2nd AFIR Colloquium 1991, 4: 277-292

Translating Traditional Asset Allocation into a Quantitative Model
The Risk is Getting it Wrong

Richard Noble

INVESCO MIM Management Limited, 11 Devonshire Square, London EC2M 4YR, United Kingdom

Summary

Traditional asset allocation is centred upon making the right bets in a portfolio. To help make these decisions fund managers are increasingly developing quantitative models to improve their consistency and boost marketability. Unfortunately quantitative tools in the wrong hands can lead to models based on dubious relationships. This gives quant a bad name and means that good work is often ignored. Here we set out some of those concerns and present guidelines on how they can be avoided.

Résumé

Traduire la Répartition des Actifs Traditionelle en Modèle Quantitatif
Le Risque C'est de se Tromper

La répartition traditionnelle des actifs consiste à faire les bons paris dans un portefeuille. Afin de faciliter ces prises de décisions, les gestionnaires de fonds mettent de plus en plus souvent au point des modèles quantitatifs pour améliorer leur régularité et augmenter leur négociabilité. Malheureusement, entre de mauvaises mains, les outils quantitatifs peuvent conduire à des modèles basés sur des relations douteuses. Cela fait une mauvaise presse à cette méthode et signifie que souvent un bon travail passe inaperçu. Nous présentons ici quelques-unes des ces inquiétudes ainsi que des conseils pour éviter ces problèmes.

THE ISSUES

What is it about quantitative techniques that invariably provokes such heated debate among asset allocators? A "quant" in a traditional fund management house seems to get the same regard as a wing forward on a rugby team. It's his job to be a backup to every player on the pitch and when play goes wrong its his fault for not being there. However when the results are good he is taken for granted and he never gets the praise he deserves. The whole "black box" mystique that surrounds quantitative models often obscures the fact that quantitative models are merely translations of a fund manager's assumptions about market movements and price formations into a systematic or computerised application. When models fail to deliver the expected results it's quantitative techniques in general that are questioned and not the model–builder's assumptions or methodology.

This talk addresses the question of why it is that so many asset allocation models <u>do</u> go wrong; it describes a range of common mistakes made by managers building quantitative tools; and it introduces some guidelines for creating more realistic tools. At the risk of being overly simplistic, we would like to suggest three broad areas that create problems for model builders or users:

1. Not taking the statistics far enough (Half–hearted Quant).
2. Not understanding the range of judgemental inputs that quantitative models demand.
3. Believing that models will add more value than they really can.

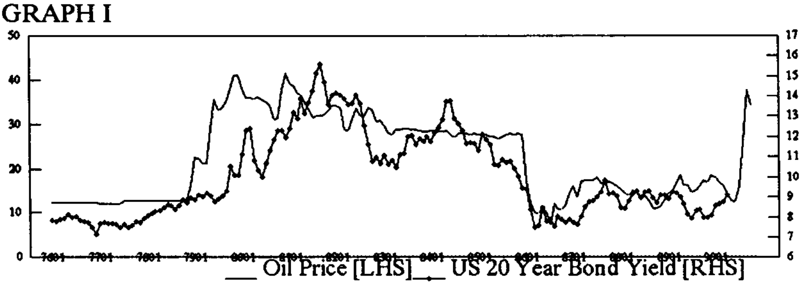
Let's take a look at how easy it is to be foiled by these common errors and how widespread their presence is among traditional fund managers in the early throes of developing quantitative strategies.

WHAT HAPPENS WHEN WE DON'T TAKE THE STATISTICAL ANALYSIS FAR ENOUGH?

Even the most traditional asset allocators begin with certain assumptions about market movements that form the foundation of their investment philosophy. Many of these assumptions have been fortified by graphs or simple statistical analyses that attest to key relationships between external factors and the markets. But have the analysts really been rigorous in their analysis of these relationships to justify the implied strength of these relationships?

Let's examine some relatively standard practices used by fund managers today. Nothing could be more familiar than the friendly Datastream terminal for analysing macroeconomic phenomena and the markets. It is one of the few international databases that can provide good quality macroeconomic data and graphs at the depth required. The graph below (GRAPH I) is a standard Datastream output comparing oil prices to US bond yields.

GRAPH I



The most immediate conclusion one would be tempted to draw would be that bond yields appear to be linked to oil prices and therefore we are likely to see a rise in yields to about 14% in the near future.

A more adventurous fund manager will have remembered his maths from school days and be tempted to produce a more precise statistical assessment of this relationship. Again, with similar ease, Datastream can perform a regression on the data and present figures similar to those listed under the Datastream column in TABLE I.

TABLE I

	DataStream Regression	Lotus-123 Regression	Stats Regression	Adjusted Regression
Constant	NA	5.97	5.97	10.23
X Coefficient	18.4%	18.1%	18.1%	0.2%
T-Statistic	***18.4***	NA	18.0	***0.1***
R^2	***66.1%***	65.6%	65.6%	***0.0%***
DW Statistic	NA	NA	0.21	1.67
Oil Price	34.10	34.10	34.10	34.10
Predicted Yield	NA	12.14	12.14	10.28

Period = 01/76 to 04/90
Independent Variable = Oil price
Dependent Variable = US 20 Year Bond Yield

Here we are in danger of making a classic error. Not only do we have a graph that suggests a strong relationship but we have a very high T-Statistic from the regression analysis apparently confirming it. Just like the graphical approach above the fund manager is likely to stop here, when he thinks he has an answer. He stops because he wants to keep the model simple and clear, but in fact he is stopping right in the middle of his work.

The disadvantage of using the Datastream framework is that the statistical information is incomplete and we have no way of determining if there are any flaws in our analysis. Further, we may fail to appreciate that the type of data we may be using may also distort the analysis. The US long bond yields available on Datastream are in fact monthly averages and as such are not appropriate for correlations against month end oil price data.
Downloading more suitable data into a Lotus spreadsheet would be the most obvious option, and another path commonly pursued. But once again, while one has a better means to actually manipulate the underlying data, there are still no new statistical diagnostics (See TABLE I, Lotus 123).

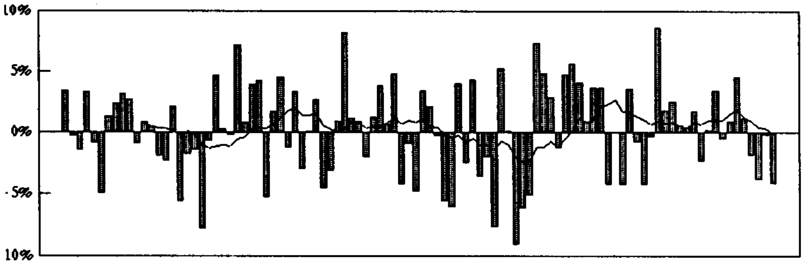
The most appropriate approach is to explore the relationship more carefully with a reasonably sophisticated stats package. By using the SPSS stats package we are now provided with one further key piece of information, the Durbin–Watson statistic. A regression employing perfectly stationary data should produce a DW of 2. In our example here, the DW is a very low .21, implying that the forecast errors are autocorrelated through time and that we have violated an important underlying assumption required in a regression analysis. Any suggestion of a strong correlation is therefore suspect and any errors in the forecast run the risk of being wrong for long periods at a time.

How is it that an error this simple yet so significant can be frequently found in investment management? Perhaps the major culprit is a fundamental misunderstanding of data analysis. A fund manager is really seeking to confirm his views with analysis and because he does not have a thorough statistical background will believe/disbelieve results at the first signs of success/failure in his analysis. Unfortunately initial results can be very misleading as the real relationships are often counter–intuitive.

As a further example, probably no assumption is more misapplied by the investment community than the notion of market trends. A traditional manager often sees his task as seeking out trends and taking advantage of their persistence. Graphs like the one below (GRAPH II) provide him with the reassurance that trends do exist and can be identified. What has been charted here is the monthly change in the S&P Composite over 310 months. The smoothed black line is the 12 month moving average and clearly suggests the bull and bear market phases that have occurred in the market over this period. Can we not conclude that we are about to enter a bear market phase?

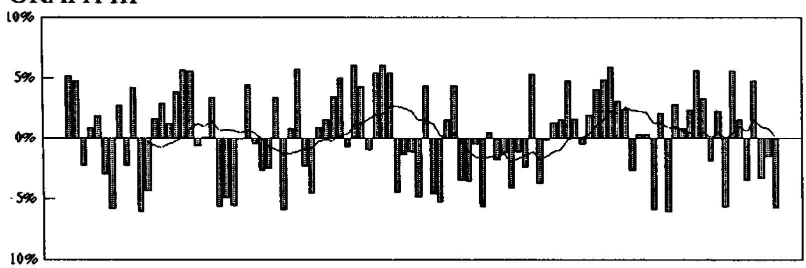
What the fund manager fails to appreciate is that moving averages by definition will trend (in Quant jargon this is referred to as autocorrelation), and therefore cannot be used to determine whether the underlying data actually has a trend.

GRAPH II



To press this point further, note the "trends" in the time series below (GRAPH III). While we may appear to have the same type of information about "bull" or "bear" phases, let us hasten to mention that the data series is in fact derived from totally random numbers. How confident do we feel about that "trend" now?

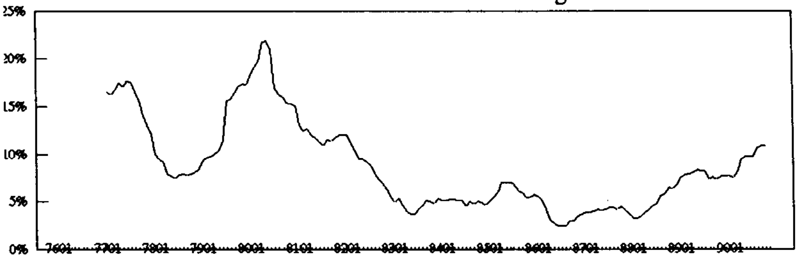
GRAPH III



Let us look at one final example of how one can put misplaced faith in trends. GRAPH IV shows year–on–year change in inflation in the UK. A common notion is that one can derive information about expected moves in inflation from examining the trend in this year–on–year change. But if we can remember from our lesson on moving averages (and year–on–year change is not unlike a twelve month moving average), we know that 11/12ths of next month's number comes from historical data and only 1/12th comes from new information. In other words, the level of year–to–year change today tells us next to nothing about the level of year–to–year change in the future.

GRAPH IV

UK RPI Year–on–Year Change



Let's return to our oil price and bond yield problem with the low Durbin Watson statistic and all its implications. How then can we effectively deal with autocorrelated data or residuals in the regression framework? The simplest solution is to transform the data so that it is no longer non–stationary. But let's say that we would like to preserve the longer term framework. One solution is to correct for it by employing the Cochrane–Orcutt procedure (also found in SPSS) that adjusts regressions with autocorrelated residuals. (TABLE I Adjusted Regression). Note that by using this method, we now have results that show no relationship and are in direct contrast with those reported earlier. The R^2 is 0.0% and T–Stat is 0.1. So, on the one hand we have a graph in our Datastream regression showing a relationship with a strong appeal backed by a bogus statistical analysis, and on the other, an analytical framework, the adjusted regression, that demonstrates that the relationship doesn't hold. Clearly this analysis needs to be taken further.

A sensible next step is to split the data into two parts. The first is a seven year window pre the 1985 oil price fall and the second pre the 1990 oil price rise. TABLE II shows the results.

These results suggest that the relationship between the oil price and bond yields has changed since 1985 from a weak negative correlation (Oil price up, yields down) to a stronger positive correlation (what we would expect). This helps to explain why the initial test was so indifferent. We now see there was a change in the sign of the relationship during this period. The more reasonable explanation about the relationship between the oil price and bond yields has materialised after adding the extra dimension of time.

TABLE II

	Nominal Oil Price	Nominal Oil Price
Period	01/78–12/84	05/83–04/90
Constant	11.40	8.91
X Coefficient	–2.0%	4.7%
T–Statistic	–0.69	1.88
R^2	0.6%	4.2%
DW Statistic	1.58	1.93
Rho	0.96	0.96
Indep Var	Oil Price	
Dep Var	US 20 Year Bond Yields	

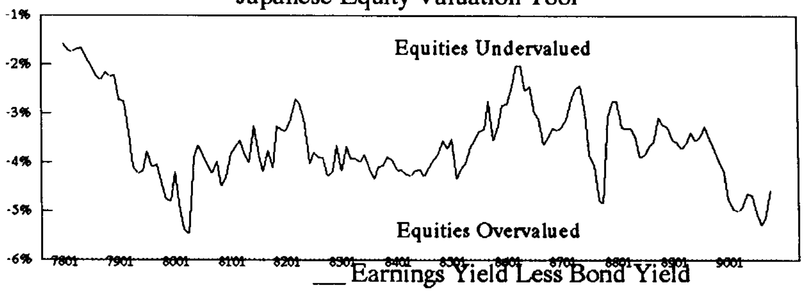
If time is apparently such an important consideration in assessing relationships correctly, then clearly the ideal regression would be to use a moving window of data, enabling us to understand the exact nature or the relationship at every point in time in the analysis.

GRAPH V shows a traditional measure of equity market valuation, the yield gap. In this case the yield spread between earnings and bonds was chosen. When this spread is very negative earnings yields are typically low relative to bonds, meaning equities are typically expensive and vice–versa. On this basis we should today be slightly more cautious of Japanese equities than we were in October 1987. But just as in the oil example above this is an IN–SAMPLE view, in other words we are using events of the recent past to make judgments about the relationships over an earlier period. A more appropriate way to determine how important this factor has been through time is by calculating OUT–OF–SAMPLE regression coeffients using OLS regression with a moving seven year window (84 monthly observations). The dependent variable is monthly returns on the

Japanese equity market less the return from cash in that month and the independent variable is the yield gap as defined above (See GRAPH VI).

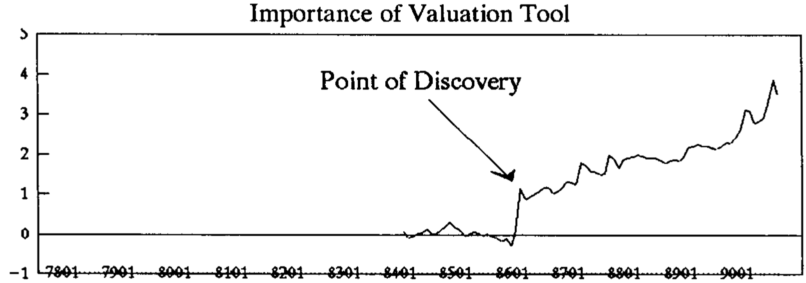
GRAPH V

Japanese Equity Valuation Tool



GRAPH VI

Importance of Valuation Tool



An OUT–OF–SAMPLE regression coefficient means that the model was estimated in one period and used in another. We can therefore ask the question "If today was 31st September, 1987 would the yield gap have been identified as an indicator using only information available at that time?". This is the true test for a model as it assumes no knowledge about future events and is an important point often overlooked in building a model. The problem is that if the model was estimated and tested over the same data then the results are bound to look good.

The results above show that the yield gap is a more important indicator today than at any other time. We can also see that it was not until the spring of 1986 that this factor would have been a useful measure to follow and indeed would not have been identified until then. In addition we would be more confident in taking an underweight position in Japanese equities now than at any time in the past.

ARE WE PREPARED FOR THE REQUIRED JUDGMENTS?

In the previous section we looked at some examples of relationships that are used in day to day fund management. In both cases the objective was to gain a better understanding about the real relationship and to try and understand how some common assumptions might be unsupportable. In this section we examine an even more unsettling problem. The fact is that there are any number of judgmental decisions that go into building quantitative models. Only by thoroughly understanding the implications of these judgments can the user become comfortable with the methodology selected.

In all the examples above a seven year window (84 monthly observations) was used to estimate the model coefficients. This period was chosen because it was believed to be long enough to minimize estimation error but short enough to be sensitive to market changes. In addition it roughly conforms with the length of the most recent business cycle and therefore should always reference all phases of market (and economonic) conditions. But would everyone be similarly comfortable with this view? We accepted the timeframe after making further tests using 3 year and 10 year rolling windows to assess the robustness of the relationship to the window size. The difference was alarming, in the case of the longer window the yield gap was not identified as an important factor until 1988, and therefore did not pick up the October 1987 market fall, while the shorter window proved to give volatile coefficients. Our solution struck a balance.

But another manager might make an equally convincing case that the best approach would be to use a growing window of data. In this case, while the model may not be sensitive to permanent structural changes, the model–builder might be more comfortable that all structural changes at least have been included.

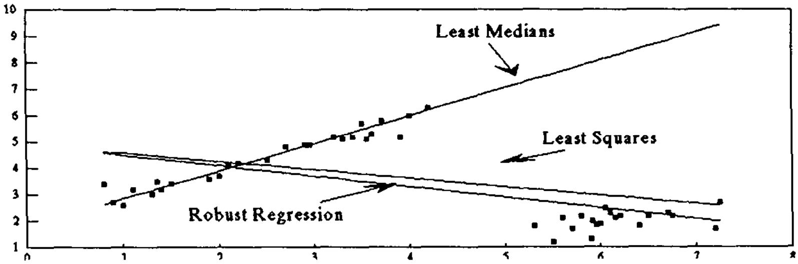
Probably no judgmental decision has more bearing on the stability of a model than that of how to treat outliers. Definitions of an outlier differ but typically it is an observation that falls outside the expected range, eg., if we are looking at capital returns on UK equities the value for October 1987 was −26.6% and this was seven standard deviations away from the average returns for the previous five years (as measured by the FT All Share Index). Now if we had a model that perfectly forecast this event is it still an outlier? On the one hand it was an extreme and unusual event and unlikely to happen again (judgment called for) while on the other it was perfectly forecast and therefore should not be considered unusual. An outlier is therefore a judgmental decision.

Outliers can arise because market returns are typically not normally distributed as there are often too many extreme observations. It is these data points which when combined with OLS regression lead to incorrect and non−robust models. This instability is detected in October 1987 when most models show a sudden large change in coefficients when this observation is included in the analysis. Clearly this is undesirable as it increases the chance of model misestimation. A way out of this is to use some form of robust regression that automatically assigns a lower weight to such observations.

Judgment however is also in the choice of method. As an extreme example GRAPH VII shows three different regression techniques which all lead to different models.

It is clear from the graph that sign of the relationship is determined by the choice of regression tool and not by the data! Each technique carries its own set of assumptions. The Least Median Line is highly robust to extreme points and clusters of points because it is only concerned with the median error while the Robust Regression Line iteratively re−weights observations so that each point has a similar influence in determining the line. We must ask which of these methods is desirable and is the most representative of the real relationship.

GRAPH VII



So far we have talked about relationships and how they can change over time, what has not been mentioned is the issue of what factors we should include. Invariably a tug-of-war takes place between including factors that are consistent with one's underlying investment strategy and factors that simply show good predictive power from an empirical perspective.
Often there will be a range of acceptable factors with only a very subtle difference between them. Because the multivariate regression framework cannot deal easily with collinearity (when two or more variables contain similar information), a judgment must then be made. Let's use the Japanese equity value indicator as an example. In TABLE III we have four variables which appear to be effective measures of value for Japanese equities.

TABLE III

		R^2	X Coef T-Statistic	Prediction for Oct 1990
1	Earnings to Bond yield	15.9%	3.9	-3.6%
2	Earnings to Cash rate	14.0%	3.7	-4.9%
3	Dividend to Bond yield	4.7%	2.0	-1.2%
4	Dividend to Cash rate	6.1%	2.3	-2.8%

Each of these factors passes the test for inclusion in the model (the sign is correct and T-Stat is high). Clearly the earnings yield model is a more

powerful series than dividend yield, but what is less clear is whether it is more appropriate to use cash rates or bond yields for the comparision. The worry is that this choice has a crucial effect on the forecast, and yet the statistics give us no clue as to which would be the more suitable choice.
Once again it points out how enormously important it is that the fund manager fully appreciates the implications of choosing one variable over another from the perspective of investment philosophy and objectives.

HOW MUCH VALUE DO QUANT MODELS REALLY ADD?

A quantitative approach to fund management is appealing because it is easy to demonstrate if the method has worked in the past and identify why it has worked. A fund manager will often be presented with numbers from "Computer simulated" backtests telling stories of untold riches to be made by using such and such a method. It is stories such as these that are very misleading because in reality we should not expect quantitative models to significantly outperform a good fund manager. After all, a computer only responds to what it is told. The real merit in this type of work is the ability to test numerous and complicated relationships while remaining objective and aloof of market psychology.

Building a model that has its foundations in questionable assumptions has already been shown to be problematic. Even more unsettling though is how easy it is to create backtests for any one of these models that provide the desired results. Misconstruction of the backtest can often be the culprit for misplaced expectations in quantitative models. The issue here is of the testing framework. Much quantitative research is tested on the basis of its ability to outperform a benchmark and therefore is tautologous because the testing criteria is the same as the reporting criteria. Compare this with the framework that has been suggested in this paper and you see that even if a factor shows up as important it does not imply that it always adds value.

Table IV shows an example of a backtest using the Japanese yield gap introduced above (without using the results of the regression analysis). The results clearly state the model adds value over time. In this test when the spread is −3.6% the suggested asset allocation to Japanese equities is 50% (the benchmark), when higher equities are overweight and vice−versa. But who chose this magic −3.6%? The important point is that it is arbitrary and has been put at a level that seems reasonable with hindsight (IN−SAMPLE). It is likely that a different level would have been chosen if we had done this analysis in say 1985. The backtest results are sensitive to this choice and therefore it is important to ensure these judgments are made independently from the backtesting framework.

TABLE IV

01/78 – 09/90	Total Return	per annum Return
Benchmark (50% Bonds, 50% Equity)	282%	11.1%
Model Portfolio	453%	14.3%
Added Value (%age over benchmark)	45%	2.9%
Information Coefficient (IC)	47.3%	
T−Statistic of IC	3.94	
Trading rule : Equity weight is proportional to spread. Neutral at −3.6%, 100% Equity at −1.1%, 100% Bonds at −6.1%		

To deal with this problem we must make use of the regression framework used to derive the OUT−OF−SAMPLE Importance indicator. The regression was set up as follows:

$$R_{t+1} = a + b \times F_t + e_t$$

R_t = Return on market at time t
a = Constant
b = X−Coefficient
F_t = Level of Factor at time t
e_t = Error in Forecast at time t

Thus for each month we estimate the model using data historical to that month and use the level of the valuation factor to forecast market moves one month ahead. We can then regress our OUT–OF–SAMPLE forecast returns with actual returns. The resulting correlation coefficent is called an information coefficient and is therefore devoid of any unfair backtesting bias (See TABLE IV).

TABLE IV also demonstrates another feature of numbers and that is the phenomon that percentage changes do not add up. The added value suggested here of 45% will often be misrepresented and quoted as 171% (453% – 282%)! The difference is one of basis of quotation. The higher figure tells us how much more money we have today compared with our 1978 starting value while the lower 45% is compared with what we would have had had we not taken any bets away from the benchmark. Clearly the higher number is not what we want to hear.

SUMMARY

There are many potholes on the road to Quant, many of which are not fully understood by both the people building models and those commissioning them. To have any chance of success there are a number of guidelines that must be followed.

All the inputs must make sense, they must relate to what a fund manager looks at day to day and the output must be clear and easy to understand. In addition anyone using the model must be able to make a clear link between what a factor is saying and what the model is saying. For example the Japanese stock valuation tool is suggesting equities will fall because the yield spread is too low.

Changes in structure and unforseeable shocks are frequent to international capital markets. The model must therefore be dynamic so that it can evolve with these changes. There will also be times when a manager will have a very strong unquantifiable view on a market (perhaps because there is insufficient historical data available) so a model must have a suitably open architecture be able to incorporate these views.

Once these objectives have been determined an environment where everyone is openminded enough to entertain new ways of looking at old problems must evolve. It is no use a manager insisting that such and such a factor is important when the quant has exhausted every avenue of possibility, but rather it is much more constructive to reassess what it is that he is trying to achieve because often there will be other ways of looking at the same thing.

The techniques used in this paper go a long way to meeting these requirements and avoiding the problems that make other models fail. But the ultimate success of this work really lies with the people responsible for both model creation and implementation. What should not be underestimated is the range of skills that are required for the effort. The ideal candidates should be conversant in the investment objectives of the fund manager (to understand his needs); the economist (to understand what might drive markets); maths (to understand the games numbers can play); stats (to be fluent in the techniques and their limitations); computers (because this is where the work is done) and teaching (to tell everyone in layman terms what has been done). Too often quantitative efforts simply cobble together people with good computing and statistical skills and assign them the task of building models to the investment strategist's specifications. The point that is generally missed is that unless equal importance is given to all of the required skills, it is unlikely that quantitative tools can be built that are methodologically correct, appropriate, and intellectually honest.