

Envision by Quant Insight: financial data visualisation, pattern identification and fair valuation

Professor Michael Hobson*

Astrophysics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE

Trinity Hall, Cambridge, CB2 1TJ and

Quant Insight Ltd, 22 Wycombe End, Beaconsfield, Buckinghamshire, HP9 1NB

Krishnan Sadasivam,[†] Mahmood Noorani,[‡] and Nikhil Agarwal[§]

Quant Insight Ltd, 22 Wycombe End, Beaconsfield, Buckinghamshire, HP9 1NB

I. INTRODUCTION

In recent years, the quantity and quality of the financial data on which asset managers may base investment decisions has grown exponentially. Unfortunately, rather than assisting in such decision making, this explosion in data volumes has complicated the process considerably, since it is almost impossible to extract the most important financial information pertinent to the investment decisions at hand. In particular, on a daily basis asset managers need to identify the macro drivers of financial markets and securities. These drivers encompass a range of factors, such as economic growth, monetary policy, impact of quantitative easing, risk aversion, credit spreads, commodity prices, and many more. Moreover, each of these macro drivers may be characterised by several hundred or more individual macro factors, which may have varying degrees of correlation between them. In addition, there may exist considerable overlap between the factors characterising different drivers. This wealth of data contains too much information for asset managers to digest each day without proper tools and techniques.

Quant Insight (QI) have therefore developed a powerful new tool, called Envision, to assist asset managers to develop such an understanding in a quick, straightforward and automated manner. In particular, Envision enables asset managers to visualise financial data, identify patterns and determine fair valuations. The resulting understanding of which factors are driving any security helps to avoid trade selection errors and maximises the value of the managers view by identifying the appropriate trades. Moreover, such an understanding also reveals the residual, unintended or implicit macro exposures within a portfolio, and identifies how best to mitigate them.

The purpose of this paper is to present a brief account of the methodology employed by Envision, and show some typical results obtained in applying our approach to real financial data.

II. PRINCIPAL COMPONENT ANALYSIS

At its heart, Envision employs a novel version of a mathematical technique called principal component analysis (PCA) to accommodate the large number of macro factors that are potentially relevant in driving a given security, many of which may share a high degree of collinearity. PCA performs an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The PCA transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. In general, there are as many principal components as there are original variables, and together they retain all the information present in the data. Typically, however, one need keep only a handful of the first few principal components, since they are able to account for the vast majority of the variability in the data. Thus, as well as producing uncorrelated linear combinations of the original variables, PCA provides a natural means of vastly compressing the data with almost no loss of information. This idea is illustrated in three dimensions in Figure 1.

In its standard form, the PCA technique is widely used in the scientific community, most notably in image processing and bioinformatics, and has also already been applied in different parts of the financial industry, such as the interest rate market and risk management. It has not, however, been applied to the fair valuation process.

Moreover, Envision employs a novel version of PCA, which is ideally suited to this application for easier interpretation of the large number of potentially relevant macro factors driving a given security.

III. PRINCIPAL COMPONENT REGRESSION

Once the final set of principal components has been determined, Envision then performs a linear regression of the observed behaviour of the security under consideration against the first few components. This yields

* mph@mrao.cam.ac.uk

† ks@quant-insight.com

‡ mn@quant-insight.com

§ na@quant-insight.com

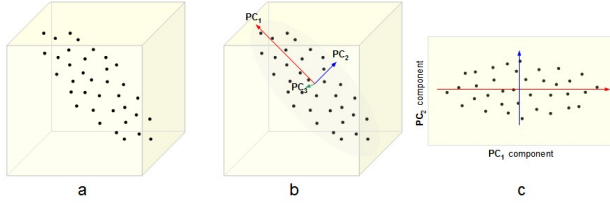


FIG. 1: Illustration of PCA in three dimensions: (a) the original data; (b) the principal components; (c) compression into the plane of the first two components.

estimates of the unknown coefficients in the model defined by the linear combination of this subset of principal components. This principal component regression (PCR) technique overcomes the multicollinearity problem that would arise in a standard linear regression analysis, where one attempts to regress a security directly against a large number of macro factors that might potentially drive its value. In the standard approach the large number of variables vastly increases the chance of overfitting the model. Since PCR uses only a subset of all the principal components for regression, it achieves dimensionality reduction through substantially lowering the effective number of parameters characterizing the underlying model, and may thus be considered as a form of regularization procedure. This concentration of most of the signal into a few principal components increases the signal-to-noise ratio in the fit and stabilises the solution. Moreover, the use of our novel PCA approach to select the principal components to be used for the regression leads to a very efficient, accurate and easily interpreted representation of the behaviour of the security under consideration.

IV. GOODNESS OF FIT

To provide a measure of how well the observed outcomes are recovered by the model, Envision calculates the R^2 -statistic, which is based on the proportion of the total variation of outcomes explained by the model. The value of R^2 provides a measure of the global fit of the model. Specifically, R^2 lies in the range $[0, 1]$ and represents the proportion of variability in the outcome that may be attributed to some linear combination of the explanatory variables. In such models, R^2 is often interpreted as the proportion of response variation "explained" by the regressors in the model. Thus, $R^2 = 1$ indicates that the fitted model explains all variability in the outcome, while $R^2 = 0$ indicates no 'linear' relationship between the response variable and regressors. Thus, an interior value such as $R^2 = 0.7$ may be interpreted as follows: seventy percent of the variance in the response variable can be explained by the explanatory variables; the remaining thirty percent can be attributed to unknown, lurking variables or inherent variability.

V. ENVISION METHODOLOGY

We now give an overview of the complete methodology used for the PCA model fair value analysis performed by Envision, which is intended to provide an introduction to its work-flow and we present some illustrative results in Section VII.

A. Identifying the driving factors

The first step is to identify the appropriate macro-economic factors that might drive the value of the security of interest. This may seem a relatively simple task, but the list of factors needs to be both comprehensive and relevant to the security under consideration. Thus each security has its unique set of macro factors, which are then categorized into factor 'baskets' (or macro drivers), as listed in Table I for ease of viewing and analysis.

TABLE I: Typical factor 'baskets' used in Envision.

Inflation expectations	Commodities Agriculture
Global Growth	Country Growth
Risk Aversion	Global QE
Country Real Rates	Metals
Global Real Rates	Energy
Country Sovereign Credit	Corporate credit
FX	Global sovereign risk

Examples of underlying macro factors contained in some of the 'baskets' listed in Table I are given in Table II, although the precise factors present will depend on the security under consideration.

TABLE II: Typical underlying factors in 'baskets'.

Inflation Expectation	Metals
US10Y Inflation expectation	CRB Rind
US5Y Inflation expectation	Iron Ore
US2Y Inflation expectation	Copper
Corporate Credit	Global QE
US HY	USD 1y5y Rate Nvol
Itraxx Xover	EUR 1y5y Rate Nvol
Itraxx japan	JPY 1y5y Rate Nvol
Fin Sub Index	GBP 1y5y Rate Nvol

Given so many macro factor series one needs to standardize the data across the different series so that they become comparable before they input to the PCA process. Also, one wishes to investigate certain timeframes (short term, longer term, etc.) and how the standardised data move across those to give a sense of regime.

B. PCA Process

To accommodate the large number of macro factors that are potentially relevant in driving a given security, many of which may share a high degree of collinearity, we input the standardised data described above for each factor into a PCA process. At the end of the complete PCA process, we retain only the first few principal components from this. Thus we obtain a small set of mutually uncorrelated macro ‘super’-themes that explain most of the variability across all the macro factors first considered.

C. Regression Process

The set of principal components are then used to perform a time-series principal component regression (PCR) against the security of interest. To give an example of this process, suppose the price time-series of the security in question is denoted by $y(t)$ and the (typically three) final principal components time-series produced by the two-step PCA process outlined above are $P_1(t)$, $P_2(t)$ and $P_3(t)$. After performing the PCR, one obtains a model for the security time-series given by

$$y_{\text{mod}}(t) = \text{constant} + \hat{\beta}_1 P_1(t) + \hat{\beta}_2 P_2(t) + \hat{\beta}_3 P_3(t), \quad (1)$$

where $\hat{\beta}_i$ ($i = 1, 2, 3$) are best-fit coefficients obtained in the PCR.

If we compare this modelled time-series with the original data $y(t)$, then we define the residual $\epsilon(t) = y(t) - y_{\text{mod}}(t)$, which should be predominantly noise and hence mean reverting. Figure 2 shows an example of a mean-reverting residual. Thus, a positive residual implies that the actual market value is higher than the predicted one and hence there should be a downward movement, and *vice versa*.

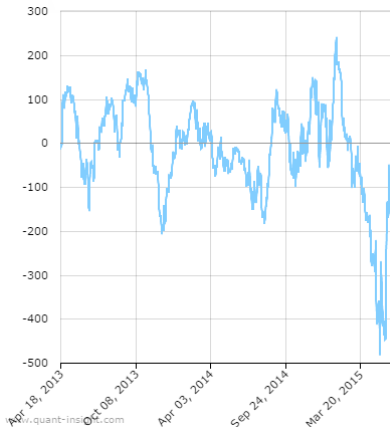


FIG. 2: Example of regression residual $\epsilon(t)$.

VI. SUMMARY

So to summarize

- Envision successfully takes care of issues related to standardization of series, non-stationarity and multi collinearity
- Envision has stable principal components (limited switching)
- Envision has stable factor sensitivities which are easier to interpret

VII. ILLUSTRATIVE RESULTS

We now present some illustrative results produced by Envision which demonstrate it to be a powerful tool for analysing drivers and sensitivities, and for making decisions on possible trades.

We take as an example the US 10yr swap (mid-July 2015). From Jan 2015, US 10yr yields had been ‘out of regime’, as signalled by the very low R^2 -values obtained for the regression. Thus, a large and fairly exhaustive set of underlying macro factors could not explain the variation of US 10yr yields for most of 2015. The rolling R^2 -statistic for the period Apr 2013 – Jul 2015 is shown in Figure 3, in which the drop in the R^2 -statistic in late 2014 and early 2015 is indicated by the red arrow. After Apr 2015, however, the R^2 -statistic began to rise and reached close to $R^2 = 0.6$ in Jul 2015, as highlighted by the red circle in the figure; this indicated that the US 10yr swap yields were moving back into a well-defined macro regime.

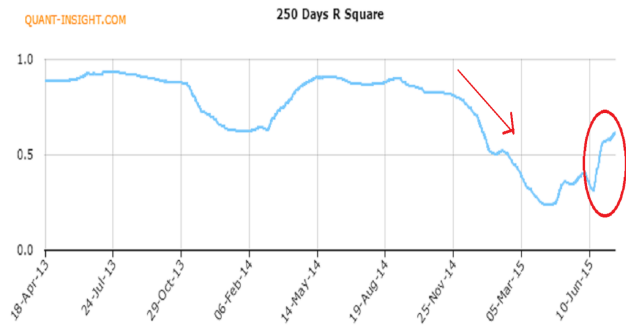


FIG. 3: R^2 -statistic for US 10yr swap over 250 days.

What did this new emerging regime look like? The corresponding macro factor sensitivity grid at the final time point is shown in Figure 4. This shows the macro factor associations now being picked up by the principal component regression analysis. These associations appear reasonably intuitive: 10yr TY yields rise with stronger data, higher inflation expectations, and higher energy prices (in that order of importance), whereas 10y yields fall if VIX rises, but that sensitivity is small.

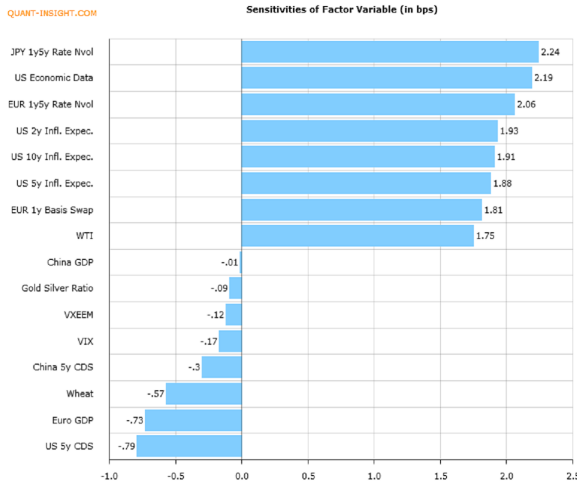


FIG. 4: US 10yr swap sensitivity bar plot.

One of the changes highlighted in Figure 4 is the stronger association with inflation expectations. The effect on the US 10yr yield resulting from a one-sigma shift in inflation expectations (as measured by 2yr, 5yr, 10yr zero coupon inflation swaps) is shown in Figure 5. It thus appeared that US 10y yields were starting to respond in the normal way to growth, inflation expectations, energy and risk premia. This marked a profound change from the behaviour of US 10y yields in the first half of 2015.

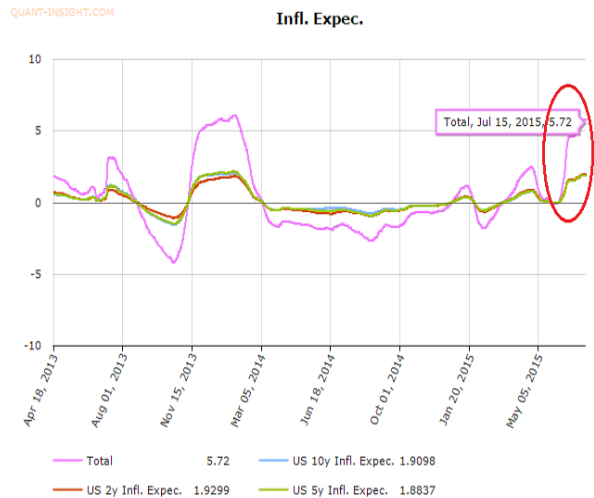


FIG. 5: US 10yr inflation expectations sensitivity.

Turning to valuation, the US 10yr also appeared to be cheap, as illustrated by the residual plot in Figure 6, which indicates that a possible 25bps could be picked up if data remained here and the actual price converged to the model.

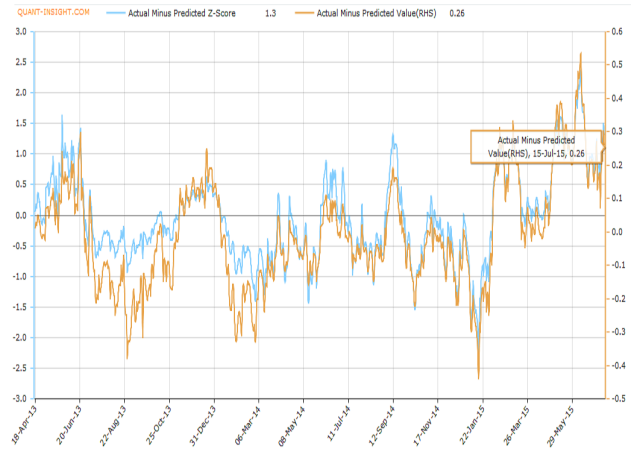


FIG. 6: US 10yr residual from the principal component regression.

Thus, the ‘**bottom line**’ of the analysis is that Envision suggested the US 10yr was a macro trade again and that it was a good receive. Indeed, the final outcome matched these expectations as in shown in Figure 7.



FIG. 7: US 10yr outcome.

VIII. CONCLUSIONS

Quant Insight (QI) have developed a powerful new tool, called Envision, to assist asset managers to identify the macro drivers of financial markets and securities in a quick, straightforward and automated manner. In particular:

- Envision is a unique quantitative tool for **deriving the drivers of market securities**. It provides a rigorous mathematical solution for understanding the sensitivity of markets to various factors that helps in trade selection, portfolio construction and risk management.
- Envision **generates a valuation** based on unique and daily updated factor-based models.
- Envision shows the factor sensitivities of portfolios, indices, sectors and ETFs, and can be used to **manage risk and check exposures**.