

# On statistical arbitrage under a conditional factor model of equity returns.

Trent Spears<sup>\*†</sup>, Stefan Zohren, Stephen Roberts

September 6, 2023

## Summary

We consider a conditional factor model for a multivariate portfolio of United States equities in the context of analysing a statistical arbitrage trading strategy. A state space framework underlies the factor model whereby asset returns are assumed to be a noisy observation of a linear combination of factor values and latent factor risk premia. Filter and state prediction estimates for the risk premia are retrieved in an online way. Such estimates induce filtered asset returns that can be compared to measurement observations, with large deviations representing candidate mean reversion trades. Further, in that the risk premia are modelled as time-varying quantities, non-stationarity in returns is *de facto* captured. We study an empirical trading strategy respectful of transaction costs, and demonstrate performance over a long history of 29 years, for both a linear and a non-linear state space model. Our results show that the model is competitive relative to the results of other methods, including simple benchmarks and other cutting-edge approaches as published in the literature. Also of note, while strategy performance degradation is noticed through time – especially for the most recent years – the strategy continues to offer compelling economics, and has scope for further advancement.

**Keywords:** Statistical arbitrage, state space models, dynamic factor models, machine learning.

## 1 Introduction

Statistical arbitrageurs seek trading opportunities whereby an asset price (or return) is dislocated from a measure of fair value to which it is assumed it will revert. Such dislocation needs to be sufficiently large so as to be profitably traded – at least, in expectation. These strategies are documented as originating within the financial services in the late 1970s and early 1980s, within both the hedge fund and investment banking industries [1]. To this day, statistical arbitrage strategies are of popular practical application, and strategy techniques and insights have inspired a deep academic literature. A comprehensive review, for work dated up till 2016, can be found in [2].

Indeed, much of the academic literature on statistical arbitrage ultimately concerns estimating a fair value model for an asset spread or portfolio, and/or modelling and analysing the stochastic dynamics of the dislocation about fair value. This paper is mostly concerned with the former. To our knowledge, we are the first authors to present and assess a conditional factor model for estimating fair value in the context of statistical arbitrage. The model assumes that future returns are predictable given (i) the asset-wise factor exposures for the universe of assets of interest, and (ii) a latent factor vector that is estimable by regression in the

---

<sup>\*</sup>University of Oxford, Oxford-Man Institute of Quantitative Finance.

<sup>†</sup>Corresponding author. E-mail: trent@robots.ox.ac.uk

asset cross-section. Further, we assume the dynamic nature of the factor vector, suggesting a state space framework for the joint modelling of factors and asset returns.

Our fair value model is closest in spirit to existing work that presents explanatory models of returns based on Principal Component Analysis (PCA) [3, 4]. In these papers, the principal components (and their loadings) that are deemed to have sufficiently high explanatory power are an analogue to our model’s factor exposures (and factor vectors). However the papers differ significantly in many key workflow assumptions relative to ours, most importantly with respect to data set construction, asset feature modelling, and the trading strategy assumed. We posit, critically, that the details of each of these steps can have a significant impact on the estimated performance statistics. Hence we clearly describe our methods in the interest of ease of interpretability and reproducibility, and present a range of performance data based on competing sets of reasonable assumptions. Owing to transparency, we also find two key, well-known works with which we can directly compare our results; this is less commonly found possible for academic quantitative finance papers that depend on strategy backtests. In any case, we further describe these comparable papers, and other recent work in the spirit of ours, in Section 2 below. This summary mostly covers work not contained in the aforementioned review [2], due to their recency.

In Section 3, we present further the state space framework as it relates to the conditional factor model. We model returns as a noisy observation of the latent factor vectors, and describe the well-known Kalman Filter algorithm for the case of a linear model, and the Unscented Kalman Filter algorithm for the case the model specified is non-linear. In Section 4 we review details of the dataset construction, and the factor dynamics model. We also discuss the particulars of the trading strategy, that amounts to the management of a long-short beta neutral equity portfolio. We present the specification of the trading rule, realistic transaction cost assumptions, the hedging of market risk, and the calculation of strategy returns and bankroll management.

Our results are presented in Section 5. We show that the conditional factor state space model is useful for trading mid-frequency statistical arbitrage in US equity markets. This is true with respect to a meaningful benchmark model and a simpler ordinary least squares-based approach, and to the results given in other well-known academic studies calculated over similar data sets. We conclude in Section 6, and given that the model presents as a compelling starting point for building a strategy at larger scale, we include notes for both academic and practitioner future work.

## 2 Existing literature

We collect and describe studies relating to statistical arbitrage trading, for work mostly post the review period of [2]. Unsurprisingly, many recent methods for approaching the problem subsume modern machine learning techniques. These approaches mostly concern pairs trading. For example, surrendering performance interpretability, the work of [5] presents a spectrum of techniques – Principal Component Analysis (PCA) for dimension reduction of raw data, clustering as a strategy for pairs identification, and forecasting with deep learning seeking better trade entry points. Similarly [6], but without the lattermost step. Many authors appeal to dimension reduction when a large universe of portfolio assets are under consideration; this often amounts to an application of PCA. (Non-PCA based) Factor models are an alternative with a strong theoretical grounding, and are presented (amongst many other introductory ideas) in the well-known work of [7]. Such models are often characterised by a collection of hand-constructed theoretically-justified features called ‘factors’ that correlate to price or returns time series of the target assets. An insightful empirical demonstration of both applied PCA and factor models can be found in [3]; also of note, the paper jointly trades a multivariate portfolio of assets – beyond targeting pairs only. The approach of [8] is to model returns via a statistical factor model, with an Ornstein-Uhlenbeck process for the error term. Novelty is introduced at the level of trade execution, formulated via a stochastic control model, though only applied on synthetic

data. Approaches to optimizing mean-reverting multivariate portfolios, subject to either budget or leverage constraints, and influenced by Modern Portfolio Theory, are found in [9, 10]. An approach to multivariate statistical arbitrage modelling beyond factor models is found in [11], that presents a framework for both linear and non-linear relationship modelling via vine copulas.

State space models combine knowledge of the data process with noisy measurement data; a focus of this work is to consider their utility for statistical arbitrage applications. Hence our eventual contribution – a filtering approach to multivariate statistical arbitrage, scalable to large portfolios. Indeed, state space models for finance have been reviewed in [12, 13]. The articles include applications to stochastic volatility modelling, and the modelling of the term structures of commodity prices and interest rates. Applications to pairs trading are relevant, and were introduced by the seminal work of Elliott [14], who models a latent pair spread, and noisy observations thereof, in a state space setting. Their results are presented for synthetic data. The work of [15, 16] extends the research by introducing time-varying parameters to the model, hence improving model flexibility, and testing the theory in realistic online trading settings. The Elliott model is also generalised somewhat for richer spread dynamics in [17], which designs a novel trading strategy based on model-estimated probabilities of large dislocations mean reverting. Further generalisations are apparent in the recent article [18] which presents a quasi-Monte Carlo estimation algorithm for the state space framework presented. This allows for estimating models with richer specifications, designed to express stylized features of financial data, including non-Gaussianity and heteroskedastic variance.

It is natural to incorporate (dynamic) factor models into a state space framework, as in [19], which also models stochastic volatility within the state variable; similarly [20]. Arguments for dynamic factor modelling for statistical arbitrage are given by [4], though the approach does not model volatility (beyond a white noise process). PCA for factor mining, and prediction (of note, at the level of asset prices, rather than returns) is a heavily exploited feature of the trading strategy.

The studies related to statistical arbitrage are summarised in Figure 1 below, and include information on the underlying market data analysed, by asset class, time scale and date range. ‘SS’ denotes modelling in a state space framework, ‘Target’ is given by ‘pair’ for pair-trading examples, and ‘MV’ for multi-variate portfolios, while ‘Dim. red.’ denotes a dimension reduction technique is a feature of the modelling approach.

Author (Year)	SS	Target	Dim. red.	Asset class	Date range	Time scale
[2] (2017)		pair		-	-	-
[5] (2020)		pair	✓	Commodities	2009–2019	intraday – 5min
[6] (2023)		MV	✓	Equities – Global	2011–2021	daily
[7] (2004)		pair	✓	-	-	-
[3] (2010)		MV	✓	Equities – US	1997–2007	daily
[8] (2019)		MV	✓	-	-	-
[9] (2018)		MV		Equities – US	2012–2014	daily
[10] (2019)		MV		Equities – US	2010–2014	daily
[11] (2016)		MV		Equities – US	1992–2015	daily
[14] (2005)	✓	pair		-	-	-
[15] (2011)	✓	pair		Equities – US, Commodities	1980–2008	daily
[16] (2010)	✓	pair		Equities – US	1997–2005	daily
[17] (2016)	✓	pair		Equities – US, Brazil	2011–2013	daily
[18] (2021)	✓	pair		Equities – US, Taiwan, Hong Kong	2012–2019	daily
[4] (2016)	✓	MV	✓	Equities – US	1989–2011	daily

Figure 1: Literature themed around statistical arbitrage, toward applications of state space models for multivariate portfolios.

### 3 A dynamic model of returns and factors

*State space modelling.* Consider the case of discrete time increments  $k = 0, 1, \dots$ . Then a linear Gaussian state space model is given by the time-varying system of equations:

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \boldsymbol{\epsilon}_k^x, \quad (1)$$

$$\mathbf{y}_k = \mathbf{B}_k \mathbf{x}_k + \boldsymbol{\epsilon}_k^y. \quad (2)$$

Here  $\boldsymbol{\epsilon}_k^x$  and  $\boldsymbol{\epsilon}_k^y$  model independent zero-mean white-noise processes with respective covariances  $\Psi_k$  and  $\Omega_k$ , assumed to be known functions of time. The matrices  $\mathbf{A}_k$  and  $\mathbf{B}_k$  are also assumed to be known functions of time. We call equations (1) and (2) the ‘state’ and ‘measurement’ equations, respectively. The state  $\mathbf{x}_k$  is assumed to be a latent random variable, of which  $\mathbf{y}_k$  denotes a noisy measurement at time  $k$ . Given the system parameters, probabilistic beliefs about the state can be inferred, and these beliefs can be updated given measurement data.

*Kalman filter.* By the classic Kalman filter (KF) algorithm, given an initial state  $\mathbf{x}_0$ , and the state and measurement equations above, the filter distribution for  $\mathbf{x}_k | \mathbf{y}_{1:k}$  at time  $k > 0$  can be shown to be  $\sim MVN(\boldsymbol{\mu}_k^f, \Sigma_k^f)$  where

$$\boldsymbol{\mu}_k^f = \boldsymbol{\mu}_k^p + \mathbf{G}_k (\mathbf{y}_k - \mathbf{B}_k \boldsymbol{\mu}_k^p), \quad (3)$$

$$\Sigma_k^f = (\mathbf{I} - \mathbf{G}_k \mathbf{B}_k) \Sigma_k^p. \quad (4)$$

These equations describe the mean and covariance in terms of the ‘Kalman gain’  $\mathbf{G}$  given by:

$$\mathbf{G}_k := \Sigma_k^p \mathbf{B}'_k (\mathbf{B}_k \Sigma_k^p \mathbf{B}'_k + \Omega_k)^{-1}.$$

Also, the predictive distribution for  $\mathbf{x}_{k+1} | \mathbf{y}_{1:k}$  at time  $k$  can be shown to be  $\sim MVN(\boldsymbol{\mu}_k^p, \Sigma_k^p)$  where

$$\boldsymbol{\mu}_k^p = \mathbf{A}_k \boldsymbol{\mu}_k^f, \quad (5)$$

$$\Sigma_k^p = \mathbf{A}_k \Sigma_k^f \mathbf{A}'_k + \Psi_k. \quad (6)$$

*Unscented Kalman Filter.* Generalising equations (1) and (2), we consider the non-linear system dynamics

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k; \mathbf{A}_k) + \boldsymbol{\epsilon}_k^x, \quad (7)$$

$$\mathbf{y}_k = g(\mathbf{x}_k; \mathbf{B}_k) + \boldsymbol{\epsilon}_k^y. \quad (8)$$

Here  $f$  and  $g$  may be non-linear functions of the state and parameters. An extension of the Kalman filter was developed based on the unscented transform [21], a method for calculating statistics of a random variable that has undergone a non-linear transformation. The idea behind the algorithm – called the Unscented Kalman Filter (UKF) – is to first draw sigma points, which are sample points drawn from the support of an underlying probability distribution. The sigma points are passed through a non-linear function, so that by the unscented transform the means and covariances of the predictive and filter distributions are estimable; indeed, they are calculable as functions of (weighted) sample means and (cross-)covariances of the transformed points.

In keeping with the approach of [22], so for  $N$  the state dimension, at initialisation we define vectors of weights  $\mathbf{w}^c, \mathbf{w}^m$  of size  $(2N + 1)$ , and select a set of sigma points denoted  $\mathcal{X}$ <sup>1</sup>. Each sigma point is

---

<sup>1</sup>There are various suggestions within the literature for how one might choose the weights and sigma points; we keep with the presentation of [22] for our implementation.

transformed on passing through  $f$ , yielding the set of transformed sigma points  $\mathcal{Y} = f(\mathcal{X})$ . Then the predictive distribution for  $\mathbf{x}_{k+1} | \mathbf{y}_{1:k}$  at time  $k$  can be shown to be  $\sim MVN(\boldsymbol{\mu}_k^p, \Sigma_k^p)$  where

$$\boldsymbol{\mu}_k^p = \sum_{i=0}^{2N} w_i^m \mathcal{Y}_{i,k}, \quad (9)$$

$$\Sigma_k^p = \sum_{i=0}^{2N} w_i^c (\mathcal{Y}_{i,k} - \boldsymbol{\mu}_k^p) (\mathcal{Y}_{i,k} - \boldsymbol{\mu}_k^p)^T + \Psi_k. \quad (10)$$

Next we calculate  $\mathcal{Z} := g(\mathcal{Y})$ , and  $\hat{\mathbf{y}}_k^- := \sum_{i=0}^{2N} w_i^m \mathcal{Z}_{i,k}$ . Then the filter distribution for  $\mathbf{x}_k | \mathbf{y}_{1:k}$  at time  $k$  can be shown to be  $\sim MVN(\boldsymbol{\mu}_k^f, \Sigma_k^f)$  where

$$\boldsymbol{\mu}_k^f = \boldsymbol{\mu}_k^p + \mathbf{G}_k (\mathbf{y}_k - \hat{\mathbf{y}}_k^-), \quad (11)$$

$$\Sigma_k^f = \Sigma_k^p - \mathbf{G}_k \mathbf{P}_{y,k} \mathbf{G}_k^T. \quad (12)$$

In analogy with the Kalman gain defined above, here we have  $\mathbf{G}$  given by:

$$\mathbf{P}_{y,k} := \sum_{i=0}^{2N} w_i^c (\mathcal{Z}_{i,k} - \hat{\mathbf{y}}_k^-) (\mathcal{Z}_{i,k} - \hat{\mathbf{y}}_k^-)^T, \quad (13)$$

$$\mathbf{P}_{xy,k} := \sum_{i=0}^{2N} w_i^c (\mathcal{Y}_{i,k} - \boldsymbol{\mu}_k^p) (\mathcal{Z}_{i,k} - \hat{\mathbf{y}}_k^-)^T, \quad (14)$$

$$\mathbf{G}_k := \mathbf{P}_{xy,k} \mathbf{P}_{y,k}^{-1}. \quad (15)$$

*A state space conditional factor model.* We consider a conditional factor model, as inspired by the Arbitrage Pricing Theory (APT) of Ross [23]. For a single-period investment setting, we assume the predictive model

$$\mathbf{r}_{k+1} = \mathbf{X}_k \mathbf{f}_k + \boldsymbol{\epsilon}_k^r, \quad \boldsymbol{\epsilon}_k^r \sim N(\mathbf{0}, \mathbf{D}_k). \quad (16)$$

Here  $\mathbf{r}_{k+1}$  is the  $n$ -dimensional random vector containing the cross-section of excess returns for  $n$  equities over the next time increment  $[k, k+1]$ ;  $\mathbf{X}_k$  is a non-random  $n \times m$  matrix of observable factor exposures estimated at  $k$ , for  $m$  the number of explanatory asset factors;  $\mathbf{D}_k$  is a covariance matrix for the returns, assumed to be diagonal; and  $\mathbf{f}_k$  is an  $m$ -dimensional latent factor vector that we estimate at each time step by cross-sectional regression.

We write the state space conditional factor model

$$\mathbf{f}_{k+1} = g(\mathbf{f}_k; \boldsymbol{\theta}_k^f) + \boldsymbol{\epsilon}_k^f, \quad (17)$$

$$\mathbf{r}_{k+1} = \mathbf{X}_k \mathbf{f}_k + \boldsymbol{\epsilon}_k^r, \quad (18)$$

assuming dynamics in the latent factor vector. The state equation implies the dynamic predictability of  $\mathbf{f}$  via some known function  $g$ . Recently, non-linear models for factor evolution have been explored in [24].

Of course, a popular modern choice for  $g$  might be one of the various machine learning-based time series models. Such models could depend on a recent history of the state variable, rather than just the current state. In the case of the factor model above, we can expand the state space to include lags in  $f$ . Consider the one dimensional case where

$$f_k = g(f_{k-1}, \dots, f_{k-M}; \mathbf{w}) + v_k.$$

The function  $g$ , for example, can be estimated by a time series model, with  $v_k$  being its aleatoric uncertainty estimate. For this single factor case, per [22], the state space model can be written:

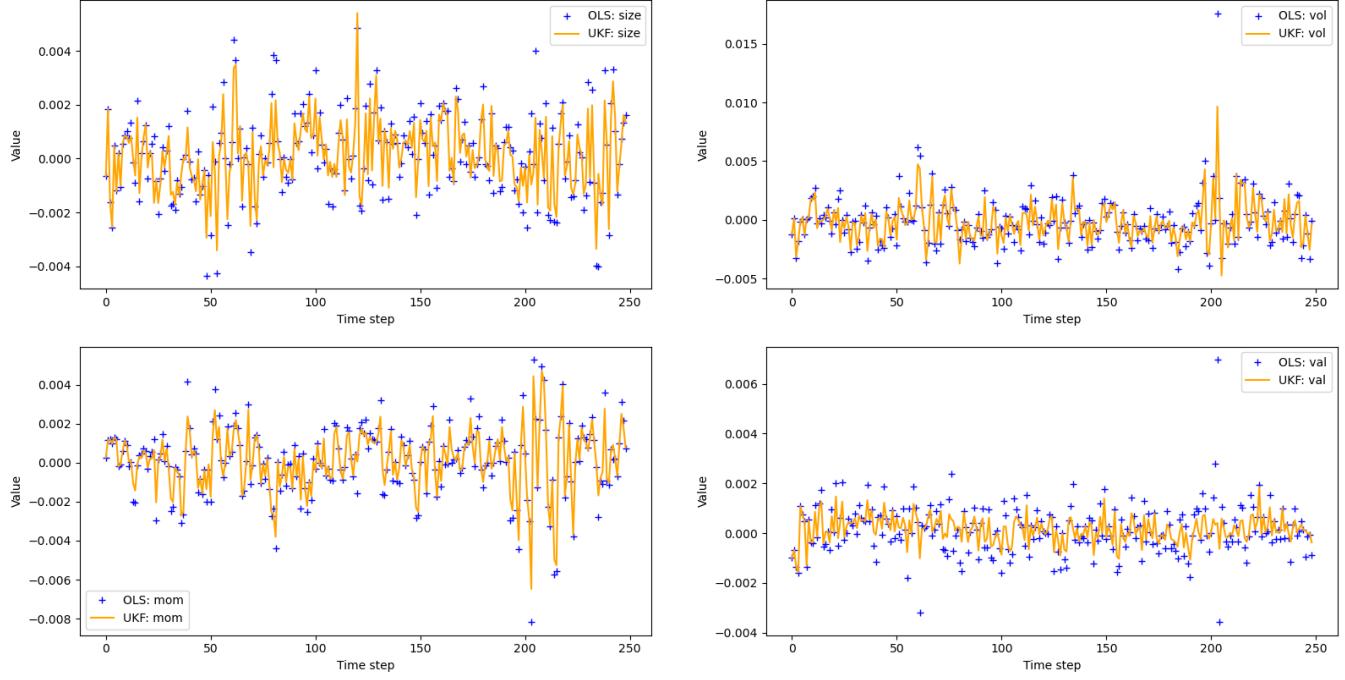


Figure 2: Factor risk premia as estimated by a conditional factor model using (i) ordinary least squares ('OLS'); and (ii) an underlying non-linear state space model ('UKF'). Four factors are displayed – size, volatility, momentum and value, for the year 1997 (which was arbitrarily chosen).

$$\begin{aligned}
f_{k+1} &= g(f_k; \mathbf{w}) + \mathbf{B}\epsilon_k^f, \\
\begin{bmatrix} f_{k+1} \\ f_k \\ \vdots \\ f_{k-M+2} \end{bmatrix} &= \left[ \begin{array}{c} g(f_k, \dots, f_{k-M+1}; \mathbf{w}) \\ \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} f_k \\ \vdots \\ f_{k-M+1} \end{bmatrix} \end{array} \right] + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} v_k, \\
r_{k+1} &= [1 \ 0 \ \dots \ 0] f_k + \epsilon_k^r.
\end{aligned}$$

We demonstrate an application of the model for four latent factors, assumed uncorrelated, showing a sample of model output in Figure 4 above. This compares with a predictive model for returns as a function of factors estimated by ordinary least squares. We provide further details of the model specification as part of the following section on experimental methods and strategy.

## 4 Experimental methods and strategy

Our statistical arbitrage scheme has three non-trivial components – data set construction and feature creation, data modelling, and specification of the trading strategy. Each of these components can have a significant bearing on performance results. In Section 3 we introduced the data model, and we present details of the

particulars here. We also address the other 2 components in this section, commenting on the data and the trading strategy.

*Data.* The data set construction is the same as used in [25], with the exception that the period is daily rather than bi-monthly. From that work, recall that the data is sourced from CRSP and IBES databases, with access granted via the Wharton Research Data Service. The raw data is daily US equity market data for the years 1992-2021 inclusive. On each trading day, we select the top  $N$  greatest stocks across all available US equities sorted by market capitalisation, for  $N = 2000$ . The data set is filtered for USD denominated common stock only – there are no closed end funds, REITs, ETFs, unit trusts, depository receipts, warrants etc. We also collected the S&P500 (excess) return time series, stock market capitalisations, and earnings per share, adjusted for splits.

In that we experiment with factor models, the data set requires the creation of our 5 chosen risk factors: market beta, size, volatility, momentum and value. Calculating size is relatively straightforward in CRSP: the calculation is the product of shares outstanding and the daily close price. Value was determined by analysing IBES data, and could be merged back to the CRSP database based on the CUSIP identifier. We note that risk factor construction is often non-trivial. Momentum was codable using CRSP data and the definitions of [26]; we calculated the daily compounded 12 month returns sans the most recent 1 month of data. We created the market Beta and volatility factors using the relevant CRSP data and the details of [27]. Indeed, over a two year rolling window we estimate a collection of linear functions by regressing each asset's daily excess returns against the S&P500 excess returns. The market Beta is the regression gradient, while the volatility factor is set to be the regression mean-square error.

*State space model specification.* The state space models described in Section 3 require the specification of the state transition function  $g$ , and the state and measurement covariances, denoted  $\mathbf{F}$  and  $\mathbf{D}$ . In the case of the linear model, we choose  $g$  to be the identity function<sup>2</sup>, and estimate a diagonal state covariance for  $f_{k+1} - g(f_k; \theta_k^f)$  based on a rolling (lookback) window of size 20. For the non-linear model, we estimate a multi-layer perceptron model for  $g$ , that takes the previous history of  $f$  on a window of size 10, and that predicts  $f$  at the next time-step, and well as the aleatoric uncertainty of  $f$ , which we set as an estimate of  $\mathbf{F}$ . To be clear, this estimate is a weighted average of the individual uncertainty estimates collected when calculating the state update equation for each sigma point. For reference, an image of the neural network architecture is given by Figure 3. This model is in the spirit of [28], and in this simplified setting we implement a multi-layer perceptron with 32 nodes followed by a drop out layer with rate 0.1, and L2 regularisation with parameter fixed at  $10^{-5}$  chosen by cross-validation on data from the earliest years. Finally, for both the linear and the non-linear model, our estimate for  $\mathbf{D}$  is a diagonal covariance gleaned from estimating Eq. (16) via ordinary least squares regression.

*Benchmark models.* We utilise a benchmark model with the trading strategy mechanics as outlined below, but with the key difference that the spread  $S$  is defined by the rolling 10 day sum total returns for each asset against the mean total return over a rolling 10 day window. This benchmark is discussed in [29], and is intuitive for considering reversals of returns. To be clear, this benchmark strategy is with reference to the total return series for each stock only; that is, it makes no reference to any underlying factors.

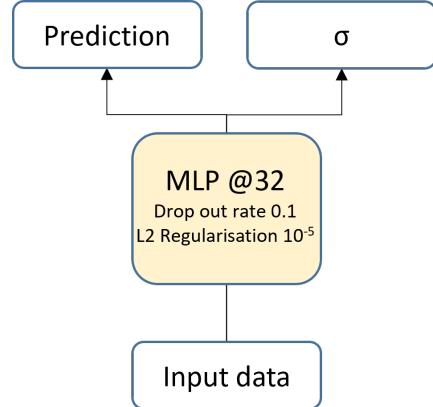


Figure 3: Suggestive neural network architecture for a non-linear latent factor vector state evolution model. The model outputs state prediction and aleatoric uncertainty estimates given an input historic time series.

<sup>2</sup>The choice of  $g$  is in analogy with the state space model underlying the pairs trading application of [16].

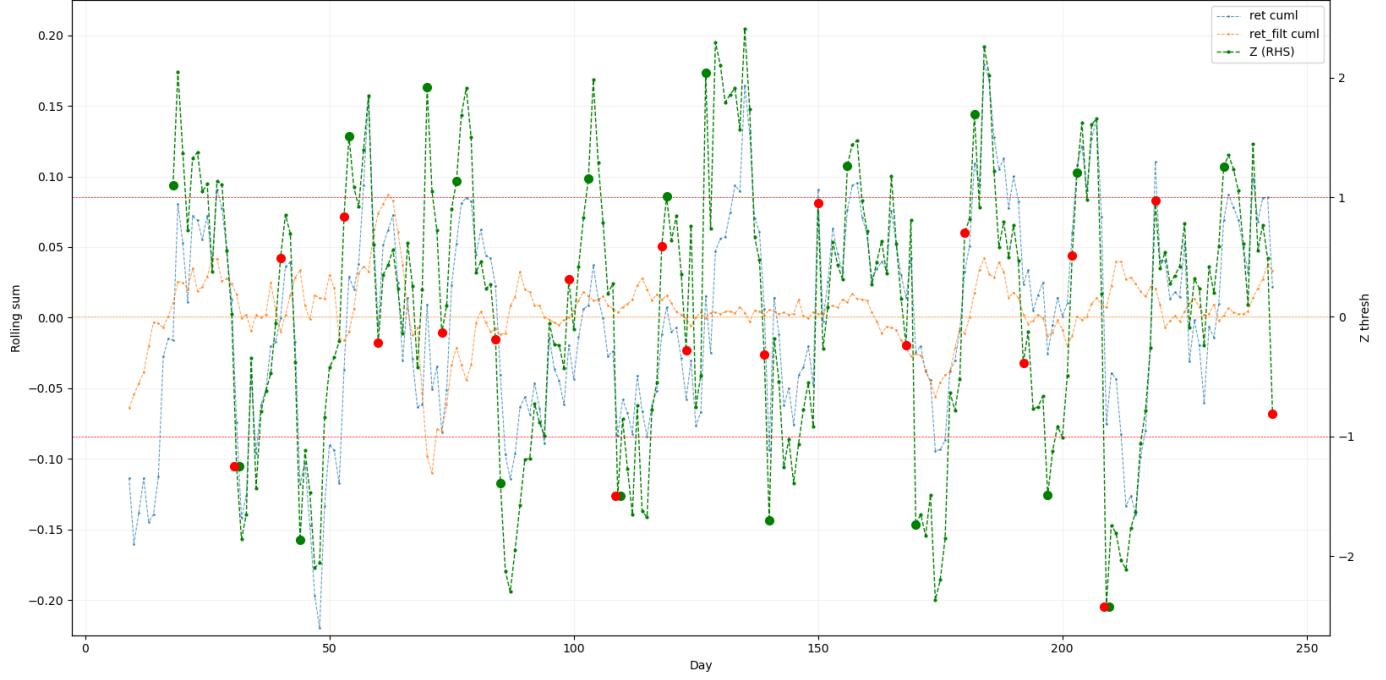


Figure 4: Trades are entered when the Z score deviation of the difference between returns and filtered returns passes a high (absolute) threshold; these points are indicated by the large green dots. Trades are closed at the first time the deviation has crossed the origin, as indicated by the large red dots.

*Trading strategy.* The state space model-inspired trading strategy is a long/short strategy, hedged to be beta neutral. The strategy depends on the spread defined by the rolling sum total returns in excess of the rolling sum filtered total returns – as estimated using the models of Section 3 – for each asset. That is, for each time  $k$ ,  $r_k$ ,  $X_k$  and  $X_{k-1}$  are observed, so the filter estimate for  $f_{k-1}$ , denoted by  $f_{k-1}^{(\text{filt})}$ , is calculable. Hence we can calculate  $r_k^{(\text{filt})} = X_k f_{k-1}^{(\text{filt})}$ . The spread, or (sum) error deviation,  $\sum(r_k - r_k^{(\text{filt})})$  is the target for our trading strategy, and we implicitly assume that it is mean-reverting.

Being arbitrary, the rolling window size for calculating the spread, denoted  $WS$ , is a target parameter for performance sensitivity analysis. Given a trading threshold  $L(S)$ , a trade is entered long (short) the first time the Z score of spread divergences is below (above) the negative (positive) values of  $L(S)$ . The trade is exited the first time the Z score reverts to, or through, the level 0, with trades ‘closed out’ on the last trading day of the year otherwise. The quantity  $L(S)$  should represent a sufficiently large deviation such that expected excess returns are not dominated by transaction costs. All position sizes long or short are set to a constant notional value. To assist intuition, the mechanics of the strategy are outlined diagrammatically in Figure 4 above, for the choice  $L/S = 1.0/1.0$ , for a target asset over the course of a year.

There are two other minor assumptions to the trading strategy that we state for completeness. Firstly, when entering a trade we assume that we can execute at the market close price. Secondly, at time  $t$ , if our database contains no close price for a stock at  $t + 1$ , we assume that time  $t$  is the last trading day for the stock, and that this is known to market. Hence, we close any existing position in the stock at the time  $t$  close price, and do not enter a new trade at  $t$  in the event that the trade entry criteria is met.

*Transaction costs.* In our primary experiments transaction costs are set to 5 basis points (bps) per trade entry and per trade exit, making 10 bps for a round-trip trade in a single underlying (be it a single stock or index). This is consistent with the works of [11, 3] and chosen for ease of comparison; these works are important benchmarks for multivariate statistical arbitrage in US equities. Further, we assume that we are able to borrow to short at all times, and that the transaction costs subsume any cost of borrowing. This deviates to some extent from reality, where not all stocks can be shorted all the time, and where some stocks incur a ‘hard to borrow’ fee, particularly when in high demand. Unless specified otherwise, all returns discussed for stocks and indices are total returns, and short positions pay the risk-free rate assuming funding is received at the time of shorting, and able to be invested in the riskless asset. We test the robustness of our strategies to alternative levels of transaction costs; this is discussed further in the Results section below.

*Hedging market risk.* We hedge our trading strategy to be beta neutral. Our simple approach is to calculate the market beta of our portfolio of long and short positions each trade period, and to maintain a position of dollar value  $P$  in the S&P500 index, where

$$P = -\beta_{port}\Pi,$$

thus providing the hedge. The term  $\Pi$  denotes the market dollar value of the portfolio to be hedged (and is defined further below). A positive value of  $P$  indicates a net long postion in the index as the hedge, while a negative value indicates a short position. The term  $\beta_{port}$  denotes the weighted-average beta of the portfolio given by

$$\beta_{port} = \frac{\boldsymbol{\beta} \mathbf{D}^T}{\Pi}.$$

The term  $\mathbf{D}$  is an  $N$ -vector of position sizes whose  $i^{th}$  entry, consistent with the trading strategy notes above, takes the value 1, -1 or 0 for a long, short or nil position in stock  $i$ . The term  $\boldsymbol{\beta}$  is an  $N$ -vector of corresponding stock beta estimates.

Implicit in the hedging strategy is the assumption that we can trade long or short positions in the S&P 500 index and earn or pay the total return of the index each period. In reality, the investor would take a position in the futures contract written on the index, which has a myriad of practical implications. Regardless, we do not consider these further, and so, for example, we ignore the spot / futures basis that could induce additional slippage.

*Strategy returns.* We assume a 100% leverage portfolio so that for each dollar invested in the portfolio, a position of up to a dollar’s worth long and / or a dollar’s worth short can be entered. Therefore, the market cost for  $l$  longs and  $s$  shorts is  $\Pi = \max(l, s)$ . We calculate (excess, unhedged) portfolio returns for a time period  $[k, k + dt]$  thus:

$$\Delta\Pi_k/\Pi_k = \frac{\sum_i(R_{i,k}^l - r_k dt) + \sum_j(R_{j,k}^s - r_k dt)}{\Pi_k} + \frac{r_k dt S_k}{\Pi_k} - \frac{TC \cdot \mathbf{1} \cdot (\Delta\mathbf{D}_k)^{|\cdot|}}{\Pi_k}.$$

Here  $R_{i,k}^l$  ( $R_{j,k}^s$ ) denotes the return over the period for a long (short) position held,  $r_k$  denotes the (annualised) risk-free rate of return,  $TC$  denotes the transaction cost, and  $\mathbf{X}^{|\cdot|}$  defines an operator returning a vector of elementwise absolute values of the input  $\mathbf{X}$ . Hence the terms on the right-hand side of the summation account for, respectively, the excess return of the long and short positions, the risk free rate earned on the proceeds of the short sales deposited in a money account, and the transaction costs for changes to the portfolio holdings<sup>3</sup>.

---

<sup>3</sup>The calculation is similar to that found in [3], and is equal when  $\Pi = s$ . Otherwise, our return is slightly more conservative for the case  $\Pi = l$ , though this has little impact on the final value calculated.

## 5 Results

In this section, we present performance statistics for the state space factor models of Section 3, relative to our benchmarks and the long-only market portfolio. For readers inclined, it is possible to replicate the results given the descriptions of the data and methodology as outlined in Section 4.

We find that the multivariate statistical arbitrage strategy has shown reasonable performance through time, which is consistent with the expectations for a mid-frequency strategy with reasonable capacity. Performance is best during the first half of our 29 year sample, and we conjecture that the impressive returns have diminished due to increased market competition and efficiency through time. Also, we find sufficient differences between strategy performance that supports the use of a state space conditional factor model for multi-variate statistical arbitrage, instead of other less complex - albeit relatively powerful - models. Despite all strategies showing relative underperformance in the most recent years, practitioners may find latitude for improvement by further modifying or optimizing components of the strategy; we briefly describe some preliminary ideas for research direction at the end of this section.

### 5.1 Strategy performance

*The earlier years: 1993–2007.* Given the data set and assumed fair value models, per the methods described in Section 4 there is an implied sensitivity of any estimated performance statistics to the key assumptions underlying the trading strategy. Not the least, this includes the setting of parameters denoting (i) the trade entry threshold Z score; (ii) the size of the returns window; and (iii) the level of transaction costs. We set the parameters of (i) and (ii) based on strategy performance over the ‘earlier years’ of 1993–2007. We assume a transaction cost level of 5 bps for each trade in all results that follow, unless otherwise specified. Choosing the parameters in this way also allows for a reasonable performance comparison between our strategy relative to results given by other well known academic works. Finally, the parameter choices are maintained when we assess strategy aggregate performance, and performance conditioned on later years only.

For various combinations of Z score-based entry thresholds for the long and short legs (denoted L and S), and various window sizes (denoted WS), we show average Sharpe ratio and maximum drawdown statistics in Figure 5, on the next page. Drawdowns are calculated on a 252 day rolling window. We find utility in solving for the trade entry parameters for the long and short legs separately, making no assumption that the same cut off should be optimal for both. We test the entry level to balance entering a sufficient number of trades, but not entering too early (or late) such that the trade is not held too long (or missed entirely). As expected, we find firstly that as the threshold decreases, the number of trades increases to the extent that small thresholds have bad economics owing to large turnover and cumulative transaction costs. Window size is similarly tested, for sizes of 1, 5 and 10, corresponding to daily, weekly and fortnightly rolling windows.

Per Figure 5, we notice two clear clusters of data, whereby (i) one exhibits the highest average Sharpe ratio, but also exhibits amongst the worst for maximum drawdowns; and (ii) the other is a cluster with good average Sharpe ratio, but amongst the strategies with the better of the maximum drawdowns. It is reasonable to suggest that different investors might have different preferences for either cluster, so we extend our study for recent years with respect to them. The first cluster is characterised by models with a window size of 5, and long/short entry thresholds in  $\{0.5, 1.0, 1.5, 2.0 : L < S\}$ . The second cluster is the same except that the window size is 10. For both clusters, the state space models tend to dominate both of the benchmark models with respect to both statistics of interest. However, the performance between the state space models is less noticeable, though is such that the KF tends to have the higher Sharpe, for worse drawdowns, relative to the UKF, and for all choices of cutoff. Regardless of an investors preference, these differences are relatively small for this experiment.

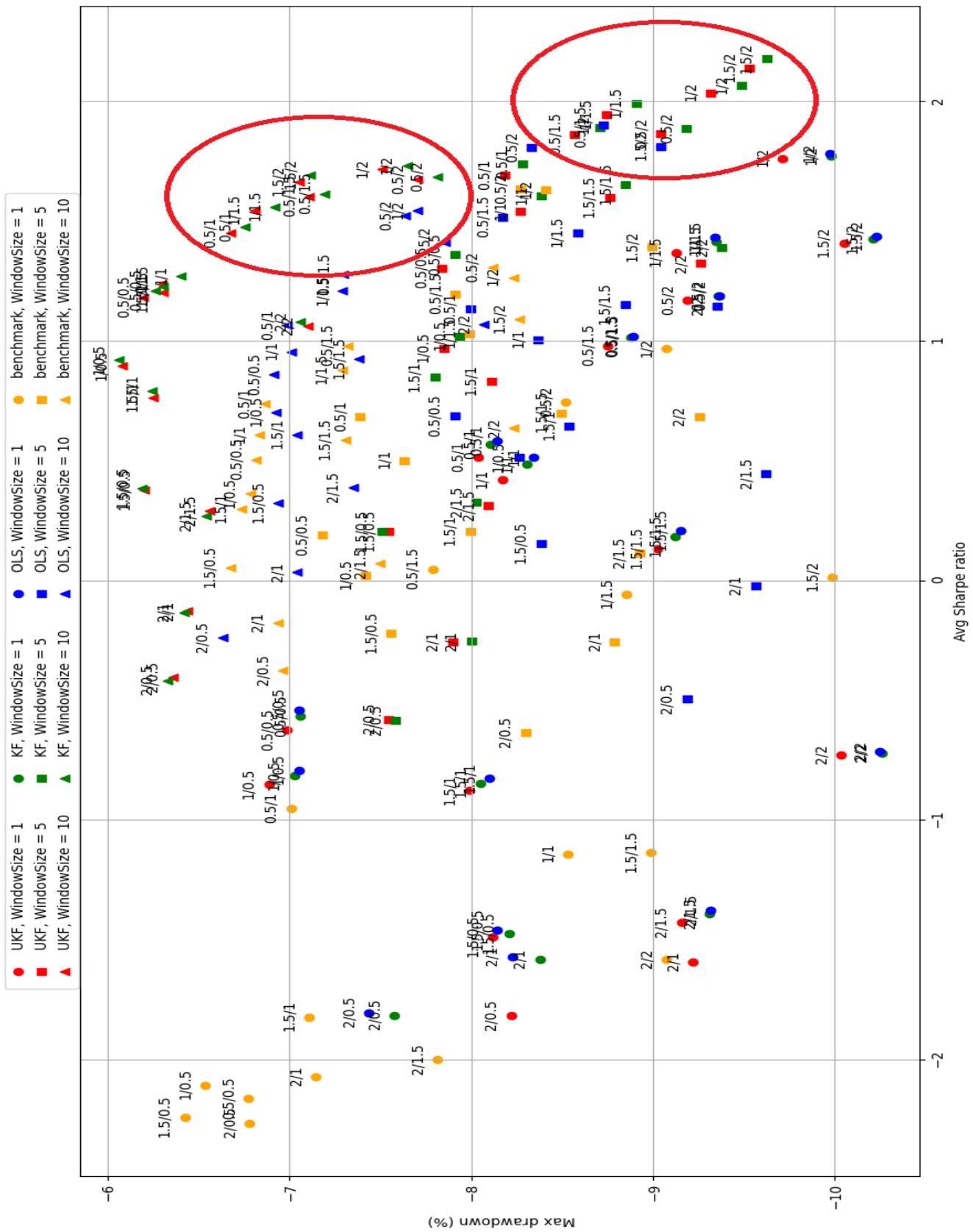


Figure 5: Average Sharpe ratio over the 15 years from 1993-2007 versus maximum drawdown for a collection of models underlying the trading strategy. 11

Strategy	Date range	Avg ann Sharpe [3, 11]	Sharpe Ours	TR SPX	Notes (Models for cited work)
Ours	1993-2021	-	<b>1.29</b>	0.87	-
[3]	1997-2007	1.44	<b>1.61</b>	0.47	PCA based
	2003-2007	0.90	<b>1.01</b>	0.78	PCA based
	1997-2007	1.10	<b>1.61</b>	0.47	ETF based
	2003-2007	<b>1.51</b>	1.01	0.78	ETF based – incorporating additional volume feature
[11]	1992-2015	0.75-1.12	<b>1.67*</b>	0.73*	Vine based. Range across 4 alt. trade selection methods.

Figure 6: Tabulated average annual Sharpe ratios by date range, for our conditional factor model (hedged UKF, L/S = 1.5/2.0, WS=5) and two leading external methods. Note I: All methods assume transaction costs of 5bps per trade. Note II: \* results mean ‘excluding 1992 due to lack of data’.

One also notices from Figure 5 that the strategies that have a lower entry threshold for the longs, and relatively larger entry threshold for the shorts, tend to exhibit the better performance statistics. It is the case that the longs outperform the shorts for this data sample<sup>4</sup>, with the shorts also more likely to have a negative year of performance. The effect of increasing an entry threshold is to reduce the number of trades entered on that long or short leg; this reduces relatively positive or negative returns, hence the observation. We leave the question of why the longs may more often outperform the shorts as future work, and include no assumption as to what leg may outperform for any period within our backtest; this could possibly be a source of improvement of the strategy alpha.

Our results can be reasonably compared with other well-known results from the literature. Our approach has a similar data set to the others (daily US equities data, albeit with different universe sizes), but most notably different models for fair-value – in [3] factors are modelled based on Principal Component Analysis, with an additional ETF-based model, and in [11] asset relationships are modelled with Vine copulas. Our trading strategy is similar to [3], though we do not refine a trade exit threshold; we have less in common with [11], who refine their trading rules with many more steps relative to our approach. Importantly, across all papers, the transaction cost assumption is equal. For the papers cited, it is easiest to compare average annual Sharpe ratios over the date ranges given in the cited paper. We tabulate the results of this comparison in Figure 6 above, and also include the total (excess) return of the S&P 500 over the same period. Our results are similar in magnitude to [3], and we conjecture that daily US equity mean reversion could be captured by strategies using a variety of underlying fair value models over the considered date range. Superficially, there appears to be some outperformance in our approach, though this claim would benefit from comparison given outcomes of a controlled experiment. On the other hand, though the average Sharpe ratios appear similar, our methods do show sufficient variation in Sharpe ratio year-on-year. For example, the PCA based method shows severe performance degradation, and negative Sharpes, in later years. On the other hand, our approach had no negative years between 1997 and 2007. The outperformance of our approach is more striking against [11], and with many degrees of freedom in the formulation of the copula-based strategy it is hard to conjecture exactly why this may be the case.

*Aggregate performance.* In Figure 7 below, we depict the annual Sharpe ratio distribution over 29 years by way of box and whisker plots, for the 4 underlying models. Further, we show the performance across 4 sets of parameter choices, which include window sizes of 5 and 10, and long/short entry thresholds of 0.5/2.0 and 1.5/2.0. We also show a plot for the Sharpe ratio of total (excess) return of the S&P 500.

<sup>4</sup>Recently, other authors have analysed long-short equity factor portfolios, albeit in an alternate strategy setting, and made the same outperformance observation [30].

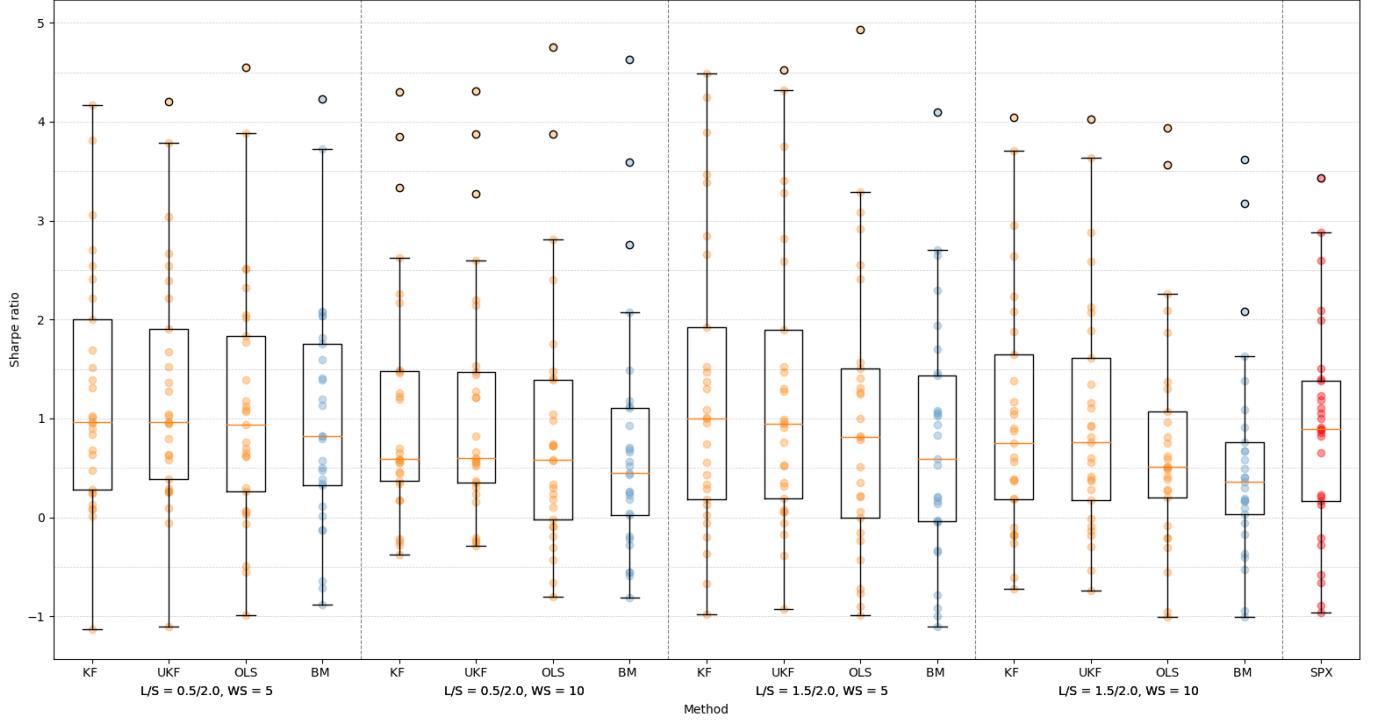


Figure 7: Sharpe ratio distribution over 29 years for 4 models underlying the trading strategy outlined in Section 3, grouped by long/short entry Z score criteria, and lookback window size for cumulative returns. The performance of the S&P 500 with respect to total (excess) return is also included for reference.

Modelwise, the KF and the UKF outperform relative to OLS and BM across all but one groupwise comparisons of 1st, 2nd and 3rd quantiles, and averages, of annual Sharpe ratios. On the other hand, any UKF under/outperformance over the KF is not apparent, so that the additional computational complexity of the UKF is not justified for this particular application. If stronger non-linear relationships between variables were observed, or expected, given additional training data, the UKF may still find use. The OLS approach induces good results, with the (U)KF appearing to be a worthwhile refinement, that is, the introduction of uncertainty quantification has a benefit. The nature of the outperformance appears to vary depending on the parameter settings of the trading strategy.

In Figure 8, we show the results of an exploration of robustness of our strategy to the transaction cost assumption. As mentioned, we implemented our strategies assuming a 5 bps (10 bps round-trip) transaction cost for each asset traded. We show Sharpe ratio percentile levels for transaction costs of 0, 5, 10 and 15 bps for the UKF 0.5/2.0 across window sizes. The yellow highlighted data are for the 5 bps transaction cost level. Clearly this is a sweet spot – and while the quoted performance can sustain a slightly higher

		Transaction costs (bps)			
		0	5	10	15
WS 5	25th	1.11	0.39	-0.08	-0.76
	50th	1.39	0.99	0.57	-0.09
	75th	2.75	1.91	1.14	0.70
WS 10	25th	0.80	0.35	-0.10	-0.66
	50th	1.09	0.59	0.26	-0.17
	75th	1.90	1.47	1.05	0.83

Figure 8: Sharpe ratio percentiles by transaction cost and window size levels for UKF 0.5/2.0.



Figure 9: Top: Percentage of the stock universe ( $N=2000$ ) invested in, long or short, at a given time. Bottom: Strategy rolling window (252 day) drawdown versus the long only market portfolio; transaction cost = 5bps.

transaction cost assumption, there is an obvious performance decay such that at 15 bps of transaction costs, the median annual Sharpe ratio over the sample realises negative. Finally, we should mention that the cost of short selling may have been significantly more than estimated in this work, if the capacity to borrow to sell was even available, particularly so for the earlier years of study. On the other hand, practitioners may have strategies for cheapening transaction costs when trading. Given these latter points, and other cited academic studies, we think our primary transaction cost assumption is reasonable.

The bottom panel of Figure 9 shows rolling 252 day drawdown for 2 UKF based strategies differing on window size. In the years up till 2007 the strategies largely avoid the worst of the drawdowns realised by the market portfolio, which has drawdowns touching -10% or worse on 3 occasions. In these cases, our strategies realise drawdowns between -7 and -5 %, with the exception of the strategy with window size of 5 that realises a -9.5% drawdown on one of the three occasions. Post-2007, the two severe market drawdowns of 2008 and 2020, both being worse than -20%, are matched by strategy drawdowns about -16 and -12%, respectively. The first of these drawdowns, within the years of the Great Financial Crisis, was to the chagrin of many real-world strategies, such that we find it less of a concern that our strategy did not register greater outperformance during this period. For intuition, we also depict the proportion of assets under consideration that are invested in for all times, in the top panel of Figure 9. We notice the obvious empirical fact that with a larger window size, more assets are held. There is an annual seasonality apparent, whereby at the end of each year positions are closed out, and at the start of the year there are spikes for new trades re-entered, as expected by the definition of our trading strategy.

It is reasonable to assume that the beta neutral statistical arbitrage strategy has zero or low (absolute) correlation of returns to the long only market portfolio. Hence, in theory, a blended investment taking a position in the strategy and the long only portfolio could show an outperformance (with respect to Sharpe ratio)

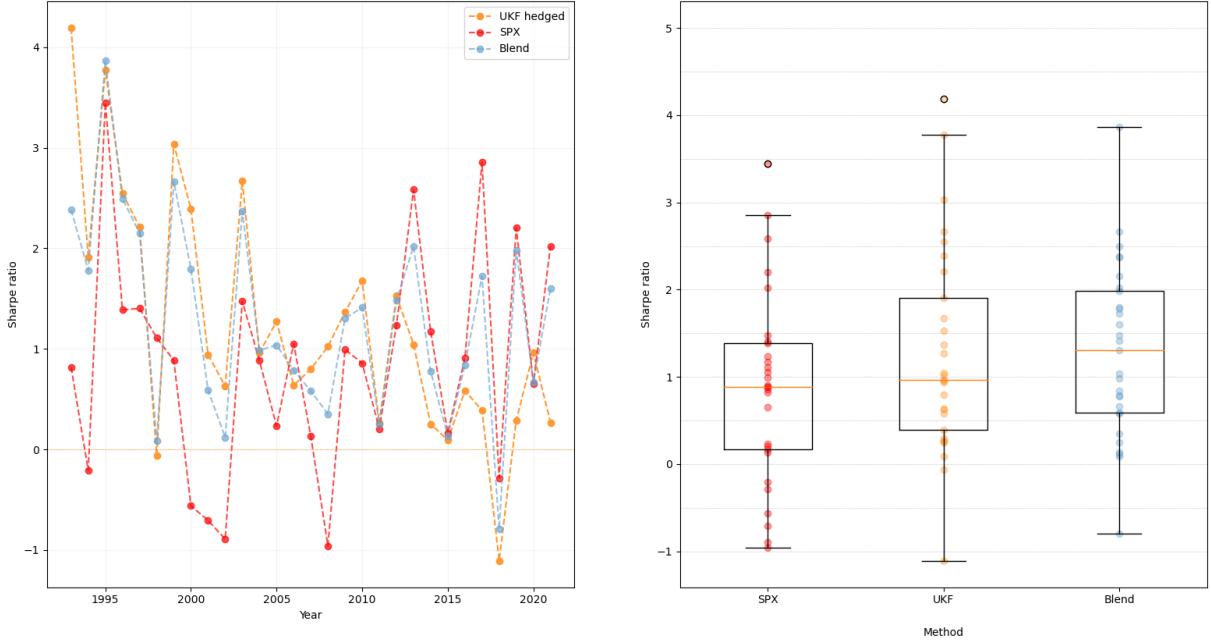


Figure 10: Sharpe ratio distribution over 29 years for a UKF model-based trading strategy as outlined in Section 3, the S&P 500 with respect to total (excess) return, and a portfolio of a weighted combination of the two.

relative to both [31]. We implement such a blended strategy by weighting the statistical arbitrage strategy and the long only portfolio as

$$w_i = \frac{S_i}{S_L + S_{LS}}, \quad S_i := \frac{\mu_i}{\sigma_i^2}, \quad i \in \{L, LS\},$$

where the weights  $w$  depend on the mean  $\mu$  and variance  $\sigma^2$  of the long portfolio  $L$  or long-short portfolio  $LS$  returns, optimizing for maximum Sharpe ratio [31, 28]. The true values of the mean and variance parameters are unknown, and for demonstration purposes we take the straightforward approach of setting the means and variances equal in the first year, and equal to their historical values on a rolling window of size (up to) 10 otherwise. We find that the results are not particularly sensitive to (reasonable) values of the choice of rolling window size. Images summarising the Sharpe ratio performance are given in Figure 10. Of note, given the right-hand side chart, we notice considerable performance improvement with respect to the blended strategy for the first, second and third quartiles, supporting the theoretical result, and offering a practical hint for deploying such a strategy.

*Recent performance:* 2008–2022. We notice from the left-hand side image of Figure 10 that the strategy has not performed so well in the most recent years, particularly for the years 2015–2019. We conjecture that this could be due to a lower volatility market regime during this time, increased sophistication of market participants and with it competition for alpha / eroded edge, or otherwise due to limitations derived from some underlying assumption of our strategy. Overcoming these challenges is likely of interest to the motivated academic or practitioner, and we offer some avenues for future work in the following conclusion.

## 6 Conclusion

We study multivariate statistical arbitrage under a conditional factor model augmented in a state space framework for US equities trading. Our study shows the modelling approach can yield compelling performance statistics over a 29 year period, an absolute basis and relative to a reasonable benchmark, and relative to a model estimating returns as a function of factors via ordinary least squares. However, for our experimental set up, we did not find that the non-linear state space model was justified relative to the linear case; though it was no worse, any outperformance benefit does not contrast well with the increased modelling complexity introduced. When we compare our results to existing attempts in the literature, we find that our approach compares well. We also show empirical evidence that blending our investment capital between the long-short strategy and the long-only market portfolio as a function of each strategies estimated mean and variance yields a strategy with improved annual Sharpe ratio, consistent with theory. All results are with respect to reasonable level of transaction costs.

There is the scope for academic and practitioner work on the strategy presented in this paper, particularly with respect to exploring the potential for improvement on the period of strategy underperformance in recent years. We conclude with some final remarks:

- Regarding the data underlying the quantitative strategy, the addition of alternative data beyond factor data and returns time series may improve the predictive power of the model. With respect to the dataset we constructed, and by the recent work of [32], it is the case that estimating submodels over clusters of data – for example, stocks grouped by sector – could lead to significant performance improvement.
- The recent work of [24, 33] consider alternative specifications of dynamic factor models for investing, including those with non-linearities at the level of the measurement equation. Such expressions could be handled by our UKF-based framework, and present as a reasonable target for improving performance.
- Recall our brief discussion on related literature in Section 2, whereby almost all work cited made use of data on a daily time scale. The literature would be bolstered by analysis completed on alternative time frames, particularly on intraday time scales. This is arguably harder, not the least in terms of data set curation. On the other hand, the recent work of [34] provides evidence of intraday return predictability as a function of factors using machine learning; this could support refining reactive entry and exit rules within the trading strategy.
- Modelling techniques subsuming regime filters could support our statistical arbitrage trading strategy, with recent evidence to this end for a pairs trading application given in [35].
- Finally, and requiring more expert / domain knowledge, it would be interesting to understand trades entered that are directly mappable to a market phenomena, such as a stock going ex-dividend, or being influenced by index rebalancing. This is interesting in the context of determining whether the quantitative strategy’s profitable trades are explainable in terms of a well-known market effect, or else is supportive of the underlying factor-based theory.

## Acknowledgements

The authors would like to thank the Oxford-Man Institute of Quantitative Finance for its generous support. SR would like to further thank the Royal Academy of Engineering.

## References

- [1] Ed Thorp. Statistical Arbitrage – Part II. *Wilmott*, pages 48–49, 11 2004.
- [2] Christopher Krauss. Statistical Arbitrage Pairs Trading Strategies: Review and Outlook. *Journal of Economic Surveys*, 31(2):513–545, 2017.
- [3] Marco Avellaneda and Jeong-Hyun Lee. Statistical arbitrage in the US equities market. *Quantitative Finance*, 10(7):761–782, 2010.
- [4] Sergio M. Focardi, Frank J. Fabozzi, and Ivan K. Mitov. A new approach to statistical arbitrage: Strategies based on dynamic factor models of prices and their performance. *Journal of Banking & Finance*, 65:134–155, 2016.
- [5] Simão Sarmento and Nuno Horta. Enhancing a Pairs Trading strategy with the application of Machine Learning. *Expert Systems with Applications*, 158, 5 2020.
- [6] Federico Gatta, Carmela Iorio, Diletta Chiaro, Fabio Giampaolo, and Salvatore Cuomo. Statistical arbitrage in the stock markets by the means of multiple time horizons clustering. *Neural Computing and Applications*, 2023.
- [7] Ganapathy Vidyamurthy. *Pairs Trading: Quantitative Methods and Analysis*. John Wiley & Sons, Hoboken, N.J., 2004.
- [8] Jorge Guijarro-Ordonez. High-dimensional Statistical Arbitrage with Factor Models and Stochastic Control. *Applied Mathematical Finance*, 26(4):328–358, 2019.
- [9] Ziping Zhao and Daniel P. Palomar. Mean-Reverting Portfolio With Budget Constraint. *IEEE Transactions on Signal Processing*, 66(9):2342–2357, 2018.
- [10] Ziping Zhao, Rui Zhou, and Daniel P. Palomar. Optimal Mean-Reverting Portfolio With Leverage Constraint for Statistical Arbitrage in Finance. *IEEE Transactions on Signal Processing*, 67(7):1681–1695, 2019.
- [11] Johannes Stübinger, Benedikt Mangold, and Christopher Krauss. Statistical arbitrage with vine copulas. *Quantitative Finance*, 18(11):1831–1849, 2018.
- [12] Delphine Lautier, Alireza Javaheri, and Alain Galli. Filtering in Finance. *Wilmott*, pages 2–18, 05 2003.
- [13] Paresh Date and Ksenia Ponomareva. Linear and non-linear filtering in mathematical finance: a review. *IMA Journal of Management Mathematics*, 22(3):195–211, 2011.
- [14] Robert Elliott, John Van Der Hoek, and William Malcolm. Pairs trading. *Quantitative Finance*, 5(3):271–276, 2005.
- [15] Kostas Triantafyllopoulos and Giovanni Montana. Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science*, 8:23–49, 2011.
- [16] Theodoros Tsagaris. *Adaptive Regression Methods with Application to Streaming Financial Data*. PhD thesis, Imperial College, 2010.
- [17] Carlos Eduardo de Moura, Adrian Pizzinga, and Jorge Zubelli. A pairs trading strategy based on linear state space models and the Kalman filter. *Quantitative Finance*, 16(10):1559–1573, 2016.
- [18] Guang Zhang. Pairs trading with general state space models. *Quantitative Finance*, 21(9):1567–1587, 2021.
- [19] Yufeng Han. Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model. *Review of Financial Studies*, 19(1):237–271, 2006.
- [20] Siddhartha Chib, Federico Nardari, and Neil Shephard. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371, 2006.
- [21] Simon Julier and Jeffrey Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92: 401 – 422, 04 2004.
- [22] Eric A. Wan and Rudolph van der Merwe. The unscented kalman filter for nonlinear estimation. In *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, pages 153–158. IEEE, 2000.
- [23] Stephen A. Ross. The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, 13(3):341–360, 1976.
- [24] Shihao Gu, Bryan Kelly, and Dacheng Xiu. Autoencoder Asset Pricing Models. *Journal of Econometrics*, 222 (1, Part B):429–450, 2021.
- [25] Trent Spears, Stefan Zohren, and Stephen Roberts. View Fusion Vis-à-Vis a Bayesian Interpretation of Black–Litterman for Portfolio Allocation. *The Journal of Financial Data Science*, 5(3):23–49, 2023.
- [26] Clifford S. Asness, Tobias J. Moskowitz, and Lasse Heje Pedersen. Value and Momentum Everywhere. *The Journal of Finance*, 68(3):929–985, 2013.

- [27] Petter N. Kolm and Gordon Ritter. On the Bayesian interpretation of Black–Litterman. *European Journal of Operational Research*, 258(2):564–572, 2017.
- [28] Trent Spears, Stefan Zohren, and Stephen Roberts. Investment Sizing with Deep Learning Prediction Uncertainties for High-Frequency Eurodollar Futures Trading. *The Journal of Financial Data Science*, 3(1):57–73, 2021.
- [29] Andrew W. Lo and A. Craig MacKinlay. When are Contrarian Profits Due to Stock Market Overreaction? *The Review of Financial Studies*, 3(2):175–205, 1990.
- [30] David Blitz, Guido Baltussen, and Pim van Vliet. When Equity Factors Drop Their Shorts. *Financial Analysts Journal*, 76(4):73–99, 2020.
- [31] John S. Brush. Comparisons and Combinations of Long and Long/Short Strategies. *Financial Analysts Journal*, 53(3):81–89, 1997.
- [32] Clint Howard. Less is More? Reducing Biases and Overfitting in Machine Learning Return Predictions. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4497739](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4497739), 2023. [Online; last accessed 1-Aug-2023].
- [33] Paolo Andreini, Cosimo Izzo, and Giovanni Ricco. Deep Dynamic Factor Models. In *Working Papers 2023-08*. Center for Research in Economics and Statistics, 2023. URL <https://ideas.repec.org/p/crs/wpaper/2023-08.html>. [Online; last accessed 1-Aug-2023].
- [34] Saketh Aleti, Tim Bollerslev, and Mathias Siggaard. Intraday Market Return Predictability Culled from the Factor Zoo. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4388560](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4388560), 2023. [Online; last accessed 1-Aug-2023].
- [35] Robert J. Elliott and Reza Bradrania. Estimating a regime switching pairs trading model. *Quantitative Finance*, 18(5):877–883, 2018.