



Heriot-Watt University
Research Gateway

A Non-parametric Test and Predictive Model for Signed Path Dependence

Citation for published version:

Dias, FS & Peters, GW 2020, 'A Non-parametric Test and Predictive Model for Signed Path Dependence', *Computational Economics*, vol. 56, pp. 461–498. <https://doi.org/10.1007/s10614-019-09934-7>

Digital Object Identifier (DOI):

[10.1007/s10614-019-09934-7](https://doi.org/10.1007/s10614-019-09934-7)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Computational Economics

Publisher Rights Statement:

© The Author(s) 2019

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



A Non-parametric Test and Predictive Model for Signed Path Dependence

Fabio S. Dias¹  · Gareth W. Peters²

Accepted: 3 October 2019
© The Author(s) 2019

Abstract

While several tests for serial correlation in financial markets have been proposed and applied successfully in the literature, such tests provide rather limited information to construct predictive econometric models. This manuscript addresses this gap by providing a model-free definition of signed path dependence based on how the sign of cumulative innovations for a given lookback horizon correlates with the future cumulative innovations for a given forecast horizon. Such concept is then theoretically validated on well-known time series model classes and used to build a predictive econometric model for future market returns, which is applied to empirical forecasting by means of a profit-seeking trading strategy. The empirical experiment revealed strong evidence of serial correlation of unknown form in equity markets, being statistically significant and economically significant even in the presence of trading costs. Moreover, in equity markets, given a forecast horizon of one day, the forecasting strategy detected the strongest evidence of signed path dependence; however, even for longer forecast horizons such as 1 week or 1 month the strategy still detected such evidence albeit to a lesser extent. Currency markets also presented statistically significant serial dependence across some pairs, though not economically significant under the trading formulation presented.

Keywords Empirical asset pricing · Serial correlation · Time series momentum · Econometric forecasting · Quantitative investment strategies

The authors gratefully acknowledge the helpful comments of two anonymous referees, the support given to FSD by Lloyds Banking Group and the support given to GWP by ARC Grant Proposal IDDP180100597 as a Partner Investigator. *The views expressed here are solely those of the authors and do not represent the views of Lloyds Banking Group.*

✉ Fabio S. Dias
fabio.dias@stalwart.it

Gareth W. Peters
garethpeters78@gmail.com

¹ Department of Statistical Science, University College London, London, UK

² Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, UK

1 Introduction

The arrival of new information into asset price formation has been subject of extensive discussion. If such information is always fully reflected and incorporated into market prices, asset returns should be largely unpredictable and in the long run driven by compensation for taking market risk. Such a perspective on the driver for returns on investment is most directly justified in settings in which there exists a competitive double auction market with informed rational market participants that will not make irrational errors on a systematic basis, and instead will exploit any known informational inefficiency until its exhaustion. This is typically formulated in financial mathematics pricing theory through concepts such as the Efficient Market Hypothesis (EMH), von Neumann Morgenstern rationality and other related financial assumptions on markets and agent behaviours as captured in, for instance, the technical discussions on such assumptions in Chapters 9 and 14 of Platen and Heath (2006).

Initial work on the EMH focussed heavily on statistical properties of stock prices (Fama 1963, 1965a, b), arguing that in strongly efficient markets asset prices fluctuate in unpredictable and completely random ways. The seminal work in Fama (1970) has subsequently provided the definition of three forms of financial market efficiency: weak, where information contained in past prices is fully reflected in current prices; semi-strong, where all public information available (including past prices) is fully reflected in current prices; and strong, where all public and private information available is fully reflected in current prices. Such school of thought has evolved considerably and several different tests have been developed; however, it is still common practice to focus on the serial dependence of asset price changes to test for market efficiency, with a notable recent example seen in Urquhart and McGroarty (2016). These tests look for violations of the weak form of financial market efficiency that, if clearly present, could be used as evidence that the EMH does not hold in any of its forms.

Results from tests on serial dependency have been mixed. Nevertheless, the fact that a specific test does not detect serial dependence in a given price series does not guarantee in itself that serial dependence is not present: it is only evidence that a particular form of dependence is not present, but there might still be another form of serial dependence that was undetected due to test misspecification. This motivates a continuous research on finding potential new models that can cope with such alleged dependence and, ideally, use the detected dependence to make informed predictions of future market behaviour.

It is a stylized fact that asset returns do not exhibit linear serial correlation (Cont 2001); however, a number of different tests have been proposed in the literature which have detected statistically significant evidence that asset returns exhibit some form of serial correlation which is not necessarily linear. Section 3 of Lim and Brooks (2011) provides an extensive review of the most well-known methods. One of the most popular tests is the variance ratio test (VR), which postulates that, under the absence of serial correlation, the variance of the k -period return should be equal to k times the variance of the one-period return; therefore, the ratio of such variances

should not be statistically different from one under the null hypothesis of absence of serial correlations. Nevertheless, this has been found not to hold across several global equity markets (Griffin et al. 2010).

Other well-known tests aiming to reject the hypothesis of the absence of serial correlations in financial markets include the automatic Box–Pierce test in Lim et al. (2013), long-memory tests in Chung (1996), Hurst Exponent tests in Qian and Rasheed (2004) and tests on the frequency domain that can be found in Lee and Hong (2001). However, despite several tests demonstrating the presence of serial correlation in financial markets and the potential for predictability, there aren't many studies in the literature developing actual predictive models and validating them in empirical financial data.

Notably, Moskowitz et al. (2012) defines a linear regression of the return of an asset scaled by its ex-ante volatility against the sign of its return lagged by some amount of time and shows empirically there is strong evidence of predictability when a lagged 1-year return is used to predict the return 1 month ahead. Similar studies of this nature can also be found in Baltas and Kosowski (2013) where the authors establish relationships between univariate trend-following strategies in assets in futures markets and commodity trading advisors in order to examine questions of capacity constraints in trend following investing. Additionally, Corte et al. (2016) found evidence of negative dependence in daily returns of international stock, equity index, interest rate, commodity and currency markets, with such negative dependence becoming even stronger when the returns are decomposed into overnight (markets closed for trading) and intraday (markets open for trading). These studies though lack a stronger theoretical basis, being instead more focussed on the empirical aspects.

This paper provides theoretical contributions to the literature by proposing in Sect. 2 a non-parametric definition of signed path dependence and demonstrating the sufficient and necessary conditions for it to be present in covariance stationary time series. Further, Sect. 3 proposes a formal inference procedure to detect serial correlation of unknown form based on a hypothesis testing formulation of signed path dependence, which is validated on experiments on synthetic data in Sect. 4. This paper provides empirical contributions to the literature in Sect. 5 by using the test previously defined to detect evidence of serial correlation in a number of equity index and foreign exchange markets. Additionally, based on the test statistic proposed, a predictive model is defined and used to feed trading strategies whose out-of-sample performance is analysed, gross and net of transaction costs. Section 6 concludes the work. The proofs for all theorems and lemmas given in this paper can be found in the Technical Appendix (other than for commonly known theorems, which are just enunciated for the sake of completeness and clarity).

It is noted that this paper uses cross-validation and bootstrap techniques to calibrate and test the predictive model proposed. A comprehensive review of these techniques can be seen in Kohavi (1995). Numerous studies have used similar techniques to extract serial dependence and predict future behaviour in market prices: Qian and Rasheed (2007), Choudhry and Garg (2008) and Huang et al. (2005) to name a few. This paper contributes to such area of research by developing its own classifiers which are bespoke to markets that have signed path dependence and hence will have better classification performance than models such as Support Vector Machines, Genetic

Algorithms for trading rules and others, given that these aim to be more generic and hence less sensitive to peculiar market features. After all, it is expected that a test whose null hypothesis is more specific to a particular problem in financial time series is likely to have more power and more accurate estimation properties than a test based on wider machine learning tools that will consequently imply a wider scope for its equivalent null hypothesis.

This paper focuses solely on both how to detect serial correlation in a formal statistical manner and how to model it, paving the way to build predictive models that can be used for the purposes of general market forecasting/risk management or even for the development of trading strategies. As such, technical discussions around assumptions on the price discovery process which might or might not generate price predictability are left outside of the scope of this paper. We also remark that some studies claim that apparent trend-following profits might actually not be arising out of positive serial correlation in returns but possibly out of intermediate horizon price performance (Novy-Marx 2012) or exogenous factors such as the presence of informed trading (Chena and Huainan 2012) or imbalances in liquidity and transaction costs (Lesmonda et al. 2004). In such cases, the use of time series models will not yield significant benefits to an investor, as claimed in Banerjee and Hung (2013). Further, as with any time series, its behaviour can abruptly change due to exogenous reasons; it is a common problem that can possibly reduce the usability of time series models, even though the tuning scheme based on a rolling observation window used in the calibration of the predictive model hereby proposed can alleviate this problem by readapting the model to most recent data and leaving obsolete data out of the fitted model.

2 A Definition of Signed Path Dependence and Implications to Time Series Models

In this section we propose an extension of the concept of *time series momentum* introduced in Moskowitz et al. (2012). In Moskowitz et al. (2012) it has been shown that for several markets the financial return in a particular month was positively correlated with the sign of the cumulative financial return of the previous 12 months. Empirical experiments demonstrated that, when employed together with a suitable leveraging strategy based on volatility scaling, such a phenomenon could be used to build investment strategies that offered a risk-reward ratio superior to the one offered by conventional equity investments. Our statistical framework is more general so that it provides a model-free hypothesis test for the presence of significant correlation, positive or negative, between the financial return in a given forecasting horizon against the sign the cumulative financial return over any arbitrary lookback window and use such alleged presence to predict the sign of future market returns. We also provide a method to detect the optimal lag to be used for forecasting and analyse our definition from the point of view of a number of parametric econometric models.

In most of this work, we will consider a *single* asset whose price is observed at different time points $T \subseteq \mathbb{Z}$, interpreted as unit intervals of observation. We will

denote the log-price of the asset at time $t \in T$ by $p(t)$, the log-return by $r(t)$, as usual defined as $r(t) := p(t) - p(t - 1)$.

We further define *cumulative log-returns* over longer time scales as follows:

- (i) $r_{-n}(t) = p(t) - p(t - n) = \sum_{i=0}^{n-1} r(t - i)$,
the (n unit intervals) *look-back cumulative log-return* at time t .
- (ii) $r_{+h}(t) = p(t + h) - p(t) = \sum_{i=1}^h r(t + i)$,
the (over a horizon of h unit intervals) *look-ahead cumulative log-return* at time t .

As it is common, by slight abuse of notation, we will denote by the above objects simultaneously random variables from which they may be obtained as observations/realisations, with arguments t, n, h being the only quantities considered fixed and not random. When expectations are taken, they are taken over the whole generative process. In the definitions below, we will condition on some observational knowledge at a time point t ; at this state, $r_{-n}(t)$ is constant (not random) for any n , while r_{+h} is a random variable since the future price $p(t+h)$ is unknown. Additionally, throughout this manuscript we shall be referring to \bar{r} as the estimated long term mean log-return of the asset being modelled and we shall refer to $s(t + 1) = \text{sgn}(r(t + 1) - \bar{r})$ where sgn is the sign function. In the context of estimation, $\widehat{s(t + 1)} \in \{-1, 1\}$ is the predicted realisation of $s(t + 1)$ given the observed realisation of the log-return time series up until time t .

Definition 1 generalizes the concept of time series momentum by the use of conditional expectations of cumulative returns:

Definition 1 (*Signed Path Dependence*) We say the considered asset (or the associated process) has *signed path dependence with memory n at forecast horizon $h > 0$* —or simply *dependence*($-n, +h$)—if, for all t where $r_{-n}(t) \neq n\mathbb{E}[r(t)]$ it holds that

$$\text{sgn}(r_{-n}(t) - \mathbb{E}[r_{-n}(t)]) (\mathbb{E}[r_{+h}(t) | r_{-n}(t)] - \mathbb{E}[r_{+h}(t)]) \neq 0 \quad (2.1)$$

Definition 1 intuitively says that, in a process with signed path dependence, knowledge of the look-back cumulative innovation over n time intervals in the past allows us to guess the sign of the look-ahead cumulative innovation h intervals in the future. If the equation is positive, such sign will be the same, and if negative, the sign will be the opposite. We intentionally look only at two cumulative innovations and only at the sign of the future one as to keep the definition intentionally parsimonious; this will allow us to check whether there is serial correlation without having to specify the exact quantitative nature of such a dependence. While being less strong for forecasting than an explicit forecasting model, we intentionally abstain from a more parametric approach as we first want to answer the question whether there is any form of dependence that could be forecast instead of right away making a forecast of a specific form. However, such definition still allows one to forecast future innovations with more accuracy than making a random guess (by accurately forecasting their signs) and such forecasting exercise on its own might after all be the best forecast one can make about future innovations if they are highly dynamic and seemingly unpredictable.

Formally, we would like to point out that the definition of signed path dependence depends on the look-back and look-ahead horizons n and h , as well as $p(t)$ considered

as a random variable. n and h are parameters of either definition, and while $p(t)$ is unobserved we will see that “having serial dependence” is a property that can be estimated from observations of $p(t)$ via predictive strategies. Further properties and basic observations pertaining to signed path dependence of a time series are listed in Remark 1 below.

- Remark 1** (i) The definition could also have been made as a function of prices only at three time points, namely at times t , $t + h$, and $t - n$.
- (ii) The same asset can have both positive *dependence* $(-n, +h)$ and negative *dependence* $(-n', +h')$ simultaneously as long as the pair (n, h) is different from the pair (n', h') .
- (iii) As the definition requires the inequality to hold for all time points t , it is possible that it won't hold across the entire process; however, one can also define a local version of dependence where the inequality would hold only in a localised version of the process.
- (iv) The definitions make no assumption on the probability distribution of the returns of the asset or the rate of decay of statistical dependence of two points with increasing time interval.
- (v) The definitions are independent of the time scale the estimation is done: given three time points, it does not matter for the definition whether the observations have finer resolution or not. This means parameters for high frequency data and low frequency data can be estimated using the same algorithm and even the same time series: for example, if one has a high frequency data series but sets the forecasting horizon h to a sufficiently large number, one will be in essence making low frequency forecasts.

In this manuscript, we theoretically validate our definition of signed path dependence by verifying the conditions for its presence in the class of linear processes that admit a Wold representation. A key assumption in the models subsequently defined is that the input series is covariance stationary. Such property defined as follows:

Definition 2 (*Covariance Stationarity*) A time series $r(t)$ is covariance stationary if $\mathbb{E}[r(i)r(i+p)] = C_p$ for all $i > 0$, $i + p > 0$ where C_p is a constant number.

We now provide the necessary and sufficient conditions for positive or negative path dependence in the case of covariance stationary processes. Our characterization will build on the the explicit classification of such processes given by Wold's representation Theorem 1 Wold (1954).

Theorem 1 (Wold Representation Theorem) *Every covariance stationary time series $r(t)$ can be written as the sum of two time series, one deterministic and one stochastic, in the following form:*

$$r(t) = \sum_{j=0}^{\infty} b_j \varepsilon(t-j) + \eta(t), \quad (2.2)$$

where ε_t is an uncorrelated innovation process, $b_j \in \mathbb{R}$, and $\eta(t)$, is a pure predictable time series, in the sense that $P[\eta(t+s)|r(t-1), r(t-2), \dots] = \eta(t+s)$, $s \geq 0$ where $P[\eta(t+s)|x]$ is the orthogonal projection of $\eta(t+s)$ on x .

Remark 2 All linear processes have a Wold representation. Furthermore, one may consider this theorem as an existence theorem for any stationary process.

The following theorem gives the conditions for negative or positive path dependence for processes admitting a Wold representation:

Theorem 2 (Signed Path Dependence in Covariance Stationary Processes) *Let $r(t)$ be a purely non-deterministic covariance stationary process with Gaussian innovations that has a Wold representation $r(t) = \sum_{j=0}^{\infty} b_j \varepsilon(t-j)$ with uncorrelated innovations such that $\varepsilon(t) \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.*

If $\Psi(h, t) := \sum_{i=1}^h \sum_{j=h}^{\infty} b_{i+j-h} \sigma_{t-i-j+h+1}^2$ then the following characterization holds:

- (i) *The process $r(t)$ has positive dependence $(-n, +h)$ if and only if $\Psi(h, t) > 0$ for all t .*
- (ii) *The process $r(t)$ has negative dependence $(-n, +h)$ if and only if $\Psi(h, t) < 0$ for all t .*

All above statements are independent of n .

Remark 3 Notice that the assumption of Gaussian innovations in Theorem 2 was only critical in the proof to ensure that the distribution of the cumulative increments was still Gaussian. Therefore, this assumption can be relaxed to an assumption that partial sums of the innovations of the process follow a stable law, i.e. their distribution is such that a linear combination of variables with this distribution has the same distribution up to location and scale parameters, and a similar derivation of the same properties can be constructed that allows for heavy tailed and skewed innovations.

Based on the results of Theorem 2, it becomes straightforward to infer in any covariance stationary fitted time series model whether there is positive or negative (or none) serial dependence in the sense of Definition 1 by finding its equivalent Wold decomposition and verifying if the fitted parameter values satisfy the derived conditions on the characteristic polynomial of AR and MA roots.

Moreover, notice that the definition of signed path dependence concerns only deviations around the process unconditional expectation—i.e. the deterministic part of the process. Also, if the process is linear, its innovations will have the same variance. As such, we can state the following corollary:

Corollary 1 (Signed Path Dependence in Linear Covariance Stationary Processes) *Let $r(t)$ be a linear covariance stationary process whose innovations follow a distribution such that a linear combination of variables with this distribution has the same distribution up to location and scale parameters. If $r(t)$ has signed path dependence of sign $\zeta \in \{-1, 1\}$ with lookback n_l and horizon h for some $n_l > 0$, then $r(t)$ will have signed path dependence of the same sign ζ with lookback n and horizon h for any integer $n > 0$.*

Remark 4 Also notice that if these sufficient conditions in Theorem 2 are not satisfied for all values of $n \in \mathbb{N}^+$ but only a finite subset of these, then the presence of signed path dependence needs further refinement. In this way we comment that these conditions are sufficient but not necessary.

To facilitate understanding, we conclude the section by enunciating two lemmas that link the concept of Signed Path Dependence to known parametric time series models. These lemmas demonstrate one-to-one correspondences between parameter ranges and the presence of signed path dependence. Beyond providing theoretical evidence and hence validation for our definition, these lemmas will allow us to generate data with known signed path dependence by simulating a parametric process with suitable parameter values and subsequently use in the synthetic validation experiment given in Sect. 4. The first lemma provides sufficient and necessary results for the simplest non-trivial case of time series models, the AR(1) processes. The lemma suppresses the constant term by assuming that the variable being modelled has been measured as deviations from its mean.

Lemma 1 (Signed Path Dependence in a Stationary AR(1) Process of Constant Volatility) *Assume that $r(t)$ is a stationary AR(1) process given by*

$$r(t) = \varphi \cdot r(t - 1) + \varepsilon(t) \tag{2.3}$$

with i.i.d. innovations $\varepsilon(t) \sim \mathcal{N}(0, \sigma^2)$ and $|\varphi| < 1$. Then:

- (i) *The process $r(t)$ has positive path dependence if and only if $\varphi > 0$.*
- (ii) *The process $r(t)$ has negative path dependence if and only if $\varphi < 0$.*

In particular, if $\varphi = 0$ the process has no signed path dependence, and all statements above are independent of n and h .

And the second lemma derives conditions for Autoregressive Fractionally Integrated Moving Average (ARFIMA) models. These models generalize ARMA models by introducing long-memory features considering a non-integer differencing parameter. Note that the ARFIMA (=sFARIMA) models include the ARIMA models (as ARFIMA(p, d, q) with d integer), the ARIMA models include the ARMA models (as ARMA(p, q) = ARIMA($p, 0, q$)), and the ARMA models include the AR models (as AR(p) = ARMA($p, 0$)), including the AR(1) of the first lemma.

Lemma 2 (Signed Path Dependence in an ARFIMA(p, d, q) Process) *Assume that $r(t)$ is a purely non-deterministic ARFIMA (p, d, q) process with i.i.d. Gaussian innovations given by $\Phi(B)(1 - B)^d r(t) = \Theta(B)\varepsilon(t)$, where Φ and Θ are polynomials of degree p resp. q in the backshift operator B , with i.i.d. innovations $\varepsilon(t) \sim \mathcal{N}(0, \sigma^2)$. Let $\sum_{j=0}^{\infty} b_j \varepsilon(t - j)$ be the Wold representation of the ARMA process $\frac{\Theta(B)}{\Phi(B)}\varepsilon(t)$ and define $\Psi(h) = \sum_{i=1}^h \sum_{j=h}^{\infty} b_{i+j-h}$. It holds that:*

- (i) *If $d > -0.5$, the process $r(t)$ has positive path dependence for all $h \in \mathbb{N}^+$ and $n \in \mathbb{N}^+$ if and only if $d\Psi(h) > 0$; and*
- (ii) *If $d > -0.5$, the process $r(t)$ has negative path dependence for all $h \in \mathbb{N}^+$ and $n \in \mathbb{N}^+$ if and only if $d\Psi(h) < 0$.*

Additionally, if $p = q = 0$, it holds that:

- (i) The process $r(t)$ has positive path dependence if and only if $d > 0$.
- (ii) The process $r(t)$ has negative path dependence if and only if $-0.5 < d < 0$.

Notice that ARFIMA models can demonstrate signed path dependence if they are at least in fact in the sub-class of invertible processes and may or may not contain long memory. In the case of reversal properties of an ARFIMA model we learn that invertible models are also of interest and again these sub-class of models may or may not contain long memory properties. Most interestingly, we see that in the special case of only fractional differencing of the process, i.e. no AR and no MA components of the process, we can only have positive path dependence in the ARFIMA process if the process is invertible and it may or may not be stationary or contain long memory. In the case of the ARFIMA model with no AR and no MA components of the process, then negative dependence can only occur in sub-processes which have the properties that they don't have long memory, furthermore, they must be invertible and stationary.

3 Statistical Inference Under Signed Path Dependence

The models in the previous section assumed some structure for the serial correlation present in the series. We now derive a non-parametric hypothesis test for signed path dependence under certain regularity conditions. Such regularity conditions are listed in Assumption 1. The test proposed assumes these conditions to hold over the entire series, but one can allow for a localised version of the test by segmenting the series being tested and testing each segment. The inference procedure proposed aims to test the following assumption (the single sided version of the test would swap the inequality sign for a greater than or less than sign depending on whether positive or negative path dependence is being tested):

H0 (Absence of Linear Correlations between Signs and Returns) given n , for all $k \leq n$ the sign of the sum of the previous k observations has no linear correlation with the observation 1-step ahead, i.e. $\mathbb{E}[\text{sgn}(r_{-k}(t))r(t+1)] = 0$ for all $k \leq n$.

H1 (Presence of some Linear Correlation between Signs and Returns) $r(t)$ are associated such that $\mathbb{E}[\text{sgn}(r_{-k}(t))r(t+1)] \neq 0$ for some value of $k \leq n$.

Before formally stating the not too restrictive classical regularity conditions, a definition is needed:

Definition 3 (*Lindeberg Condition*) A time series $r(t)$ is said to satisfy the Lindeberg condition if for all $\epsilon > 0$ the following holds:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_{|x| > \epsilon \sqrt{n}} x^2 df_i = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[(r(i))^2 \mathbb{I}\{|r(i)| > \epsilon \sqrt{n}\} \right] = 0 \tag{3.1}$$

where f_i is the density function of $r(i)$ and \mathbb{I} is the indicator function.

The intuition behind the Lindeberg condition is that, in the series being tested, the individual contribution of any observation to the sum of the variances of all observations should be arbitrarily small for a sufficiently large series.

As such, the regularity conditions are given by the following assumption:

Assumption 1 The time series being tested $r(t)$ is comprised of deviations from the long-term mean of a stochastic process with a well-defined mean, hence having zero mean by definition, is covariance stationary in the sense of Definition 2, satisfies the Lindeberg condition and weighted partial sums of the process are assumed to follow asymptotically a Gaussian law.

The assumption that the Lindeberg condition is satisfied is very general and few processes of interest will fail to satisfy this condition so it is not overly restrictive in any sense. Also notice that, as the series being tested is assumed covariance stationary in the sense of Definition 2, Theorem 2 will guarantee that if this series has positive or negative path dependence for a given forecast horizon h , it will have the same property (i.e. it will be always positive or always negative) for all values of the memory n , and at all time points t . So, one can fix the forecast horizon $h = 1$ and test for signed path dependence at that forecast horizon by aggregating the time series over t (for example, testing with daily data if the forecast horizon h equals to one day, testing with weekly data if it equals to 1 week, and so on so forth).

Before proposing our test statistic, we will state a result that will base the derivation of the asymptotic distribution of our test statistic under the null hypothesis for the class of processes which are covariance stationary. For proof, see H Diananda (1954).

Theorem 3 (Central Limit Theorem for Covariance Stationary m -Dependent Variables) *Let $R = r(1), r(2), \dots, r(n)$ be random variables of zero mean and finite variances such that $r(t)$ is uncorrelated with $r(t+i)$ for all $i > m$, m fixed. If the variables are covariance stationary and satisfy the Lindeberg condition then, as $n \rightarrow \infty$, $\frac{\sum_{t=1}^n r(t)}{\sigma\sqrt{n}} \xrightarrow{D} \mathcal{N}(0, 1)$ with $\sigma^2 = \mathbb{E}[(r(t))^2] + 2\sum_{k=1}^m \mathbb{E}[r(t)r(t+k)]$.*

As such, the asymptotic distribution of our test statistic is given by the following theorem:

Theorem 4 (Critical Rejection Value of the Test Statistic) *Given an arbitrary positive value of n , define $d(t) = r(t)\omega(t)$ where $\omega(t) = \frac{\sum_{i=1}^n \text{sgn}(\sum_{k=t-i}^{t-1} r(k))}{n}$ and let $\bar{d} = \frac{1}{s-n} \sum_{j=n+1}^s d(j)$. Let $r(t)$ be a time series that satisfies Assumption 1 and define a random vector $R := (r(1), r(2), \dots, r(s))$ where s is a given constant. When $s \rightarrow \infty$, under the null hypothesis of Absence of Linear Correlations between Signs and Returns the test statistic $H(R) \xrightarrow{D} \mathcal{N}(0, 1)$ with*

$$H(R) = \frac{\bar{d}}{\sqrt{\frac{\sum_{\tau=-n}^n \hat{\gamma}(\tau)}{s-n}}} \tag{3.2}$$

where

$$\hat{\gamma}(\tau) = \frac{1}{s-n} \sum_{j=|\tau|+1}^{s-n} (d(j) - \bar{d})(d(j - |\tau|) - \bar{d}) \tag{3.3}$$

Therefore, for a given significance threshold α , the null hypothesis will be rejected as true if $|H| > \Phi^{-1}(1 - \frac{\alpha}{2})$ where $\Phi^{-1}(x)$ is the inverse standard normal cumulative distribution function of x .

Remark 5 There are situations where one would be interested only to test for a specific sign of the signed path dependence. For example, if the aim is to test the suitability of trend-following strategies, one would be interested to test only for positive signed dependence. Equally, if the aim is to test the presence of mean reversion in a particular series, one would be interested to test only for negative signed dependence. These situations can be accommodated by making the test proposed in Theorem 4 one-sided and rejecting the null hypothesis if $H > \Phi^{-1}(1 - \alpha)$ for a positive sign dependence test or if $H < \Phi^{-1}(\alpha)$ for a negative sign dependence test.

Remark 6 The test statistic defined in Eq. 3.2 can also be interpreted as a correlation coefficient if normalised by the sum of all absolute returns one-step ahead instead of normalised by the sample variance, i.e. if calculated as $\rho = \frac{\sum_{j=n+1}^s d(j)}{\sum_{j=n+1}^s |r(j)|}$. Such coefficient will be bounded between -1 and 1 , with its sign indicating the sign of the path dependence being estimated and its absolute value indicating the strength of such dependency. Sect. 5 calculates this coefficient for a number of financial assets and forecast horizons and compares it against the Pearson correlation of the future return against the average return over the previous n days. We chose comparing our measure against the Pearson correlation to provide a meaningful comparison between signed path dependence and classical dependence in correlation metrics. The results show that this coefficient was able to detect a significantly positive serial correlation in situations where its Pearson counterpart was not significantly different than zero.

Remark 7 Theorem 4 can be easily extended with relaxation of the Lindeburg condition to admit covariation stationary processes which admit an analogous α -stable limit theorem result, interested readers are referred to derivations of such results in Peters and Shevchenko (2015).

As the test given by Theorem 4 relies on the asymptotic distribution of the test statistic under the null hypothesis, one might want to investigate how large the sample should be to obtain convergence. In this subsection, an alternative test is proposed that has the same asymptotic distribution for the test statistic under the null hypothesis. This way, one can run both tests under the same simulation experimental conditions whilst increasing the sample size of the simulated data to assess the power of the test, also obtaining an indication of the speed of convergence of the asymptotic test given by Theorem 4. Such test is described in Algorithm 1. The test relies on an empirical percentile bootstrap to get a confidence interval for the test statistic under the null hypothesis of no signed path dependence and can therefore be used to check whether

the value of test statistic calculated in the given sample is outside of this confidence interval, rejecting the null if this is the case.

The bootstrap test proposed is adequate to check for convergence of the asymptotic test statistic given by Theorem 4 because of the result given by Lemma 3:

Lemma 3 (Convergence of the Bootstrap Test Statistic Distribution under the Null Assumption) *Let \widehat{h}_α be the $1 - \frac{\alpha}{2}$ empirical percentile of the bootstrap test statistic used to reject the null hypothesis of the test procedure given by Algorithm 1 at significance threshold α . For large samples ($s \rightarrow \infty$), $\widehat{h}_\alpha \rightarrow \Phi^{-1}(1 - \frac{\alpha}{2})$, where Φ^{-1} is the inverse CDF of a standard normal distribution. Therefore, for large samples, the critical rejection value of the bootstrap test statistic given by Algorithm 1 converges to the critical rejection value of the test given by Theorem 4.*

It is worth noting that the rejection of the null hypothesis of the test given by Theorem 4 also implies that one can build a predictive model for the sign of return one step ahead based on the sign of the cumulative returns observed on the previous n steps. This is further explored in Sect. 5. The section compares the performance of a predictive model based on signed path dependence to make buy and sell decisions against the performance of an investment strategy known as Buy and Hold, which in essence is always buying and hence always estimating $s(\widehat{t + 1}) = 1$. Theorem 5 gives the mathematical construction of the predictive model to be studied.

Algorithm 1 Bootstrap procedure to test for signed path dependence

Input: $R = (r(1), r(2), \dots, r(z))$ a sample of length z of observed log-returns, a significance threshold α and a number of simulations W

Output: TRUE if the null hypothesis of no signed path dependence is rejected with Type I error probability α and FALSE otherwise

- 1: Generate matrix $U \in [0, 1]^{W \times z}$ with each $U_{w,t}$ drawn i.i.d. according to a discrete uniform distribution on the set $\{1, 2, \dots, z\}$
 - 2: Generate W random series $r'_w(t)$ of length s such that $r'_w(t) := r(U_{w,t})$
 - 3: Construct an array B of length W so that $B(1 \leq w \leq W) = H(r'_w)$ where $H(r'_w)$ is the test statistic given by Equation (3.2) calculated over series r'_w
 - 4: Construct \widehat{F}_B the empirical cdf of B
 - 5: Evaluate $\widehat{h}_\alpha \leftarrow \inf\{x \in \mathbb{R} : \widehat{F}_B(x) \geq 1 - \frac{\alpha}{2}\}$
 - 6: Return TRUE if $|H(R)| \geq \widehat{h}_\alpha$ or FALSE otherwise
-

Theorem 5 (Mathematical Construction of the Predictive Model) *Obtain $s(\widehat{t + 1})$ by performing the following steps:*

1. Given a log-return sample $r(1), r(2), \dots, r(t)$ estimate the sample mean as $\bar{r} = \sum_{i=1}^t r(i) / t$.
2. Apply the test given by Theorem 4 to the given sample to detect if there is path dependence of any sign and let $\hat{\zeta} \in \{-1, 1\}$ be the sign of the path dependence detected.
3. If $\hat{\zeta} = -1$ choose an appropriate value of n such that $\mathbb{E}[sgn(r_{-n}(t) - n\bar{r}) r(t + 1) - \bar{r}] < 0$ and set $s(\widehat{t + 1}) = -sgn(r_{-n}(t) - n\bar{r})$
4. Otherwise choose an appropriate value of n such that $\mathbb{E}[sgn(r_{-n}(t) - n\bar{r}) r(t + 1) - \bar{r}] > 0$ and set $s(\widehat{t + 1}) = sgn(r_{-n}(t) - n\bar{r})$

Also define the prediction loss of a predictive strategy as $L = -s(\widehat{t+1})r(t+1)$. Note that the prediction loss of a Buy and Hold strategy is trivially given by $L_{B\&H} = -r(t+1)$. For the aforementioned strategy we have that $\mathbb{E} \left[L_{s(\widehat{t+1})} \right] < \mathbb{E} [L_{B\&H}]$ if the null hypothesis of the test given by Theorem 4 is not true for the demeaned returns $r(1) - \bar{r}, r(2) - \bar{r}, \dots, r(t) - \bar{r}$.

We conclude the section by describing a predictive algorithm that attempts to forecast the sign of future asset returns using the method described in Theorem 5. Notice that the method given by Theorem 5 dynamically decides whether the prediction should be made expecting a continuation of sign ($\hat{\zeta} = 1$) or reversal of sign ($\hat{\zeta} = -1$). Given that the particular case of trading strategies that only assume continuation in sign (*trend-following*) receives special attention in Finance, see CSzakmary et al. (2010). Our algorithm will cater for both cases (only continuation assumed or dynamic decision between continuation and reversals). To achieve that, we define two different predictors in Definitions 4 and 5:

Definition 4 (Moving Average (MA) Classifier) An $MA(n, \bar{r})$ predictor receives a log-return sample $r = (r(1), \dots, r(n))$ and predicts that $s(\widehat{n+1}) = \text{sgn}(r_{-n}(n) - \bar{r})$.

Definition 5 (Dynamic Moving Average (DMA) Classifier) A $DMA(n, \bar{r}, \hat{\zeta})$ predictor receives a log-return sample $r = (r(1), \dots, r(n))$ and predicts that $s(\widehat{n+1}) = \hat{\zeta} \times \text{sgn}(r_{-n}(n) - \bar{r})$ where $\hat{\zeta} \in \{-1, 1\}$.

In order to apply these methods for prediction, parameters n , \bar{r} and $\hat{\zeta}$ have to be estimated. Naturally, \bar{r} can be estimated as the sample mean log-return. The remaining parameters can be estimated by weighted least squares of an adequately defined error function. A simple suitable error weighted function can be given as $\varepsilon(t) = r(t) \left(\widehat{s(t)} - s(t) \right)$. Also, it is clear from Theorem 5 that predictions made by the aforementioned predictors will not be useful in a process that has no signed path dependence. Therefore, before attempting to apply any of these methods, the test defined in Theorem 4 is applied to the return series used to fit the model to ascertain the presence of signed path dependence, given that in the absence of this property no prediction should be attempted. This yields Algorithm .

Algorithm 2 Simplified predictive method based on signed path dependence

Input: $R = (r(1), r(2), \dots, r(z))$ a sample of length z of observed log-returns in chronological order, a significance threshold α and a test lookback horizon n'

Output: 0 if no prediction was made for $s(z+1)$, $s(\widehat{z+1})$ otherwise

- 1: **if** $\left| \frac{\bar{d}}{\sqrt{\frac{\sum_{\tau=-n'}^{n'} \widehat{r}(\tau)}{s-n'}}} \right| > \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ **then**
- 2: Compute $\bar{r} \leftarrow \frac{1}{z} \sum_{i=1}^z r(i)$
- 3: Compute $\hat{n}, \hat{\zeta} \leftarrow \arg \min_{\substack{n \in \mathbb{N}^+, n < z \\ \zeta \in \{-1, 1\}}} \sum_{i=n+1}^z (r(i) (\zeta \times \text{sgn}(r_{-n}(i-1) - \bar{r}) - \text{sgn}(r(i) - \bar{r})))^2$
- 4: Return $\hat{\zeta} \times \text{sgn}(r_{-n}(z) - \bar{r})$
- 5: **else**
- 6: Return 0
- 7: **end if**

Remark 8 Algorithm 2 implements the DMA predictor, but it can be easily modified to implement the MA predictor by fixing $\hat{\zeta} = 1$ and performing a single sided test in Step 1.

As with any statistical model, there are multiple ways to estimate its parameters and weighted least squares is just one simplified possibility. The empirical tests presented in Sect. 5 have been performed using a more sophisticated machine learning algorithm, whose detailed description is outside of the scope of the main text of this manuscript. A description of the machine learning algorithm in the form of commented pseudocode is available in the Technical Appendix and the R implementation of the machine learning algorithm used in the empirical experiments of this manuscript is available in the Supplementary Material.

While the use of such machine learning method is not required to establish the main results in the paper, we have chosen it for our empirical application because the use of Train and Test sets reduces potential overfitting that could come out of a simple weighted least squares procedure, and the use of a rolling model tuning stage produces better estimates of how accurately the predictive model will perform in a practical scenario where the model gets updated as new data arrives with the passage of time.

4 Validation by Simulation

In this section we validate by simulation the properties of the test given in Sect. 3 by simulating synthetic time series where parameters are chosen so that positive or negative signed path dependence is guaranteed in the simulated series. As part of the validation exercise, the test defined in Sect. 3 is applied to a number of different synthetic series and used to detect the simulated dependence. Such results can be used to infer the robustness of the test statistic against the following factors:

1. Different assumptions in the functional form of the input time series;

2. Strength of the dependence present in the input time series, including the effect of long-range dependence;
3. Time varying variance in the input time series when the required assumption of covariance stationarity is met; and
4. Time varying variance in the input time series when the required assumption of covariance stationarity is not met.

Robustness against the aforementioned factors in a controlled environment can give assurances for the suitability of the test in empirical applications, where the actual functional form of the input time series is not known and the required assumption of covariance stationarity might not necessarily be met. Moreover, comparing different levels of Type I and Type II error probabilities across all validation scenarios can give useful insights on which properties have higher impact on the accuracy of the test and any potential biases.

We start our synthetic examples by the simplest case of the AR(1) models. As per Lemma 1, we know that the sign of the dependence for an AR(1) process depends entirely on the sign of the parameter ϕ , and all series $r_j(i)$ have positive path dependence when $\phi_j > 0$ while all other series will not have so. As such, the first synthetic validation experiment performed by us is given by the steps below:

Experiment 1 *Perform the following steps to illustrate Lemma 1:*

1. Generate $P = 1000$ AR(1) series of length $z = 1000$ so that $r_j(i) = \phi_j r_j(i - 1) + \varepsilon_j(i)$ with $r_j(1) = \varepsilon_j(1)$, $0 < i \leq z$, $0 < j \leq P$, $\phi_j = 4 \left(\frac{j}{5P} - \frac{P+1}{10P} \right)$ and $\varepsilon_j(i) \sim \mathcal{N}(0, 1)$
2. Fix the significance threshold α and apply the single sided version of the asymptotic test to detect positive path dependence as per Remark 5 to all the P series.
3. Use the known presence of positive dependence to calculate the type I and type II error probabilities respectively by counting how many of the series where $j \leq \frac{P}{2}$ had the null hypothesis rejected and how many of the series where $j > \frac{P}{2}$ had the test failing to reject the null hypothesis.
4. Perform Steps 2 and 3 using the bootstrap test defined in Algorithm 1 to detect positive path dependence (as a control case).
5. Repeat the procedure for different values of α .

Figure 1 shows the behaviour of type I and type II errors for the bootstrap and asymptotic tests as a function of the significance threshold for the procedure given by Experiment 1. As expected by Lemma 3, the two tests converged and the type I error probability of the asymptotic test was not statistically different than the one of the bootstrap test. Both tests had their type I error probabilities below the identity line suggesting that the test is being conservative, though not statistically incorrect. The type II error probabilities of both tests are also not statistically different than each other and always lower than 50%, suggesting that even though the test is conservative, for the AR(1) series constructed it will fail to reject the null when the alternative hypothesis truly holds only in the minority of the cases.

The theoretical model has also been validated by simulating an ARFIMA(0,d,0) series. Lemma 2 gives the condition under the forecast horizon $h = 1$ where the sign

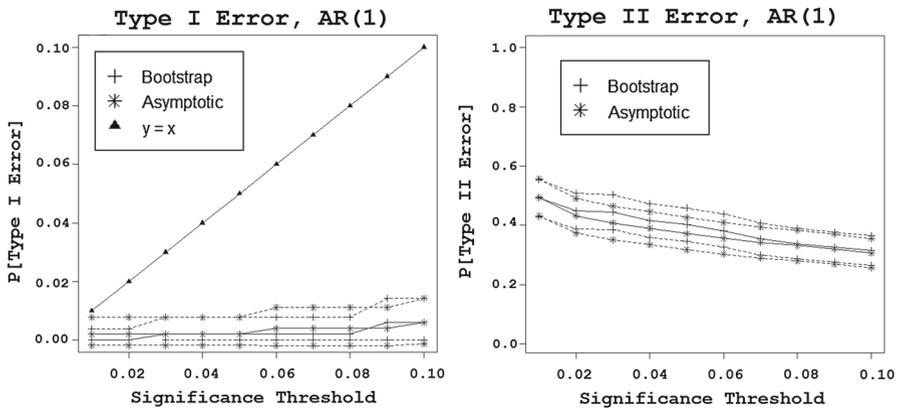


Fig. 1 Performance of the tests in the simulated AR(1) series. The solid lines represent the simulated error probabilities and the dashed lines represent a 95% confidence margin of the error probability of each test, obtained via a Poisson approximation for the count of false acceptances/rejections. The chart on the left shows the simulated type I error probabilities of the asymptotic and bootstrap statistics as functions of the significance threshold when these tests are applied to the generated series. The black straight line corresponds to the identity line, where the test type I error is exactly equal to the desired significance threshold. The chart on the right shows the simulated type II error probabilities of the two tests as functions of the significance threshold with the dashed lines being the 95% confidence margin

of the path dependence for this model class is known for the whole series: the sign of the parameter d . To keep the simulation procedure simple, we have reused the procedure given by Experiment 1 and in Step 1 we constructed the series based on a single parameter $-0.1 \leq d \leq 0.1$, with the only modification being that the series generated in Step 1 are ARFIMA(0, ϕ , 0) series (instead of AR(1)).

Figure 2 shows the behaviour of type I and type II errors for the bootstrap and asymptotic tests as function of the significance threshold for the simulated ARFIMA(0, d , 0) series. The type I error probability of the asymptotic test is not statistically different from the one of the bootstrap test; however, both tests have their type I error probabilities well below the identity line suggesting that the test is being very conservative, though not statistically incorrect. The type II error probabilities of both tests are also not statistically different than each other and always lower than 25%, suggesting that even though the test is conservative, in the large majority of the cases it will not fail to reject the null when the alternative hypothesis truly holds. The synthetic experiment for the ARFIMA(0, d , 0) model yielded conclusions similar to the ones obtained for the AR(1) model, with the difference being that both tests were considerably more powerful. This suggests the long-memory feature of the ARFIMA model introduces a stronger signed path dependence to the modelled time series than the one introduced by an AR(1).

Finally, to ascertain some boundary conditions for the applicability of our tests, we have created a synthetic MA(n) series and analysed the behaviour of the test by modifying it so that it has a) a very weak correlation structure; b) a non-constant volatility that follows a stationary GARCH process; or c) a non-constant and non-stationary volatility that is driven by a two-state Markov-switching process.

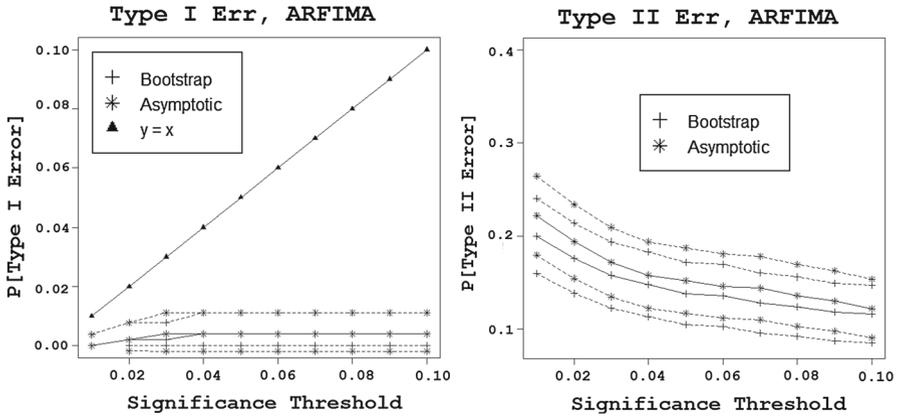


Fig. 2 Performance of the tests in the simulated ARFIMA(0,d,0) series. Lines have the same meaning as in Fig. 1

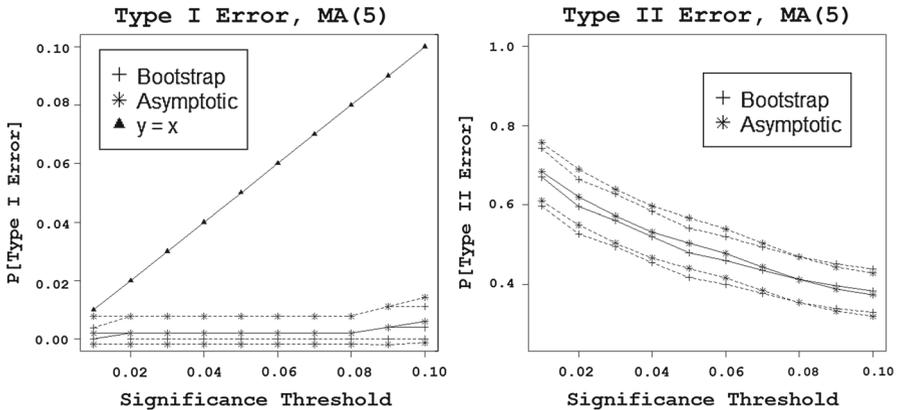


Fig. 3 Performance of the tests in the baseline simulated MA(5) series. Lines have the same meaning as in Fig. 1

To build the baseline MA(n) series we note that, as per Theorem 2, when the forecast horizon $h = 1$ we have that the sign of the sum of the MA coefficients will determine the sign of the path dependence of the model. We again have reused the procedure given by Experiment 1 but in Step 1 we generated MA(5) series based on a single parameter $-0.1 \leq \phi \leq 0.1$ so that the coefficients of all 5 lags were equal to $\frac{4}{5}\phi$. The synthetic experiment for the MA(5) model yielded conclusions similar to the ones obtained for the AR(1) model, with the difference being that the test was slightly less powerful. Figure 3 shows the behaviour of type I and type II errors for the bootstrap and asymptotic tests as function of the significance threshold for the simulated MA(5) series. Like in the AR(1) case the tests converged, with both the type I and type II error probabilities of these tests not being statistically different amongst themselves. Additionally, the type II error probabilities of both tests was still lower than 50% when the significance threshold was 0.05 or higher.

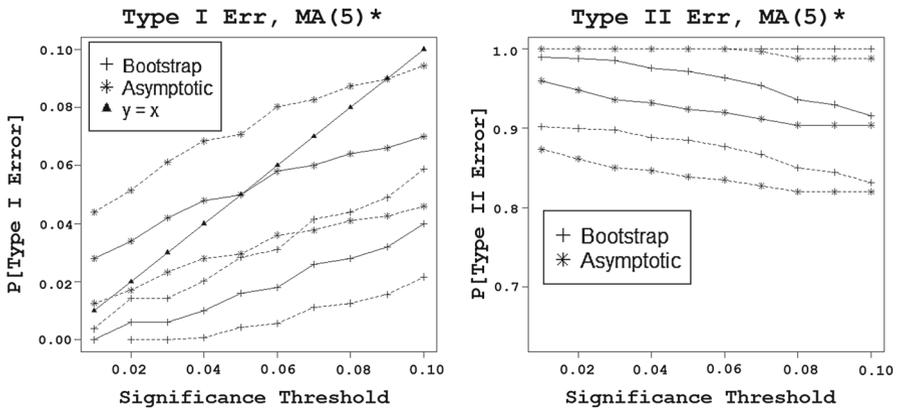


Fig. 4 Performance of the tests in the stressed MA(5). Lines have the same meaning as in Fig. 1

After generating the baseline case, we created a stressed MA(5) simulation where series length is reduced considerably and the correlation structure is weakened. In this stressed MA(5) simulation, the procedure given by Experiment 1 was reused but in Step 1 we generated MA(5) series based on a single parameter $-0.1 \leq \phi \leq 0.1$ so that the coefficients of lags 1, 3 and 5 were equal to ϕ and the coefficients of lags 2 and 4 were equal to $-\phi$. In this situation, the average sum of all MA coefficients when the alternative hypothesis holds was still positive, equal to 0.05, though the alternating signs of the lags weaken the overall positive dependence. Further, the procedure generated series of shorter length $z = 100$. Figure 4 shows the behaviour of type I and type II errors for the bootstrap and asymptotic tests as function of the significance threshold for the stressed MA(5) series. The type I error probability of the asymptotic test is always greater than the one of the bootstrap test for a given significance threshold; when the significance threshold is lower than 0.05 it is also above the identity line and for a significance threshold lower than 0.015 the identity line falls below the 95% confidence interval of the error probability for the asymptotic test though stays very close to its lower boundary. At the same time, the type II error probability of the asymptotic test is always lower than the one of the bootstrap test for a given significance threshold, though it is always higher than 90%, suggesting the test will conservatively fail to reject the null in many genuine cases that the alternative hypothesis truly holds. Of course, the reduced power is a consequence of the intentionally weak correlation structure imposed.

Despite being subject to series of very weak correlation structure, the asymptotic test still behaved reasonably, with adequate conservatism for all but very low significance thresholds ($\alpha \leq 0.02$, when the lower 95% confidence bound goes above the identity line). The bootstrap test remained adequate even for this low significance threshold, though at the cost of an even reduced power.

All simulations so far assumed constant volatility. In empirical applications, volatility is known vary over time and to exhibit serial dependence (Cont 2001). To verify the effect of a non-constant volatility, we have changed the baseline MA(5) case so that the error variance follows a GARCH(1,1) process with coefficients $\alpha = 1$ (the constant

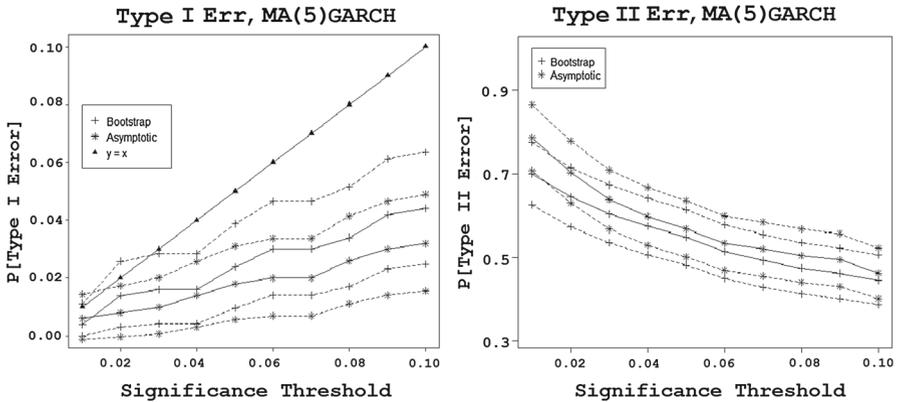


Fig. 5 Performance of the tests in the MA(5)-GARCH(1,1) series. Lines have the same meaning as in Fig. 1

term), $\beta = 0.89$ (the term that multiplies the lagged squared error) and $\gamma = 0.1$ (the term that multiplies the lagged variance). The choice of GARCH parameters ensured the series was still unconditionally covariance stationary. All other parameters of the baseline MA(5) model remained unchanged.

We note that, while the power of the test reduced slightly compared to the baseline case, comparing Figs. 4 and 5 one can see that the impact of a GARCH volatility in the power of the test was smaller than the impact of reduced correlation structure. The test remained very adequate even in the presence of GARCH volatility and Type I errors stayed either within or under the identity line for all significance thresholds, indicating adequate conservatism.

Finally, to ascertain a boundary scenario where the main assumption of covariance stationarity is not strictly satisfied, we have changed the baseline MA(5) case so that the error variance follows a Markov Switching between two equally probably GARCH(1,1) states. Such framework is widely popular in financial market research and is accepted to be a relevant deviation from time series models that has been extensively observed in empirical applications. In one of the states, the error variance follows a GARCH(1,1) process with coefficients $\alpha = 1$, $\beta = 0.89$ and $\gamma = 0.1$ and in the other state the error variance follows a GARCH(1,1) process coefficients $\alpha = 3$, $\beta = 0.95$ and $\gamma = 0.049$. All other parameters of the baseline MA(5) model remained unchanged.

We again note that, while the power of the test reduced slightly compared to the baseline case, comparing Figs. 4 and 6 one can see that the impact of a GARCH volatility in the power of the test was smaller than the impact of reduced correlation structure. The test remained very adequate even in the presence of Regime Switching GARCH volatility and Type I errors stayed either within or under the identity line for all significance thresholds, indicating adequate conservatism.

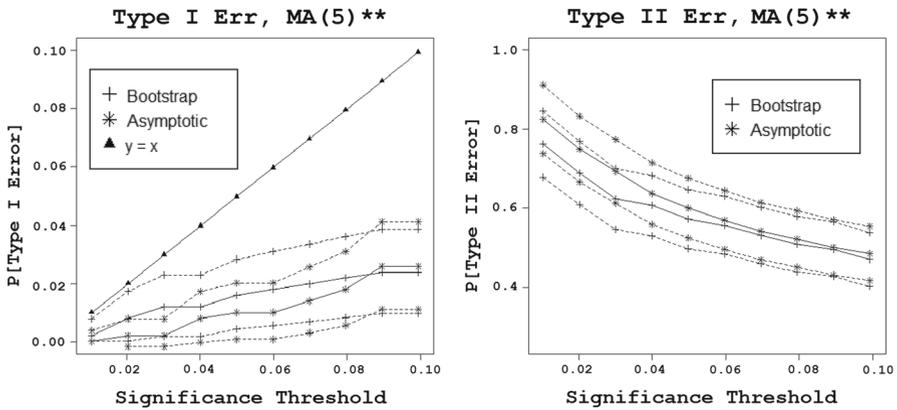


Fig. 6 Performance of the tests in the MA(5)-RSGARCH(1,1) series. Lines have the same meaning as in Fig. 1

5 Empirical Applications

This section applies to real financial data the test given by Theorem 4 and validates the performance of trading strategies created based on Theorem 5. The dataset used comprised of the time series of the prices for the MSCI Equity Total Return Indices (in US Dollar, price returns plus dividend returns net of withholding taxes) for 21 different countries and the nominal exchange rates of 16 different currencies, all of them traded against the US Dollar. Table 1 lists the equity indices and currencies studied.

For all the 37 financial instruments the daily log-returns (defined as the logarithm of the closing price at the end of day $d+1$ divided by the closing price at the end of day d) were obtained based on the price series from 31-Dec-1998 to 31-May-2017, yielding 4803 observations per instrument (and 177,711 observations in total). To conduct further analyses at different levels of observational noise, several other sampling frequencies were studied. The same log-returns were aggregated to weekly log-returns (defined as the logarithm of the closing price at the end of the Friday of week $w+1$ divided by the closing price at the end of the Friday of week w), yielding 961 observations per instrument (and 35,557 observations in total); the returns were also aggregated to monthly log-returns (defined as the logarithm of the closing price at the end of the last business day of month $m+1$ divided by the closing price at the end of the last business day of month m), yielding 221 observations per instrument (and 8177 observations in total). Each of the MSCI Index studied currently has at least one tradable Futures contract linked to it and many of them have also shorthable ETFs linked to them, which means the any predictive model resulting out of the experiments performed in this paper can also be used to guide real equity investment strategies.

In the case of exchange rates, to accurately reflect the total returns of the investment strategies proposed, the results of the experiment include the financial returns arising out of the interest rate differential between the issuing country and the US (also known as the “rollover interest”) for each open position and also include the interest paid on the cash deposit held as margin for the currency trade.

Table 1 Different correlation metrics for equity indices in the first half of the sample

Country	Daily		Weekly		Monthly	
	Rho	Pearson	Rho	Pearson	Rho	Pearson
Australia	0.0193	-0.0248	0.0067	-0.0108	-0.0303	-0.0747
Brazil	0.0851***	0.0218	0.0292	0.0072	0.1305**	-0.0567
Canada	0.0210	-0.0185	-0.0213	0.0120	0.1414**	0.1321*
Chile	0.1642***	0.0978***	0.1012**	0.0449	0.1274	0.1054
China	0.0871***	0.0316*	0.0916**	0.0560	0.1914***	0.0459
Colombia	0.2021***	0.0972***	0.2152***	0.1301***	0.0590	-0.0083
Czech	0.0557***	0.0051	0.0989***	0.0469	0.0447	-0.0353
France	-0.0327**	-0.0710***	-0.0271	-0.0278	0.1684***	0.0308
Greece	0.0809***	-0.0129	0.0445	0.0261	0.2160***	0.0845
Hongkong	0.0418***	-0.0045	0.0773**	0.0326	0.1407*	0.0200
Hungary	0.0671***	0.0250	0.0699**	0.0262	0.0303	-0.0038
India	0.1186***	0.0540***	0.1379***	0.0750*	0.1145**	0.0011
Japan	-0.0177	-0.0271*	-0.0190	-0.0211	0.2458***	0.2195**
Mexico	0.0669***	-0.0045	0.0079	-0.0164	0.0253	-0.1339*
New Zealand	-0.0118	-0.0358**	0.0437	-0.0388	-0.0086	0.0376
Russia	0.0661***	-0.0003	0.0537**	-0.0162	-0.0572*	-0.0973
Singapore	0.0628***	0.0276*	0.0774*	0.0497	0.2127**	0.0586
South Africa	0.0334**	0.0004	0.0266	0.0171	-0.0114	-0.1463*
Taiwan	0.0546***	0.0255	0.0123	0.0038	0.0852**	0.0143
UK	-0.0450***	-0.1026***	-0.0249	-0.0532	0.1575**	0.0848
USA	-0.0405***	-0.0668***	-0.0456	-0.0959**	0.0603*	0.0377

*Significantly different than zero at 10% *p* value

**Significantly different than zero at 5% *p* value

***Significantly different than zero at 1% *p* value

Before testing any trading strategies, the first half of the sample described in the second paragraph of this section (corresponding to observations from 31-Dec-1999 to 31-Mar-2008) was used to calculate the test statistic defined in Equation 3.2 and detect the presence of significant serial dependence in individual assets. Tables 1 and 2 show the values of the serial correlation coefficient calculated as Remark 6 for forecast horizons of one day, 1 week and 1 month ($n = 5$ in all cases).

The statistical significances in the tables were obtained using the test procedure described in Theorem 4. For control, the tables also display the Pearson correlation coefficient between the average return of the n previous periods and the 1-step ahead return. The significances of the Pearson correlation were estimated using a Fisher transform and its limiting Gaussian approximation. The rationale for applying the test only on the first half of the sample is to ensure there is no fitting bias in the trading model, whose performance is evaluated using the second half of the sample and whose equities and currencies traded are selected based on the test results.

Table 2 Different correlation metrics for currencies in the first half of the sample

Country	Daily		Weekly		Monthly	
	Rho	Pearson	Rho	Pearson	Rho	Pearson
AUDUSD	0.0107	-0.0125	0.0347	-0.0042	0.0108	-0.0086
BRLUSD	0.0791***	0.0539***	0.1283***	0.0866**	0.2025**	0.1113
CADUSD	-0.0291*	-0.0319*	0.0041	0.0280	0.0252	0.0033
CHFUSD	-0.0279*	-0.0156	0.0123	0.0024	-0.0169	-0.0516
CZKUSD	0.0116	0.0152	0.0407	0.0391	0.0409	-0.0537
EURUSD	-0.0112	0.0015	0.0312	0.0502	0.1393**	-0.0149
GBPUSD	-0.0193	-0.0033	0.0121	-0.0375	-0.0559*	-0.0052
HUFUSD	-0.0196	-0.0299*	-0.0239	-0.0012	0.0046	-0.0210
INRUSD	-0.0651*	-0.0657***	0.2311***	0.1995***	0.2657***	0.0942
JPYUSD	-0.0107	-0.0271*	-0.0247	-0.0245	0.0504	0.0075
MXNUSD	-0.0077	-0.0392**	0.0531	-0.0274	-0.0674	-0.1037
NOKUSD	0.0077	-0.0055	-0.0049	-0.0025	0.0069	-0.0667
NZDUSD	0.0043	-0.0199	0.0142	0.0046	0.1042	-0.0051
PLNUSD	0.0576***	0.0275*	0.0292	0.0357	-0.0802	-0.0653
SEKUSD	-0.0011	-0.0165	0.0330	0.0355	0.0411	0.0272
ZARUSD	0.0005	-0.0519***	0.0868*	0.0467	0.1157	0.0712

*Significantly different than zero at 10% p value

**Significantly different than zero at 5% p value

***Significantly different than zero at 1% p value

It can be seen that the correlation coefficient proposed in this paper and its associated test are more powerful than the Pearson estimate. Notably, only in the case of New Zealander equities that the Pearson estimate was able to detect significant correlation without our proposed method not simultaneously finding any significant correlation, whilst for many equity indices only our proposal could find a significant coefficient when the Pearson estimate was not significantly different than zero. Another interesting finding is that even though most of the equity indices had positive dependence, this was not always the case. For example, US, UK and French equities bucked the trend and exhibited significant negative dependence at a 1-day forecast horizon and Russian equities had significant negative dependence at a 1-month forecast horizon. Interestingly enough, the markets that had significant negative 1-day dependence had significant positive 1-month dependence and the market that had significant negative 1-month dependence had significant positive 1-day dependence.

It can also be seen that the procedure proposed in this paper detected dependence in a mix of emerging and developed market currencies, with significant 1-day dependence being found in the Brazilian Real, Indian Rupee, Polish Zloty, Canadian Dollar and Swiss Franc in addition 1 month-dependence being found in the Euro and British Pound. Our proposal detected dependence in more currencies than the Pearson estimate could, though the Pearson estimate detected 1-day dependence on the Japanese Yen and Mexican Peso that was undetected by our proposal. However, in most cases that

the dependence found is statistically significant it is of lower magnitude than with Equities and therefore less pronounced and likely to be less exploitable for economic purposes.

To validate the power of our predictive framework, the machine learning version of Algorithm 2 (described in the “Technical Appendix”) has been applied to the data described in the beginning of this section and the predictions were translated into investment decisions: buy (when the predicted sign is positive) and sell (when the predicted sign is negative). The trading strategies were tested assuming the following: (i) there are no short sale restrictions—which is most likely the case if only futures are used to trade; and (ii) if a shortable ETF is used, its borrow cost is no more than the interest on the cash received from the short sale proceeds. Initially the markets are assumed frictionless but this assumption is relaxed at the end of this section by performing a sensitivity analysis to transaction costs.

To ascertain the out-of-sample performance of the trading models proposed, five different performance metrics were calculated for all strategies and compared against the same values of a Buy and Hold benchmark portfolio. The Buy and Hold benchmark portfolio consisted of an equal weighted long position across all 21 equity indices in the case of equity strategies and an equal weighted long position across all 16 foreign currencies in the case of currency strategies. As the rollover rates are being included in the calculation of the returns, the currencies benchmark can be interpreted as a global unhedged money market investment portfolio from a US investor’s perspective.

The performance metrics used were:

- Average monthly out-of-sample returns of a portfolio that invested equally weighted across all equities and currencies according to the underlying classification strategy, per asset class per classifier;
- Standard Deviation of the out-of-sample monthly returns of the said portfolio;
- Sharpe Ratio, not adjusted by the risk-free rate: simple average of the out-of-sample monthly returns divided by their respective standard deviations; and
- Maximum Drawdown: the maximum loss from a peak to trough of the cumulative monthly portfolio returns before a new peak is attained.
- Alpha, not adjusted by the risk-free rate: the intercept of a linear regression between the returns of the strategy and the returns of the Buy and Hold portfolio.

The returns used to calculate the Sharpe Ratio and Alpha were not adjusted by the risk-free rate to keep the implementation simple and avoid an extra choice of variable (i.e. the choice of what constitutes a risk-free rate). The lack of adjustment does not change the conclusions of the present section of the paper given that these performance metrics are being used solely to compare the performance of the different strategies against the Buy and Hold without making a more general assessment of the economic value of such strategies. All performance measures are also accompanied by their standard errors calculated using the leave-one-out Jackknife method. Additionally, the Alpha was tested for statistical significance and the Type I Error probability of the Hypothesis Test that the Alpha is greater than zero is shown on all comparative performance tables.

The trading strategy given by Theorem 5 has been empirically tested using the MA and DMA predictors to obtain out-of-sample predictions for the period corresponding

Table 3 Comparative performance of the strategies applied to equity indices on daily data, gross of transaction costs

Strategy	B&H	MA	DMA
Monthly avg return	0.31% (0.60%)	1.07% (0.24%)	1.25% (0.22%)
Monthly std deviation	6.27% (0.68%)	2.48% (0.56%)	2.26% (0.59%)
Sharpe ratio	0.0502 (0.0999)	0.4312 (0.0833)	0.5516 (0.0867)
Max drawdown	56.2% (19.2%)	6.3% (2.9%)	3.7% (2.4%)
Alpha	N/A	0.0111 (0.0023)	0.0128 (0.0020)
<i>p</i> value (Alpha > 0)	N/A	< 0.0001	< 0.0001

Standard errors in parenthesis

to the months from April-2008 to May-2017, equivalent to the second half of the entire dataset used in this study. As our predictive method only predicts one step ahead, after a prediction is made the in-sample period gets the actual return observed one step ahead appended to it whilst the first in-sample point is discarded, so that in the next iteration the algorithm is over the new in-sample period. This is repeated until the end of the out-of-sample period. Tables 3, 4, 5, 6, 7 and 8 list the returns gross of transaction costs; however, we later incorporate transaction costs into our analysis.

The strategies proposed obtained their best performance when trading equities using a 1-day forecast horizon. As Table 3 shows, the MA strategy was able to get a monthly return that was more than three times the return of the Buy and Hold strategy, with a monthly standard deviation that was almost one third and, more importantly, a maximum drawdown that was less than one sixth, meaning that the tail behaviour was even better than the average behaviour. The resulting Sharpe Ratio, when annualised, is greater than one, a level taken as very good for the practitioners. The DMA strategy, that takes both trend-following and contrarian positions, performed even better, with an 18bps gain in expected return over the MA and a reduced maximum drawdown of just 3.7%, a fraction of the Buy and Hold maximum drawdown. The resulting Sharpe Ratio, when annualised, is very high and close to two. This can be seen as evidence that, on a 1-day trading interval, the introduction of contrarian positions in the DMA strategy can have a significant improvement over the pure trend-following MA strategy. However, when applied to foreign exchange markets, both the MA and DMA trading strategies had returns close to zero with very small standard errors. This is driven by the fact that not many currencies had displayed significant correlation (as per Table 2) and the trading model only enters into positions when the correlation detected is statistically significant.

As demonstrated on Table 5, there was a reduction in performance, gross of transaction costs, when the forecast horizon lengthened from 1 day to 1 week. However, as Table 9 shows, there was a commensurate reduction in trading activity, meaning that depending on the level of transaction costs, it might be more advantageous to trade using a 1 week forecast horizon (as opposed to a 1 day forecast horizon). Whilst the average monthly return of the strategies is no longer statistically different than the Buy and Hold, they have considerable lower risk which translates into a significantly positive Alpha. The DMA trading strategy did not have a significant difference in

Table 4 Comparative performance of the strategies applied to currencies on daily data, gross of transaction costs

Strategy	B&H	MA	DMA
Monthly avg return	- 0.10% (0.28%)	- 0.03% (0.04%)	0.00% (0.03%)
Monthly std dev	2.93% (0.25%)	0.43% (0.04%)	0.33% (0.04%)
Sharpe Ratio	- 0.0355 (0.0964)	- 0.0776 (0.0984)	- 0.0208 (0.0990)
Max Drawdown	28.5% (12.1%)	6.7% (3.5%)	3.7% (2.5%)
Alpha	N/A	- 0.0004 (0.0004)	- 0.0001 (0.0003)
p value (Alpha > 0)	N/A	0.8296	0.6044

Standard errors in parenthesis

Table 5 Comparative performance of the strategies applied to equity indices on weekly data, gross of transaction costs

Strategy	B&H	MA	DMA
Monthly avg return	0.34% (0.61%)	0.35% (0.22%)	0.28% (0.17%)
Monthly std dev	6.33% (0.66%)	2.28% (0.41%)	1.77% (0.36%)
Sharpe ratio	0.0533 (0.0993)	0.1539 (0.0894)	0.1595 (0.0757)
Max drawdown	56.2% (19.2%)	8.9% (5.7%)	6.6% (2.9%)
Alpha	N/A	0.0040 (0.0020)	0.0031 (0.0016)
p value (Alpha > 0)	N/A	0.0240	0.0245

Standard errors in parenthesis

Table 6 Comparative performance of the strategies applied to currencies on weekly data, gross of transaction costs

Strategy	B&H	MA	DMA
Monthly avg return	- 0.09% (0.27%)	- 0.03% (0.03%)	0.01% (0.03%)
Monthly std dev	2.82% (0.25%)	0.52% (0.07%)	0.32% (0.05%)
Sharpe ratio	- 0.0318 (0.0969)	- 0.0573 (0.1034)	0.0047 (0.0993)
Max drawdown	28.5% (11.4%)	1.9% (1.1%)	1.6% (0.7%)
Alpha	N/A	- 0.0003 (0.0005)	0.0000 (0.0003)
p value (Alpha > 0)	N/A	0.7682	0.5112

Standard errors in parenthesis

risk-adjusted performance compared to the MA strategy, with both obtaining similar Sharpe Ratios and Alphas. This can be seen as evidence that, in longer forecast horizons, equity markets have mostly positive dependence and trend-following trading strategies alone will capture this dependency, with the inclusion of contrarian trading not adding economic value.

In the case of weekly currency trading, Table 6 shows similar performance when compared to the daily strategy (gross of transaction costs), though the DMA strategy had a slight improvement. Like in the case of the daily frequency, the Alphas generated by currency trading are not significantly different than zero, though the maximum

Table 7 Comparative performance of the strategies applied to equity indices on monthly data, gross of transaction costs

Strategy	B&H	MA	DMA
Monthly avg return	0.30% (0.60%)	0.34% (0.27%)	0.34% (0.23%)
Monthly std dev	6.27% (0.68%)	2.36% (0.38%)	2.05% (0.23%)
Sharpe ratio	0.0473 (0.1002)	0.1174 (0.0910)	0.1165 (0.0905)
Max drawdown	56.2% (19.2%)	11.4% (7.9%)	8.2% (5.5%)
Alpha	N/A	0.0041 (0.0024)	0.0039 (0.0021)
<i>p</i> value (Alpha > 0)	N/A	0.0458	0.0305

Standard Errors in parenthesis

Table 8 Comparative performance of the strategies applied to currencies on monthly data, gross of transaction costs

Strategy	B&H	MA	DMA
Monthly avg return	-0.10% (0.28%)	-0.01% (0.04%)	0.00% (0.03%)
Monthly std dev	2.94% (0.25%)	0.53% (0.09%)	0.35% (0.07%)
Sharpe ratio	-0.0334 (0.0970)	-0.0113 (0.1035)	-0.0065 (0.1064)
Max drawdown	28.5% (12.1%)	5.1% (4.1%)	3.5% (2.1%)
Alpha	N/A	-0.0001 (0.0005)	0.0000 (0.0003)
<i>p</i> value (Alpha > 0)	N/A	0.5856	0.5621

Standard Errors in parenthesis

drawdown of the DMA strategy was much lower than the one of the Buy and Hold, indicating potential use of the model for risk management and forecasting purposes.

As shown in Table 7, the equity strategies did not provide monthly returns statistically different than the one of the Buy and Hold though they did have much reduced risk, again translating into a significantly positive Alpha. This means all equity trading strategies applied to all forecast horizons proposed produced significantly positive Alphas. As it can be seen on Table 9, there was a reduction in portfolio turnover when lengthening the forecast horizon from 1 week to 1 month (from about 75% of the portfolio per month to about 25% of the portfolio per month). This reduction is most likely not enough to justify the reduction in performance for investors with access to relatively sophisticated execution methods and low execution costs, but is likely to justify the implementation of a smart-beta/semi-passive strategy, as it will yield a similar return to Buy and Hold but with a considerable reduction in market risk and limited portfolio turnover.

We remark that the currency strategies with monthly trading (shown in Table 8) also did not produce an Alpha statistically different than zero, though the significant reduction in maximum drawdown was another feature present in this case. Therefore, we conclude that the use of the strategies proposed in currency markets is not likely to add economic value on its own, but there is potential risk management and forecasting value for the underlying correlation model.

Table 9 Average monthly turnover per strategy (times the portfolio size)

Equities						Currencies					
Daily		Weekly		Monthly		Daily		Weekly		Monthly	
MA	DMA	MA	DMA	MA	DMA	MA	DMA	MA	DMA	MA	DMA
9.81	11.32	0.94	0.61	0.30	0.18	2.06	0.92	0.55	0.31	0.11	0.05

All the results so far are gross of transaction costs. The previous steps have avoided embedding the costs as part of the model given that the actual transaction costs an investor has will depend massively on a large numbers of factors that are particular to each individual investor: whether market or limit orders are used, the order execution algorithm, the latency between the server issuing the order and the exchange receiving the order, the size of the order, the choice of broker, the level of commissions and rebates are all factors affecting the final cost.

In Anand et al. (2012) it is presented a thorough descriptive analysis of trades executed in the US equity markets by 664 institutions that traded 100 times or more per month during the time period from January 1, 1999 to December 31, 2005. The authors found that the difference between the costs of the lowest quintile and the highest quintile was as high as 69 basis points. They also found that institutions in the lowest quintile were able to get negative execution shortfalls, meaning their technology, order execution policy and broker arrangements was able to add value to the gross returns, instead of removing it. Similar findings are made by Brogaard et al. (2014) where the authors claim the lowest quartile of execution costs by institutional investors in UK equities was only 0.4 basis points whilst the top quartile had costs of 27.8 basis points. Therefore, instead of fixing a single cost number, we have chosen a set of possible costs per transaction and combined it with the out-of-sample metrics provided in Tables 3, 4 5, 6 7 and 8 to determine what would be return of a given strategy if it was subject to a given level of transaction costs. To make the numbers easier to compare between themselves and with Buy and Hold, we have rescaled the net returns for all strategies by the ratio of the strategy's monthly standard deviation and the Buy and Hold's monthly standard deviation—i.e. all returns, net of transaction costs, are risk adjusted so that all of them have the same market risk as the Buy and Hold strategy.

The level of transaction activity for each strategy was also calculated as this is an additional input needed to calculate the net returns. Table 9 shows how many times the portfolio was entirely rebalanced over a month. It can be seen that the Daily Equities MA strategy had the highest turnover, trading on average 11.319 times the entire portfolio in a typical month. On the other hand, the Monthly FX DMA strategy had the lowest turnover, trading only 5% of the portfolio per month. FX strategies naturally had lower turnovers as the test identified less evidence of serial correlation in currencies than in equities, so it traded them less.

Table 10 shows the returns of the strategies applied to equities. The daily strategies have the best returns for investors that can achieve low trading costs. In particular, if one can achieve trading costs of 5bps per trade (a level that can be achieved for the most

Table 10 Annualised return, net of costs, per equity trading strategy per level of costs

Cost per transaction	3 bps	5 bps	10 bps	15 bps	20 bps
Daily EQ MA	24.58%	18.16%	2.71%	-11.93%	-25.80%
Daily EQ DMA	31.79%	23.55%	3.83%	-14.70%	-32.11%
Weekly EQ MA	10.96%	10.31%	8.71%	7.12%	5.53%
Weekly EQ DMA	11.50%	10.96%	9.62%	8.28%	6.94%
Monthly EQ MA	9.04%	8.88%	8.46%	8.05%	7.63%
Monthly EQ DMA	10.87%	10.75%	10.46%	10.16%	9.87%
Annualised Return of Buy and Hold gross of any transaction costs					3.87%

Table 11 Annualised return, net of costs, per currency trading strategy per level of costs

Cost per transaction	2 bps	4 bps	6 bps	10 bps	15 bps
Daily FX MA	-6.10%	-9.47%	-12.82%	-19.47%	-27.71%
Daily FX DMA	-2.67%	-4.60%	-6.53%	-10.37%	-15.15%
Weekly FX MA	-2.66%	-3.37%	-4.08%	-5.50%	-7.27%
Weekly FX DMA	-0.48%	-1.13%	-1.77%	-3.05%	-4.65%
Monthly FX MA	-0.55%	-0.69%	-0.84%	-1.13%	-1.50%
Monthly FX DMA	-0.33%	-0.43%	-0.53%	-0.73%	-0.98%
Annualised Return of Buy and Hold gross of any transaction costs					-1.16%

sophisticated investors), the EQ DMA strategy would have obtained an annualised return of more than 20% for the same level of risk that the Buy and Hold would have obtained an annualised return of 3.87%. However, even with trading costs as high as 20bps per trade (a level that is attainable even for savvy retail investors), trading the DMA strategy on a monthly basis can achieve an annualised return in the region of 10% for the same level of risk of the Buy and Hold.

Table 11 shows the returns of the strategies applied to currency trading. Given that no strategy could have a positive Alpha, it is also seen that no strategy could have a positive annualised return net of costs, though the Buy and Hold return was also negative for the period. Therefore, these strategies when applied to currencies cannot provide absolute returns on their own. Nevertheless, it is also worth noting that the lower maximum drawdown indicates some potential use of the underlying model for Global Fixed Income portfolios, where the foreign carry interest is also a component of the return, and the strategies would be acting as currency risk mitigants.

6 Concluding Remarks

This research has proposed a novel framework to deal with unstructured serial dependence by introducing a measure of dependence that only focuses on the behaviour of the sign of changes in the process. This definition also has the power to simplify the interpretation of models with several parameters like ARFIMA by establishing

boundaries in the values of the parameters that can be interpreted as simply positive or negative dependence.

The definition was used to propose a statistical inference framework to detect unstructured serial dependence, which was validated both via simulation in synthetic models and in empirical data. In the case of empirical data for equity indices, the framework was able to detect significant dependence that could not be detected using a linear correlation model and was also able to feed a quantitative trading model that provided risk adjusted returns which were significantly superior to the ones of a Buy and Hold strategy and also robust to the presence of transaction costs.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Technical Appendix

A Proofs of All Theorems in the Main Paper

The derivation of necessary conditions for signed path dependence in several parametric models depends on the lemma below:

Lemma 4 *Let $x_{1 \leq i \leq n}$ be independent and normally distributed random variables with zero mean and variance s_i^2 . For a fixed j , $1 \leq j \leq n$, given weights $b_{1 \leq i \leq n}$ we have that*

$$\mathbb{E} \left[x_j \mid \sum_{1 \leq i \leq n} b_i x_i \right] = b_j \frac{s_j^2}{\sum_{1 \leq i \leq n} b_i^2 s_i^2} \sum_{1 \leq i \leq n} b_i x_i \tag{A.1}$$

Lemma 4 allows us to validate the definition of positive dependence in a number of well known time series model classes, as long as the innovations are Gaussian. We note that extensions can be developed for other classes of driving white noise sequences, including heavy tailed cases.

Proof Define $z = \sum_{1 \leq i \leq n} x_i$. Since the x_i are independent, z is normally distributed with zero mean and variance $\sum_{1 \leq i \leq n} s_i^2$, hence (x_i, z) are jointly normally distributed with correlation $\rho_{x_i, z} = \frac{s_i}{\sqrt{\sum_{1 \leq i \leq n} s_i^2}}$. From this result it follows that

$$\mathbb{E} \left[x_j \mid \sum_{1 \leq i \leq n} x_i \right] = \frac{s_j^2}{\sum_{1 \leq i \leq n} s_i^2} \sum_{1 \leq i \leq n} x_i,$$

which implies the equality that was to be proven. □

We now provide all the subsequent proofs.

Proof of Theorem 2. Observe the following identity

$$\begin{aligned}
 r_{+h}(t) &= \sum_{i=1}^h \sum_{j=0}^{\infty} b_j \varepsilon(t+i-j) \\
 &= \sum_{i=0}^{h-1} \left(b_i \sum_{j=1}^{h-i} \varepsilon(t+j) \right) + \sum_{i=h}^{\infty} \sum_{j=1}^h b_{i+j-h} \varepsilon(t-i-j+h+1),
 \end{aligned}
 \tag{A.2}$$

implying that

$$\mathbb{E}[r_{+h}(t) | r_{-n}(t)] = \sum_{i=h}^{\infty} \sum_{j=1}^h b_{i+j-h} \mathbb{E}[\varepsilon(t-i-j+h+1) | r_{-n}(t)].
 \tag{A.3}$$

Similarly, the following identity holds

$$\begin{aligned}
 r_{-n}(t) &= \sum_{i=0}^{n-1} \sum_{j=i}^{\infty} b_j \varepsilon(t-i-j) \\
 &= b_0 \varepsilon(t) + (b_0 + b_1) \varepsilon(t-1) \\
 &\quad + \dots + \varepsilon(t-n+1) \sum_{i=0}^{n-1} b_i + \varepsilon(t-n) \sum_{i=1}^n b_i + \dots \\
 &= \sum_{i=0}^{n-1} \left(\sum_{j=0}^i b_j \right) \varepsilon(t-i) + \sum_{i=n}^{\infty} \left(\sum_{j=i-n+1}^i b_j \right) \varepsilon(t-i) \\
 &= \sum_{i=0}^{\infty} \left(\varepsilon(t-i) \sum_{j=\max(0, i-n+1)}^i b_j \right).
 \end{aligned}
 \tag{A.4}$$

Which implies that

$$\begin{aligned}
 &\mathbb{E}[r_{+h}(t) | r_{-n}(t)] \\
 &= \sum_{i=h}^{\infty} \sum_{j=1}^h b_{i+j-h} \mathbb{E} \left[\varepsilon(t-i-j+h+1) \left| \sum_{i=0}^{\infty} \left(\varepsilon(t-i) \sum_{j=\max(0, i-n+1)}^i b_j \right) \right. \right].
 \end{aligned}
 \tag{A.5}$$

Now, using the result from Lemma 4 yields the following expression

$$\mathbb{E}[r_{+h}(t) | r_{-n}(t)] = \frac{\sum_{i=1}^h \sum_{j=h}^{\infty} b_{i+j-h} \sigma_{t-i-j+h+1}^2}{\sum_{i=0}^{\infty} \sigma_{t-i}^2 \left(\sum_{j=\max(0, i-n+1)}^i b_j \right)^2} r_{-n}(t).
 \tag{A.6}$$

Now define $\Psi(h, t) = \sum_{i=1}^h \sum_{j=h}^{\infty} b_{i+j-h} \sigma_{t-i-j+h+1}^2$. As the denominator of the expression in Equation A.6 will always be positive, it follows that the sign of $\mathbb{E}[r_{+h}(t)|r_{-n}(t)]$ will be entirely determined by the sign of $r_{-n}(t) \Psi(h, t)$. As such, if $\Psi(h, t) > 0$ for all t the process will have positive sign dependence and if $\Psi(h, t) < 0$ for all t the process will have negative sign dependence. \square

Proof of Lemma 1. From Equation 2.3 we have that

$$r(t) = \sum_{j=0}^{\infty} \varphi^j \varepsilon(t-j). \tag{A.7}$$

Now, let $\Psi(h, t) = \sigma^2 \sum_{i=1}^h \sum_{j=h}^{\infty} \varphi^{i+j-h}$. It trivially follows that if $\varphi > 0$ then $\Psi(h, t) > 0$ and hence from Theorem 2 it follows that the process will have positive path dependence irrespective of n and h . Equally, if $\varphi = 0$ then $\Psi(h, t) = 0$ and as such the process will have no signed path dependence.

Also notice that

$$\Psi(h, t) = \sigma^2 \sum_{i=1}^h \frac{\varphi^i}{1-\varphi} = \sigma^2 \varphi \frac{(1-\varphi^h)}{(1-\varphi)^2}. \tag{A.8}$$

As whenever $0 > \varphi > -1$ we have that $1 - \varphi^h > 0$, it follows that whenever $0 > \varphi > -1$, $\Psi(h, t) < 0$ irrespective of h and therefore, from Theorem 2 it follows that the process will have negative path dependence. \square

Proof of Lemma 2. According to Granger and Joyeux (1980), an ARFIMA (p, d, q) process can also be written as

$$\Phi(B)r(t) = \Theta(B) \left(\sum_{k=0}^{\infty} \frac{\Gamma(k+2d)}{\Gamma(k+1)\Gamma(2d)} \varepsilon(t-k) \right) \tag{A.9}$$

and therefore we have that

$$\begin{aligned} r(t) &= \frac{\Theta(B)}{\Phi(B)\Gamma(2d)} \sum_{k=0}^{\infty} \frac{\Gamma(k+2d)}{\Gamma(k+1)} \varepsilon(t-k) \\ &= \sum_{k=0}^{\infty} \frac{\Gamma(k+2d)}{\Gamma(k+1)} \frac{\sum_{j=0}^{\infty} b_j \varepsilon(t-j)}{\Gamma(2d)} \end{aligned} \tag{A.10}$$

If $d > -0.5$ then $\sum_{k=h}^{\infty} \frac{\Gamma(k+2d)}{\Gamma(k+1)} > 0$ and $\text{sgn}(\Gamma(2d)) = \text{sgn}(d)$. Therefore, using the same derivation as per the proof of Theorem 2, it follows that given $h \in \mathbb{N}^+$ the sign of $d\Psi(h)$ will determine whether the process has positive or negative path dependence for all n .

Additionally, when $p = q = 0$, $\Phi(B) = \Theta(B) = 1$ and hence the process will have positive path dependence for all h and n whenever $d > 0$ and negative path dependence whenever $-0.5 < d < 0$. \square

Proof of Theorem 4. As $\omega(t + 1) = \frac{1}{n} \sum_{k=1}^n \text{sgn}(r_{-k}(t))$, we have that

$$\mathbb{E}[d(t)] = \sum_{k=1}^n \mathbb{E}[\text{sgn}(r_{-k}(t-1))r(t)] \tag{A.11}$$

and hence under the null hypothesis it follows that $\mathbb{E}[d(t)] = 0$.

Further, given that $r(t)$ is assumed covariance stationary, by Theorem 2 we know that the value of $\text{sgn}(r_{-k}(t)) \mathbb{E}[r(t+1)|r_{-k}(t)]$ is the same for all k and hence if

$$\text{sgn}(r_{-k}(t)) \mathbb{E}[r(t+1)|r_{-k}(t)] = 0 \tag{A.12}$$

for all $k \leq n$, then the equality will also hold for all $k > n$. Therefore, under the null hypothesis we have that

$$\mathbb{E}[\text{sgn}(r_{-k}(t))r(t+1)] = 0 \tag{A.13}$$

for all values of k which implies that $d(t)$ is uncorrelated with $d(t+i)$ for all $i > n$, hence being n -dependent.

Additionally, as $0 < |\omega(t)| \leq 1$, we have that

$$\begin{aligned} \mathbb{E}\left[(d(i))^2 \mathbb{I}\{|d(i)| > \epsilon\sqrt{n}\}\right] &= \mathbb{E}\left[(r(i)\omega(i))^2 \mathbb{I}\{|r(i)| > \frac{\epsilon}{|\omega(i)|}\sqrt{n}\}\right] \\ &\leq \mathbb{E}\left[(r(i))^2 \mathbb{I}\{|r(i)| > \frac{\epsilon}{|\omega(i)|}\sqrt{n}\}\right] \end{aligned} \tag{A.14}$$

implying that $d(t)$ also satisfies the Lindeberg condition given that $r(t)$ is assumed to satisfy it.

Therefore, as per Theorem 3, we have that under the null hypothesis, when $s \rightarrow \infty$ the statistic defined in Equation 3.2 will converge in distribution to a standard normal normal variable as $\hat{\gamma}$ defined in Equation 3.3 is the sample estimator of the variance of \bar{d} . □

Proof of Lemma 3. Let h_α be the $1 - \frac{\alpha}{2}$ quantile of the true distribution of the test statistic given by Equation (3.2) under the null hypothesis stated in the second paragraph of Section 3. As per Efron and Tibshirani (1993) (p. 187), $\mathbb{P}[h_\alpha > \hat{h}_\alpha] = \frac{\alpha}{2} + \frac{c}{\sqrt{s}}$ where c is a constant value. Hence when $s \rightarrow \infty$ we have that $\mathbb{P}[h_\alpha > \hat{h}_\alpha] \rightarrow \frac{\alpha}{2}$.

Now, as per the Theorem 4 the true distribution of the test statistic given by Equation (3.2) converges to a standard normal as $s \rightarrow \infty$. Therefore $h_\alpha \rightarrow \Phi^{-1}(1 - \frac{\alpha}{2})$ and hence $\mathbb{P}[\hat{h}_\alpha < \Phi^{-1}(1 - \frac{\alpha}{2})] \rightarrow \frac{\alpha}{2}$, meaning that $\hat{h}_\alpha \rightarrow \Phi^{-1}(1 - \frac{\alpha}{2})$. □

Proof of Theorem 5. If the null hypothesis of the test given by Theorem 4 is not true then there is a value of $\hat{\zeta} \in \{-1, 1\}$ such that $\hat{\zeta} = \text{sgn}(\mathbb{E}[\text{sgn}(r_{-n}(t) - n\bar{r})r(t+1)] - \bar{r})$. In that case, we have that if $\hat{\zeta} = -1$ then $\mathbb{E}[-\text{sgn}(r_{-n}(t) - n\bar{r})r(t+1)] > \bar{r} \implies \mathbb{E}[L_{\widehat{s(t+1)}}] < \mathbb{E}[L_{B\&H}]$ as $\mathbb{E}[L_{B\&H}] = -\bar{r}$. Equally, if $\hat{\zeta} = 1$ then $\mathbb{E}[\text{sgn}(r_{-n}(t) - n\bar{r})r(t+1)] > \bar{r} \implies \mathbb{E}[L_{\widehat{s(t+1)}}] < \mathbb{E}[L_{B\&H}]$. □

B Machine Learning Estimation Method

This section describes a machine learning method that reads a data stream of log-returns in chronological order $t = 1, 2, \dots$ and for each $r(t)$ read it returns either $s(\widehat{t + 1})$ or zero if it cannot make a reliable prediction for this sign, which could happen either because there is not enough observed information up until time t to make such prediction or because the test defined in Theorem 4 could not find evidence of signed path dependence in the series provided up until that point.

The section alternates between pseudocode, in smaller font to make it distinguishable, and descriptive text. Reading this section requires some knowledge of object-oriented programming concepts as the pseudocode listed comprises of an object-oriented implementation aimed to be ported into languages such as C++ and Python. All required variables are listed and commented in the pseudocode.

A single class object is described in this section. Such object, once fully initialised, should be run by sequentially invoking method *Classify* every time a new data point is observed, obtaining either $s(\widehat{t + 1})$ or zero. For a practical application, the output of the method *Classify* is translated into a trading strategy with three decisions: buy, sell or take no position.

All class variables only require private visibility and are listed as follows:

- className* the name of the classifier being implemented (either MA or DMA)
- R* a list that will contain a sample of observed log-returns for a given asset at a given trading frequency
- P* a list that will contain positive integer numbers candidate values for the parameter n of the classifier being implemented
- z an integer number corresponding to the desired learning sample size, assumed even for simplicity
- α the significance threshold for the application of the signed path dependence test
- w an integer number corresponding to a retuning interval
- v an integer number corresponding to a retraining interval
- n' an integer number corresponding to a test lookback horizon
- n_{max} the optimal value of the classifier lookback for prediction
- bTrade* indicating whether the asset is worth trading or not
- \bar{r} the long term unconditional mean of the return process
- ζ the sign of the path dependence being assumed by the predictive method
- countSinceTune* storing how many data points have been read since the last time the classifier was tuned
- countSinceTrain* storing how many data points have been read since the last time the classifier was trained

The class object contains four different methods: Initialise, Classify, Predict and Update.

Method **Initialise** is implemented as below:

```

procedure INITIALISE( $P = (n_1, n_2, \dots, n_N), z > 0, 1 > \alpha > 0, w > 0, v > 0, n' > 0, className$ )
  this.R  $\leftarrow \emptyset$             $\triangleright R$  is initialised as the empty set and will be built as new data is observed
  this.P  $\leftarrow P; this.z \leftarrow z; this.\alpha \leftarrow \alpha; this.w \leftarrow w; this.v \leftarrow v; this.className \leftarrow className$ 
  this.countSinceTune  $\leftarrow 0; this.countSinceTrain \leftarrow 0; this.n' \leftarrow n'$ 
end procedure
    
```

Description of method **Initialise**: This method is the class constructor and will be called whenever an instance of the class is created. It stores into internal memory all parameters that are required by subsequent methods and initialises other relevant internal variables. Creating an instance of this class will always require the following parameters:

- *className*: one of MA or DMA, representing the classifier being implemented by this instance of the class—notice that it is a decision of the user whether only sign continuation is to be assumed or if a dynamic decision between continuation and reversals is needed.
- *P*: a list of integers containing acceptable values for the parameter *n* of the classifier being implemented (the actual choice of *n* will be performed by method *Update*).
- *z*: an integer corresponding to the desired learning sample size, to be used by methods *Classify* and *Update*.
- α : the significance threshold to be used by method *Update*.
- *w*: an integer corresponding to a retuning interval, to be used by method *Update*.
- *v*: an integer corresponding to a retraining interval, to be used by method *Update*.
- *n'*: an integer corresponding to a test lookback horizon, to be used by method *Update*.

Method **Classify** is implemented as below:

```

function CLASSIFY(r (t))
  Append r (t) to the end of list this.R
  if length(this.R) > this.z then
    Drop the first element of list this.R
  end if
  if length(this.R) = this.z then
    Call this.Update()
    if bTrade = TRUE then
      return this.Predict(this.R, this.F, this.nmax)
    end if
  end if
  return 0
end function

```

Description of method **Classify**: This is the only public method of the class. It should be called every time *t* a new data point has become available and returns either $\widehat{s}(t + 1)$ or zero, which also can be taken as a trading decision: -1 meaning *negative/sell*; 1 meaning *positive/buy*; 0 meaning *no predicted sign/take no position*. The only parameter required to run this method is:

- *r* (*t*): the observed data point (log-return).

The method stores *r* (*t*) in an internal class variable, in sequential order so that it can be used in subsequent iterations and updates of the model. However, if after storing *r* (*t*) the internal storage has $z + 1$ points (with *z* given during initialisation), the method will drop the oldest (first) data point of the internal storage to ensure that the learning sample size stays constant at *z*.

If there are less than *z* data points in the internal storage, the method simply returns zero, not attempting any prediction until more data is received to fill up the internal storage. Otherwise, it calls the internal method *Update* which will learn the optimal values for parameters *n* and \bar{r} given the data in the internal storage and save to internal memory the results of the test defined in Theorem 4.

Finally, if the test results saved in memory inform that there is no significant path dependence in the series saved to the class internal storage, the method will again simply return zero. Otherwise, it will return the result of calling the internal method

Predict giving the whole data in the internal storage as the parameter R and the parameters n and \bar{r} as obtained in the method *Update*.

Method **Update** is implemented as below:

```

procedure UPDATE()
  this.countSinceTune ← this.countSinceTune + 1
  this.countSinceTrain ← this.countSinceTrain + 1
  if this.countSinceTune =  $w$  then                                ▷ The tuning stage is performed inside this block
    this. $\bar{r}$  ←  $\frac{1}{z} \sum_{i=1}^z this.R(i)$ 
    if this.className = "MA" and  $\frac{\bar{d}}{\sqrt{\frac{\sum_{\tau=-this.n'}^{this.n'} \hat{\gamma}(\tau)}{s-this.n'}}} > \Phi^{-1}(1 - this.\alpha)$  then
      this.bTrade ← TRUE
      this. $\zeta$  ← 1
    else if this.className = "DMA" and  $\left| \frac{\bar{d}}{\sqrt{\frac{\sum_{\tau=-this.n'}^{this.n'} \hat{\gamma}(\tau)}{s-this.n'}}} \right| > \Phi^{-1}\left(1 - \frac{this.\alpha}{2}\right)$  then
      this.bTrade ← TRUE
      this. $\zeta$  ← 1 if  $\frac{\bar{d}}{\sqrt{\frac{\sum_{\tau=-this.n'}^{this.n'} \hat{\gamma}(\tau)}{s-this.n'}}} > \Phi^{-1}\left(1 - \frac{this.\alpha}{2}\right)$  or -1 if  $\frac{\bar{d}}{\sqrt{\frac{\sum_{\tau=-this.n'}^{this.n'} \hat{\gamma}(\tau)}{s-this.n'}}} < \Phi^{-1}\left(\frac{this.\alpha}{2}\right)$ 
    else
      this.bTrade ← FALSE                                ▷ The asset is not worth trading
    end if
    this.countSinceTune ← 0
  end if
  if this.countSinceTrain =  $v$  then                                ▷ The training stage is performed inside this block
     $\bar{r}_1$  ←  $\frac{2}{z} \sum_{i=1}^{z/2} this.R(i)$ 
    this.nmax ←  $\arg \max_{n \in P} \frac{2}{z} \sum_{i=\frac{z}{2}+1}^z \exp(this.R(i) \times this.Predict((this.R(i - z/2), \dots, this.R(i - 1)), \bar{r}_1, n))$ 
    this.countSinceTrain ← 0
  end if
end procedure
  
```

Description of method **Update**: This is an internal method that performs the learning process of the class, split into two stages:

- *Tune*: This stage decides if any reliable prediction can be made and sets other relevant internal variables of the class. To enhance runtime performance, the *Update* method runs this stage only once every w iterations (with w given during initialisation) and relies on results from previous runs (stored in internal memory) when this stage is not run. This stage uses all data points in the internal storage of the class to decide if the asset is worth trading by checking for significant path dependence as follows:
 - If the classifier being implemented by this instance of the class is the MA classifier, the test defined in Theorem 4 is applied to all stored data points to look for significant positive path dependence only, and the asset is deemed worth trading if such dependence is found.
 - Alternatively, if the classifier being implemented by this instance of the class is the DMA classifier, the test defined in Theorem 4 is applied to all stored

data points to look for significant path dependence of any sign which, if found, would make the asset worth trading.

If the asset is deemed worth trading, the method estimates its long term mean return (and stores it internally for future usage) as the observed mean return over all data points in the internal storage of the class.

- *Train*: This stage decides the actual value of n that should be used for classification. To enhance runtime performance, the *Update* method runs this stage only once every v iterations (with v given during initialisation) and relies on results from previous runs (stored in internal memory) when this stage is not run. This stage splits the data points in the internal storage into two subsamples of equal length: *training* (the first chronological half of the stored list) and *validation* (the second chronological half of the stored list, henceforth referred to as $R_{validation}$). Using the mean return over the *training* sample as an input to the classifier being trained, it finds which value of n produces the highest classification gain over the *validation* sample, with the gain defined as:

$$G(n) = \exp \left(\sum_{r(t) \in R_{validation}} r(t) \times \widehat{s}(t) \right) \quad (\text{B.1})$$

where $\widehat{s}(t)$ is obtained by method *Predict* using $R = (r(t-1), r(t-2), \dots, r(t-z/2))$, \bar{r} equal to the mean return over the *training* sample and n as the lookback parameter.

All learned parameters are stored in the private class variables for subsequent usage.

Remark 9 The gain function chosen has the advantage of being linear on the returns, which means it can also be used to evaluate the returns of a portfolio comprised of several assets invested according to the underlying classification strategy.

And, finally, method **Predict** is implemented as below:

```
function PREDICT( R,  $\bar{r}$ , n )
  return this. $\zeta$   $\times$  sgn  $\left( \sum_{i=0}^{n-1} (R(\text{length}(R) - i) - \bar{r}) \right)$ 
end function
```

Description of method **Predict**: This is an internal method that performs the actual prediction as a direct one step ahead implementation of Definition 4 if *className* = "MA" or as per Definition 5 if *className* = "DMA", returning either -1 or 1 . It receives three inputs:

- R : a list of log-returns in chronological order, assumed to be the most recent returns up to the point that the method *Predict* was called.
- \bar{r} : the long term mean log-return of the asset.
- n : the value of the lookback parameter.

In case this instance of the class is a DMA classifier, the value of $\hat{\zeta}$ will have already been obtained and stored internally by method *Update* by the time this method is called.

In the empirical experiment of Sect. 5, the tuning stage of the algorithm was done every 6 months, which meant a retune was done after six new monthly data points, 24 new weekly data points and 126 new daily data points. The training stage was done monthly, which meant a retrain was done for every new point for monthly data, after 4 new points for weekly data and after 21 new points for daily data.

References

- Anand, A., Irvine, P., Puckett, A., & Venkataraman, K. (2012). Performance of institutional trading desks: An analysis of persistence in trading cost. *The Review of Financial Studies*, 25, 557–598.
- Baltas, A. N., & Kosowski, R. (2013). Momentum strategies in futures markets and trend-following funds. Paper presented at the 2012 finance meeting EUROFIDAI-AFFI
- Banerjee, A. N., & Hung, C. H. D. (2013). Active momentum trading versus passive “naive diversification”. *Quantitative Finance*, 13, 655–663.
- Brogaard, J., Hendershott, T., Hunt, S., & Ysusi, C. (2014). High-frequency trading and the execution costs of institutional investors. *The Financial Review*, 49, 345–369.
- Chena, Y., & Huainan, Z. (2012). Informed trading, information uncertainty, and price momentum. *Journal of Banking and Finance*, 36, 2095–2109.
- Choudhry, R., & Garg, K. (2008). A hybrid machine learning system for stock market forecasting. *World Academy of Science, Engineering and Technology: International Journal of Computer and Information Engineering*, 2, 689–692.
- Chung, C. F. (1996). Estimating a generalized long memory process. *Journal of Econometrics*, 73, 237–259.
- Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223–236.
- Corte, P. D., Kosowski, R., & Wang, T. (2016). Market closure and short-term reversal. Paper presented at the Asian Finance Association (AsianFA) 2016 conference
- CSzakmary, A., Shen, Q., & Sharmac, S. C. (2010). Trend-following trading strategies in commodity futures: A re-examination. *Journal of Banking and Finance*, 34, 409–426.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. London: Chapman and Hall.
- Fama, E. (1963). Mandelbrot and the stable paretian hypothesis. *Journal of Business*, 36, 420–429.
- Fama, E. (1965a). The behavior of stock market prices. *Journal of Business*, 38, 34–105.
- Fama, E. (1965b). Random walks in stock market prices. *Journal of Business*, 21, 55–59.
- Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25, 383–417.
- Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–29.
- Griffin, J. M., Kelly, P. J., & Nardari, F. (2010). Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. *Journal of Economic Surveys*, 23, 3225–3277.
- H Diananda, P. (1954). The central limit theorem for m-dependent variables asymptotically stationary to second order. *Mathematical Proceedings of the Cambridge Philosophical Society*, 50, 287–292.
- Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers and Operations Research*, 32, 2513–2522.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, 12, 1137–1143.
- Lee, J., & Hong, Y. (2001). Testing for serial correlation of unknown form using wavelet methods. *Econometric Theory*, 17, 386–423.
- Lesmonda, D. A., Schill, M. J., & Zhou, C. (2004). The illusory nature of momentum profits. *Journal of Financial Economics*, 71, 349–380.
- Lim, K. P., & Brooks, R. (2011). The evolution of stock market efficiency over time: A survey of the empirical literature. *Journal of Economic Surveys*, 25, 69–108.
- Lim, K. P., Luo, W., & Kim, J. H. (2013). Are us stock index returns predictable? Evidence from automatic autocorrelation-based tests. *Applied Economics*, 45(8), 953–962.

-
- Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, *104*, 228–250.
- Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics*, *103*, 429–453.
- Peters, G. W., & Shevchenko, P. V. (2015). *Advances in heavy tailed risk modeling: A handbook of operational risk*. New York: Wiley.
- Platen, E., & Heath, D. (2006). *A benchmark approach to quantitative finance*. New York: Springer.
- Qian, B., & Rasheed, K. (2004). Hurst exponent and financial market predictability. Paper presented at the 2004 IASTED conference on financial engineering and applications
- Qian, B., & Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence*, *26*, 25–33.
- Urquhart, A., & McGroarty, F. (2016). Are stock markets really efficient? Evidence of the adaptive market hypothesis. *International Review of Financial Analysis*, *47*, 420–429.
- Wold, H. (1954). *A study in the analysis of stationary time series*. Stockholm: Almqvist and Wiksell Book Co.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.