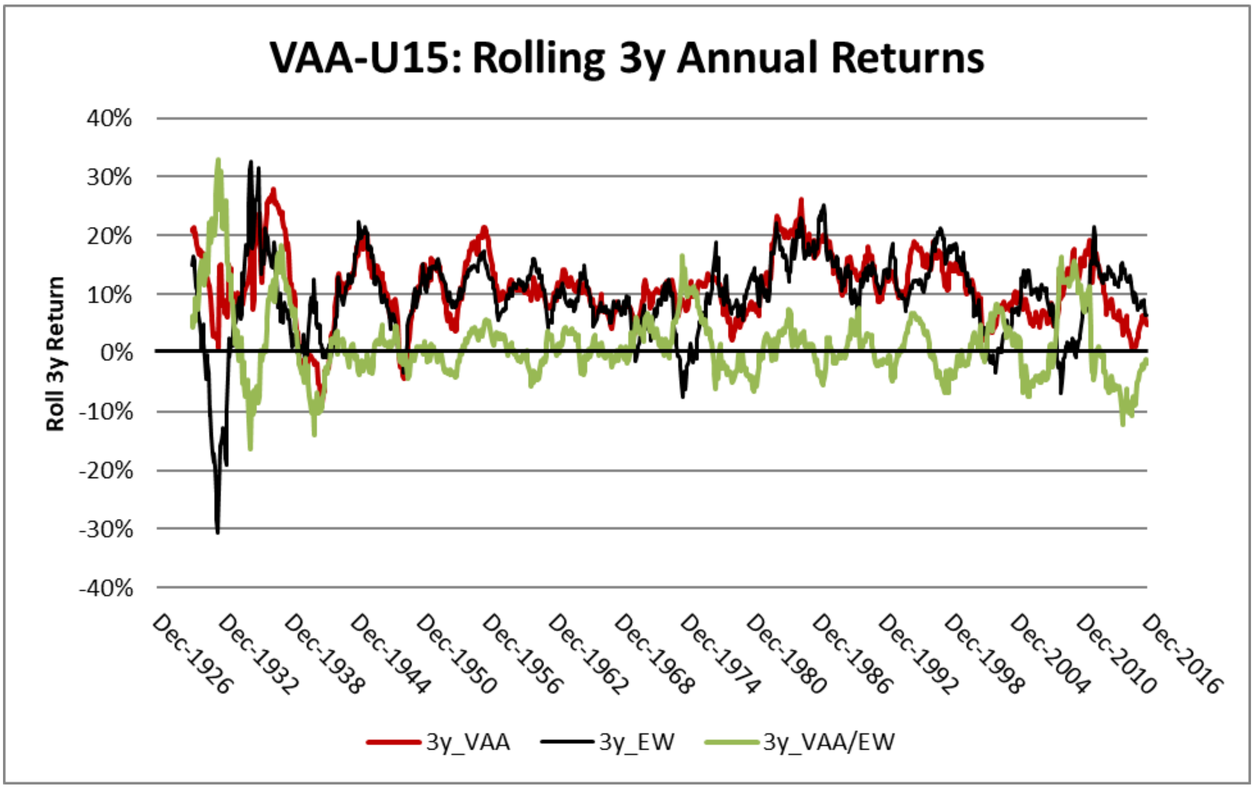

Fig. 18. Rolling 3y Annual Returns for VAA-U15, EW and VAA/EW for FS (Dec 1926 – Dec 2016)

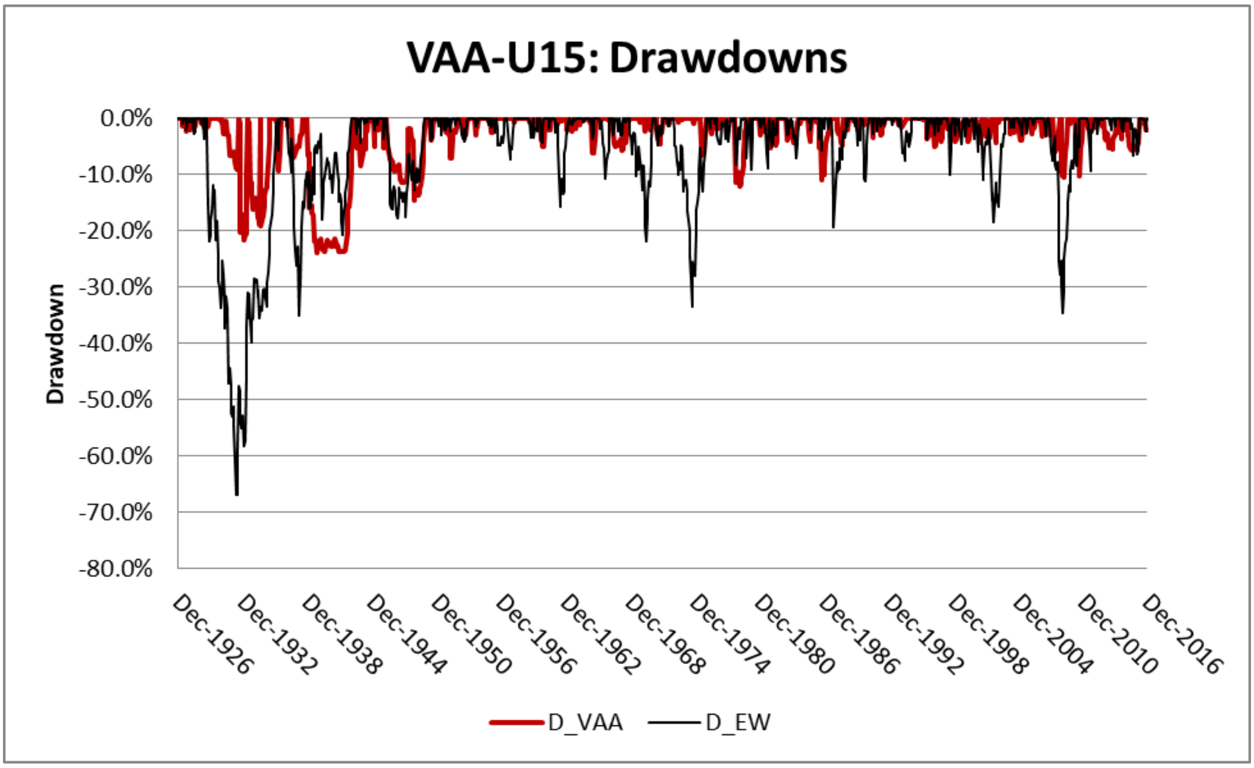

Fig. 19. Drawdowns for VAA-U15, EW and VAA/EW for FS (Dec 1926 – Dec 2016)

12. The VAA-U6 universe

The fourth and last universe is a long backtest (from Dec 1926) with a small universe, using the Ibbotson (Ibb) and Fama French (FF) dataset. We will use a risky universe of 6 assets, consisting on the four US factors Size x Value from FF plus two US bonds (30d T-Bill, and LT Corp) from Ibb, so N=6. As with VAA-U15 (see section 11), we will use 30d T-Bill, IT Gov and LT corp from Ibb as Cash, similar to SHY, IEF and LQD for the shorter dataset (see Sections 9 and 10 above). So, all assets refer to the US. We will therefore denote our VAA strategy with this N6+3 US universe as VAA-U6. Our in-sample (IS) period is from (ult.) Dec 1926 – Dec 1970 (44 years), while our out-of-sample (OS) period is Dec 1970 - Dec 2016 (46 years). We use index data (EOM Total Returns) from Ibb/FF, so not corrected for ETF fees etc.

In-sample (IS)

We compute RAD on IS over top T=1..6 and breadth B=1..6 for a total of 6x6=36 values (scenarios), using our 13612W momentum filter. Dual is the dual momentum strategy computed for various T=1..6 using our own momentum filter (13612W) for both absolute and relative momentum. From Table 13, the best VAA score is RAD=4.9% for T=6 and B=1, and for Dual RAD=2.2% for T=6. We can see more detail behind these T/B values when we draw the R/D frontier as function of T and B, as in Fig. 20 and 21. We include (as with VAA-G4) again the (100% cash) point B=0 in Fig. 20 to show the knee more clearly and leave out some overlapping VAA-T and VAA-B lines.

							T
RAD	4.9%	1	**2**	**3**	**4**	**5**	**6**
	0	2.3%	2.3%	2.3%	2.3%	2.3%	2.3%
	1	2.8%	4.0%	4.7%	4.3%	4.7%	4.9%
	2	0.0%	0.6%	0.0%	0.7%	1.6%	3.0%
	3	0.0%	0.0%	0.0%	0.0%	0.1%	2.1%
	4	0.0%	0.0%	0.0%	0.0%	0.0%	2.2%
	5	0.0%	0.0%	0.0%	0.0%	0.0%	2.2%
B	6	0.0%	0.0%	0.0%	0.0%	0.0%	2.2%
	Dual	0.0%	0.0%	0.0%	0.0%	0.0%	2.2%

Table 13. RAD for VAA-U6 on IS (Dec 1926-Dec 1970) for T/B=1..6

As we see from and Table 13 and Fig. 20, 21 (red line), the best point is T/B=6/1. The VAA-T6 line (so with all 6 assets included when b=0, and therefore no relative momentum) in Fig. 20 is less efficient than VAA-T2 and T4, except when D<30% where R/D=7/25% wins at B=1. The red VAA-T6 line in Fig. 20 has less clear "bend and knee" than in our global (G12/4) universes, but there is still a knee at B=1. The Dual line is less efficient in terms of return/risk to all VAA-T lines (Fig. 20) and to VAA-B1 (Fig. 21). From Fig. 21 we see that VAA-B1 for T>1 is also optimal in terms of return/risk to all other points with D<50%. The best Dual strategy (see also table 13) is T=6 (left point of the black Dual line).

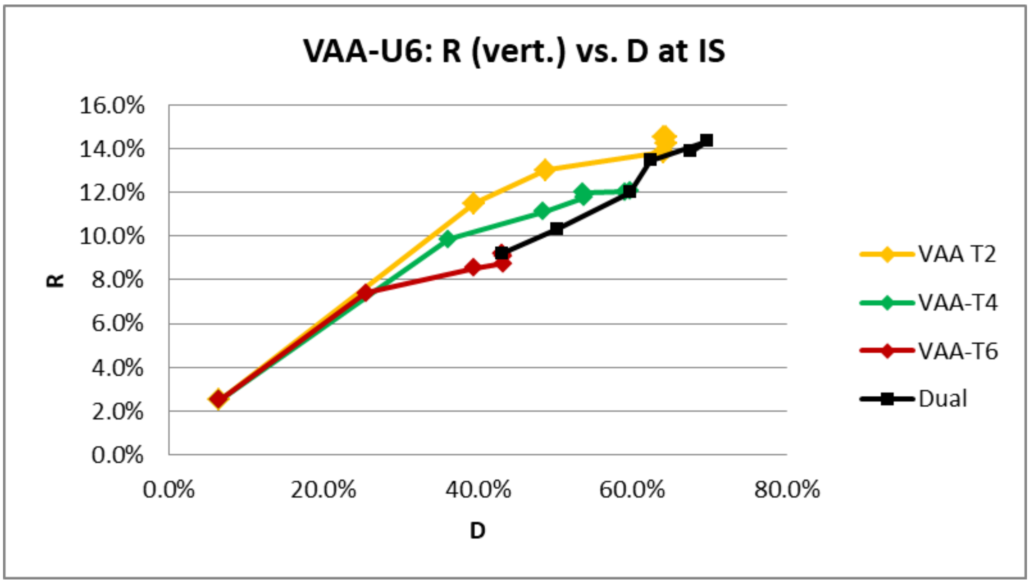

Fig. 20. VAA-U6: R/D frontier on IS (Dec 1926 - Dec 1970) for VAA-T

From now on we will refer to VAA-U6 as the optimal strategy with T/B=6/1. Notice that relative momentum (eg. between the 4 factors) is disabled since T=N=6. Only absolute momentum (at the universe level) is relevant in a very protective way (B=1, so CF=100% when any of the 6 assets becomes bad), like with the small global strategy VAA-G4 (see Section 10).

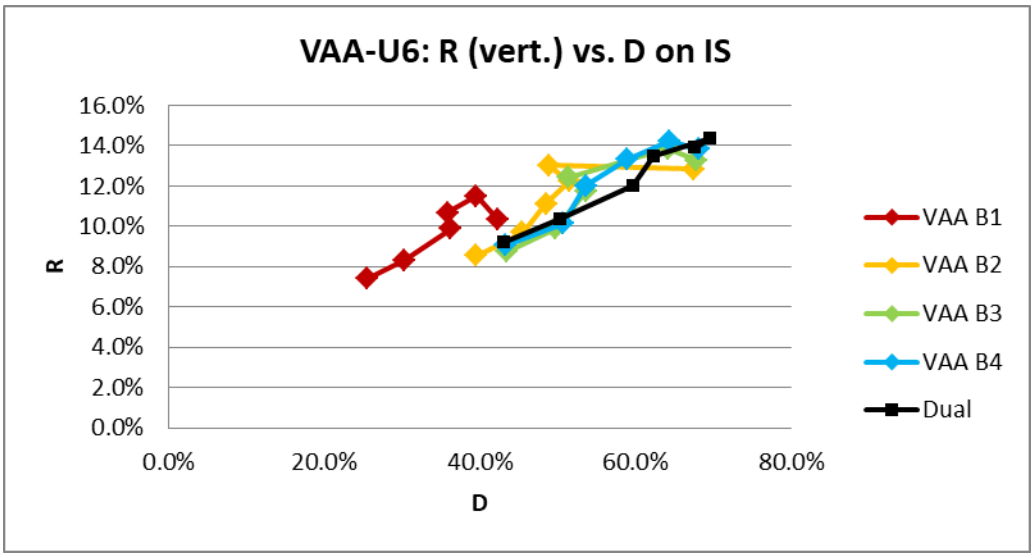

Fig. 21. VAA-U6: R/D frontier on IS (Dec 1926 – Dec 1970) for VAA-B and Dual

In Table 14, we show the familiar performance indicators for VAA-U6 (with optimal T/B=6/1) at IS (Dec 1926- Dec 1970). As we see in Table 14, the return R=7.4% is smaller than that of EW (R=9.1%), 60/40 (7.6%), SPY (9.6%), and Dual6 (9.2%), while the max drawdown D=25% is less than Dual (43%) and less than half of EW (71%) and 60/40 (62%) and 1/3 of that of SPY (83%). Sharpe for VAA-U15 (0.42) is better than of those of EW (0.34), 60/40 (0.35) and SPY (0.30), but not better than Dual (0.49). MAR of VAA-U6 (0.29) is nearly three times that of EW (0.13), 60/40 (0.12), and SPY (0.11), and better than Dual (0.21). Since max Drawdown D for EW, 60/40 and SPY are all larger than 50%, RAD=0%, in contrast to our VAA-U6 which has RAD=4.9%, while for EWC and Dual RAD is 2.7% and 2.2%, resp. This is all in-sample (IS), so possibly with datasnooping for parameters T and B.

IS	R	V	D	Sharpe	MAR	RAD
VAA	7.4%	10.7%	25.4%	0.42	0.29	4.9%
EW	9.1%	18.7%	71.0%	0.34	0.13	0.0%
EWC	2.9%	2.1%	4.4%	0.00	0.66	2.7%
60/40	7.6%	13.4%	62.1%	0.35	0.12	0.0%
SPY	9.6%	22.1%	83.4%	0.30	0.11	0.0%
Dual6	9.2%	13.0%	43.2%	0.49	0.21	2.2%

Table 14. Performance indicators for VAA-U6 at IS (Dec 1926-Dec 1970)

Out-of-Sample (OS)

The performance indicators for OS (Dec 1970 – Dec 2016) of VAA-U6 (with T/B=6/1) are shown in Table 15. As we can see in Table 15, the return R=10.8% for VAA-U6 out-of-sample is slightly better than EW (10.6%, the benchmark), 60/40 (9.4%) and SPY (10.4%). D is much better (12% vs. 39%, 30%, and 51% for EW, 60/40, SPY, resp. Sharpe (0.48) is nearly double that of EW, 60/40 and SPY, while MAR is three times as good. RAD=9.3% for VAA-U6, while RAD of EW, and 60/40 are roughly half of that, with RAD=0% for SPY at OS, since D>50%. Dual is better than VAA on R (9.2% vs 7.4% for VAA) but worse on V, D (43% vs. 25% for VAA), Sharpe, MAR and RAD (8.4% vs. 9.3% for VAA). The picture for the last decade (RS: Dec 2006 - Dec 2016) is similar to that for OS, although with a lower return R. Dual is slightly better than VAA on R, V, D, and RAD (7.1% vs. 6.1% on VAA). This is all out-of-sample (OS), so without datasnooping for parameters T and B.

OS	R	V	D	Sharpe	MAR	RAD	**RS**	R	V	D	RAD
VAA-U6	10.8%	8.1%	12.0%	0.48	0.90	9.3%		6.9%	9.5%	10.5%	6.1%
EW	10.6%	12.2%	39.4%	0.31	0.27	3.7%		6.6%	12.6%	39.4%	2.3%
EWC	6.8%	4.7%	8.3%	0.00	0.83	6.2%		3.9%	4.5%	5.0%	3.7%
60/40	9.4%	9.5%	29.9%	0.27	0.31	5.4%		6.0%	8.8%	29.9%	3.5%
SPY	10.4%	15.1%	51.0%	0.24	0.20	0.0%		7.0%	15.3%	51.0%	0.0%
Dual6	10.8%	9.1%	18.1%	0.44	0.60	8.4%		8.0%	9.2%	9.9%	7.1%

Table 15. Performance indicators for VAA-U6 at OS (Dec 1970-Dec 2016) and RS (Dec 2006-Dec 2016)

In Table 16 we show the main statistics for the Full Sample (FS: Dec 1926 - Dec 2016, so 90 years), including the yearly Total Transaction costs (TTC) and the average Cash Fraction (CF), both over FS. While return R=9.1% are slightly less than EW (9.9%), SPY (10.0%), and Dual (10%), it is slightly better than 60/40 (8.5%). The max Drawdown (D=25%) is nearly three times smaller than that of EW (71%), 60/40 (62%), and SPY (83%). All the return/risk stats like Sharpe, MAR and RAD are better than EW, 60/40, and SPY, with RAD=6% while zero for EW, 60/40 and SPY. Notice the D=25% (on FS) in the early years (see Fig. 24), which is large for VAA but small compared to D=62-83% for EW, 60/40, and SPY. Dual is slightly better at R and Sharpe, but clearly worse on D, MAR and RAD (2.4% vs. 6.0% for

VAA). The TTC shows a monthly turnover of 50% while average cash is CF=58%, both high and similar to other VAA models. The TTC and average CF of Dual are roughly half of VAA.

FS	R	V	D	Sharpe	MAR	RAD	TTC	CF
VAA-U6	9.1%	9.5%	25.4%	0.44	0.36	6.0%	1.2%	57.7%
EW	9.9%	15.7%	71.0%	0.32	0.14	0.0%	0.0%	0.0%
EWC	4.9%	3.7%	8.3%	0.00	0.59	4.4%	0.0%	0.0%
60/40	8.5%	11.6%	62.1%	0.31	0.14	0.0%	0.0%	0.0%
SPY	10.0%	18.8%	83.4%	0.27	0.12	0.0%	0.0%	0.0%
Dual6	10.0%	11.2%	43.2%	0.46	0.23	2.4%	0.6%	24.5%

Table 16. Performance indicators for VAA-U6 at FS (Dec 1926 - Dec 2016)

Below (see Fig. 22, 23, and 24) we also give the (log) equity line, rolling 3y annual returns and drawdowns for VAA-U6, EW and the relative price of VAA/EW for the full sample (FS). Notice that the relative price in Fig. 22 is decreasing from 1933-1969, and mostly flat after that, indicating that returns (not risks) in the first period are less and in the last period similar to the benchmark (EW). The rolling 3y returns (Fig. 18) of VAA/EW are more than -20% in 1935, and 1941, and between +10% and -10% most of the other years with a +30% score during the Great depression. Max drawdowns (Fig. 24) for VAA-U6 are 25% before WW2 and 12% in all other years.

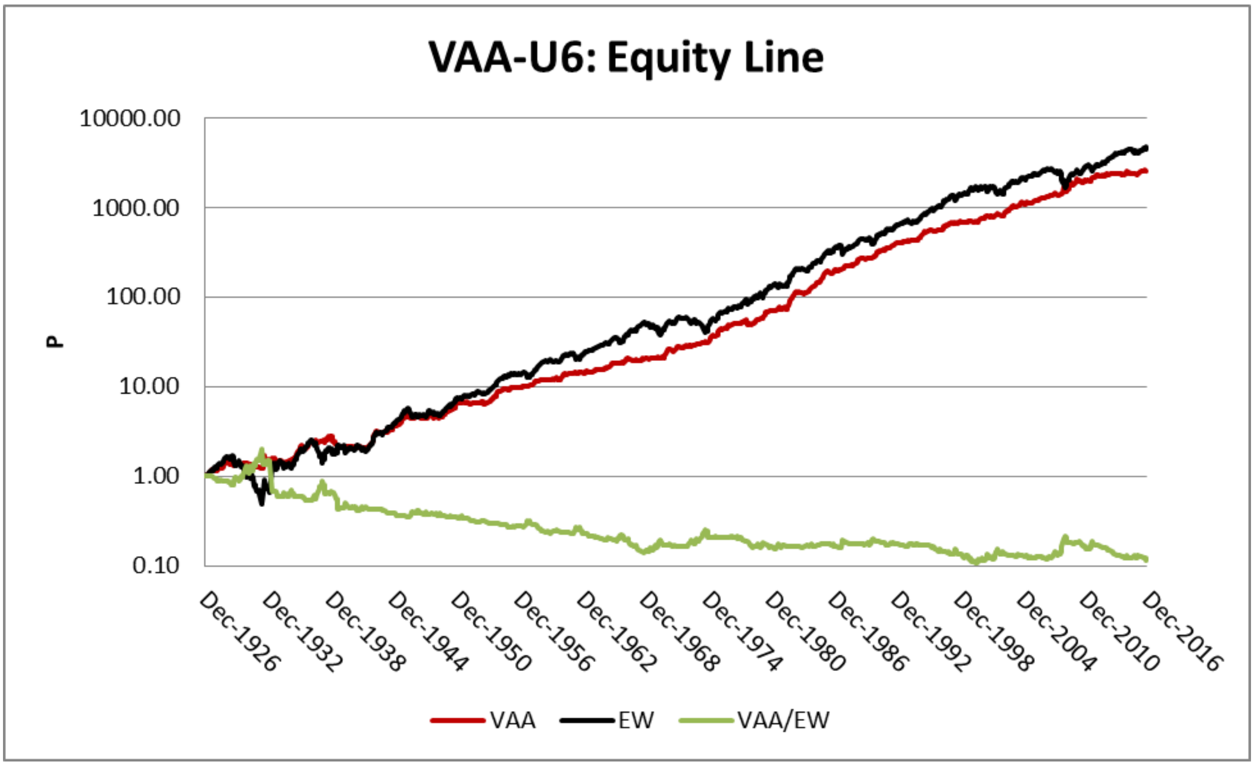

Fig. 22. Equity line for VAA-U6, EW and VAA/EW on FS (Dec 1926 – Dec 2016), log scale

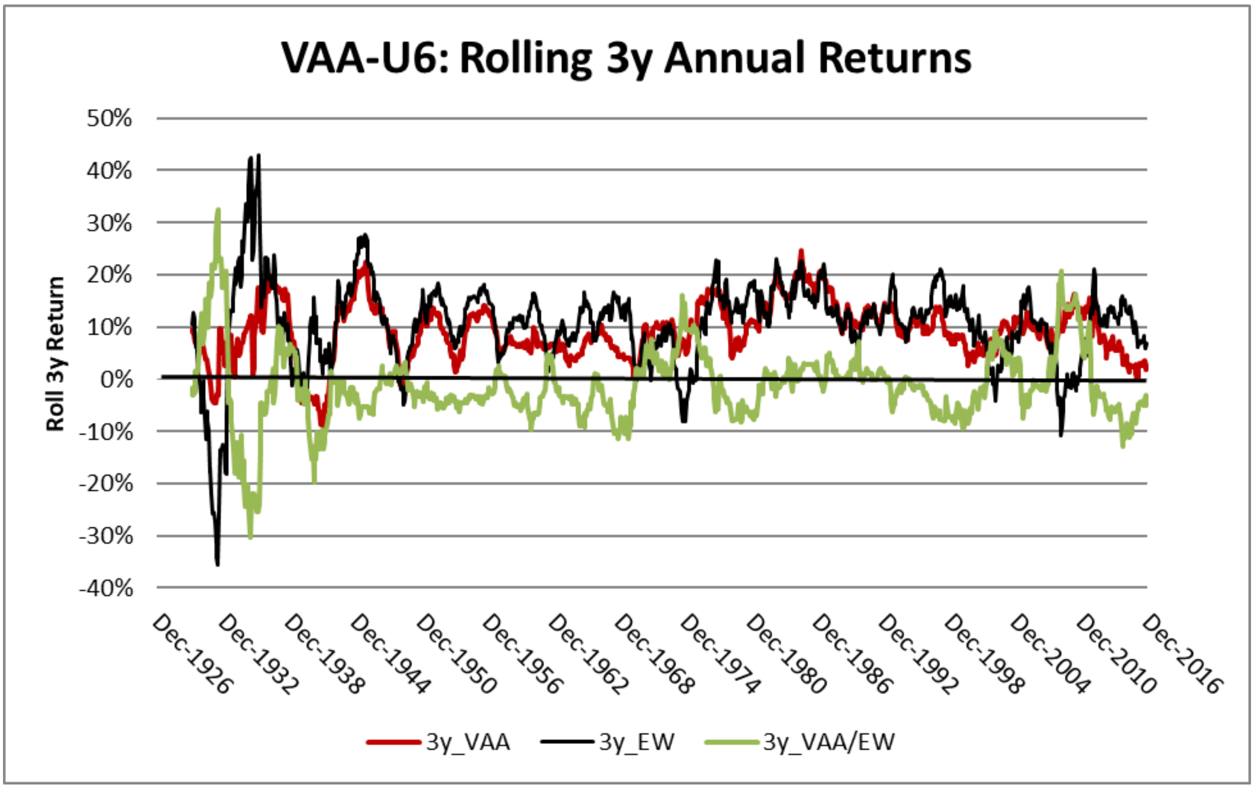

Fig. 23. Rolling 3y returns for VAA-U6, EW and VAA/EW on FS (Dec 1926 – Dec 2016)

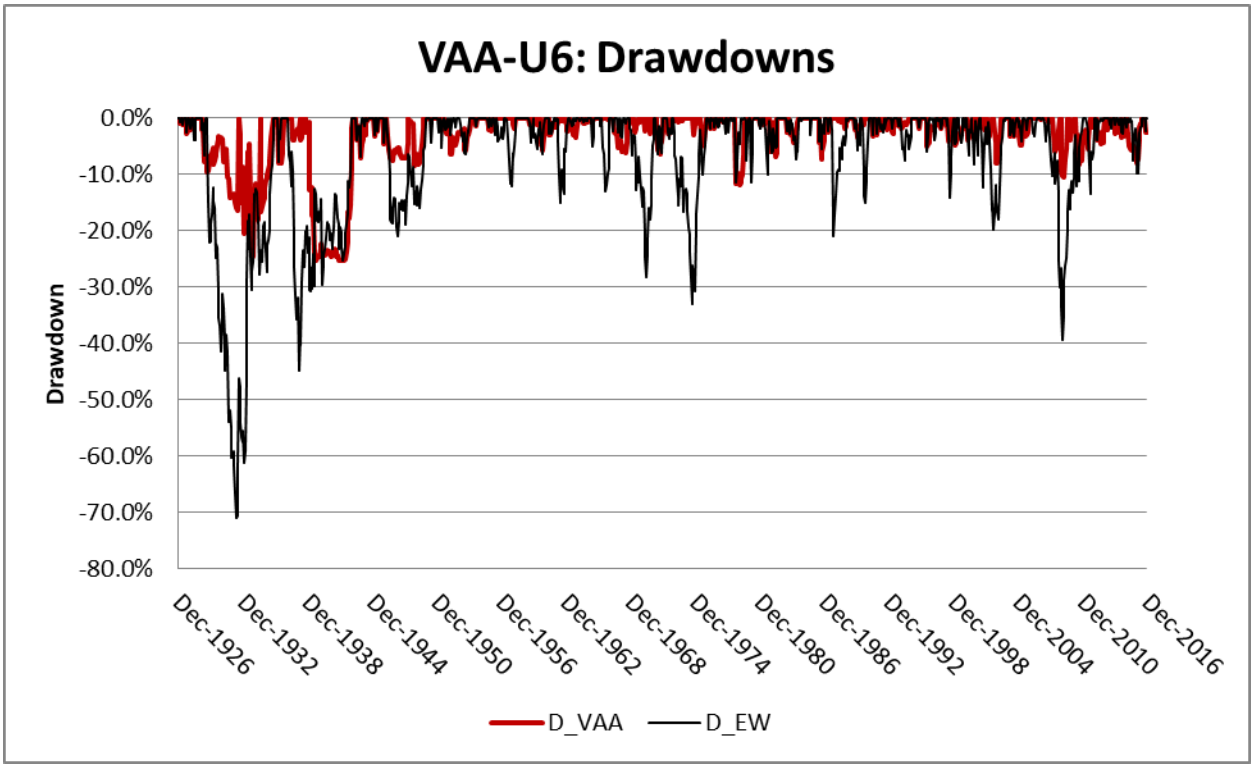

Fig. 24. Drawdowns for VAA-U6, and EW on FS (Dec 1926 – Dec 2016)

13. Summary and concluding remarks

In this paper, we have tested a dual momentum strategy (dual: with both absolute and relative momentum) with vigorous crash protection, called Vigilant Asset Allocation (VAA). However, instead of absolute momentum (trendfollowing) at the individual asset level, for VAA we use *breadth momentum* at the universe level for crash protection, as we did with PAA. Compared with PAA, however, we now use the number of bad assets (with non-positive momentum) in the universe (of size N) relative to a *breadth protection threshold* B (or breadth B, for short, with B<=N) as a more granular crash indicator. As we demonstrated, in-sample optimization of this breadth B often leads to 100% crash protection when only one or a limited number of all assets in the universe are bad.

We have shown that with this aggressive crash protection and ditto momentum filter (13612W), we arrived out-of-sample (OS) at returns above 10% with max drawdowns of less than 15%, even over periods of more than 40 years (Dec 1970 – Dec 2016). In Table 17 we show the OS results for our four universes, with 12 and 4 global assets (VAA-G12 and VAA-G4, resp.) and with 15 and 6 US assets (VAA-U15 and VAA-U6, resp.).

OS 1993-2016	R	V	D	Sharpe	MAR	RAD	N	T	B	TTC	CF
VAA-G12	10.5%	10.3%	13.0%	0.51	0.81	8.9%	12	2	4	1.4%	59%
VAA-G4	16.0%	11.7%	10.4%	0.92	1.54	14.1%	4	1	1	1.3%	57%
EWC	5.3%	4.4%	4.7%	0.0	1.11	5.0%	3	-	-	0%	100%
60/40	8.2%	8.7%	29.4%	0.34	0.28	4.8%	2	-	-	0%	40%
SPY	9.1%	14.6%	50.8%	0.26	0.18%	0.0%	1	-	-	0%	0%

OS 1970-2016	R	V	D	Sharpe	MAR	RAD	N	T	B	TTC	CF
VAA-U15	11.7%	8.5%	12.1%	0.57%	0.96	10.1%	15	4	3	1.4%	61%
VAA-U6	10.8%	8.1%	12.0%	0.48	0/90	9.3%	6	6	1	1.2%	58%
EWC	6.8%	4.7%	8.3%	0.0	0.83	6.2%	3	-	-	0%	100%
60/40	9.4%	9.5%	29.9%	0.27	0.31	5.4%	1	-	-	0%	40%
SPY	10.4%	15.1%	51.0%	0.24	0.21	0.0%	1	-	-	0%	0%

Table 17. Out-of-sample results for two global (VAA-G) and two US (VAA-U) universes

As above, Return (CAGR), Volatility and max Drawdown are all based on end-of-month (EOM) TR index prices (so including dividends etc.) with one-way transaction costs of 0.1% and only corrected for ETF fees in the VAA-G universes. The Sharpe ratio is defined relative to the returns of the equal weight of the three-bond cash universe (EWC), MAR=R/D, and RAD is our own Return Adjusted for Drawdowns measure (see section 6).

Dual momentum is applied to the assets using the responsive 13612W (average *annual* returns over past 1,3,6, and 12 months) momentum filter for both absolute (trend following) and relative (cross sectional) momentum. Traditionally, absolute momentum implies trendfollowing where bad assets are replaced by cash as a form of crash protection. With VAA, we apply crash protection on the universe, using the number of bad assets b of the risky universe relative to breadth (so b/B) as the cash fraction (CF).

With VAA, this breadth momentum driven crash protection might give rise to extensive trading since we do not necessarily replace individual assets by cash, as with traditional dual momentum. Therefore, we used for VAA a so-called "Easy Trading" (ET) approach where we limited cash to multiples of 1/T, where T is the number of top assets used in relative momentum. The result is similar to traditional dual momentum, where we replaced risky assets by cash, but now much earlier, depending on the (rounded) ratio b/B.

"Cash" itself is also based on a (simplified) momentum model, using the same 13612W filter for relative momentum only for picking the best bond of our cash universe with the highest momentum, without regard to sign (so no absolute momentum). The cash universe for the shorter, global backtest from Dec 1970 (VAA-G) encompasses BIL, IEF, and LQD proxies, while the longer, US backtest uses similar indices (T-Bill, IT Gov, and LT Corp) from Ibbotson.

Using ET, the cash fraction CF in the above four VAA models is based the two parameters T and B. The top number T equals the maximum number of best risky assets allocated and is well known from dual momentum. The breadth parameter B is new for VAA. Both parameters B and T are estimated for each universe (by an in-sample optimization of RAD), see Table 17. Notice that the optimal parameters T (top) and B (breadth) increases with the universe size N, with T and B mostly of similar magnitude except for the N=6 universe (where T=6 and B=1), which seems an exception (due to too little diversification?).

Both small universes (VAA-G4 and VAA-U6) obtained optimal in-sample (IS) results with B=1 and therefore go completely to cash when only one of the (four or six) risky assets turns bad. For the VAA-G12 universe with 12 risky assets, B=4 and T=2 proved to be optimal resulting in a cash fraction CF=100% when *four* or more of the twelve assets are bad, while for traditional dual this only happens when *all* risky assets are bad. Both examples demonstrate the responsiveness of our crash protection in all four VAA models compared to the traditional dual approach.

Therefore, we were out of the market (in cash) in all VAA strategies for roughly 60% of the time. However, we still arrived out-of-sample at nearly offensive (R>10%) returns with very defensive drawdowns (D<15%). This demonstrates the responsive timing of breadth momentum in our VAA strategies and our motto "winning more by losing less". We also surpassed Antonacci's highly successful GEM model in return and return/risk with our VAA-G4 strategy but with more turnover and higher cash fraction. This is the price we pay for our aggressive strategy.

With such high cash fractions, the strategy of the cash universe becomes important, even more so with the current low and increasing yields of the coming years. However, similar rising yields were also present from 1950 and in the seventies (1970 – 1982) when eg. both VAA-G universes clearly outperformed EW and EWC (with double RAD's), while the same holds in the years after 1982 with the cash tailwind (decreasing yields).

Our four VAA strategies show a remarkable characteristic in our breadth momentum approach. This is the “knee” in the R/D space, going from low breadth B to higher values. Before the knee, return increases with B with nearly stable drawdown while after the knee, drawdown increases while returns are nearly stable. Also, the optimal breadth parameter B turns out to be the same for both VAA-G universes in both the in-sample and the out-of-sample periods, which demonstrates robustness. This all demonstrates some remarkable timing characteristic of our breath momentum approach for various regimes and universes. Against our expectations, it could be all just luck, but if not, future research should try to answer why. Why does this breath momentum work, what is the role of the universe used, is there a relationship with diversification and correlation between assets, and can we repeat the knee with other universes? Is breadth B constant or should we use some adaptive walk-forward scheme? Enough questions, enough to be done.

14. Literature

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

Antonacci, G., 2013b, Risk Premia Harvesting Through Dual Momentum SSRN 2042750

Antonacci, G., 2014, Dual Momentum Investing, McGraw Hill (book)

Asness, C.S., A. Frazzini, R. Israel, and T.J. Moskowitz, 2014, Fact, Fiction and Momentum Investing. Journal of Portfolio Management, Fall 2014, SSRN 2435323

Beekhuizen, P. and W.G. Hallerbach, 2015, Uncovering Trend Rules, SSRN 2604942

Faber, M. T., 2007, A Quantitative Approach to Tactical Asset Allocation, Journal of Wealth Management, Spring 2007. Updated in Faber (2013).

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

Faber, M. T., 2013, A Quantitative Approach to Tactical Asset Allocation, SSRN 962461. Update of Faber (2007).

Faber, M.T., 2017, The Idea Farm (subscription), http://www.theideafarm.com/about/

Faber, Nathan, 2015, The Search for Crisis Alpha: Weathering the Storm Using Relative Momentum, ThinkNewfound.com (paper)

French, K.R., 2017, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Goldberg, L.R., and O. Mahmoud, Drawdown from Practice to Theory, arxiv.org/pdf/1404.7493.pdf

Harvey, C.R., and Y. Liu, 2013, Backtesting, SSRN 2345489

Harvey, C.R., and Y. Liu, 2014, Evaluating Trading Strategies, SSRN 2474755

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

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

Jones, W., 2017, https://allocatesmartly.com/list-of-strategies/

Keller, W.J. and A. Butler, 2014, A Century of Generalized Momentum; From Flexible Asset Allocations (FAA) to Elastic Asset Allocation (EAA), SSRN 2543979

Keller, W.J., A. Butler, and I. Kipnis, 2015, Momentum and Markowitz: A Golden Combination, SSRN 2606884

Keller, W.J., and J.W. Keuning, 2016, Protective Asset Allocation (PAA), SSRN 2759734

Levine, A. and L. H. Pedersen, 2015, Which Trend is Your Friend?, SSRN 2603731

Moskowitz, T., Y.H. Ooi, and L.H. Pedersen, 2011, Time Series Momentum, Working Paper nr. 79, The Initiative on Global Markets, University of Chicago.

Newfound, 2015, Two Centuries of Momentum, Thinknewfound.com (paper)

Paulsen, D. and J. Söhl, 2016, Noise Fit, Estimation Error and a Sharpe Information Criterion, SSRN-2735087

TrendXplorer, 2016, http://indexswingtrader.blogspot.nl/2016/10/prospecting-dual-momentum-with-gem.html

TrendXplorer, 2017, http://indexswingtrader.blogspot.nl/2017/03/index-mapping-for-etf-proxies.html

Zakamulin, V., 2015, Market Timing with Moving Averages: Anatomy and Performance of Trading Rules, SSRN 2585056