

Constructing Smart Portfolios From Data Driven Quantitative Investment Models

Chetan Saran Mehra
Electronics and Computer Science
Southampton University
Southampton, UK
Email: csm1g11@ecs.soton.ac.uk

Adam Prugel-Bennett
Enrico Gerding
Valentin Robu
Electronics and Computer Science
Southampton University
Email: {apb, eg, vr2} @ecs.soton.ac.uk

Abstract—In this paper we present a smart portfolio management methodology, which advances existing portfolio management techniques at two distinct levels. First, we develop a set of investment models that target regimes found in the data over different time horizons. We then build a meta-model which uses the Kelly criterion to determine an optimal allocation over these investment strategies, thus simultaneously capturing regimes operating in the data over different time horizons. Finally, in order to detect changes in the relevant data regime, and hence investment allocations, we use a forecasting algorithm which relies on a Kalman filter. We call our combined method, that uses both the Kelly criterion and the Kalman filter, the K2 algorithm. Using a large-scale historical dataset of both stocks and indices, we show that our K2 algorithm gives better risk adjusted returns in terms of the Sharpe ratio, better average gain to average loss ratio and higher probability of success compared to existing benchmarks, when measured in out-of-sample tests.

I. INTRODUCTION

Portfolio managers have access to large amounts of financial time series (FTS) data, which is rich in structure and information. Such structure at different time horizons, exhibits different characteristics [1]. Specifically, on short time horizons (i.e. a few days to a few weeks) we typically observe negative autocorrelation and, on longer time horizons, positive autocorrelation in the data. In this paper we focus on the structure found in time series data at different time horizons, capturing the two most ubiquitous features found in FTS i.e. time series momentum, which is related to positive autocorrelation in returns, and mean reversion, which is related to negative autocorrelation in returns.

The key challenge towards building our smart portfolio is to, first, identify and model the relevant data regimes operating at different time frames and convert them to an investment strategy targeting each regime separately. Regimes in financial time series can change over a period of time i.e. they are non-stationary. This has implication for a model, as it may stop being profitable once the regime it is targeting has stopped or evolved into another one over a period of time.

Changing regimes or those evolving into other regimes is one of the key reasons why we should have several independent models targeting relevant regimes at that point in time. Although there is considerable research done about building portfolios of securities, there is little research on portfolio of

investment strategies focused specifically on structure found in time series data at different time horizons.

Previous research has focused on designing strategies on separate aspects of time series data. Two of the more popular approaches are time series momentum [2] to capture a shifting mean or drift in the data and statistical arbitrage through pairs trading [3] to capture mean reversion. However, little work is available about simultaneously capturing different aspects of the data at different time horizons and then combining them, hence capturing and extracting better returns from the data to build a “smart portfolio” over different time frames. Note that our research is methodologically different from previous works, where researchers have looked at cross sectional evidence (especially in the case of momentum), by taking a multivariate approach.

In this paper we present five quantitative models, four of these models are implemented on stocks and equity index futures on different time horizons, to capture relevant aspects of the structure found in the data at different horizons. The fifth model which is our key contribution, is called the K2 model. In this model we use the Kelly Criterion [4] and the Kalman Filter [5] to identify the optimal allocation to these strategies. The Kelly Criterion is a method of finding the optimal investment size when one plays a game of chance with positive expectations, such that it maximises log expected wealth, whilst protecting from total loss of wealth. The Kalman Filter is a predict-correct-update algorithm which is used to forecast linear dynamical systems which, as we show, can be used to detect changes in the data regime.

To summarize, the main aims of our work are:

- 1) To develop a set of quantitative models that utilize the structure found in FTS over different time horizons to build a smart portfolio. By “smart”, here we mean using the structure found at different time frames in FTS to its full potential in order to build a consistent quantitative portfolio management strategy.
- 2) We combine these strategies into a portfolio using a combination of the Kelly Criterion and Kalman filters, which we call the K2 model. The K2 algorithm can be seen as a meta model, whose objective is to generate better risk adjusted returns (Sharpe ratio), while at the same time avoiding investments in strategies that are loss

making.

The remainder of the paper is structured as follows. In the next section we review related work on portfolio construction. This is followed by the Model section, where we detail all the models that we have built and how the portfolios have been constructed. In the Experiment Data section, we present the data set used for our experiments. In the Experiment Results section we present an analysis of the results for all five models, both in-sample and out-of-sample, showing that our method produces a better out-of-sample Sharpe ratio than any of the benchmarks. Finally in conclusion we propose future work and improvements to the model.

II. RELATED WORK

The goal of portfolio management and portfolio construction is to obtain the best forecast to build a profitable portfolio, while achieving a balance between risk and return. To this end, portfolio managers have used tools such as Mean Variance Optimisation (MVO) introduced by Markowitz [6], Capital Asset Pricing Model (CAPM) [7], a multi-factor version of the Arbitrage Pricing Theory (APT) [8] [9], the Black-Litterman model (BL) [10] and portfolio construction from sorts (c.f. Chriss & Almgren [11]), depending on the requirements of the portfolio manager. Another approach was introduced by Thorp and Browne [12] [13] [14] who previously adapted the Kelly criterion for the stock market and his work is the inspiration for our model. Thorp's approach was initially developed outside the mainstream portfolio management literature, and has its origins in betting games. We propose a practical improvement of Thorp's model using a Kalman filter in order to avoid investing in strategies that are not performing well because the data regime has changed. In essence, a smart portfolio should be efficient in terms of extracting value out of the market, whatever the current regime in the data.

There has been considerable previous research on stock market predictability, mean reversion and momentum. The momentum literature focuses on the relative performance of securities in the cross section [2] [15]. However, some debate remains whether the momentum evidence was conclusive as Balvers et al. found conflicting evidence in the same data (c.f. [16], [17]). Time series momentum research has been covered by prominent behavioural and rational asset pricing theories [18] which focus on single asset behaviour. Similarly rational theory of momentum [19] focuses on single time series rather than cross sectional work. There is also a stream of research in the artificial intelligence community on time series data with algorithms such as Anticor [20] and the Online Passive-Aggressive algorithm [21]. In contrast to these works, our work is based purely on time series instead of looking at a cross section of securities. We focus on capturing structure in securities prices at different time horizons and combine them to capture as much structure in the data as possible and combine them to extract better risk adjusted returns.

III. QUANTITATIVE INVESTMENT MODELS

In this section we present the four models that we have developed for both stocks and index futures. The four models

focus on stocks and index futures. For global equity indices we build two models: a time series momentum model and a pairs model. For European equities and UK equities we build one model each, we build a pairs model for the European stocks and a long only model for UK stocks. We conduct diagnostics tests to understand the nature of our data and found, as expected based on previous literature, that distributions of the in-sample data are fat tailed and fail the test of normality.

A. Index Pairs Model

The objective of the index pairs model is to capture mean reversion in related time series data that have dislocated from their long-term mean. The strategy uses index futures data from some of the most actively traded indices in Asia, Europe and the Americas. The index pairs model is designed as difference between logarithm of daily closing price of two time series, such as $\log(CACIndex) - \log(DAXIndex)$, leading to a third time series which we call the *spread*. All pairs are set up according to time zones such that European indices are only paired with European indices and the same for the Americas and Asian regions, so all trading begins and ends at the same time. The pairs model is designed to identify and utilize short term dislocation from their historic relationship of the *spread*.

To build the model, we start with the distributions of the log difference of pairs, and find most pairs to be reasonably well behaved. We then run a regression analysis on the two times series that comprise all the pairs, and store the regression residuals, where by regression residuals we mean the unexplained part of the regression. These residuals are of interest to us as they show the nature of their relationship between two time series. We then conduct the Runs test on the residuals to check the degree of serial correlation in the residuals, and expected mean reversion in the residue. Pairs of stocks that have their regression residuals cross the mean many times are of most interest to us, as they are likely to be the most stable and consistent mean reverting relationships. We then model the *spread* using the Ornstein Uhlenbeck (OU) model as shown in Eq.(1) and identify a threshold for λ to initiate our trades.

$$dS_t = \lambda (\mu - S_t) dt + \sigma dw_t \quad (1)$$

Here S_t is the spread at time t , μ is the mean, λ is the speed of mean reversion, σ is volatility, w_t is a stochastic term. In the in-sample data we find the optimal value of the lambda to trigger our trade. Here we are optimizing for two key factors maximum success rate and maximum average gain, while identifying the optimal threshold for λ . To identify the appropriate window size, we test our model on several window sizes ranging from 10 day to 50 days in increments of 5 days. Empirically, we found the 10 day window is the most stable for this model. In our tests we find that the 10 day window to be optimal as it provides the most stable performance for this model. All trades have equal capital allocation in the portfolio, e.g. if we have 100 trades then each trade has 1/100 allocation

Algorithm 1 *Mean reversion*: λ = speed of mean reversion, Γ = trade threshold, μ = spread mean, Input X: Log Spread, Output: Investment decision.

```

1: for X do
2:   if  $\lambda > \Gamma$  & Intrade = False then
3:     Sell      Intrade = True
4:
5:   else
6:     if  $\lambda < -\Gamma$  & Intrade = False then
7:       Buy      Intrade = True
8:
9:     end if
10:  end if
11:  if  $\lambda = 0$  or  $X = \mu$  then
12:    Close trade Then      Intrade = False
13:
14:  end if
15: end for

```

with a ceiling of 10%. We present the algorithm in more detail in Algorithm.(1)

B. Stock Pairs Model

The stock pairs model uses the same methodology as the one used for index pairs earlier, albeit with a little more structure. We consider a set of 250 of the largest stocks by market capitalisation, and group them into the relevant sectors and sub-sectors. For example in the telecommunication sector we put all the telecommunication companies from France, Germany, Spain, Netherlands, Italy etc. in the same sector and create pairs such as $\log(\text{FranceTelecom}) - \log(\text{DeutscheTelekom})$. Failing such a categorisation stocks are removed from the test data set. We go through the same methodology as we did for index pair selection and model the *spread* using the OU model and use the same framework to build this model, Algorithm (1).

In our in-sample tests we find that the optimal window size is 20 days, we identify this through empirical testing on several window sizes from 10 to 50 days. Pairs that do not perform well in the in-sample period are removed. We found that pairs of stocks from the same country have the shortest holding period and smaller returns than inter-country pairs, which are more volatile e.g. EDF and GDF Suez in France. We give equal capital allocation to each trade in the portfolio, just as in the index pairs model, however for this model we use trailing stop of 15% on every trade, which means we close out a trade at a loss if it loses 15% of more.

C. Momentum Model

The time series momentum model is designed to capture drift in securities over long horizon in index futures. To identify long run dependence in returns we test for positive autocorrelation in monthly returns of all the indices. We find that most of the series show positive autocorrelation up to

Algorithm 2 *Momentum*: LMA = long moving average, SMA = short moving average, Input X: Log of index futures, Output: Investment decision.

```

1: for X do
2:   if  $SMA > LMA$  & Intrade = False then
3:     Buy      Intrade = True
4:
5:   else
6:     if  $SMA < LMA$  & Intrade = False then
7:       Sell      Intrade = True
8:
9:     else
10:      if  $SMA = LMA$  & Intrade = True then
11:        Close trade Then      Intrade = False
12:
13:      end if
14:    end if
15:  end if
16: end for

```

five periods. Although Moskowitz et al. [2] uses 12 month change to identify momentum, we use the simple moving average crossover methodology, when the short moving average (SMA) is above long moving average (LMA) we *Buy* futures and when When SMA is less than the LMA we *Sell* the futures. To build our time series momentum model we use a fixed length SMA of one month vs. LMA of three months cross over to identify the beginning and end of trends, we outline the model in algorithm.(2). We give equal capital allocation to each trade in the portfolio, this is a constant exposure model. By constant exposure we mean that the model is always invested in the market.

D. Long-only Model

The Long only model is built on the UK FTSE 100 constituents. The objective of this model is to profit by buying stocks that are the weakest performers among their peers in the same sector. The premise for this model is that stocks within the same sector have similar volatilities and returns, and they do not stray too far away from their peers unless their is a major company specific event. For this model we use both market open and market close data of in-sample period, giving us twice the data points and capturing more structure in the data generating process. To build this strategy, first, we divide all the stocks into their respective sectors or sub-sectors, such as Telecom, Banks, Insurance, Oil, etc. Thus, we obtain a total of 10 such segmented groups.

We analyse the correlation between stocks that are in the same sector and, as expected we find that they are strongly correlated with each other. The long only strategy is a constant rebalancing strategy i.e. we rebalance at every time step and maintain equal capital allocation to all positions in each sector. We maintain constant exposure in the model, i.e. we are always invested, we do not have any stop loss on this strategy.

In more detail we present an example of the model. At each time step (market open and market close) we rank the

stocks in each sector by their returns in the previous period ($t-1$). At time step (t) we buy stocks with the lowest returns i.e. stocks that are ranked in the lowest half of the sector, rounded to the lowest round number. For example, if we have 5 stocks in a group A, B, C, D and E, after ranking them by their returns, we get the following order: E is ranked 1, C is ranked 2, B is ranked 3, D is ranked 4, and A is ranked 5. So for the next period we would buy D and A as they are the weakest performers. We repeat this at every time step (open and close) maintaining a constant position in the weakest performers in the previous time period at all times, as referenced in algorithm.(3). Owing to high correlation we expect the weakest performers to catch up with their peer group.

Algorithm 3 *Long Only: Sort* = Sorting Algorithm, Input: Matrix of returns = Z , Output: Investment decision.

```

for each time step do
  for each sector do
    Sell previous position
    Ranked = Sort(total  $N$  stocks) based on  $Z$ 
    Buy  $\lfloor (N/2) \rfloor \sim N$  stocks
  end for
end for

```

IV. K2, THE META MODEL

In this section we present K2, our meta model, which is the key contribution of this paper. The K2 model uses both Kalman Filters and the Kelly Criterion [12] to make allocations to the four models, constructing a smart portfolio from these models. The Kalman filter plays an important role for K2. Whenever the individual investment models (presented above) are operating on non-stationary data, they are vulnerable to shifts in data regime, rendering them ineffective and potentially turning them from profitable to loss making strategies. When such a situation arises the Kalman filter helps us avoid investing in a strategy that is loss making, allowing K2 to take “smarter” capital allocation decisions. We now briefly present the Kelly Criterion and the Kalman Filter.

A. Kelly Criterion

The Kelly Criterion is based on Information Theory and was initially used for repeatable betting games. Subsequently Thorp et al. [22], Browne [14] expanded it to blackjack and then the stock market by extending Kelly and Breiman’s [23] work to continuous form. The Kelly Criterion has many desirable features, which make it suitable for our approach. First, Kelly maximises log of wealth without risking bankruptcy Fig. 1; second, it minimises the expected time to reach target wealth; and finally Kelly is myopic i.e. we only need to consider our current investment opportunity [22].

The Kelly Criterion has a few negative properties as well, Firstly the size of bet can get very large when the probability of loss is very small. This also has the side effect of making returns volatile, impacting risk adjusted returns, even though maximising wealth. Second, Kelly assumes that wealth is

infinity divisible; and finally, good estimates probabilities of profit/loss are important for the model to work well.

Although Kelly was originally applied to binomial games with well defined outcomes, Thorp et.al [22] have furthered the work for financial markets and portfolios. The Kelly equation for the discrete case equation and equal win and loss sizes is:

$$\log G(f) = p \log(1 + f) + q \log(1 - f) = E \log(X) \quad (2)$$

Here the $\log G(f)$ signifies capital growth, which we want to maximise, p is the probability of winning, q is the probability of losing and f is the Kelly fraction of your capital that is bet at each time step. We can adapt this for unequal payoff as follows:

$$\log G(f) = p \log(1 + fb) + q \log(1 - fa) = E \log(X) \quad (3)$$

Next we introduce a and b in the equation, which signify the size of win and size of loss. Thorp [13] adapted this for stock market investments incorporating the risk free rate r .

$$\log G = \frac{1}{2} \log(1 + (1 - f)r + f(\mu + \sigma)) + \frac{1}{2} \log(1 + (1 - f)r + f(\mu - \sigma)) \quad (4)$$

By Taylor expansion of $G(f)$ for $f = 0$ and taking the limit we get the continuous form of Kelly fraction.

$$f^* = \frac{\mu - r}{\sigma^2} \quad (5)$$

Where μ is the mean return, σ^2 is the variance, σ is the standard deviation, r is the risk free rate of return and f^* is the optimal Kelly bet for the four models. This is very close to the Sharpe ratio, the only difference is that in the denominator we have variance where as the Sharpe ratio has standard deviation. We will be using this equation to calculate Kelly fraction for our model as well. However, our model returns will be filtered using the Kalman Filter which we present in the next sub section.

B. Kalman Filter

The Kalman filter is a recursive predict-correct-update algorithm devised by R.E. Kalman [5] [24]. The Kalman Filter is a Markov model that predicts the future state of variable at time step $t + 1$. We briefly recall the steps of Kalman filtering for linear state-space models. We set the following notations:

$$\begin{cases} \theta_0 \sim N(m_0, V_0) \\ \theta_t = \mathbf{A}_t \theta_{t-1} + w_t, & w_t \sim N(0, W_t) \\ z_t = \mathbf{H}_t \theta_t + v_t, & v_t \sim N(0, V_t) \end{cases} \quad (9)$$

$z_t \in \mathbb{R}^m$ stands for the observation vector, $\theta_t \in \mathbb{R}^p$ is a hidden random vector, \mathbf{H}_t is the observation matrix; and \mathbf{A}_t is the system matrix, that predicts our position at the next time step; and are of size respectively $(m \times m)$ and $(p \times p)$, are to be specified, and V_t and W_t are the observation and evolution covariance matrices of size $(m \times m)$ and $(p \times p)$ respectively.

In the following sections, we assume that $V_t = V$ and $W_t = W$, for any t . They are estimated from available in sample

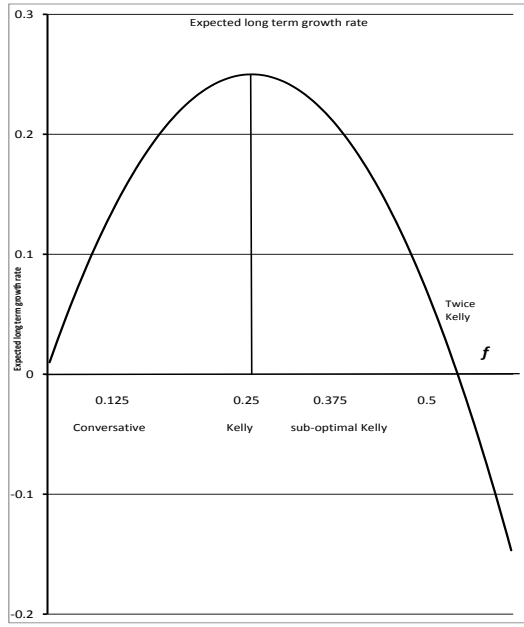


Fig. 1. Expected long-term rate as a function of investment function.

data. The Kalman filter recursively estimates the internal state of the process θ_t given the sequence of noisy observations z_t . We denote by $\theta_{t|T}$ the estimate of the state at time t given observations up to and including time T , and by $P_{t|T}$ the associated error covariance matrix. This can be summed up by the system of equations:

$$\begin{cases} \hat{\theta}_{t|t-1} = \mathbf{A}_t \hat{\theta}_{t-1|t-1} + w_t & (10) \end{cases}$$

$$\begin{cases} \mathbf{P}_{t|t-1} = \mathbf{A}_t \mathbf{P}_{t-1|t-1} \mathbf{A}_t' + \mathbf{W}_{k-1}, \\ y_t = z_t - \mathbf{H}_t' \hat{\theta}_{t|t-1}, & (11) \end{cases}$$

$$\begin{cases} \mathbf{S}_t = \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t' + V_t, & (12) \end{cases}$$

$$\begin{cases} \mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t' \mathbf{S}_t^{-1}, & (13) \end{cases}$$

$$\begin{cases} \hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + \mathbf{K}_t r_t, \\ \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} & (14) \end{cases}$$

Equation (10) gives the predicted state at step t and equation (11) the innovation residual. S_t Eq.(12) is the innovation covariance and compares the real error against prediction. K_t Eq.(13) is the the Kalman gain which moderates the prediction based on the accuracy of the last time step $t-1$ and Eq.(14) represents the new estimation of error for the next time step.

C. Design of K2 Model & Benchmarks

In this section we present the design of our K2 meta model. One of the key challenges for us is that the data regime can change, rendering our model loss making. When a model becomes loss making we want K2 to be efficient enough not to invest in that model, yet while the model is profitable we want to be able to maximise our gain. To avoid investing in strategies when they are loss making, we use the Kalman filter. When the Kalman Filter forecasts a negative return we make no allocation to that model, and when we have a positive forecast we make the optimal investment to that

strategy Eq.(5). However there is a difference between Kelly calculated from the raw data generated by the four models and one calculated by K2. We calculated our optimal weights using the Kelly formula on data that has been forecast using the Kalman filter.

The following steps are used to calculate the K2 allocations to the four individual investment models:

- 1) We take the monthly returns for the four models and build an equity curve starting at the base of 100.
 - 2) Then we take this equity curve for each of the models and fit them to a Kalman filter.
 - 3) We then take the Kalman filtered forecast to assess whether we will make an investment in a particular model, if our forecast is “negative” for a particular model allocation is “zero” to that model and if its “positive” we make an allocation.
 - 4) After filtering the data in the previous step, we now have “actual model returns” when there is a “positive” forecast and “zero” when there is a “negative forecast”.
 - 5) Now that we have taken account of the zero allocation we now calculate our new Kelly fraction using Eq.(5) for all four models. Here μ is the annualised compounded return, σ^2 is the annualised variance, r is annualised risk free rate.
 - 6) We now rescale and convert our Kelly fractions to percentages with constraint that they are not higher than the non-rescaled Kelly to allocate capital to our model.
- More specifically, we do not calculate Kelly from actual model returns but after they have been forecast by the Kalman filter. When the Kalman filter forecasts negative return we attribute that period a return of “zero” for that model, and when we have a positive forecast we have the actual strategy returns, we step through the algorithm 4. Thus, this smart allocation is able to bring about added efficiency to the portfolio by avoiding loss making models and using capital in a smarter way.

We also implement four other models that we use as benchmarks. We use mean variance optimisation (MVO) [6], Equally weighted portfolio (1/n), Rescaled Kelly (Re. Kelly) [12] [23] [14], Optimal Kelly [25] and K2. All the models are implemented in an *online* setting, that is, whenever any new data is available all models receive it and recompute and update.

Therefore, our smart portfolio management algorithm operates at two levels. At the first level, we are capturing the different regimes found in the data at different time horizons. At the second level, we are making an efficient capital allocation decision by avoiding potentially losing investments, resulting in a model that has a better chance of overall positive returns.

Algorithm 4 *K2*: KF = Kalman Filter, MR = Model Returns Matrix, NR = New Returns Matrix, K = Kelly calculation
Output: *K2* Allocation matrix for each model.

```

for  $i = 1$ : time steps do
  for  $j = 1$ : number of models do
    if  $KF(MR(i, j)) \leq 0$  then
       $NR(i, j) = 0$ 
    else  $NR(i, j) = KF(MR(i, j))$ 
    end if
     $K2(i, j) = K(NR(i, j))$ 
  end for
  if  $sum(K2(i, :)) > 100$  then
     $K2 = K2 / sum(K2)$ 
  end if
end for

```

V. EXPERIMENTAL DATA

In this section we present the data that was used for experimental validations of our models. This data comes from equity markets, more specifically from individual stocks and index futures, where a futures contract tracks the movements of an index such as FTSE 100. Firstly for we use global equity indices with liquid futures markets, for the Americas we use Bovespa (Brazil), S&P/TSX60 (Canada), S&P500 (USA), Nasdaq100 (USA), for Europe we use FTSE100 (UK), EuroStoxx (Pan Europe), DAX (Germany), CAC40 (France), IBEX (Spain), S&P/MIB (Italy), AEX (Netherlands), SMI (Switzerland) and for Asia we use Nikkie250 (Japan), S&P/ASX200 (Australia) and Hang Seng (Hong Kong). This data is used for two models: the times series momentum and the pairs model. Second, we use stocks from the European markets, focusing on 250 of the most liquid European stocks denominated in Euros and constituents of the FTSE100 index, we build one model for European stocks and another model for UK stocks. Hence we are able to capture the interconnect- edness between stocks as well the indices they are part of, as shown in Fig.2.

For model development and testing we have daily frequency data, market open and market close from January 2005 to Dec 2012. Data from 2005 to 2009 being the in-sample period and 2010 to 2012 being the out-of-sample period for validation. All securities in the in-sample period are included in the out-of-sample period. We use close data for all our models except one where we use both open and close data points. Our data is synchronised for holidays and gaps i.e. when there is a holiday we use the previous days data point so all data is aligned in time.

VI. EXPERIMENTAL RESULTS

In this section we present an analysis of our results from the experiments we conducted. We present both in-sample and out-of-sample results, where we show that our *K2* model has a better Sharpe ratio when compared to benchmarks. Our measure of better performance is the Sharpe ratio, for the period 2005 to 2007 we use 3% as our average risk free rate

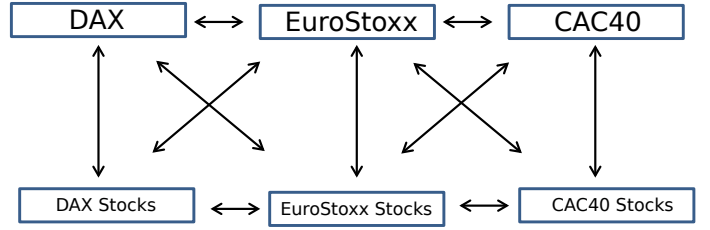


Fig. 2. Inter connected data that impacts the index and other stocks

and from 2008 to 2012 we use 0.50% as the risk free rate, even though at some instances the risk free rate was negative. For the out-of-sample validation tests we keep the same parameters and make no changes to any of the models. First we will present the results of the individual models that target different data regimes and present their results. Subsequently, we will present the results of models that combine the underlying models and discuss their performance.

We also implement four other models that we use as benchmarks. We use mean variance optimisation (MVO) [6], Equally weighted portfolio (1/n), Rescaled Kelly (Re. Kelly) [22], Optimal Kelly [25]. All the models are implemented in an online setting, that is, whenever any new data is available all models receive it and recompute and update their estimates.

1) *Individual models*: Below we present the results of the individual models. In Table I we present the in-sample results of the individual models and observe that both the pairs models have considerably lower volatility and much higher Sharpe ratio, when compared to the equity long and index momentum model. Index momentum model has particularly low Sharpe ratio, well below 1. In the out-of-sample results (Table II), we find that most of the models have lower performance than in the in-sample period, except stock pairs. However, the index momentum model is clearly the worst performer, with negative returns and high volatility. This is a typical situation where the data changes and the investment model starts to lose money, that we aim to avoid.

TABLE I
INDIVIDUAL MODELS IN-SAMPLE

Model	Return	Volatility	Sharpe
Index Pairs	21.3%	6.35%	2.89
Stock Pairs	15.9%	3.79%	3.41
Equity long	33.7%	17.43%	1.76
Index Mom	12.9%	12.68%	0.78

TABLE II
INDIVIDUAL MODELS OUT-OF-SAMPLE

Model	Return	Volatility	Sharpe
Index Pairs	15.22%	6.79%	2.17
Stock Pairs	16.30%	3.61%	4.37
Equity long	24.89%	14.61%	1.67
Index Mom	-2.82%	11.37%	- 0.29

2) *K2 and benchmarks*: We now compare the results of our meta model *K2* with four other benchmarks we have mean variance (MV) [6], equally weighted (1/n), and Optimal Kelly

(OK) and Rescaled Kelly (RK). In the in-sample period we find Optimal Kelly has the highest return but the lowest Sharpe ratio, MVO has the highest Sharpe Ratio, however rescaled Kelly and K2 are not too far behind. We also find that in the in-sample period K2 does not have the highest return but has the lowest volatility of all other benchmarks, as it misses out some opportunities while trying to avoid losses.

TABLE III
K2 IN-SAMPLE COMPARISON

In-sample	Ann. return	Ann. Volatility	Sharpe ratio
K2	16.0%	3.27%	3.99
Re. Kelly	17.7%	3.46%	4.25
Optimal Kelly	26.9%	13.94%	1.72
MVO	19.2%	3.75%	4.31
Eq. Wted	20.7%	5.34%	3.32

TABLE IV
K2 OUT-OF-SAMPLE COMPARISON

Out-of-sample	Ann. return	Ann. Volatility	Sharpe ratio
K2	14.7%	3.33%	4.28
Re. Kelly	15.6%	3.55%	4.25
Optimal Kelly	24.9%	14.61%	1.67
MVO	14.6%	3.59%	3.93
Eq. Wted	13.0%	4.65%	2.68

TABLE V
KEY STATISTICS

Out-of-sample	Gain/Loss ratio	Success Prob.
K2	5.46	0.94
Rescaled Kelly	4.99	0.92
Optimal Kelly	1.49	0.72
MVO	2.99	0.92
Eq. Wted	1.70	0.78
Optimal Kelly	1.49	0.72

In the out-of sample-period we find that K2 does not have the highest returns when not adjusted for risk. Optimal Kelly is still the best performer, followed by rescaled Kelly. Performance when adjusted for risk and measured through the Sharpe ratio as shown in Fig. 3, we find that K2 has the highest Sharpe ratio and it is the only model that has a higher Sharpe ratio in the out-of-sample period. K2's high Sharpe ratio is high in the out-of-sample period because, it is able to avoid investing in volatile loss making periods in the underlying models Fig.5, as is the case with the index momentum model, which makes losses in the out-of-sample period. This has two benefits, first we avoid losses reducing volatility and second we make better allocation to capital by reallocating capital to profitable models.

If we look at the out-of-sample equity curve graph Fig.6, Optimal Kelly clearly beats all other models, Rescaled Kelly is second and K2 is third but only by a very small margin. Furthermore in the out-of-sample period K2 has the highest probability of success when compared to all the benchmarks. Additionally K2's ratio of average gain to average loss is the highest as well as shown in Fig.4, primarily because K2 is able to avoid many negative (i.e. loss making) periods in the underlying four models. Although Optimal Kelly clearly

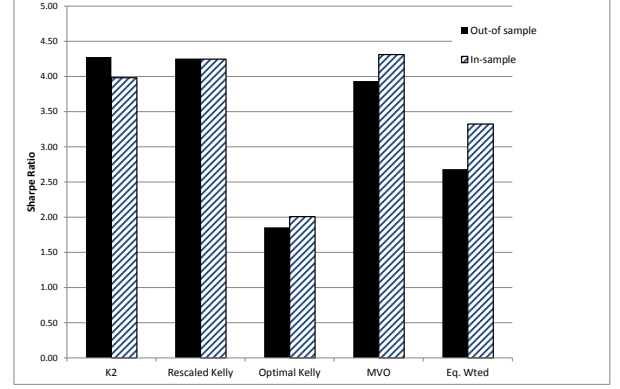


Fig. 3. Sharpe Ratio, out-of-sample in solid black.

maximises wealth both in-sample and out-of-sample, in our case it picks the best performer at each time step giving practically zero weights to the other 3 models. We do not think achieving slightly lower growth is a significant disadvantage in real life as extra growth can be achieved through use of leverage, however the same cannot be said for Sharpe ratio or volatility.

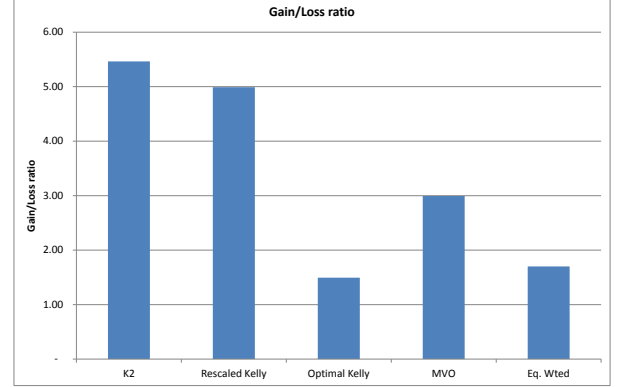


Fig. 4. Gain/Loss ratio

VII. CONCLUSIONS

In this paper, we develop a smart portfolio algorithm that gives better risk adjusted returns than existing investment models. Our work takes a two-step approach to this problem.

First, using existing literature on portfolio management, we build a set of four quantitative investment models, each of which is able to take advantage of a regime operating in the data, for a given time horizon. However, we show that when the data driving a model changes, the model can become loss making.

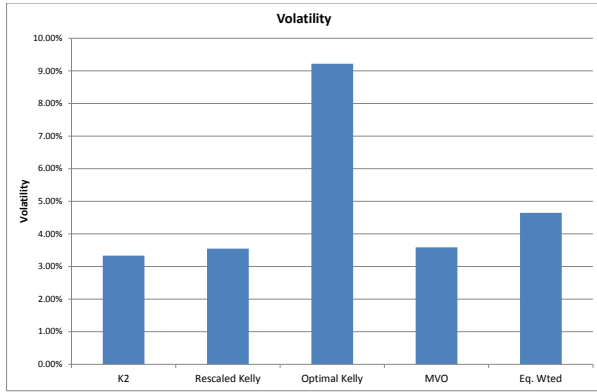


Fig. 5. Volatility out-of-sample

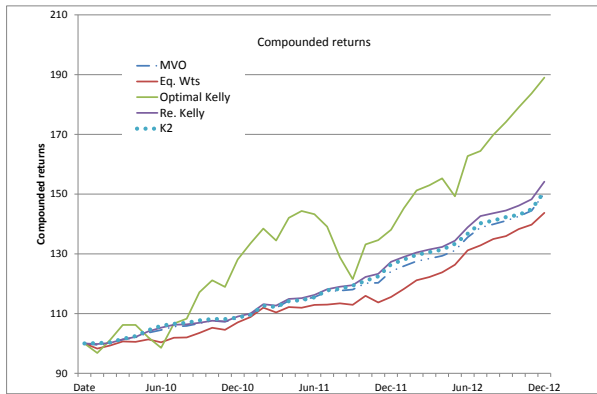


Fig. 6. Equity curve

Second, we aimed to build a model that would be efficient in allocating capital in light of data being non-stationary and evolving into different regimes as time passes. In order to achieve this, we propose a novel, meta-level investment model (which we call the K2 algorithm) that combines a capital allocation model based on the optimal Kelly criterion with a forecasting algorithm using a Kalman filter.

Using large-scale dataset of stock and index futures data, we show that our K2 algorithm achieves better risk adjusted returns, as measured by the Sharpe Ratio than any of the individual benchmarks we considered. Thus, although the K2 algorithm achieves somewhat lower overall returns (because it sometimes fails to invest in profitable opportunities), it performs a smarter overall capital allocation, because it manages to avoid loss-making investments.

In future work, we aim to explore other algorithms that capture further aspects of data structure, including using cross sectional models when taking portfolio allocation decisions.

ACKNOWLEDGMENT

The authors would like to thank Beining Shang and Lampros Stavrogiannis for their useful comments.

REFERENCES

- [1] R. Cont, "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative Finance Volume 1 (2001)* 223236, vol. 1, pp. 223–236, 2001.
- [2] J. T. Moskowitz, Y. H. Ooib, and L. Pedersenb, "Time series momentum," *Journal of Financial Economics*, vol. 104, pp. 228–250, 2011.
- [3] E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst, "Pairs trading: Performance of a relative-value arbitrage rule," *Review of Financial Studies*, vol. 19, pp. 797–827, 2006.
- [4] J. Kelly, "A new interpretation of information rate," *Information Theory*, vol. 2, pp. 185–189, 1956.
- [5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME—Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.
- [6] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, pp. 77–91, 1952.
- [7] W. F. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The Journal of Finance*, pp. 425–442, 1964.
- [8] S. Ross and R. Roll, "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, vol. 13, pp. 341–360, 1976.
- [9] —, "An empirical investigation of the arbitrage pricing theory," *The Journal of Finance*, vol. 35, pp. 1073–1103, 1980.
- [10] F. Black and R. Litterman, "Global portfolio optimization," *Financial Analysts Journal*, vol. September/October, pp. 28–43, 1992.
- [11] N. A. Chriss and R. Almgren, "Portfolios from sorts," *Working paper*, 2005.
- [12] E. O. Thorp, "Portfolio choice and the kelly criterion," *Stochastic Optimization Models in Finance*, pp. 599–619, 1971.
- [13] —, "The kelly criterion in blackjack sports betting and the stock market," in *The 10th international conference of gambling and risk taking Montreal*, 1997.
- [14] S. Browne, "Survival and growth with liability: Optimal portfolio strategies in continuous time," *Mathematics of Operations Research*, vol. 22, 2, pp. 468–493, 1998.
- [15] C. Asness, T. J. Moskowitz, and L. Pedersen, "Value and momentum everywhere," *The Journal of Finance*, vol. 68, pp. 929–985, 2013.
- [16] Y. Balvers, R. J. & Wu, "Momentum and mean reversion across national equity markets," *Journal of Empirical Finance*, vol. 13, pp. 24–48, 2006.
- [17] G.-A. L. . P. de Garcia. F. and C. J, "Mean reversion in stock market prices: New evidence based on bull and bear markets," *Research in International Business and Finance*, vol. 24 Issue 2, pp. 113–122, 2010.
- [18] H. Hong and J. Stein, "A unified theory of underreaction, momentum trading and over reaction in asset markets," *Journal of Finance*, vol. 54, pp. 2143–2184, 1999.
- [19] L. Liu and L. Zhang, "Momentum profits, factor pricing, and macro-economic risk," *Review of Financial Studies*, vol. 21, pp. 2417–2448, 2008.
- [20] A. Borodin, R. El-Yaniv, and V. Gogan, "Can we learn to beat the best stock," *Journal of Artificial Intelligence Research*, vol. 21, pp. 579–594, 2004.
- [21] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and y. Singer, "Online passive-aggressive algorithms," *Journal of Artificial Intelligence Research*, vol. 7, pp. 551–585, 2006.
- [22] L. MacLean, E. O. Thorp, and W. T. Z. W. T., *The Kelly Capital Growth Investment Criterion*, L. MacLean, E. O. Thorp, and W. T. Z. W. T., Eds. World Scientific Press, 2011.
- [23] L. Breiman, "Optimal gambling systems for favorable games," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 65–78, 1961.
- [24] N. Mahler, *Modeling the S&P 500 index using the Kalman filter and the LagLasso*, *Machine Learning for Signal Processing, 2009. MLSP 2009. IEEE International Workshop*, 2009.
- [25] C. Whitrow, "Algorithms for optimal allocation of bets on many simultaneous events," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 56, pp. 607–623, 2007.