ASSET MANAGEMENT ASSIGNMENT #2
TACTICAL ASSET ALLOCATION

Feng Lin

Fulvio Mazza

Matteo Rubino

Maurice Kees

Michel Mustapha Raggio

Paolo Brandellero

Abstract/Aim

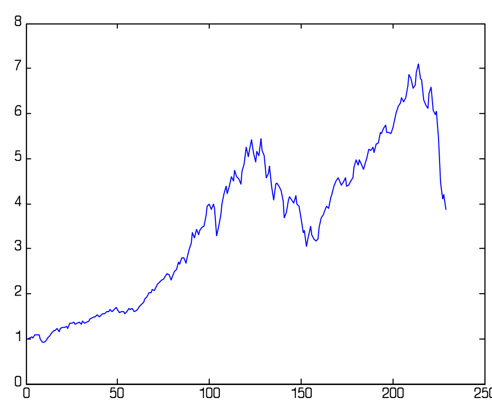
The purpose of our work is to create a tactical asset allocation model with a three months length of investment using data from 1990 to 2008 of 48 different US industries, trying to build a long/short financial product able to get a sustainable return. We have to keep in mind that this return has the goal to beat the benchmark that is represented by the market index.
We divide this dataset in two parts: an in-sample period (from 1990 to 2001) and an out-of sample period (from 2002 to 2008). In this way we use the first part of the dataset trying to see which signal, among all the sixteen available, would have been more significant in explaining these returns. From these forecasts we then assign weights to each characteristic and develop a final mean variance efficient portfolio. On the other side, we used the out of sample database to back-test the result of this portfolio comparing the result with our benchmark.

Data description

The panel data considered is composed by monthly returns for 48 different assets for a time horizon of 18 years. For all these returns we have Accounting, Valuation, Expectation, Momentum and Size Factors. Each of this category of signals has to be verified through univariate regressions. Respectively we have: **imv** as Size Factor, **ep** and **bm** as valuation factors, **mom1 mom3** and **mom12** as momentum factors, **roa, roe, leverage, cur_rat, cash_rat, ass_turn, eq_turn** as accounting factors, **rev_ratio FY1_3mch** and **FY1_12mch** as expectation factors

The benchmark

We are trying to get sustainably higher than average returns. In order to see what the average returns are, we need to have a measure of the average. For this we created a benchmark.
In our case the benchmark is constructed from a weighted average of the industries. It is rebalanced every month to offset the difference in returns among industries, so that there is a new weight every month



Picking the right signals

In order to get good and consistent results it is of great importance to pick the right signals that can forecast returns. In order to select for useful signals that can consistently predict future returns we look first at the economic rationale for which a signal should be able to forecast returns. Then we look at the ability to forecast returns by using univariate regressions on these signals in a screening period.

Economic rationale

Sounded economic theory has to be the first criteria to use in order to discriminate among all the factors. Also according to models like the Fama-French three factors model and Carhart 4-factors model, we started to outline in our mind which signals could have been effective in explaining our returns as a final multivariate model. For example, Fama and French adopted a three factors model where the factors chosen were the market risk premium, the market risk premium based on size, the market risk premium based on the book to market ratio, therefore they used **size factors and valuation factors**; while Carhart adopted a four factors model where adding also the **momentum factor**.

Therefore, our theoretical starting point was to pick up not more than two screening factors for each category of factors in order to be effective (thinking that each of these categories could have had a value added) and also in order to try to avoid the multi-collinearity issues among explanatory variables (then we will show that we will use a **precise procedure** to eliminate some significant variables that could be collinear). For instance, we supposed that **imv**, which represents the small cap effect therefore we would have seen a positive relation with our returns. About valuation factors we hypothesized that **ep** would have been positive related with returns, while about momentum factors we would have imagined that **mom_12** would have been positive correlated with returns more than mom_3 and mom_6. For the last two categories of factors: accounting factors and expectation factors we can say that we thought to find a positive relation between each factor and our dependent variable.

Statistical rationale

A model like $\boldsymbol{Y_i = \beta X_i + u_i}$ represents our univariate regression, so substituting each time a different screening factor in our Xi, we will see different values of Betas assessing if our screening factor is significant or not. We can see from the next chart that we have 9 significant signals out of 16. The statistics used to choose the right signals are: the p-value (we used a confidence level<0.1) and the

relative size of the percentage of positive and significant betas versus the percentage of negative and significant betas depending on the theoretical relation that we expected to find. For example, for **ep** we have a significant p-value (0,0653) and a value of 42,86% for the percentage of Pos e Sign Betas against a value of 10,71% Neg e Sign Betas. So, we take the first statistics as a way to demonstrate the overall significance of the info signals, while the second statistics as a way to confirm our prevision about the right sign that our independent variable should have. So in this case **ep** is significant and also positive in accordance with the economic theory. However, if we look at **imv** we see a contradiction: we have a negative beta while we should find a positive beta. This is only a matter of sampling and data characteristics because when we will show the result of our dynamic model (next paragraph) we will see the right relation working (a positive beta in normal economic conditions or booming conditions and a negative signal in conditions of recession).
For all the other signals then reported the statistics are in line with the economic theory mentioned in the previous paragraph.

IN-SAMPLE	1	2	4	6	7	8	13	14	15
Panel regression	**Imv**	**ep**	**mom1**	**mom12**	**Roa**	**roe**	**eq_tur**	**rev_ratio**	**FY1_3mch**
% Positive beta	25,00%	78,57%	75,00%	71,43%	75,00%	75,00%	67,86%	89,29%	82,14%
% Negative beta	75,00%	21,43%	25,00%	28,57%	25,00%	25,00%	32,14%	10,71%	17,86%
% Significant beta	57,14%	53,57%	39,29%	75,00%	50,00%	35,71%	17,86%	35,71%	50,00%
% Pos & Sig beta	14,29%	42,86%	35,71%	57,14%	39,29%	32,14%	17,86%	35,71%	46,43%
% Neg & Sig beta	42,86%	10,71%	3,57%	17,86%	10,71%	3,57%	0,00%	0,00%	3,57%
mean R2	34,66%	34,90%	34,29%	36,87%	34,48%	34,28%	34,28%	34,31%	34,33%
P-Value	0,0390	0,0653	0,0015	0,0319	0,0503	0,0044	0,0158	0,0000	0,0001
Mean Coeff. x 100	-0,3720	0,3943	0,4059	0,9191	0,2806	0,3008	0,3166	0,4708	0,4451
Median	-0,5295	0,3618	0,2838	0,5723	0,4911	0,3355	0,1536	0,5111	0,4604
1st Quartile	-0,9382	0,0684	-0,0037	-0,0737	-0,0189	0,0393	-0,0865	0,2097	0,2637
3rd Quartile	0,0413	0,7829	0,5306	1,8799	0,7938	0,6151	0,5878	0,7402	0,7166

Dynamic model

Actually, before making our final choice we got the intuition that some signals might change their behaviour according to the some industry/macroeconomic conditions. This is nothing else than the underlying idea of Dynamic Models.
Let's make a very simple example. Think about credit risk. When you buy a bond of a company not

free from the default of risk, as an investor you get paid for the counterparty risk. Still, you're bearing the risk of not being repaid the notional you invested. This means that, till time is good credit risk has a positive beta w.r.t. return, but if we are in bad times it is likely that you will lose all your money.

This fact applies for sure also to some signals, meaning that these would be actually risk factors. How do we test for this? First we find a measure to see if we are in good or bad economic times and then we construct a new signal from the combination of the old one and the measure of the state of the economy. To see if the economy is bullish or bearish, we look at the sign of the return of the benchmark over the last three months.

We wanted to test two opposite possibilities, so we declined our indicator into two models.

In the first model our assumption was that the signal worked during healthy market and that it didn't work during bad economic conditions. Practically, we kept the signal when the indicator was positive, while we set it to zero in the other case, setting to zero also the return on which we regressed the signal. This technique is equivalent to deleting entirely some periods from our database (for instance this method is used to cancel out the data from September 2001 to avoid any kind of information distortion due to 9/11 events).

In the opposite model we did the opposite. We kept the signal as is when our indicator was negative and we set to zero data when our "market health" indicator was positive.

The next step was to run again univariate panel regression on the dynamic signals.

Case 1/0 (keep the signal only during bullish market)

IN-SAMPLE	1	2	4	6	7	8	13	14	15	9
Panel regression	imv	ep	mom1	mom12	roa	roe	eq_tur	rev_ratio	FY1_3mch	leverage
% Positive beta	67,86%	57,14%	50,00%	71,43%	53,57%	60,71%	67,86%	64,29%	60,71%	60,71%
% Negative beta	28,57%	39,29%	46,43%	25,00%	42,86%	35,71%	28,57%	32,14%	35,71%	35,71%
% Significant beta	46,43%	42,86%	60,71%	46,43%	46,43%	46,43%	21,43%	21,43%	35,71%	25,00%
% Pos & Sig beta	28,57%	35,71%	32,14%	39,29%	28,57%	32,14%	3,57%	14,29%	21,43%	21,43%
% Neg & Sig beta	14,29%	3,57%	25,00%	3,57%	14,29%	10,71%	14,29%	3,57%	10,71%	0,00%
mean R2	40,97%	40,98%	41,24%	41,25%	41,15%	41,03%	40,82%	40,81%	40,89%	40,88%
P-Value	0,0407	0,0193	0,6748	0,0030	0,6280	0,0914	0,6855	0,0743	0,1551	0,0282
Mean Coeff. x 100	0,0954	0,1125	0,0281	0,1814	0,0297	0,0912	0,0138	0,0605	0,0629	0,0893
Median	0,1144	0,1389	0,0144	0,1384	0,0249	0,0929	0,0396	0,0539	0,0325	0,0489
1st Quartile	-0,0617	-0,0781	-0,2248	-0,0485	-0,1555	-0,0852	-0,0288	-0,0331	-0,1068	-0,0450
3rd Quartile	0,2723	0,2729	0,2741	0,3664	0,2717	0,3512	0,1363	0,1618	0,2091	0,2127

Case 0 /1 (keep the signal only during bearish market)

IN-SAMPLE	1	2	4	6	7	8	13	14	15	9
Panel regression	imv	ep	mom1	mom12	roa	roe	eq_tur	rev_ratio	FY1_3mch	leverage
% Positive beta	21,43%	50,00%	57,14%	78,57%	71,43%	75,00%	50,00%	71,43%	71,43%	53,57%
% Negative beta	75,00%	46,43%	39,29%	17,86%	25,00%	21,43%	46,43%	25,00%	25,00%	42,86%
% Significant beta	71,43%	60,71%	75,00%	39,29%	60,71%	60,71%	39,29%	46,43%	75,00%	78,57%
% Pos & Sig beta	10,71%	21,43%	46,43%	32,14%	42,86%	42,86%	21,43%	39,29%	53,57%	42,86%
% Neg & Sig beta	57,14%	35,71%	25,00%	3,57%	14,29%	14,29%	14,29%	3,57%	17,86%	32,14%
mean R2	43,44%	43,40%	44,16%	43,32%	43,31%	43,32%	43,12%	43,15%	43,83%	43,80%
P-Value	0,0039	0,3281	0,1443	0,0058	0,0775	0,0056	0,4999	0,0002	0,0106	0,3065
Mean Coeff. x 100	-0,1865	-0,0676	0,1459	0,1678	0,1047	0,1736	0,0332	0,1920	0,2400	0,0903
Median	-0,2683	0,0068	0,1115	0,1416	0,2238	0,1604	0,0204	0,1705	0,2854	0,1015
1st Quartile	-0,4198	-0,3254	-0,2377	0,0423	-0,0529	0,0001	-0,1818	0,0000	-0,0675	-0,4163
3rd Quartile	-0,0137	0,2272	0,5709	0,3146	0,2740	0,4615	0,2394	0,4011	0,5710	0,5546

If we compare the results with the original simple, the improvement is marginal and not enough to be confident to switch to the dynamic model.

The only finding that seems to be appreciable are the results of the beta of 'imv' in both models. In the first model the beta of 'imv' has a very low p-value, it is quite significant and it is positive almost 70% of the times. In the second model 'imv' has again a very low p-value, it is very significant and is negative more than 70% of the times. These results just confirm what we learn from the Fama-French model, i.e. that the small cap effect is indeed a risk factor.

We think that the disappointing results are due to the rough estimation of indicating the state of the economy. Official government bodies say that an economy is in recession if the economic output shrank at least two quarters in a row. But using a two quarter estimate would be a too much lag for our model (while we need a faster indicator of the macroeconomic trend since we are building a tactical portfolio). Another issue is that recession is indicated by economic output and *not* by stock prices. Stock prices are usually forecasts, so they are out of sync with the state of the economy. Actually, we also got the intuition that market health is better measured looking at volatility, but building up a dynamic model on volatility would have become a bit too complex.
Last, there's always a data-mining risk.
Due to these issues we return to our simple model.

Putting the signals together

Now we have an indication of which signals are helpful. In order to make a useful model we have to combine the signals. From the economic theory we know that we can incur in a Multi-Colinearity problem among our variables. For example, we can imagine that ROA and ROE are highly correlated. Furthermore, according with our hypothesis we can see that the plot of the ROA end ROE betas show pretty much the same pattern. For avoiding the bias that multi-colinearity would cause, we run an optimization on a top-bottom portfolio built (through rankings) using the "first-round" signals and then we look at the behaviours of the weights. According to our results we can easily see how there are several signals that are correlated with each other.

Therefore, according to the results we decided to remove from our signals 'mom1', 'roa' and 'rev_ratio', which are highly correlated respectively with 'mom12','roe' and 'eq_tur'. Once chosen the signals we ran a multivariate regression in order to verify the goodness of our information variables altogether. We tried different compositions of signals, trying to pick the one with the best statistics, but they all looked pretty similar. At the end picked as signals **imw, ep, mom12, roe and FY1_3mch**, but this choice wasn't based on strong numbers. We have to keep in mind that this choice was not strictly meaning that statistical factors are not the only criteria used, even because data mining and sampling errors are still present.

Panel regression	'imv'	'ep'	'mom12'	'roe'	'FY1_3mch'
% Positive beta	64,29%	67,86%	64,29%	64,29%	78,57%
% Negative beta	35,71%	32,14%	35,71%	35,71%	21,43%
% Significant beta	39,29%	28,57%	50,00%	39,29%	50,00%
% Pos & Sig beta	25,00%	21,43%	39,29%	25,00%	39,29%
% Neg & Sig beta	14,29%	7,14%	10,71%	14,29%	10,71%
mean R2	40,50%	40,50%	40,50%	40,50%	40,50%
P-Value	0,1120	0,0453	0,0159	0,2936	0,0034
Mean Coeff. x 100	0,1819	0,1920	0,3433	0,1491	0,3130
Median	0,2489	0,1332	0,3192	0,1637	0,3156
1st Quartile	-0,2523	-0,0701	-0,1847	-0,1562	0,0424
3rd Quartile	0,5140	0,4961	0,7849	0,5482	0,6583

The portfolio

Using the signals chosen above we use the multivariate model to forecast the returns of the individual industries.

$$\mathbf{r}_{i,t} = \beta_{i,t} X_{i,t}$$

Now we have a dataset of forecasted returns solely constructed from previous available data at that moment in time. From here we can construct a portfolio that should mimic the risk of the benchmark and optimize the returns, so that it can outperform the benchmark. We should also mention that we should optimize for transaction costs, because a rebalance can have potentially high costs, especially when a large portion of the portfolio needs to be switched. The transaction costs are 50 bps per action, in which selling and consequent buying are two actions.

So we pick the portfolio from all the industries. The portfolio has the following constraints:

- Risk_{pf} equal to risk_{bm} plus a tracking error measured as the volatility
- Highest expected return, taking into account transaction costs
- No short selling

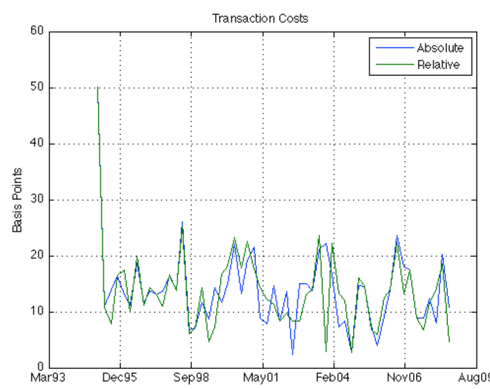
Under these constraints we build two models, one with a tracking error of 4% and one with a tracking error of 1%. The portfolio will be built in a 130/30 fashion in which the benchmark is bought and from the benchmark is only deviated in a way that the forecasted worst industries will be taken on less in the portfolio, and with the free money, the forecasted better industries are acquired. This way we end up with more or less the benchmark, but with the important difference that we replace bad industries with good ones. The amount of the difference from the benchmark is not limited, but in our model it is quite constant around 20%.

Taking into account the transaction costs is also built into the model. It is a tradeoff between transaction costs and switching to the optimal portfolio. When there is a new portfolio that has higher returns, part or even all of the extra return can be eaten by having to switch to that portfolio and having to pay the transaction costs for that.

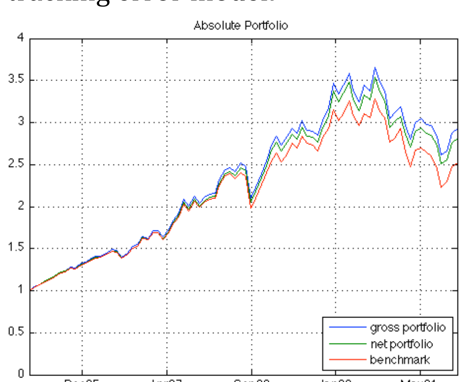
Under these constraints we build a mean variance portfolio that should get the highest possible returns with similar risk as the benchmark.

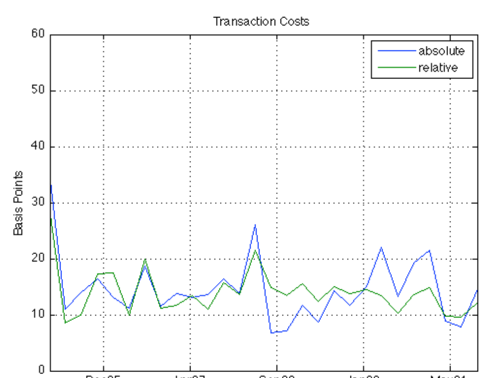
Results in-sample

4% tracking error model:



1% tracking error model:

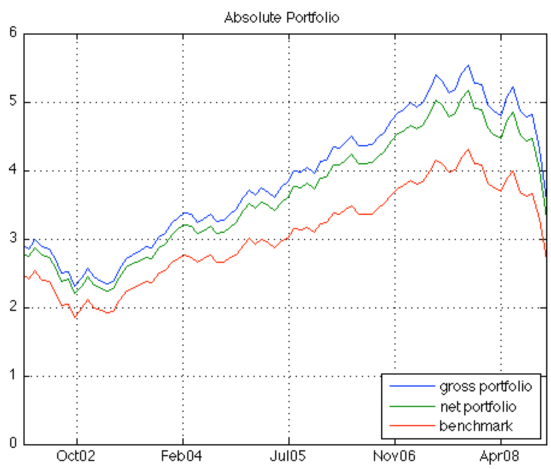




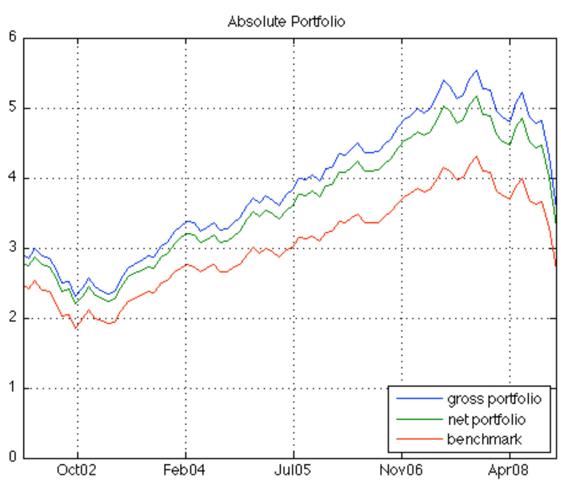
Both the model for the tracking error of 4% and 1% perform better in the screening period. The net return of the portfolio is better than the benchmark. And the difference seems to grow over time. While the transaction costs are quite significant, because the transaction costs are about 15 bps every quarter, which translates in a 60 bps on a yearly basis, our portfolio still manages to overcome the costs.

Backtesting
The previous results were the consequence of an optimizing process for that specific period. Now in order to check if the model works also under different circumstances, for which no optimization took place, we look if the model still fives higher than average returns in an out-of-sample period. This is the true test that can tell us if the model is useful.

Results out-of-sample
4% tracking error model:



1% tracking error model:



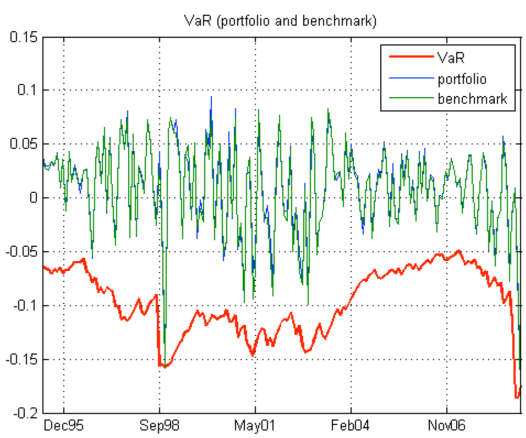
The above net results for the out-of-sample period are also better than the benchmark, again both for the model with a tracking error of 4% and 1%. The difference between the portfolio and the benchmark also seems to increase over time, except for the financial crisis. This might mean different things. It can be that our portfolio is actually a beta strategy and will only amplify the market result, but it can also mean that the signals during the crisis were meaningless and that no sound portfolio could be built.

Risks

Although the use of a 130/30 model and a tracking error were constructed to limit the risk, there are several risks for the investors in the model.

First of all if we look at the results, we see that we possibly have a beta strategy. That means that higher than average returns of the portfolio are mostly the consequence of taking more risk. And that gives nice results in good times, but it will be negative in bad times.

Second point is the nature of a 130/30 model. If only a small part is different from the benchmark, say 10%, then that means that with a tracking error of 4% all the higher risk can concentrate in the few extra stock, which means that those stocks can be enormously risky. Although that is compensated, because we invest in industries and not stocks, which are already diversified.



When we compute a VaR for the portfolio it doesn't differ from the benchmark, because when we compare the returns of the portfolio with that from our portfolio, we see that the risk profile matches that of the benchmark.

Then there are the general risks for any QEPM portfolio. The first one is that only quantitative data is used. This can be good, because there is no biased manager that can make wrong decisions. The decisions made are on which signals to use and on the investment horizon. But there might be also qualitative data that is very useful. And it is very hard and arbitrary to translate qualitative data

into quantitative numbers. For instance it is hard to say something about the management of a company in a quantitative way.
Also it only looks at historical data, so it can be that the signals were useful in the past, but that doesn't say if they also will be in the future.
Also because you need historical data, and enough of it, it takes a while before new trends pop up and find their way into new useful signals. This hampers the reactivity of the model.
And last is that because of all the available data and the data is so anonymous, it will be tempting to choose the optimal signals for a certain period. They might be the signals that are correlated with the returns during that period, but that doesn't necessarily mean that they are economically sound signals and were the drivers behind the better results. So then the model ends up with meaningless signals. This backward thinking, of first looking at statistics is called data mining.