



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Journal of Financial Economics 73 (2004) 525–565

JOURNAL OF  
Financial  
ECONOMICS

[www.elsevier.com/locate/econbase](http://www.elsevier.com/locate/econbase)

# Testing market efficiency using statistical arbitrage with applications to momentum and value strategies<sup>☆</sup>

Steve Hogan<sup>a</sup>, Robert Jarrow<sup>b</sup>, Melvyn Teo<sup>c,d,\*</sup>,  
Mitch Warachka<sup>c</sup>

<sup>a</sup> *Credit Suisse First Boston, NY, 10010, USA*

<sup>b</sup> *Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853, USA*

<sup>c</sup> *Singapore Management University, School of Business, 469 Bukit Timah Road, S(259756), Singapore*

<sup>d</sup> *FDO Partners LLC, Cambridge, MA, 02318, USA*

Received 2 June 2003; accepted 2 October 2003

Available online 8 June 2004

---

## Abstract

This paper introduces the concept of statistical arbitrage, a long horizon trading opportunity that generates a riskless profit and is designed to exploit persistent anomalies. Statistical arbitrage circumvents the joint hypothesis dilemma of traditional market efficiency tests because its definition is independent of any equilibrium model and its existence is incompatible with market efficiency. We provide a methodology to test for statistical arbitrage and then empirically investigate whether momentum and value trading strategies constitute statistical arbitrage opportunities. Despite adjusting for transaction costs, the influence of small stocks, margin requirements, liquidity buffers for the marking-to-market of

---

<sup>☆</sup>We thank Warren Bailey, Peter Carr, Amiyatosh Purnanandam, Bill Schwert, Yiu-Kuen Tse, David Weinbaum, seminar participants at Cornell University and the Singapore Management University, and an anonymous referee for many helpful comments and suggestions. Mitch Warachka is grateful for financial support from the Wharton-SMU Research Center, Singapore Management University. Melvyn Teo is indebted to the Lee Foundation for financial support. Credit Suisse First Boston does not necessarily agree with or endorse the contents of this article.

\*Corresponding author. Singapore Management University, School of Business, 469 Bukit Timah Road, S(259756), Singapore. Tel.: +65-6822-0735; fax: +65-6822-0777.

E-mail address: [melvynteo@smu.edu.sg](mailto:melvynteo@smu.edu.sg) (M. Teo).

short-sales, and higher borrowing rates, we find evidence that these strategies generate statistical arbitrage.

© 2004 Elsevier B.V. All rights reserved.

*JEL classification:* G12; G14

*Keywords:* Statistical arbitrage; Market efficiency; Momentum; Value

---

## 1. Introduction

A large number of empirical studies conclude that stock prices appear to contradict the efficient markets hypothesis. For example, [Jegadeesh and Titman \(1993\)](#) investigate a trading strategy that buys well-performing stocks and sells poor-performing stocks. They show average excess returns of 12% per year, where excess returns are defined relative to a standard capital asset pricing model. [Lakonishok et al. \(1994\)](#) reach a similar conclusion via buying value and selling glamour stocks identified with variables such as price earnings ratios, dividends, book-to-market values, cash flows, and sales growth. [Chan et al. \(1996\)](#) confirm the excess returns of portfolios formed on the basis of past returns and earnings announcements. Evidence supports the existence of many other persistent anomalies such as those pertaining to dividends, earnings announcements, shares issuances, and share repurchases. [Shleifer \(2000\)](#) provides an excellent review of this literature.

However, all existing studies are compromised by the joint hypothesis problem. [Fama \(1998\)](#) cautions against rejecting market efficiency because previous tests are contingent upon a specific model for equilibrium returns. Therefore, abnormal returns need not imply the rejection of market efficiency. Instead, they could indicate that the equilibrium model is misspecified. According to [Fama \(1998\)](#), most long-term anomalies are also sensitive to the statistical methodology utilized and therefore inferences drawn using long-term returns (i.e., over five years) are circumspect. This paper offers a methodology that addresses Fama's statistical criticisms and circumvents the joint hypothesis problem.

The methodology proposed in this paper is based on the concept of a statistical arbitrage. We define a statistical arbitrage as a long horizon trading opportunity that generates a riskless profit. As such, statistical arbitrage is a natural extension of the trading strategies utilized in the existing empirical literature on persistent anomalies. Statistical arbitrage is the time-series analog of the limiting arbitrage opportunity contained in [Ross \(1976\)](#). More important, statistical arbitrage is defined without reference to any equilibrium model therefore, its existence is inconsistent with market equilibrium and, by inference, market efficiency (see [Jarrow, 1988, Chapter 19](#)). As such, statistical arbitrage rejects market efficiency without invoking the joint hypothesis of an equilibrium model. This paper subsequently provides a statistical test to determine whether a trading strategy constitutes statistical arbitrage based on its incremental trading profits computed over short time horizons. Our test for statistical arbitrage is intended to replace the standard intercept test performed in the

existing empirical literature on market efficiency. Similar to arbitrage pricing theory's reliance on estimated covariances when determining arbitrage opportunities, our statistical test utilizes historical data to detect statistical arbitrage in the economy. We then apply this new methodology to momentum and value trading strategies. In addition, we also compute their probability of a loss, providing additional insights into their ability to eventually produce arbitrage profits.

To minimize data mining concerns, the momentum strategies we investigate are modeled after the momentum strategies tested in Jegadeesh and Titman (1993), while the value strategies are modeled after the contrarian strategies of Lakonishok et al. (1994). Our sample period is from January 1965 to December 2000 and subsumes that of both original papers. Our test of statistical arbitrage reveals that of the 16 momentum trading strategies examined, six produce statistical arbitrage at the 5% level with three more at the 10% level. The probability of incurring a loss for a momentum strategy that longs the highest return decile and shorts the lowest return decile based on six months of past returns, and holds that portfolio for 12 months, falls below 1% after just 89 months of trading. Of the 12 value strategies we examine, five strategies yield statistical arbitrages at the 5% level. In addition, after just 79 months of trading, the probability of incurring a loss on the value strategy using the past three years of sales growth with a one-year holding period falls below 1%. In summary, momentum strategies and value strategies provide strong evidence against the efficient markets hypothesis.

The trading strategies are tested under the assumption that the expected trading profits are constant over time, which we refer to as the constrained mean (CM) version of statistical arbitrage. Our first robustness check evaluates whether the assumption of constant expected profits is appropriate. We compare the constrained mean model with an unconstrained mean (UM) model that incorporates an additional parameter, allowing the expected profits to vary over time. Four independent tests confirm that the CM model is appropriate. First, the rates of change for expected profits are often not statistically different from zero. Of the 15 momentum and value strategies that are CM statistical arbitrages at the 10% or 5% level, only one registers a statistically negative rate of decay in its expected profits. Second, the insample fits as measured by the root mean squared errors (RMSEs) for the CM and UM models are indistinguishable, indicating there is no advantage to using the more complicated UM model. Third, the sum of squared normalized residuals are almost identical for both models. Fourth, likelihood ratio tests cannot reject the null hypothesis that expected profits are constant over time. Hence we find that the CM version of statistical arbitrage is the appropriate test for momentum and value trading strategies. This is because the UM model spreads the information contained in trading profits over an additional variable without offering an improved fit, thereby weakening the power of the statistical arbitrage test.

To ascertain the robustness of our statistical methodology, we perform a series of simulations incorporating deviations from our assumed process for incremental trading profits. Specifically, we consider autocorrelation, jumps, and parameter nonstationarity. Our statistical arbitrage test is then applied to these simulated

processes. The simulations indicate that our statistical test is biased toward accepting the null hypothesis of no statistical arbitrage, and thus the acceptance of market efficiency. Consequently, instances of statistical arbitrage found herein are hard to reconcile with market efficiency.

A possible concern is that market inefficiencies could disappear after including transactions costs. To investigate this issue, we compute the turnover of each portfolio and combine these results with the estimated round-trip transactions costs of [Chan and Lakonishok \(1997\)](#). We then adjust the incremental profits downward by these transactions costs. Besides incorporating transaction costs, we consider four additional market frictions as in [Alexander \(2000\)](#). We study the impact of requiring margin for both long and short positions as well as reduced proceeds from the interest earned on the margin account. We also jointly assume an additional liquidity buffer for the short position and borrowing rates that are higher than lending rates. In this realistic trading environment, our conclusions regarding the presence of statistical arbitrage remain unchanged as nine of the 11 statistical arbitrages at the 5% level remain statistical arbitrages at the 5% level. This suggests that the statistical arbitrages represent feasible and attractive trading opportunities for market participants.

A common feature of stock market anomaly studies is that small stocks are often less efficient than large stocks ([Hong et al., 2000](#); [Mitchell and Stafford, 2000](#)). Moreover, because small stocks could require greater transactions costs and are less liquid, we investigate whether our results survive their exclusion. Even after removing stocks with market equity below the 50% NYSE percentile from our sample, all but one of the momentum portfolios and seven out of 12 value portfolios still test positively for statistical arbitrage at the 5% level. Thus, our previous results do not appear to be driven by small stocks.

Lastly, we investigate the dependence of our results on the sequence of observed returns. Using a bootstrap procedure, we generate multiple trading profit sequences from the observed time series and apply our test for statistical arbitrage. This robustness test reaffirms the presence of statistical arbitrage in momentum and value strategies.

As an interesting application of the statistical arbitrage methodology, we investigate the size effect. We test the [Fama and French \(1993\)](#) Small Minus Big (SMB) factor for statistical arbitrage. The *t*-statistic for the SMB factor over our sample period suggests that one will be rewarded with positive expected profits when investing in a strategy that longs small stocks and shorts large stocks. However, we find that this SMB-based trading strategy does not constitute a statistical arbitrage, confirming the popular belief that the size effect disappeared after the early 1980s.

The remainder of this paper is organized as follows. Section 2 introduces the concept of statistical arbitrage and its relation to market efficiency. Section 3 derives statistical tests for the existence of statistical arbitrage. The data used in our empirical study are described in Section 4, and Section 5 presents our empirical results. Section 6 describes various tests to gauge robustness. Section 7 applies the statistical arbitrage methodology to the testing of the size effect. Section 8 concludes and offers avenues for further research.

## 2. Statistical arbitrage and market efficiency

This section introduces the notion of a statistical arbitrage and its relation to market efficiency. Traded in the economy are a stock (or portfolio)  $S_t$  and money market account  $B_t$  initialized at a dollar ( $B_0 = 1$ ).

Trading strategies are central to the notion of statistical arbitrage. As in the standard option pricing literature, let the stochastic processes  $(x(t), y(t) : t \geq 0)$  represent a zero initial cost, self-financing trading strategy involving  $x(t)$  units of a stock and  $y(t)$  units of a money market account at time  $t$ . This trading strategy is formulated using only available information such as past returns, firm sizes, earnings announcements, market versus book values, sales growth rates, or macroeconomic conditions. These trading strategies, by definition, must have zero initial cost, i.e.,  $x(0)S_0 + y(0) = 0$ . For emphasis, the stock could represent a (zero-cost) self-financing portfolio consisting of long and short positions in various risky assets.

Although models of market equilibrium could assist in identifying possible trading strategies, a model of market equilibrium is not required. Let the process  $V(t)$  denote the cumulative trading profits at time  $t$  that are generated by such a trading strategy  $(x(t) : t \geq 0)$ .<sup>1</sup> In the following analysis, it will be necessary to work with the discounted value of the cumulative trading profits defined as  $v(t) = V(t)/B_t$ . To illustrate these concepts and various characteristics of typical trading strategies, we consider a few examples.

**Example 1.** Consider the standard Black-Scholes economy with a nondividend paying stock price  $S_t$  that evolves according to

$$S_t = S_0 e^{\alpha t - \sigma^2 t/2 + \sigma W_t} \tag{1}$$

and a money market account  $B_t = e^{rt}$ , where  $\alpha, \sigma^2, r$  are nonnegative constants,  $\alpha > r$ , and  $W_t$  is a standard Brownian motion. Consider the self-financing trading strategy that consists of buying and holding one unit of the stock, financed with the money market account. The value of this portfolio at time  $t$  is

$$V(t) = 1 \cdot S_t - S_0 \cdot e^{rt} = S_0 (e^{[\alpha - \sigma^2/2]t + \sigma W_t} - e^{rt}) \tag{2}$$

representing the cumulative trading profits with an initial value  $V_0 = 0$ . The expectation and variance of the cumulative discounted trading profits both approach infinity as  $t \rightarrow \infty$ ,<sup>2</sup>

$$E^P[v(t)] = S_0 [e^{(\alpha-r)t} - 1] \rightarrow \infty \quad \text{and} \quad \text{Var}^P[v(t)] = S_0^2 e^{2(\alpha-r)t} [e^{\sigma^2 t} - 1] \rightarrow \infty. \tag{3}$$

In addition, the time-averaged variance  $\text{Var}^P[v(t)]/t$  converges to  $\infty$  as  $t \rightarrow \infty$ , the significance of which will become apparent when statistical arbitrage is defined.

<sup>1</sup>The process  $V(t)$  is dollar-denominated, and it is neither a cumulative excess return nor a cumulative residual with respect to an equilibrium model.

<sup>2</sup>See Casella and Berger (1990, p. 628) for the variance of a lognormal random variable.

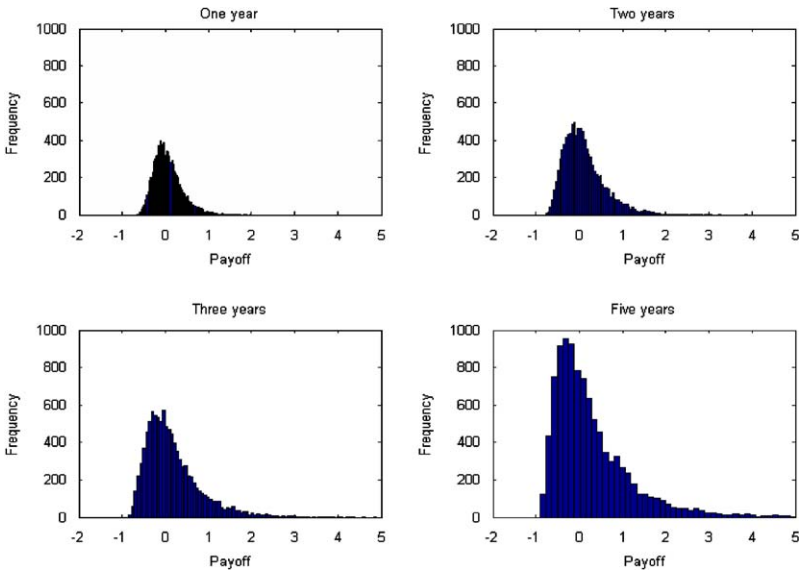


Fig. 1. Discounted cumulative trading profits from a theoretical trading strategy that invests in a risky asset whose price follows a geometric Brownian motion and borrows at the risk-free rate (Example 1 in the text). Ten thousand Monte Carlo trials are used to obtain the distribution of the payoffs from the above theoretical trading strategy. The frequency distribution of the cumulative payoffs for one, two, three, and five years are illustrated.

As a visual illustration, the discounted cumulative payoffs after one, two, three, and five years are simulated for the above trading strategy, which finances the purchase of stock by borrowing from the money market account. A total of 10,000 simulations are conducted with histograms of the payoffs provided in Fig. 1. Observe that the distribution is right-skewed, consistent with its lognormal distribution, and appears to shift to the right over time. Although the mean increases, both the variance and the time-averaged variance increase over time, with the second property the focus of two subsequent examples.<sup>3</sup> Hence Example 1 does not constitute a statistical arbitrage. Furthermore, a trading strategy that only buys stock, without borrowing, would neither be self-financing nor have a declining time-averaged variance.

Many instances of premiums or excess returns in the anomalies literature are generated by portfolios that are long and short portfolios of stocks formed on the basis of certain cross-sectional characteristics. The next example involves arithmetic instead of geometric Brownian motion, a reasonable process for a portfolio of long and short positions that is not a limited liability asset.

<sup>3</sup>The probability of a loss is easily computed by evaluating the terminal value of the discounted cumulative trading profits for each time horizon after every simulated path, to complement the visual presentation given in the histograms. This property is discussed later in more detail as the examples are intended to emphasize the strategies' time-averaged variances.

**Example 2.** Consider the discounted cumulative value of a trading strategy  $v(t)$  that evolves according to

$$v(t) = \alpha t + \sigma W_t \tag{4}$$

with  $v(0) = 0$ . The expectation and variance of the cumulative discounted trading profits are both unbounded as  $t \rightarrow \infty$ ,

$$E^P[v(t)] = \alpha t \rightarrow \infty \quad \text{and} \quad Var^P[v(t)] = \sigma^2 t \rightarrow \infty, \tag{5}$$

while the time-averaged variance is finite and equals  $\sigma^2$ .

The previous examples illustrate trading strategies that are not long horizon excess profit opportunities. This is because, as the trading strategies' expected profits increase across time, the variances of the trading profits also increase. In particular, the time-averaged variance does not disappear. In contrast, the following example captures the intuitive notion of a statistical arbitrage.

**Example 3.** Revisit the buy and hold trading strategy underlying Example 2. Suppose that the discounted trading profits over an arbitrary intermediate trading interval  $[t_{k-1}, t_k]$  can be represented as

$$v(t_k) - v(t_{k-1}) = \mu + \sigma z_k, \tag{6}$$

where  $\mu, \sigma > 0$  and  $z_k$  are independent and identically distributed random variables with zero mean and variance  $1/k$ . This trading strategy has positive expected discounted profits over every interval ( $\mu$ ), but with random noise ( $\sigma z_k$ ) appended. The variance of the noise is decreasing over time. Once again,  $v(t_0) = 0$  and the cumulative discounted trading profits at time  $t_n$  equals

$$v(t_n) = \sum_{k=1}^n [v(t_k) - v(t_{k-1})] = \mu n + \sigma \sum_{k=1}^n z_k. \tag{7}$$

Observe that  $E^P[v(t_n)] = \mu n$  and  $Var^P[v(t_n)] = \sigma^2 \sum_{k=1}^n (1/k)$  both converge to infinity. However,  $Var^P[v(t_n)]/n = \sigma^2 \sum_{k=1}^n (1/k)/n \rightarrow 0$  as  $n \rightarrow \infty$ .

Over time, analogous to cross-sectional diversification in the Ross (1976) arbitrage pricing theory (APT), the random noise in the trading strategy of Example 3 is diversifiable. This trading strategy offers a positive discounted expected profit with a time-averaged variance approaching zero and captures the intuition behind statistical arbitrage. Given these insights, we now formalize the notion of a statistical arbitrage.

**Definition 1.** A statistical arbitrage is a zero initial cost, self-financing trading strategy  $(x(t) : t \geq 0)$  with cumulative discounted value  $v(t)$  such that

1.  $v(0) = 0,$  (8)

2.  $\lim_{t \rightarrow \infty} E^P[v(t)] > 0,$  (9)

$$3. \quad \lim_{t \rightarrow \infty} P(v(t) < 0) = 0, \text{ and} \quad (10)$$

$$4. \quad \lim_{t \rightarrow \infty} \frac{\text{Var}^P[v(t)]}{t} = 0 \text{ if } P(v(t) < 0) > 0 \quad \forall t < \infty. \quad (11)$$

By definition, a statistical arbitrage satisfies four conditions: (1) it is a zero initial cost ( $v(0) = 0$ ) self-financing trading strategy, that in the limit has (2) positive expected discounted profits, (3) a probability of a loss converging to zero, and (4) a time-averaged variance converging to zero if the probability of a loss does not become zero in finite time. That is, the fourth condition applies only when there always exists a positive probability of losing money. Otherwise, if  $P(v(t) < 0) = 0$  for all  $t \geq T$  for some  $T < \infty$ , then a standard arbitrage opportunity is available. In economic terms, the fourth condition implies that a statistical arbitrage opportunity eventually produces riskless incremental profit, with an associated Sharpe ratio increasing monotonically through time. This becomes apparent in Section 3 when the convergence rate to arbitrage is studied. This is consistent with the variance of the trading strategy increasing toward infinity with time, but with a growth rate less than linear. The fourth property of statistical arbitrage essentially mitigates the concerns of Shleifer and Vishny (1997) with respect to risky arbitrage. Additional justification for the fourth condition is given in Appendix A. Overall, to generate statistical arbitrage, it is essential that a trading strategy's time-averaged variance falls over time.<sup>4</sup> Returning to Example 1, we observe that the buy and hold strategy involving the purchase of a stock financed with the money market account is not capable of generating statistical arbitrage. The same conclusion applies to Example 2. Only Example 3 is consistent with statistical arbitrage.

A standard arbitrage opportunity is a special case of statistical arbitrage. A standard arbitrage has  $V(0) = 0$ , where there exists a finite time  $T$  such that  $P(V(t) > 0) > 0$  and  $P(V(t) \geq 0) = 1$  for  $t \geq T$ . To transform the standard arbitrage opportunity into an infinite horizon self-financing trading strategy, we just invest the proceeds at time  $T$  into the money market account, i.e.,  $V(s) = V(T)B_s/B_T$  for  $s \geq T$ . Note that  $v(s) = V(T)(B_s/B_T)1/B_s = v(T)$ . Then,  $\lim_{s \rightarrow \infty} E^P[v(s)] = E^P[v(T)] > 0$  which satisfies the second condition and  $\lim_{s \rightarrow \infty} P(v(s) < 0) = P(v(T) < 0) = 0$ , which satisfies the third condition and implies that the fourth condition is not applicable. However, this paper's intent is to construct a statistical test for evaluating market efficiency with regard to persistent anomalies, trading strategies whose probabilities of losing money are always positive at a given finite time.

This definition is similar to the limiting arbitrage opportunity used to construct the Ross (1976) APT. The difference between the two concepts is that a statistical arbitrage is a limiting condition across time, while Ross's APT is a cross-sectional limit at a specific time. This difference necessitates discounting by the money market

<sup>4</sup>The self-financing condition alters the composition of the portfolio over time, creating an interaction between the self-financing and the decreasing time-averaged variance properties.



account and the normalization by time in the fourth condition. Therefore, just as Ross's APT is appropriate in an economy with a large number of assets, our methodology is appropriate for long time horizons. Investors with long but finite time horizons view statistical arbitrage opportunities as too good to pass up as they offer positive expected discounted profit, variance (per unit time) that becomes arbitrarily small, and decreasing risk of a loss. Although statistical arbitrage is defined over an infinite time horizon, there exists a finite timepoint  $t^*$  such that the probability of a loss is arbitrarily small,  $P(v(t^*) < 0) = \varepsilon$ . Thus, comparing an arbitrage opportunity at time  $t^*$  having the property  $P(v(t^*) < 0) = 0$  with a statistical arbitrage opportunity reveals they are separated only by an  $\varepsilon$  probability of a loss.

It is well known that arbitrage opportunities are incompatible with an efficient market (see Jensen (1978), for a general discussion).<sup>5</sup> Tests of market efficiency in the options market based on the absence of arbitrage are numerous. For example, Kamara and Miller (1995) study violations of put-call parity, and Bodurtha and Courtadon (1986) investigate the efficiency of foreign currency options. Tests for statistical arbitrage rejecting market efficiency can be viewed as an extension of this literature. Statistical arbitrage rejects the market as being in any economic equilibrium, an important prerequisite for an efficient market (see Jarrow, 1988, Chapter 19). A statistical arbitrage (analogous to an arbitrage opportunity in Ross's APT) will induce investors to engage in trading and, consequently, prevent the market from reaching an equilibrium. The existence of a single trader operating under a finite time horizon and willing to pursue a statistical arbitrage is sufficient for one to conclude that these trading opportunities reject market efficiency.

Our idea of extending standard arbitrage to its infinite horizon counterpart is related to a literature extending simple arbitrage in alternative ways. In the context of incomplete markets, Cochrane and Saá-Requejo (2000) invoke Sharpe ratios to find asset prices that are good deals to investors while Bernardo and Ledoit (2000) exclude investments whose maximum gain–loss ratios are too attractive. Both of these approaches also investigate trading opportunities that generalize the definition of arbitrage without specifying a particular model of market equilibrium. Carr et al. (2001) introduce the concept of acceptable opportunities that would be executed by any reasonable investor, implying their existence is incompatible with an efficient market. This paper introduces statistical arbitrage as a similar generalization of arbitrage.

This notion of statistical arbitrage can also be extended to study the traditional weak- and strong-form tests of market efficiency. In constructing the trading strategies, weak-form statistical arbitrage would utilize only publicly available information, while strong-form statistical arbitrage would utilize all available information.

Furthermore, statistical arbitrage requires no information beyond the traditional intercept test. Portfolio returns from long and short positions are easily converted

---

<sup>5</sup>Jensen (1978, p. 96) states that the existence of economic trading profits rejects an efficient market. An arbitrage opportunity (and a statistical arbitrage) satisfies his definition of economic trading profits.

into dollar-denominated trading profits with gains and losses accruing through time according to the risk-free rate, as shown in our later empirical analysis.

An important issue regarding the implementation of our trading strategies is the exclusion of doubling strategies (Duffie, 2001). In our later empirical work, \$1 is kept constant in the risky asset (long minus short portfolio position) over time. Investing the trading profits, which have positive expectation, in the riskless money market account serves to reduce the trading strategy's time-averaged variance. Appendix B proves that doubling strategies are incompatible with the second and fourth properties of statistical arbitrage in Definition 1. Moreover, our empirical study does not terminate trading at the first instance of a positive cumulative trading profit. Dybvig and Huang (1988) prove that a nonnegative lower bound on wealth is sufficient to preclude arbitrage opportunities of the doubling persuasion. To appeal to this economically sensible restriction, we record the maximum loss for trading strategies that generate statistical arbitrage in our empirical applications.

Several differences exist between the statistical arbitrage test and the traditional tests of market efficiency using a risk-adjusted alpha. First and most important, the risk-adjustment process for returns requires an underlying model of market equilibrium. This makes alpha, as a test of market efficiency, subject to the joint hypothesis critique. In contrast, the test for statistical arbitrage does not require a model of market equilibrium to generate risk-adjusted or excess returns. Instead, dollar-denominated incremental profits from a self-financing trading strategy are analyzed when testing for statistical arbitrage. Second, a linear factor model is typically assumed in the risk-adjustment process (e.g., Fama and French, 1993; Carhart, 1997). The statistical arbitrage test can be applied to any asset, including those that are not priced using linear factor models such as derivatives. Third, the alpha test on the mean is unable to detect the presence of statistical arbitrage. As demonstrated in Section 3, the statistical arbitrage test does not reduce to a *t*-ratio test. Such a reduction implicitly assumes that the rate of change in the trading profits' volatility is zero and violates the crucial fourth condition of statistical arbitrage. Fourth, a statistical arbitrage requires that the time-averaged variance of a strategy declines while the alpha test has no such condition.

### 3. Tests for statistical arbitrage

In this section, we derive statistical tests for determining whether a trading strategy yields statistical arbitrage. Two classes of statistical tests are considered: constrained mean tests and unconstrained mean tests. This distinction is important when we investigate momentum and value strategies.

Our proposed framework is an empirical methodology that offers an alternative to the standard risk-adjusted intercept test by circumventing the required specification of an equilibrium model for returns. However, as an empirical methodology, the statistical test is confined to testing whether statistical arbitrage opportunities existed *ex post* in the sample period instead of *ex ante*.

### 3.1. The statistical arbitrage methodology

To test for statistical arbitrage, a time series of dollar-denominated discounted cumulative trading profits  $v(t_1), v(t_2), \dots, v(t_n)$  generated by a trading strategy are analyzed. For a given trading strategy, let  $\Delta v_i = v(t_i) - v(t_{i-1})$  denote increments of the discounted cumulative trading profit measured at equidistant time points  $t_i - t_{i-1} = \Delta$  with  $t_i = i \times \Delta$ .<sup>6</sup>

Our test for statistical arbitrage requires an assumed process for the evolution of incremental trading profits. At first, for simplicity, we impose the assumption that the trading profits have independent increments. This assumption is later relaxed, but empirical results confirm that it does not alter our conclusions regarding market efficiency. For greater generality in presenting our results, we also allow the expected profits of the trading strategy to vary over time. Thus, we begin our analysis with the unconstrained mean model. The quantity  $\Delta$ , denoting the time between equidistant increments, can be set to one without any loss of generality with  $t_i$  understood as being  $i$ .

**Assumption 1.** Let the discounted incremental trading profits satisfy

$$\Delta v_i = \mu i^\theta + \sigma i^\lambda z_i \quad (12)$$

for  $i = 1, 2, \dots, n$  where  $z_i$  are i.i.d  $N(0, 1)$  random variables with  $z_0 = 0$ .<sup>7</sup> Both  $v(t_0)$  and  $\Delta v_0$  are zero.

A Taylor series expansion reveals that our formulation encompasses linear, and potentially quadratic, specifications, depending on the magnitude of  $\lambda$  and  $\theta$ . More important, these specifications account for the changing composition of a statistical arbitrage portfolio as trading profits are invested in the risk-free account, causing proportionally less to be invested in the risky position over time. In particular, consider an expansion of  $i^\lambda$  with respect to  $i$ , which yields  $(i+1)^\lambda = i^\lambda [1 + \lambda/i + \lambda(\lambda-1)/2i^2]$  plus higher order terms. For  $\lambda < 0$ , an important condition of statistical arbitrage, the marginal decrease in the portfolio's volatility itself declines over time. This property is consistent with the declining marginal impact of the risk-free asset. In our empirical implementation, a profitable trading strategy generates trading profits  $V(t)$  (undiscounted) that are invested in the risk-free asset while a constant amount of \$1 is invested in both the long and short portfolios over time. For statistical arbitrage opportunities, the fraction of the portfolio invested in the risky asset (long–short portfolio) equals 1 divided by 1 plus the trading profits,  $1/(1+V(t))$ , a convex decreasing function when trading profits increase over time. This feature parallels the  $i^\lambda$  (and the associated specification for the mean) component of the volatility function.

<sup>6</sup>Trading profits derived from portfolios are well represented by a normal distribution because the impact of idiosyncratic jumps are mitigated.

<sup>7</sup>Assumption 1 implies that the fourth condition in the definition of statistical arbitrage is relevant. It excludes standard arbitrage opportunities, whose existence can be tested via alternative and well-known procedures.

Moreover,  $\sigma$  need not be a constant. For example, volatility  $\sigma(t)$  could evolve as a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) process. This enhancement is not considered in our later empirical work, although the main theorem of this section remains valid with this generalization.

The expectation and variance of the discounted incremental trading profits in Eq. (12) are  $E[\Delta v_i] = \mu i^\theta$  and  $Var[\Delta v_i] = \sigma^2 i^{2\lambda}$ . For  $\lambda < 0$ , the variance of  $\Delta v_i$  decreases over time, which ultimately satisfies the fourth condition of statistical arbitrage. Also, the dynamics of the discounted incremental trading profit in Eq. (12) with  $\lambda < 0$  and  $\mu > 0$  do not advocate a trading strategy based on waiting for the volatility to decline before investing. Instead, it is optimal for investors to immediately begin trading. These investors will earn a positive expected profit while enjoying the benefit of decreasing time-averaged variance. A bootstrap procedure in Section 6 confirms that our statistical arbitrage conclusions in the later empirical study are robust to altering the return sequence.

The discounted cumulative trading profits generated by the trading strategy are

$$v(t_n) = \sum_{i=1}^n \Delta v_i \sim dN\left(\mu \sum_{i=1}^n i^\theta, \sigma^2 \sum_{i=1}^n i^{2\lambda}\right), \tag{13}$$

while the log likelihood function for the increments in Eq. (12) equals

$$\log L(\mu, \sigma^2, \lambda, \theta | \Delta v) = -\frac{1}{2} \sum_{i=1}^n \log(\sigma^2 i^{2\lambda}) - \frac{1}{2\sigma^2} \sum_{i=1}^n \frac{1}{i^{2\lambda}} (\Delta v_i - \mu i^\theta)^2 \tag{14}$$

allowing maximum likelihood estimation to generate the four required parameters. The score equations required to solve for the four parameters are found in Lemma 1.<sup>8</sup>

**Lemma 1.** *Parameter estimates are obtained by solving the following four equations for the four unknown parameters:*

$$\frac{\partial \log L(\mu, \sigma^2, \lambda, \theta | \Delta v)}{\partial \mu} : \hat{\mu} = \frac{\sum_{i=1}^n \Delta v_i i^{\hat{\theta}-2\hat{\lambda}}}{\sum_{i=1}^n i^{2(\hat{\theta}-\hat{\lambda})}}, \tag{15}$$

$$\frac{\partial \log L(\mu, \sigma^2, \lambda, \theta | \Delta v)}{\partial \sigma^2} : \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \frac{1}{i^{2\hat{\lambda}}} (\Delta v_i - \hat{\mu} i^{\hat{\theta}})^2, \tag{16}$$

$$\frac{\partial \log L(\mu, \sigma^2, \lambda, \theta | \Delta v)}{\partial \theta} : \sum_{i=1}^n \Delta v_i \log(i) i^{\hat{\theta}-2\hat{\lambda}} = \hat{\mu} \sum_{i=1}^n \log(i) i^{2(\hat{\theta}-\hat{\lambda})}, \tag{17}$$

$$\frac{\partial \log L(\mu, \sigma^2, \lambda, \theta | \Delta v)}{\partial \lambda} : \hat{\sigma}^2 \sum_{i=1}^n \log(i) = \sum_{i=1}^n \frac{\log(i)}{i^{2\hat{\lambda}}} (\Delta v_i - \hat{\mu} i^{\hat{\theta}})^2. \tag{18}$$

<sup>8</sup>The property  $(\partial/\partial x)t^{ax} = a \log(t)t^{ax}$  is used repeatedly when solving for  $\hat{\lambda}$  and  $\hat{\theta}$ .

The consistency and efficiency of these maximum likelihood estimates (MLE) are both well established (see Casella and Berger, 1990, Chapter 7). As expected, for  $\theta = 0$  and  $\lambda = 0$ , Eqs. (15) and (16) reduce to the standard normal MLE estimators

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \Delta v_i \quad \text{and} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\Delta v_i - \hat{\mu})^2. \tag{19}$$

The functional form of the coefficients in Eq. (12) allows for a large variety of potential changes to the mean and variance.<sup>9</sup> Moreover, misspecifying the incremental trading profit process serves only to increase the likelihood of accepting the null hypothesis of no statistical arbitrage. Therefore, the existence of statistical arbitrage, the subject of Theorem 1, is robust to criticisms regarding the specified stochastic process.

**Theorem 1.** *A trading strategy generates a statistical arbitrage with  $1 - \alpha$  percent confidence if the following conditions are satisfied:*

$$H1 : \hat{\mu} > 0, \tag{20}$$

$$H2 : \hat{\lambda} < 0, \text{ and} \tag{21}$$

$$H3 : \hat{\theta} > \max \left\{ \hat{\lambda} - \frac{1}{2}, -1 \right\} \tag{22}$$

with the sum of the individual  $p$ -values forming an upper bound for the test's Type I error. Therefore, the sum of the  $p$ -values associated with the individual hypotheses must be below  $\alpha$  to conclude that a trading strategy generates statistical arbitrage.

To satisfy the fourth condition of statistical arbitrage, Eq. (13) implies the parameter  $\lambda$  must be negative to ensure  $\sigma^2 \sum_{i=1}^n i^{2\lambda} / n \rightarrow \infty$ , while the second condition requires the parameter  $\mu$  to be positive. Any value of  $\theta$  ensures  $\mu \sum_{i=1}^n i^\theta > 0$  provided  $\mu > 0$ . However, the convergence of  $P(v(t) < 0)$  to zero requires

$$\frac{\sum_{i=1}^n i^\theta}{\sqrt{\sum_{i=1}^n i^{2\lambda}}} \rightarrow \infty, \tag{23}$$

with details in Appendix C. The leading order term of the sum in the numerator (using left endpoints)

$$\sum_{i=1}^n i^\theta \geq \int_1^n s^\theta ds = \frac{n^{1+\theta}}{1+\theta} - \frac{1^{1+\theta}}{1+\theta} \tag{24}$$

<sup>9</sup>The specification of the coefficients implicitly accounts for the usual rescaling of large positive numbers, such as trade volume, using a log transformation. A Taylor series expansion of  $i^\lambda$  with respect to  $\lambda$  evaluated at  $\lambda = 0$  equals  $i^\lambda = 1 + \ln(i)\lambda + \frac{1}{2}(\ln(i))^2\lambda^2$  plus higher order terms. Thus, our formulation conforms with the standard empirical practice of taking logs for large positive variables as the index  $i$  reaches values over 400 (number of months) in our later empirical implementation.

divided by the square root of the sum in the denominator (using right endpoints)

$$\sum_{i=1}^n i^{2\lambda} \leq \int_1^n (s+1)^{2\lambda} ds = \frac{(n+1)^{2\lambda+1}}{2\lambda+1} - \frac{2^{2\lambda+1}}{2\lambda+1} \quad (25)$$

equals  $n^{\theta+1}/(n+1)^{\lambda+1/2}$ . Therefore, the issue of the tail probability converging to zero is whether  $n^{\theta-\lambda+1/2}$  converges to infinity, which is the case when  $\theta > \lambda - 1/2$ . The leading order approximation is only valid in the numerator if  $\theta > -1$ .<sup>10</sup> Therefore, with a negligible loss of the endpoint  $-1$ , the third condition of statistical arbitrage requires  $\theta > \max\{\lambda - 1/2, -1\}$ .

The three parameters are tested individually with the Bonferroni inequality, which stipulates that the sum of the  $p$ -values for the individual tests becomes the upper bound for the Type I error of the joint hypothesis test.<sup>11</sup> Standard errors for the hypothesis tests in Theorem 1 are extracted from the Hessian matrix to produce  $t$ -statistics and their corresponding  $p$ -values.

As demonstrated in Theorem 1, based on the distribution of discounted cumulative trading profits found in Eq. (13), the second and third conditions of statistical arbitrage require  $\mu > 0$  and  $\theta > \lambda - \frac{1}{2}$ , while the fourth condition requires  $\lambda < 0$ . The condition  $\theta > \lambda - \frac{1}{2}$  ensures the probability of a loss converges to zero. The  $\theta$  parameters could be negative, indicating that expected trading profits are decreasing over time, provided they are not negative enough to prevent convergence to arbitrage. Thus, the third hypothesis is equivalent to a test of long run market efficiency. The existence of statistical arbitrage is not rejected if  $\theta$  equals zero, in contrast to the situation when either  $\mu$  or  $\lambda$  is zero. This leads to a constrained mean test for statistical arbitrage given in Section 3.2.

Furthermore, applying a single  $t$ -statistic to trading profits is not a valid test for statistical arbitrage. Assume  $\lambda$  and  $\theta$  are both zero, implying the cumulative trading profit  $v(t)$  in Eq. (13) is distributed  $N(\mu n, \sigma^2 n)$ . Given an estimate  $\hat{\sigma}^2 = (1/n \sum_{i=1}^n (\Delta v_i - \hat{\mu})^2)$ , Theorem 1 reduces to a  $t$ -test, with  $n - 1$  degrees of freedom, of  $\hat{\mu} = (1/n) \sum_{i=1}^n \Delta v_i$  being statistically positive

$$\hat{\mu} > t_{n-1, \alpha} \sqrt{\frac{\hat{\sigma}^2}{n}} \quad (26)$$

<sup>10</sup> For  $\theta = -1$ ,  $\sum_{i=1}^n i^{-1}$  is a harmonic series that diverges to infinity with integral representation  $\int_1^n s^{-1} ds = \ln(n)$ . Therefore,  $\theta = -1$  is the critical value that ensures the numerator approaches infinity, a condition required to ensure the probability of a loss approaches zero. Values of  $\theta$  smaller than  $-1$  result in the numerator converging to a finite number.

<sup>11</sup>  $P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$ , see Casella and Berger (1990, p. 11). The Bonferroni inequality sums the probabilities of the events in the union without subtracting the probability of their intersection to obtain an upper bound. There is no uniformly most powerful test (Definition 8.3.5, Casella and Berger, 1990, p. 365) as the conditions for applying the Neyman-Pearson Lemma (Corollary 8.3.2, Casella and Berger, 1990, p. 368) on composite hypotheses are not satisfied by the non-monotonic likelihood function in Eq. (14). A likelihood ratio test (Definition 8.2.1, Casella and Berger, 1990, p. 347) would be biased towards over-rejecting the null hypothesis of no statistical arbitrage because all three conditions in Theorem 1 have to be satisfied to reject the null.

However, this single  $t$ -test is unable to test for statistical arbitrage because the fourth property of statistical arbitrage cannot be verified. In summary, a single  $t$ -test is limited to testing only the hypothesis of positive expected trading profits and cannot detect statistical arbitrage.

Besides the existence of statistical arbitrage, the rate of convergence to arbitrage is interesting and the subject of Theorem 2, which states that convergence of the tail probability  $P(v(t_n) < 0)$  is faster than exponential. Thus, our introduction of an infinite time horizon is less onerous than it appears, as investors quickly realize an investment opportunity that is close to arbitrage.

**Theorem 2.** *The rate of convergence from statistical arbitrage to arbitrage is faster than exponential. Specifically, the convergence rate is faster than*

$$O\left(\exp\left\{-\frac{\mu^2\left(\sum_{i=1}^n i^\theta\right)^2}{2\sigma^2\sum_{i=1}^n i^{2\lambda}}\right\}\right). \tag{27}$$

The proof is contained in Appendix C. The rate of convergence depends on the squared Sharpe ratio  $\mu^2/\sigma^2$  as well as the rates of change  $\theta$  and  $\lambda$ . Trading strategies could be evaluated by the convergence rate of  $P(v(t_n) < 0)$  to zero and the desirability of trading strategies ranked according to this function.

Besides implying the Sharpe ratio increases toward infinity, we consider an investor with log utility presented with a statistical arbitrage opportunity. Such an investor would find it desirable to pursue a statistical arbitrage opportunity. Furthermore, their demand for the statistical arbitrage opportunity is unbounded, as is the case for standard arbitrage opportunities.

For any level of exogenous wealth  $W > 0$  in the future, there exists a time horizon  $t^*$  such that for all  $t_n \geq t^*$ ,<sup>12</sup>

$$E[U(W + \delta v(t_n))] > E[U(W)] \quad \forall \delta > 0. \tag{28}$$

Intuitively, as with a standard arbitrage opportunity, the investor increases his expected utility by pursuing a statistical arbitrage opportunity. Taking an infinitely large position in the zero cost self-financing asset is possible. Hence, the economy cannot be in equilibrium if such trading strategies exist. We consider the case of log utility and prove that

$$E[U(W + \delta v(t_n)) - U(W)] = E\left[\ln\left(1 + \frac{\delta}{W}v(t_n)\right)\right] > 0 \quad \forall \delta > 0, \tag{29}$$

where the equality stems from the identity  $\ln(x + y) - \ln(x) = \ln((x + y)/x) = \ln(1 + y/x)$ . Eq. (29) can be expressed in terms of  $v(t_n)$ 's normal distribution  $v(t_n) \sim^d N(\mu \sum_{i=1}^n i^\theta, \sigma^2 \sum_{i=1}^n i^{2\lambda})$ . Define  $M_{t_n}$  as  $1 + (\delta/W)v(t_n)$  to express the inequality in Eq. (29) as  $E[\ln(M_{t_n})] > 0$ . Under the conditions of statistical arbitrage in Theorem 1, since  $\delta/W$  is positive, Appendix C proves that  $P(M_{t_n} - 1 < 0) =$

<sup>12</sup>We suppress the time dependence of  $W$  for simplicity as it is not the focus of our present analysis. This level of wealth could be random, in which case our proof would first condition on this amount before continuing in its present form.

$P((\delta/W)v(t_n) < 0) = P(v(t_n) < 0) \rightarrow 0$ . Therefore, Eq. (29) is satisfied for all  $\delta$  as there exists a  $t^*$  such that for all  $t_n \geq t^*$ ,  $P(M_{t_n} > 1) = 1 - \varepsilon$  for  $\varepsilon > 0$  arbitrarily small, implying  $E[\ln(M_{t_n})] > 0$  for the same set of  $t_n$ .

### 3.2. Unconstrained versus constrained mean statistical arbitrage

The results of this section remain valid when the parameter  $\theta$  is set to zero, implying the trading strategy's expected profits are constant over time. In this constrained mean model, special cases of Lemma 1 as well as Theorems 1 and 2 immediately result. We refer to this second test as the constrained mean test for statistical arbitrage to differentiate it from the unconstrained mean test in Theorem 1.

We begin our empirical analysis with the CM test and later verify that this formulation is appropriate. A variety of robustness checks in Section 6 support the usage of the CM model versus the enhanced UM specification in the context of momentum and value trading strategies. With the parameter  $\theta = 0$ , the process in Eq. (12) reduces to

$$\Delta v_i = \mu + \sigma i^\lambda z_i, \quad (30)$$

The corresponding joint hypothesis test for statistical arbitrage is

$$H1 : \mu > 0 \quad \text{and} \quad (31)$$

$$H2 : \lambda < 0, \quad (32)$$

while the probability of loss after  $n$  periods (see Appendix C) is equal to

$$\text{Probability of loss (after } n \text{ periods)} = N\left(\frac{-\mu n}{\sigma \sqrt{\sum_{i=1}^n i^{2\lambda}}}\right), \quad (33)$$

where  $N(\cdot)$  denotes the cumulative standard normal distribution.

## 4. Data

Our sample period starts in January 1965 and ends in December 2000. Monthly equity returns data are derived from the Center for Research in Security Prices (CRSP) at the University of Chicago. Our analysis covers all stocks traded on the NYSE, AMEX, and Nasdaq that are ordinary common shares (CRSP sharecodes 10 and 11), excluding ADRs, SBIs, certificates, units, REITs, closed-end funds, companies incorporated outside the United States, and Americus Trust Components.

The stock characteristics data are derived from the CRSP-Compustat intersection. The stock characteristics we focus on are book-to-market equity, cash flow-to-price



ratio, earnings-to-price ratio, and annual sales growth. To calculate book-to-market equity, book value per share is taken from the CRSP/Compustat price, dividends, and earnings database. We treat all negative book values as missing. We take the sum of Compustat data item 123 (income before extraordinary items) and data item 125 (depreciation and amortization) as cash flow. Only data item 123 is used to calculate cash flow if data item 125 is missing. To compute earnings, we draw on Compustat data item 58 [earnings per share (basic) excluding extraordinary items], and to compute sales, we utilize Compustat data item 12 [sales(net)]. Also, all prices and common shares outstanding numbers employed in the calculation of the ratios are computed at the end of the year.

To ensure that accounting variables are known beforehand and to accommodate the variation in fiscal year-ends among firms, sorting on stock characteristics is performed in July of year  $t$  based on accounting information from year  $t - 1$ . Hence, following Fama and French (1993), to construct the book-to-market deciles from July 1 of year  $t$  to June 30 of year  $t + 1$ , stocks are sorted into deciles based on their book-to-market equity ( $BE/ME$ ), where book equity is book equity in the fiscal year ending in year  $t - 1$  and market equity is calculated in December of year  $t - 1$ . Similarly, to construct the cash flow-to-price deciles from July 1 of year  $t$  to June 30 of year  $t + 1$ , stocks are sorted into deciles based on their cash flow-to-price, where cash flow is cash flow in the fiscal year ending in year  $t - 1$  and price is the closing price in December of year  $t - 1$ . Earnings-to-price is calculated in a similar fashion. All portfolios are rebalanced every month as some firms disappear from the sample over the 12-month period.

An investment of \$1 is maintained in the portfolios at all times. The self-financing condition is enforced by investing (borrowing) trading profits (losses) generated by the various trading strategies at the risk-free rate. Risk-free rate data are obtained from Kenneth French's website.

## 5. An application to momentum and value strategies

This section presents the results of the tests for statistical arbitrage performed on momentum and value strategies under the assumption that expected incremental trading profits are constant over time. Two hypotheses are jointly tested. First, the incremental profits from the strategy must be statistically greater than zero, and second, the time-averaged variance of the strategy must decline to zero as time approaches infinity. Table 1 contains summary statistics for the incremental trading profits of the momentum and value trading strategies under investigation.

In terms of our previous notation in Section 2, the cumulative trading profit  $V(t_i)$  is a function of its past value  $V(t_{i-1})$  and the current profit earned on the \$1 invested in the long and short positions. Specifically, the trading strategy  $x(t)$  is a two by one vector,  $[1, -1]^T$  across time, while the amount invested in the money market,  $y(t)$ , simply equals the previous cumulative trading profit,  $V(t_{i-1})$ .  $L(t_i)$  and  $S(t_i)$  represent the return of the long and short portfolios, respectively. With \$1 in the

Table 1

Summary statistics for the incremental profits from momentum and value strategies

Sample period is from January 1965 to December 2000. Portfolio MOM  $x/y$  denotes a momentum portfolio with a formation period of  $x$  months and a holding period of  $y$  months. Every month, stocks are sorted based on past  $x$  months of returns into decile portfolios. The portfolio MOM  $x/y$  longs the top decile and shorts the bottom decile and holds that spread for  $y$  months as in Jegadeesh and Titman (1993). The other portfolios are value portfolios. As in Lakonishok et al. (1994), the sorting variables are past one-year book-to-market (BM), cash flow-to-price (CP), and earnings-to-price (EP) as well as past three-year sales growth (SALES). Portfolio  $xy$  denotes a value portfolio based on stock characteristic  $x$  and a holding period of  $y$  years. Every July, stocks are sorted into decile portfolios based on a particular stock characteristic. For BM, CP, and EP portfolios, the portfolio  $xy$  longs the top decile and shorts the bottom decile and holds that spread for  $y$  years. For SALES portfolios, the portfolio  $xy$  longs the bottom decile and shorts the top decile and holds that spread for  $y$  years. The risk-free asset is used to finance the portfolios.

Portfolio	Minimum	Maximum	Median	Mean	Standard deviation
MOM 3/3	-0.1931	0.1176	0.0005	-0.0026	0.0301
MOM 3/6	-0.0939	0.0950	0.0030	0.0016	0.0192
MOM 3/9	-0.0676	0.0544	0.0037	0.0022	0.0150
MOM 3/12	-0.0540	0.0398	0.0033	0.0034	0.0121
MOM 6/3	-0.2042	0.1623	0.0049	0.0021	0.0339
MOM 6/6	-0.0994	0.0524	0.0055	0.0043	0.0200
MOM 6/9	-0.0670	0.0513	0.0070	0.0055	0.0166
MOM 6/12	-0.0576	0.0475	0.0043	0.0044	0.0129
MOM 9/3	-0.2104	0.1095	0.0066	0.0036	0.0327
MOM 9/6	-0.1063	0.0621	0.0074	0.0065	0.0203
MOM 9/9	-0.0611	0.0540	0.0064	0.0053	0.0157
MOM 9/12	-0.0446	0.0378	0.0041	0.0033	0.0117
MOM 12/3	-0.1596	0.1165	0.0080	0.0061	0.0315
MOM 12/6	-0.0803	0.0718	0.0063	0.0057	0.0189
MOM 12/9	-0.0635	0.0539	0.0048	0.0037	0.0145
MOM 12/12	-0.0331	0.0315	0.0019	0.0015	0.0104
BM1	-0.1469	0.2719	0.0126	0.0139	0.0445
BM3	-0.0997	0.1768	0.0126	0.0115	0.0337
BM5	-0.0843	0.1426	0.0102	0.0104	0.0291
CP1	-0.3524	0.5473	0.0092	0.0045	0.0722
CP3	-0.2390	0.2169	0.0067	0.0042	0.0462
CP5	-0.2139	0.1229	0.0065	0.0032	0.0409
EP1	-0.3683	0.1578	0.0059	0.0006	0.0515
EP3	-0.2785	0.1169	0.0037	0.0011	0.0379
EP5	-0.1880	0.0715	0.0032	0.0011	0.0310
SALES1	-0.0650	0.1607	0.0076	0.0085	0.0285
SALES3	-0.0564	0.1022	0.0062	0.0058	0.0229
SALES5	-0.0777	0.0789	0.0047	0.0048	0.0212

these portfolios, the current trading profit at period  $t_i$  equals  $1 + L(t_i) - (1 + S(t_i)) = L(t_i) - S(t_i)$ . The previous trading profit  $V(t_{i-1})$  accumulates at the risk-free rate  $r(t_i)$  to yield  $V(t_i) = L(t_i) - S(t_i) + V(t_{i-1})[1 + r(t_{i-1})]$ . This cumulative trading profit  $V(t_i)$  is then discounted by  $B(t_i) = \exp\{\sum_{j=1}^i r(t_j)\}$  to produce  $v(t_i)$ , whose first difference  $\Delta v(t_i)$  is analyzed for statistical arbitrage.

### 5.1. Momentum strategies

The momentum strategies we focus on are modeled directly after the strategies tested in Jegadeesh and Titman (1993). Specifically we examine strategies that long stocks in the top return decile and short stocks in the bottom return decile. These strategies are based on formation periods of three, six, nine, and 12 months and are held for three, six, nine, or 12 months.

For example, consider the construction of a momentum strategy with a formation period of three months and a holding period of six months (MOM 3/6). Each month, we look back three months to find the top return decile stocks and the bottom return decile stocks. Next, we long the top return decile, short the bottom return decile, and hold this portfolio for six months. The portfolio is rebalanced monthly to account for stocks that drop out of the database.

Table 2

Constrained mean (CM) statistical arbitrage tests: Jegadeesh and Titman (1993) momentum strategies. Parameter estimates and corresponding  $p$ -values for constrained mean test of statistical arbitrage. Sample period is from January 1965 to December 2000. Portfolio  $x/y$  denotes a momentum portfolio with a formation period of  $x$  months and a holding period of  $y$  months. Every month, stocks are sorted based on past  $x$  months of returns into decile portfolios. The portfolio  $x/y$  longs the top decile and shorts the bottom decile and holds that spread for  $y$  months as in Jegadeesh and Titman (1993). The risk-free asset is used to finance the portfolio. H1 and H2 denote the  $p$ -values from statistical arbitrage tests, which test whether the portfolio's mean monthly incremental profit is positive and whether its time-averaged variance is declining over time. The sum of the H1 and H2 columns is the  $p$ -value for the statistical arbitrage test. The sum can exceed one because of the Bonferonni inequality. The  $t$ -statistic on the mean monthly trading profit is provided for comparison.

Portfolio	$\mu$ (mean)	$t$ -Statistic	$\sigma$ (standard deviation)	$\lambda$ (growth rate of standard deviation)	H1 ( $\mu > 0$ )	H2 ( $\lambda < 0$ )	Sum (H1 + H2)	Statistical arbitrage
MOM 3/3	-0.0027	-1.78	0.033	-0.0174	0.967	0.306	1.273	No
MOM 3/6 <sup>a</sup>	0.0015	1.68	0.023	-0.0350	0.056	0.156	0.212	No
MOM 3/9 <sup>a</sup>	0.0021	3.08	0.025	-0.1083	0.001	0.001	0.002	Yes**
MOM 3/12 <sup>a</sup>	0.0033	5.81	0.023	-0.1317	0.000	0.000	0.000	Yes**
MOM 6/3 <sup>a</sup>	0.0020	1.27	0.037	-0.0168	0.110	0.306	0.416	No
MOM 6/6 <sup>a</sup>	0.0042	4.45	0.037	-0.1254	0.000	0.000	0.000	Yes**
MOM 6/9 <sup>a</sup>	0.0055	6.78	0.026	-0.0908	0.000	0.004	0.004	Yes**
MOM 6/12 <sup>a</sup>	0.0042	6.92	0.030	-0.1734	0.000	0.000	0.000	Yes**
MOM 9/3 <sup>a</sup>	0.0035	2.24	0.043	-0.0532	0.013	0.075	0.088	Yes*
MOM 9/6 <sup>a</sup>	0.0065	6.53	0.027	-0.0603	0.000	0.052	0.052	Yes*
MOM 9/9 <sup>a</sup>	0.0053	6.93	0.021	-0.0549	0.000	0.064	0.064	Yes*
MOM 9/12 <sup>a</sup>	0.0031	5.78	0.022	-0.1256	0.000	0.001	0.001	Yes**
MOM 12/3 <sup>a</sup>	0.0060	3.93	0.035	-0.0206	0.000	0.266	0.266	No
MOM 12/6 <sup>a</sup>	0.0057	6.13	0.020	-0.0100	0.000	0.385	0.385	No
MOM 12/9	0.0038	5.19	0.012	0.0432	0.000	0.878	0.878	No
MOM 12/12 <sup>a</sup>	0.0015	3.00	0.010	-0.0004	0.002	0.496	0.498	No

\* Significant at the 10% level.

\*\* Significant at the 5% level.

<sup>a</sup> Point estimates of  $\mu$  and  $\lambda$  are consistent with statistical arbitrage.

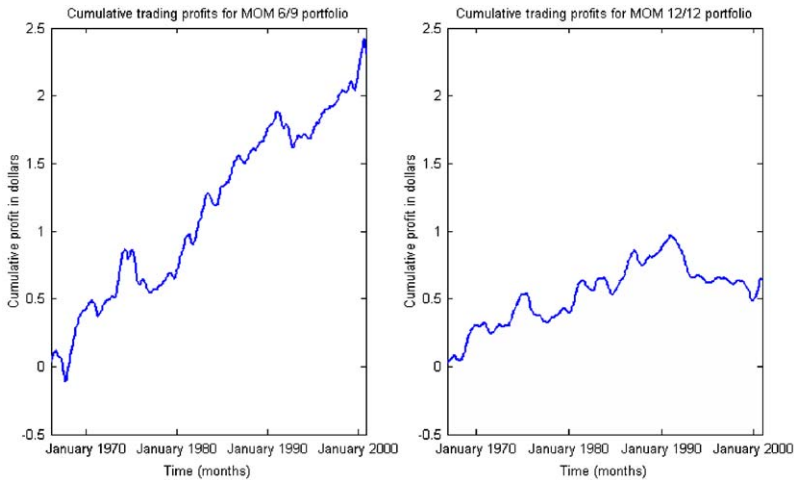


Fig. 2. Comparison of cumulative trading profits from two momentum strategies. Sample period is from January 1965 to December 2000. The momentum strategy MOM 6/9 longs the top return decile of stocks and shorts the bottom return decile of stocks in the sample based on a formation period of six months and a holding period of nine months. The momentum strategy MOM 12/12 longs the top return decile of stocks and shorts the bottom return decile of stocks in the sample based on a formation period of 12 months and a holding period of 12 months. One dollar is invested at all times in both portfolios. Any profits (losses) are reinvested in (financed with) the risk-free asset.

Given the possible permutations of formation and holding periods, we examine 16 momentum strategies in total. While there are certainly momentum strategies with other formation and holding period combinations that one can examine, doing so might expose our results to a data snooping criticism.

The results of our tests for statistical arbitrage are shown in Table 2. We also present *t*-ratio tests on the expected profits of the portfolios for comparison. Consistent with Jegadeesh and Titman (1993), the portfolios' expected profits are almost always statistically greater than zero.<sup>13</sup> Fig. 2 plots and contrasts the cumulative trading profits of two momentum strategies: MOM 6/9 and MOM 12/12. Clearly, MOM 6/9 is preferable to MOM 12/12, even though they both produce positive expected profits.

The statistical arbitrage results from Table 2 confirm the presence of market inefficiency. For 14 of the 16 portfolios, the point estimate for the mean ( $\mu$ ) is greater than zero, and the point estimate for the growth rate of the variance ( $\lambda$ ) is less than zero, consistent with statistical arbitrage. Furthermore, at the 5% level, 13 out of 16 momentum trading strategies have statistically positive estimates for  $\mu$ . However, the requirement that the time-averaged variance of the incremental profits decline over time appears harder to satisfy. Only six of the 16 trading strategies have significantly negative estimates for  $\lambda$  at the 5% level. Nonetheless, the same six trading strategies

<sup>13</sup>The mean incremental profit,  $\mu$ , is related but not identical to the usual mean returns from the trading strategy because  $\mu$  is a dollar-denominated quantity derived from a self-financing trading strategy.

are able to generate statistical arbitrage at the 5% level. Another three trading strategies yield statistical arbitrage at the 10% level. In short, roughly half of the momentum strategies examined converge to riskless arbitrages with decreasing time-averaged variances. This finding is hard to reconcile with the notion of market efficiency.

The results from Table 2 for the MOM 12/12 portfolio illustrate the point that a simple  $t$ -ratio test on the mean is not equivalent to a statistical arbitrage test. The MOM 12/12 portfolio registers a  $t$ -statistic of 3 on the mean and yet fails to be a statistical arbitrage. This is because its time-averaged variance is not decreasing, as suggested in Fig. 2. This observation is more apparent when the cumulative profits from the MOM 12/12 portfolio are compared with those of the MOM 6/9 portfolio (which tests positively for statistical arbitrage).

For comparison with Fig. 1, we use Monte Carlo simulation (once again with 10,000 trials) to generate an empirical distribution for the cumulative trading profits of the MOM 6/9 trading strategy. The empirical distributions of the payoffs are plotted in Fig. 3 over the same four horizons and are consistent with the underlying normal distribution. The results for the MOM 6/9 strategy compare favorably with those of Example 1. The mean payoffs in Fig. 3 are higher than the corresponding mean payoffs in Fig. 1, while the variances in Fig. 3 are smaller. The time-averaged

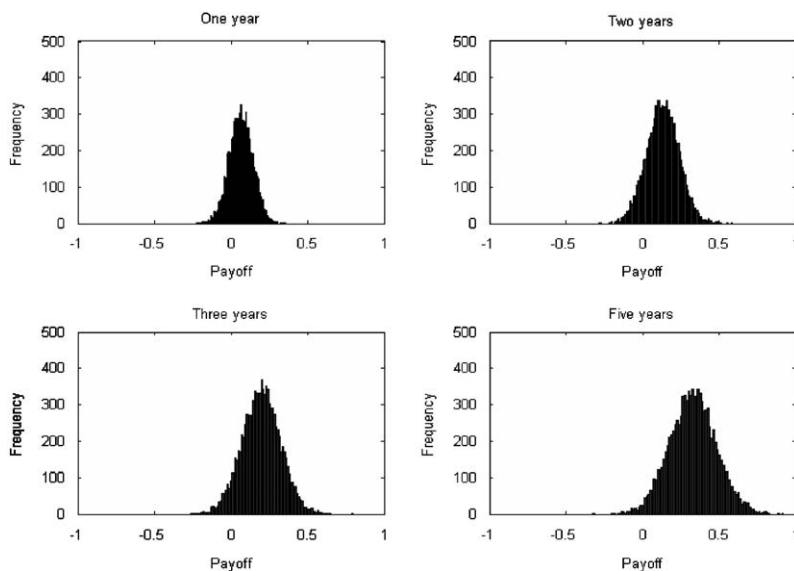


Fig. 3. Discounted cumulative trading profits from a momentum (MOM 6/9) trading strategy. The momentum strategy MOM 6/9 longs the top return decile of stocks and shorts the bottom return decile of stocks in the sample based on a formation period of six months and a holding period of nine months. One dollar is invested at all times. Any profits (losses) are reinvested in (financed with) the risk-free asset. Ten thousand Monte Carlo trials are used to obtain the empirical distribution of the payoffs from the above trading strategy. The frequency distribution of the cumulative payoffs for one, two, three, and five years are illustrated.

variance of MOM 6/9 is declining, in contrast to Example 1. In addition, the probability of a loss declines much more rapidly in Fig. 3 than in Fig. 1.

To get a sense of how fast the statistical arbitrages are converging to riskless arbitrages, we plot the probability of a loss, using Eq. (33), for two representative momentum strategies that constitute statistical arbitrages at the 5%

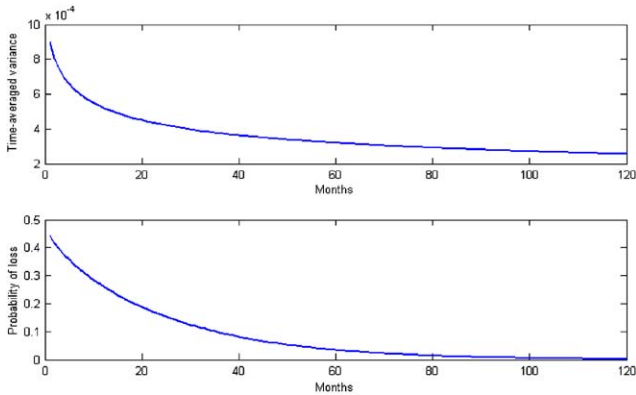


Fig. 4. Probability of a loss and time-averaged variance for the momentum strategy (MOM 6/12) that represents a constrained mean statistical arbitrage opportunity. Sample period is from January 1965 to December 2000. The momentum strategy (MOM 6/12) longs the top return decile of stocks and shorts the bottom return decile of stocks in the sample based on a formation period of six months and a holding period of 12 months. One dollar is invested at all times. Any profits (losses) are reinvested in (financed with) the risk-free asset. The probability of loss is calculated as per Eq. (33).

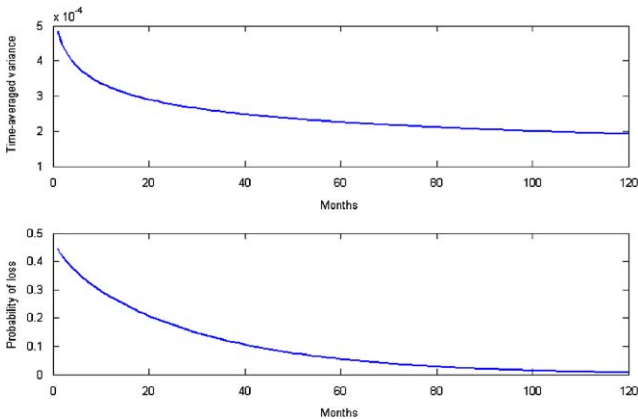


Fig. 5. Probability of a loss and time-averaged variance for the momentum strategy (MOM 9/12) that represents a constrained mean statistical arbitrage opportunity. Sample period is from January 1965 to December 2000. The momentum strategy (MOM 9/12) longs the top return decile of stocks and shorts the bottom return decile of stocks in the sample based on a formation period of nine months and a holding period of 12 months. One dollar is invested at all times. Any profits (losses) are reinvested in (financed with) the risk-free asset. The probability of loss is calculated as per Eq. (33).

level: MOM 6/12 and MOM 9/12 (see Figs. 4 and 5). The time-averaged variances of the strategies are plotted as well. For the MOM 6/12 strategy, after just 89 months, the probability of incurring a loss on this strategy falls below 1%. A casual observation of the time-averaged variance graphs in Figs. 4 and 5 reveals that the MOM 6/12 converges more rapidly to a riskless arbitrage than MOM 9/12. This observation is corroborated by the  $p$ -values and  $\lambda$  estimates reported in Table 2.

## 5.2. Value strategies

Lakonishok et al. (1994) are among the first to examine a comprehensive set of contrarian strategies. They focus on value strategies constructed from stock ratios such as book-to-market, cash flow-to-price, and earnings-to-price, and from past performance variables such as sales growth. They find evidence to suggest that such value strategies are profitable for holding periods ranging from one year to five years. Whether these value strategies contradict the concept of market efficiency remains an open question.

In this section, we test whether the strategies examined by Lakonishok et al. (1994) constitute statistical arbitrage opportunities. Specifically, we test strategies that long the top decile and short the bottom decile of stocks based on book-to-market, cash flow-to-price, or earnings-to-price ratios of the past year. We hold these portfolios for one, three, or five years. We also test a value strategy based on the past three years of sales growth. We long the bottom decile and short the top decile of stocks based on the past three years of sales growth and hold this spread for one, three, or five years.

The results in Table 3 suggest that value strategies based on past three years of sales growth are consistent with statistical arbitrage. For all holding periods considered, the sales growth-based value strategy tests positively for statistical arbitrage at the 5% level. Value strategies based on cash flow-to-price also exhibit this tendency. For holding periods of three and five years, the cash flow-to-price-based value strategies yield statistical arbitrages at the 5% level. However, their expected profits are lower than those of their sales growth-based counterparts. The value strategies based on the other variables do not yield statistical arbitrages at the 5% level. In particular, the point estimates for the variance growth rates of earnings-to-price-based strategies (holding period equals one and three years) are positive (though not significantly so) indicating that these strategies could have become riskier over time.

Figs. 6 and 7 plot the probability of loss and time-averaged variance for two representative value strategies that constitute statistical arbitrage at the 5% level. Supporting the results of Table 2, both the probability of loss and the time-averaged variance of the sales-based value strategy with a holding period of five years are lower than those of the cash flow-to-price-based value strategy with a holding period of five years.

Consistent with the results from Fama and French (1992), among others, the expected profits of strategies based on the book-to-market ratio are statistically greater than zero both from the  $t$ -ratio test results and from the H1 (which is joint

Table 3

Constrained mean (CM) statistical arbitrage tests: Lakonishok et al. (1994) value strategies  
 Parameter estimates and corresponding  $p$ -values for constrained mean test of statistical arbitrage. Sample period is from January 1965 to December 2000. As in Lakonishok et al. (1994), the sorting variables are past one-year book-to-market (BM), cash flow-to-price (CP), and earnings-to-price (EP) as well as past three-year sales growth (SALES). Portfolio  $xy$  denotes a value portfolio based on stock characteristic  $x$  and a holding period of  $y$  years. Every July, stocks are sorted into decile portfolios based on a particular stock characteristic. For BM, CP, and EP portfolios, the portfolio  $xy$  longs the top decile and shorts the bottom decile and holds that spread for  $y$  years. For SALES portfolios, the portfolio  $xy$  longs the bottom decile and shorts the top decile and holds that spread for  $y$  years. The risk-free asset is used to finance the portfolio. H1 and H2 denote the  $p$ -values from statistical arbitrage tests which test whether the portfolio's mean monthly incremental profit is positive and whether its time-averaged variance is declining over time. The sum of the H1 and H2 columns is the  $p$ -value for the statistical arbitrage test. The sum can exceed one because of the Bonferonni inequality. The  $t$ -statistic on the mean monthly trading profit is provided for comparison.

Portfolio	$\mu$ (mean)	$t$ -Statistic	$\sigma$ (standard deviation)	$\lambda$ (growth rate of standard deviation)	H1 ( $\mu > 0$ )	H2 ( $\lambda < 0$ )	Sum (H1 + H2)	Statistical arbitrage
BM1 <sup>a</sup>	0.0141	6.34	0.054	-0.0373	0.000	0.174	0.174	No
BM3 <sup>a</sup>	0.0115	6.54	0.042	-0.0427	0.000	0.166	0.166	No
BM5 <sup>a</sup>	0.0106	6.46	0.041	-0.0720	0.000	0.063	0.063	Yes*
CP1 <sup>a</sup>	0.0021	1.28	0.196	-0.2179	0.249	0.000	0.249	No
CP3 <sup>a</sup>	0.0036	1.74	0.132	-0.2272	0.046	0.000	0.046	Yes**
CP5 <sup>a</sup>	0.0037	1.39	0.131	-0.2685	0.027	0.000	0.027	Yes**
EP1	0.0021	0.25	0.015	0.2320	0.177	1.000	1.177	No
EP3	0.0013	0.57	0.026	0.0735	0.246	0.980	1.226	No
EP5 <sup>a</sup>	0.0011	0.63	0.039	-0.0494	0.268	0.079	0.347	No
SALES1 <sup>a</sup>	0.0085	5.79	0.043	-0.0856	0.000	0.019	0.019	Yes**
SALES3 <sup>a</sup>	0.0056	4.65	0.036	-0.0936	0.000	0.028	0.028	Yes**
SALES5 <sup>a</sup>	0.0046	3.83	0.042	-0.1517	0.000	0.001	0.001	Yes**

\* Significant at the 10% level.

\*\* Significant at the 5% level.

<sup>a</sup> Point estimates of  $\mu$  and  $\lambda$  are consistent with statistical arbitrage.

with H2) test results. However, the growth rates of the variance for such strategies are not statistically less than zero, at the 5% level. Therefore, book-to-market based value strategies do not constitute statistical arbitrage opportunities. This dovetails with the Fama and French (1993) notion that high book-to-market stocks could be systematically riskier than low book-to-market stocks and therefore are entitled to higher expected returns.

The reason that the cash flow-to-price and earnings-to-price strategies do not perform as well as they do in Lakonishok et al. (1994) could be the difference in the sample period. Lakonishok, Shleifer, and Vishny's sample period ends in 1989, while our sample period extends to 2000. In results not shown, plots of the cumulative discounted trading profits generated by cash flow-to-price- and earnings-to-price-based value strategies suggest that these strategies do not perform as well in the post-1990 period as in the pre-1990 period.



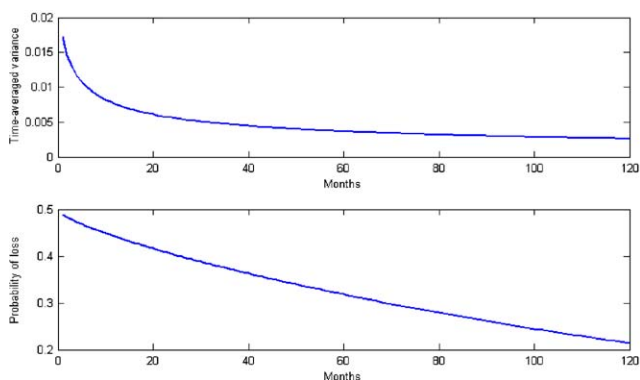


Fig. 6. Probability of a loss and time-averaged variance for the value strategy CP/5 that represents a constrained mean statistical arbitrage opportunity. Sample period is from January 1965 to December 2000. The value strategy CP/5 longs the bottom past year cash flow to price decile of stocks and shorts the top past year cash flow to price decile of stocks in the sample and holds the spread for five years. One dollar is invested at all times. Any profits (losses) are reinvested in (financed with) the risk-free asset. The probability of loss is calculated as per Eq. (33).

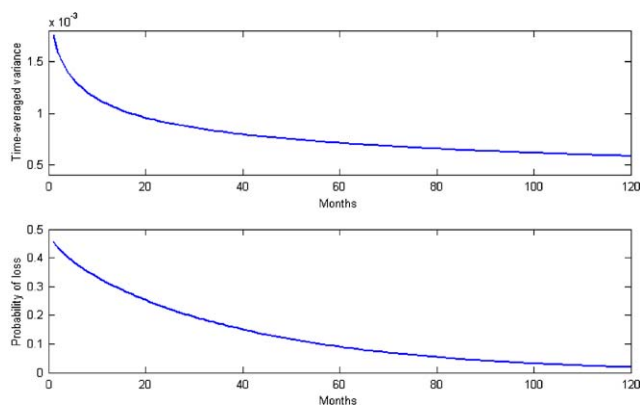


Fig. 7. Probability of a loss and time-averaged variance for the value strategy Sales/5 that represents a constrained mean statistical arbitrage opportunity. Sample period is from January 1965 to December 2000. The value strategy Sales/5 longs the bottom past three-year sales growth decile of stocks and shorts the top past three-year sales growth decile of stocks in the sample and holds the spread for five years. One dollar is invested at all times. Any profits (losses) are reinvested in (financed with) the risk-free asset. The probability of loss is calculated as per Eq. (33).

Overall, roughly half the value strategies considered (five out of 12) provide evidence against market efficiency. This result is similar to that from testing momentum trading strategies. All value and momentum trading strategies are tested on the same data set, as described in Section 4. Thus, statistical tests applied to different trading strategies are not independent, an issue that permeates existing empirical studies. Even if one ignores this dependence, out of the 28 trading

strategies we test, two trading strategies would incorrectly test positive for statistical arbitrage at the 7% level (2/28), slightly above the 5% level we consider sufficient to reject market efficiency. Consequently, Type I error in conjunction with the possibility of dependence amongst the trading strategies could explain a portion of our empirical findings.

## 6. Robustness checks

In this section, we perform a series of tests to gauge the robustness of the results from Section 5. First, we evaluate the appropriateness of the constrained mean model. Second, we test the sensitivity of the statistical arbitrage tests to deviations from our assumed incremental trading process. Third, we measure the impact of market frictions such as transactions costs, margin requirements, liquidity buffers, and a higher borrowing rate on the incremental profits from the momentum and value strategies, and retest for statistical arbitrage. Fourth, we check for the influence of small stocks on our results. Fifth, we investigate the sensitivity of our results to varying the sequence of returns.

### 6.1. Testing the constrained mean model

The first order of business is to verify whether the constrained mean model, which assumes expected trading profits are constant over time, offers a good fit for the incremental profit processes we examine. To do so, we compare measures of fit for the CM model with those of the unconstrained mean model. We also study the estimated rates of change in the expected trading profits to determine whether they are significantly negative and implement a likelihood ratio test. From a statistical perspective, introducing the  $\theta$  parameter, whose point estimates are usually negative, reduces the number of significantly positive  $\mu$  estimates. Overall, for the UM test, the point estimates of  $\mu$  increase along with their standard errors as the information contained in trading profits is spread over a fourth parameter.

The first measure of fit we examine is the average root mean squared error. The RMSE numbers are calculated by comparing the portfolio's observed incremental trading profits with those from 10,000 simulations. Each simulated time series is based on parameters estimated from the observed incremental trading profits of a given trading strategy. Specifically, we simulate an identical number of incremental trading profits as the observed sample, using each trading strategy's parameter estimates. We then compute the RMSE between the observed trading profits and those from the simulation and repeat 10,000 times with the average RMSE reported in Table 4. If the UM model offers a better fit of the data than the CM model, then the RMSE numbers for the former should be lower than those for the latter. However, Table 4, which records the RMSE numbers for both models, demonstrates that the UM model is not statistically better than its CM counterpart. In fact, the RMSE numbers are almost indistinguishable.

Table 4

Comparison between constrained mean and unconstrained mean models

A comparison of the root mean squared errors (RMSE) and sum of squared residuals (SSR) between constrained mean and unconstrained mean models of statistical arbitrage for the momentum portfolios and value portfolios in Tables 2 and 3. The RMSE is based on a Monte Carlo experiment with 10 thousand simulated incremental trading profit time series. The  $p$ -value for the  $t$ -ratio test that the incremental trading profit of the portfolios is declining over time is also presented. The likelihood ratio test evaluates the null that the rate of change of the mean is zero. The test statistics associated with the likelihood ratio tests are displayed and would have to exceed the critical value of 2.71 to reject the null at the 10% level (see footnote 15).

Portfolio	RMSE (CM)	RMSE (UM)	SSR (CM)	SSR (UM)	$p$ -value, $\theta$ (growth rate of mean) $< 0$	Likelihood ratio test
MOM 3/3	0.879	0.877	1.000	1.000	0.83	0.037
MOM 3/6	0.558	0.556	1.001	1.000	0.00	0.066
MOM 3/9**	0.433	0.432	1.003	1.003	0.14	0.017
MOM 3/12**	0.350	0.349	1.004	1.004	0.20	0.014
MOM 6/3	0.986	0.980	1.000	1.000	0.00	0.150
MOM 6/6**	0.581	0.580	1.003	1.003	0.38	0.007
MOM 6/9**	0.479	0.479	1.002	1.002	0.41	0.002
MOM 6/12**	0.371	0.370	1.005	1.004	0.25	0.012
MOM 9/3*	0.948	0.947	1.001	1.001	0.17	0.020
MOM 9/6*	0.587	0.587	1.002	1.002	0.53	0.000
MOM 9/9*	0.450	0.450	1.001	1.001	0.32	0.004
MOM 9/12**	0.335	0.333	1.003	1.003	0.01	0.049
MOM 12/3	0.909	0.907	1.001	1.001	0.07	0.040
MOM 12/6	0.542	0.542	1.000	1.000	0.14	0.017
MOM 12/9	0.417	0.415	0.999	0.999	0.01	0.047
MOM 12/12	0.297	0.293	1.000	1.000	0.00	0.124
BM1	1.279	1.277	1.001	1.001	0.79	0.044
BM3	0.919	0.919	1.001	1.001	0.69	0.010
BM5*	0.740	0.740	1.002	1.002	0.69	0.014
CP1	2.040	2.030	1.009	1.008	0.00	0.206
CP3**	1.255	1.255	1.007	1.007	0.20	0.013
CP5**	1.024	1.024	1.012	1.012	0.55	0.003
EP1	1.468	1.468	0.991	0.997	0.05	0.078
EP3	1.032	1.030	0.998	0.998	0.24	0.012
EP5	0.788	0.788	1.002	1.002	0.43	0.000
SALES1**	0.782	0.782	1.002	1.002	0.48	0.000
SALES3**	0.593	0.592	1.002	1.002	0.19	0.011
SALES5**	0.509	0.509	1.005	1.005	0.28	0.006

\* Constrained mean statistical arbitrages at the 10% level.

\*\* Constrained mean statistical arbitrages at the 5% level.

The second measure of fit we look at is the sum of normalized squared residuals (SSR). If the SSR values of the UM model are lower than those of the CM model, then we can conclude that it is the better model for the data. Yet Table 4 shows that the SSR numbers for the two models are once again almost identical.

Next, we test whether the incremental profits of the portfolios are decreasing using a  $t$ -statistic. The  $p$ -values for this test are also recorded in Table 4. We find that for

all but one of the portfolios that are CM statistical arbitrages at the 10% or 5% level, the growth rates of the incremental profits are statistically indistinguishable from zero.<sup>14</sup> Therefore, there is no need to estimate  $\theta$  and weaken the power of the statistical arbitrage test.

More formally, we employ a likelihood ratio test for the restriction  $\theta = 0$  and compare the results with a  $\chi^2_{1,0.10}$  test statistic equal to 2.71.<sup>15</sup> For all trading strategies under consideration, the null hypothesis that  $\theta = 0$  is accepted without reservation.

In summary, the UM test for statistical arbitrage increases the standard errors of the mean incremental profit estimates, without compensating with a better goodness of fit. Hence the CM model is more appropriate for modeling the observed incremental trading profits.

## 6.2. Accuracy with autocorrelation, jumps, and nonstationarity

To gauge the robustness of our assumed process in Eq. (12), simulations are conducted to investigate the impact of autocorrelation, jumps, and nonstationary parameters on inferences regarding the presence of statistical arbitrage. Using parameters  $\mu$ ,  $\sigma$ , and  $\lambda$  consistent with no statistical arbitrage and statistical arbitrage, a time series of 400 incremental trading profits is simulated, corresponding to the approximate number of months in our sample period.<sup>16</sup> The  $\phi$  parameter used to generate autocorrelation following the  $MA(1)$  process specified in Assumption D.1 of Appendix D is 0.40. This parameter corresponds to a correlation between incremental trading profits of  $0.34 \approx \phi / (1 + \phi^2)$  as seen in Appendix D, and is similar to the average estimate of our trading strategies. Jumps are simulated from a normal distribution, while nonstationary parameters are modeled using a two-state Markov chain with low and high values for  $\mu$  and  $\sigma$ . Statistical arbitrage opportunities are simulated using parameters  $\mu = 0.01$ ,  $\sigma = 0.02$ , and  $\lambda = -0.15$ . Trading profits that are not consistent with statistical arbitrage are generated by the same  $\mu$ , and  $\sigma$ , parameters but with  $\lambda = 0$ . Different values for  $\lambda$  are considered with the statistical procedure performing perfectly whenever  $\lambda > 0$ , with  $\mu \leq 0$  producing similar results. Therefore,  $\lambda = 0$  ensures the test for statistical arbitrage is as difficult as possible, and this value is used to differentiate between an economy with and without the existence of statistical arbitrage. For

<sup>14</sup>This portfolio (MOM 9/12) is a statistical arbitrage according to the UM test. Therefore, for the one instance in which trading profits are better modeled using a UM process with a significantly negative  $\theta$  estimate, the same conclusion is reached regarding market efficiency.

<sup>15</sup>For the likelihood ratio test, abbreviated LRT,

$$-2 \ln(LRT) = 2 \sum_{i=1}^n \ln(\sigma_R i^{\lambda_R}) + \sum_{i=1}^n \frac{(\Delta v_i - \mu_R)^2}{\sigma_R^2 i^{2\lambda_R}} - 2 \sum_{i=1}^n \ln(\sigma_{UR} i^{\lambda_{UR}}) + \sum_{i=1}^n \frac{(\Delta v_i - \mu_{UR} i^{\lambda_{UR}})^2}{\sigma_{UR}^2 i^{2\lambda_{UR}}},$$

and this statistic must be greater than a  $\chi^2$  variable with one degree of freedom at a given level  $\alpha$  to reject the null hypothesis that  $\theta = 0$ . The subscripts R and UR denote the restricted and unrestricted parameter estimates resulting from two separate estimations. See Theorem 8.4.1, Casella and Berger (1990, p. 380).

<sup>16</sup>No null hypotheses for the autocorrelation, jump, or Markov chain parameters are required as we are interested only in studying their impact on results derived under Assumption 1.

completeness, many other possible parameter values are investigated with little change to the final results.

As a second step, estimation involving Lemma 1, based on the assumed process in Eq. (12), is conducted on each simulated time series. A decision regarding statistical arbitrage is rendered according to Theorem 1. A total of 10,000 time series are simulated to heuristically ascertain the power of our statistical test.

The intensity parameter for jumps yields an average of one jump every three years with jump magnitudes normally distributed with zero mean and standard deviation equal to 0.05. A two-state Markov chain generates nonstationary drift and volatility parameters according to the transition matrix

$$\begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$$

for low and high  $\mu$  and  $\sigma$  values. The low and high values for  $\mu$  are 0.005 and 0.01, respectively, while  $\sigma$  have low and high values of 0.015 and 0.03. The parameter values for  $\lambda$  are once again  $-0.15$  and  $0$  to simulate circumstances with and without statistical arbitrage, respectively.

Table 5

Simulation accuracy tests with deviations from assumed process

This table reports simulation results on the accuracy of the statistical arbitrage test in the face of deviations from the assumed incremental profit process. The deviations include autocorrelation, jumps, and parameter nonstationarity. Panel A details the situations in which the original parameters are consistent with statistical arbitrage ( $\lambda$  is set to  $-0.15$ ). Panel B details the situations in which the original parameters are consistent with no statistical arbitrage ( $\lambda$  is set to  $0$ ). The parameter  $\lambda$  is the growth rate of the standard deviation of the incremental profits. The first column details the departure from the assumed dynamics in Eq. (12). Columns two and three record the number of instances in the 10 thousand simulated time series that the null hypothesis of no statistical arbitrage is correctly rejected or accepted at the 5% and the 10% levels.

*Panel A: true state = statistical arbitrage*

Type of deviation	Percentage of simulations that correctly reject the null of no statistical arbitrage	
	At the 5% level	At the 10% level
No deviation	99	99
Autocorrelation	95	98
Jumps	60	70
Nonstationary $\mu, \sigma$	71	81

*Panel B: true state = no statistical arbitrage*

Type of deviation	Percentage of simulations that correctly accept the null of no statistical arbitrage	
	At the 5% level	At the 10% level
No deviation	100	95
Autocorrelation	100	97
Jumps	99	95
Nonstationary $\mu, \sigma$	100	95

The number of conclusions that are altered by autocorrelation, jumps, or nonstationary parameters can be inferred from Table 5. The statistical test correctly accepts the null hypothesis of no statistical arbitrage with great accuracy, even when there exists a deviation from the assumed process. At the 10% level, when the underlying parameters are not consistent with statistical arbitrage, the test rejects the null hypothesis no more than 5% of the time.

Overall, the performance of the statistical test is exceptional even with deviations in the assumed process. Clearly, the presence of autocorrelation, jumps, and parameter nonstationarity bias the test toward accepting the null hypothesis of no statistical arbitrage. Hence, the formulation in Eq. (12) could be considered fail-safe in the presence of deviations. This property stems from the fourth property of statistical arbitrage. The additional volatility caused by jumps and nonstationary parameters increases the standard error of  $\lambda$ , which translates into higher corresponding  $p$ -values for H2, and consequently a higher probability of accepting the null hypothesis of no statistical arbitrage. Along with the Bonferroni inequality, these simulations demonstrate that our test for statistical arbitrage is very conservative when rejecting the null hypothesis of no statistical arbitrage, allowing one to confidently reject market efficiency.

### 6.3. Transactions costs and market frictions

A common critique of financial anomalies is that the trading profits from such anomalies disappear after adjusting for transactions costs. Therefore, we estimate the transactions costs of the portfolios in Section 5 as follows. First, we estimate the average monthly turnover for each of the portfolios by taking a ratio of the sum of buys and sells each period over two times the total number of stocks held in that period. This estimates the number of round-trip transactions as a percentage of the number of stocks held. Then we multiply this measure of monthly turnover with an estimate of the round-trip transactions cost given by Chan and Lakonishok (1997). According to Chan and Lakonishok, the average stock in the Nasdaq has a round-trip transactions cost of 1.34%.<sup>17</sup> Finally, we adjust the monthly profits downward by the transactions costs.

We also consider the impact of four market frictions on the ability of momentum and value trading strategies to yield statistical arbitrage. Our analysis follows Alexander (2000). The first three frictions consist of the margin ( $m$ ) required for both long and short positions, the additional margin known as a liquidity buffer for the marking-to-market of short positions ( $c$ ), and the levy brokerages implicitly impose on their clients by withholding a small percentage of the interest earned on their margin accounts ( $h$ ). Moreover, we also impose a borrowing rate for financing trading losses that is higher than the lending rate for investing trading profits.

---

<sup>17</sup> We also calculate transactions costs using the higher NYSE estimate of 1.55% and retest for statistical arbitrage. None of the inferences changes with the NYSE estimate for round-trip transactions cost.

More specifically, the adjusted return on the portfolio denoted  $r^{\text{adj}}$  equals

$$r^{\text{adj}} = (r^{\text{unadj}} - h)/(2m + c) \quad (34)$$

as in Eq. (3) of Alexander (2000), where  $r^{\text{unadj}}$  represents the unadjusted portfolio return used in our previous analysis. See Alexander (2000) for a detailed discussion.

Jacobs and Levy (1995) estimate the levy on the margin account to be between 25 and 30 basis points per year, while Jacobs and Levy (1997) find the liquidity buffer to be 10% of short sale proceeds. Hence, we select a liquidity buffer for short sales of 10%, and a levy on the margin account of 27.5 basis points. Furthermore, we set a conservative margin rate of 50%.<sup>18</sup> The borrowing rate is chosen to be 2% higher (per year) than the lending rate. Using AA 30 day commercial paper for financial firms available from the St. Louis Federal Reserve, we find that the maximum spread over Treasury bills of the same maturity is 1.7% from 1997 to 2002, the period for which data are available. Therefore, our choice of a 2% spread is very conservative. The higher borrowing rate applies only in those infrequent instances when the cumulative trading profit for a statistical arbitrage is negative. Thus, a higher lending rate is unlikely to have significant repercussions for our previous analysis.

After incorporating these four frictions and accounting for transaction costs, the statistical arbitrage results, showcased in Table 6, are nearly identical to previous results in Tables 2 and 3. Although the estimated  $\mu$  (mean incremental profit) parameters are smaller, of the six momentum trading strategies that originally test positive for statistical arbitrage at the 5% level, five remain statistical arbitrages at this level. For value trading strategies, four of the five strategies remain statistical arbitrages at the 5% level, and the other represents a statistical arbitrage at the 10% level. Thus, in the presence of realistic trading frictions, momentum and value trading strategies continue to test positively for statistical arbitrage.

#### 6.4. Influence of small stocks

Several tests of stock market anomalies conclude that small stocks are more inefficient than large stocks (Hong et al., 2000; Mitchell and Stafford, 2000). In addition, it is well known that small stocks are less liquid and have greater transactions costs than large stocks (Chan and Lakonishok, 1997). Therefore, it could be useful to see if our results survive the exclusion of these illiquid stocks.

To this end, we remove stocks in our sample with market capitalizations below the 50th NYSE market equity percentile (which is about \$1 billion in 2000). This reduces our sample size by more than two-thirds because there are many more small stocks than large stocks and there are relatively more small stocks traded in the Nasdaq (which are included in the CRSP database) than in the NYSE. We then retest our trading strategies for statistical arbitrage.

Results from the tests of statistical arbitrage suggest that evidence of statistical arbitrage is even greater without the bottom 50th NYSE percentile. All but one of

<sup>18</sup> See the Securities and Exchange Commission website, <http://www.sec.gov/investor/pubs/margin.htm>, for details regarding Regulation T and justification for our choice of 50%.

Table 6

Transactions costs and other market frictions adjusted, constrained mean (CM) tests of statistical arbitrage

Constrained test of statistical arbitrage after adjusting for market frictions such as transactions cost, margins, liquidity buffers, levies on margin interest, and a higher borrowing rate. The sample period and column variables are as per defined in Tables 2 and 3. All parameter estimates are calculated from portfolio incremental profits after adjusting for five types of market frictions: (1) margin ( $m$ ) required on long and short positions, (2) liquidity buffer ( $c$ ) for short positions, (3) the levy ( $h$ ) implicitly imposed on the margin interest, (4) a higher borrowing rate than lending rate, and (5) transactions costs. The values of  $h$  and  $c$  are set at 27.5 basis points and 10%, respectively, following Jacobs and Levy (1995, 1997). The value of  $m$  is set at a conservative 50%. The borrowing rate is set 2% higher than the lending rate (which equals the risk-free rate). This spread is higher than the average spread between the AA 30-day commercial paper rate and the risk-free rate from 1997 to 2002. The incremental profits are adjusted for (1), (2), (3), and (4) using the methodology of Alexander (2000) and the adjustment process is detailed in Eq. (34) of the text. Estimated transactions costs per month are then deducted from the adjusted incremental profits. Estimated transactions cost per month is estimated round-trip transactions cost times turnover where round-trip transactions cost = 1.34%, the transactions costs for the average stock traded in the Nasdaq (Chan and Lakonishok, 1997). Turnover is the number of round-trip transactions per month over the number of stocks held in the portfolio. Max loss is the minimum cumulative profit, and max drawdown is the minimum monthly incremental profit.

Portfolio	Turnover (%)	Max loss of \$1 invested	Max drawdown	$\mu$ (mean)	$\sigma$ (standard deviation)	$\lambda$ (growth rate of standard deviation)	H1 ( $\mu > 0$ )	H2 ( $\lambda < 0$ )	Sum (H1 + H2)	Statistical arbitrage
MOM 3/9 <sup>a</sup>	9.26	0.0949	0.0676	0.0006	0.0232	-0.108	0.193	0.001	0.195	No
MOM 3/12 <sup>a</sup>	7.22	0.0425	0.0540	0.0019	0.0210	-0.131	0.000	0.000	0.000	Yes**
MOM 6/6 <sup>a</sup>	14.20	0.1131	0.0994	0.0019	0.0337	-0.125	0.016	0.000	0.016	Yes**
MOM 6/9 <sup>a</sup>	9.49	0.1091	0.0670	0.0036	0.0236	-0.091	0.000	0.004	0.004	Yes**
MOM 6/12 <sup>a</sup>	7.23	0.0418	0.0576	0.0027	0.0271	-0.173	0.000	0.000	0.000	Yes**
MOM 9/12 <sup>a</sup>	7.24	0.0052	0.0446	0.0017	0.0196	-0.125	0.000	0.001	0.001	Yes**
CP3 <sup>a</sup>	2.17	0.2250	0.2390	0.0028	0.1198	-0.227	0.075	0.000	0.075	Yes*
CP5 <sup>a</sup>	1.46	0.3957	0.2139	0.0029	0.1189	-0.268	0.045	0.000	0.045	Yes**
SALES1 <sup>a</sup>	4.07	0.0817	0.0650	0.0070	0.0392	-0.086	0.000	0.019	0.019	Yes**
SALES3 <sup>a</sup>	2.41	0.0246	0.0564	0.0046	0.0324	-0.094	0.000	0.028	0.028	Yes**
SALES5 <sup>a</sup>	1.52	0.0487	0.0777	0.0038	0.0383	-0.152	0.000	0.001	0.001	Yes**

\* Significant at the 10% level.

\*\* Significant at the 5% level.

<sup>a</sup> Point estimates of  $\mu$  and  $\lambda$  are consistent with statistical arbitrage.



the momentum strategies and seven out of 12 value strategies are statistical arbitrages at the 5% level. Further, all the momentum portfolios that are statistical arbitrages for the full sample are also statistical arbitrages for the top 50th NYSE market equity percentile. The value portfolios are less robust to the variation in sample composition. The sales based strategies fail to be statistical arbitrages with the reduced sample. However there exist strategies based on cash flow-to-price, book-to-market equity, and earnings-to-price that become statistical arbitrages following the removal of small stocks. This clearly indicates that our results are not driven by small stocks.

### 6.5. Sensitivity to the sequence of returns

In response to concerns that statistical arbitrage could be an artifact of the observed sequence of returns, a further robustness check is conducted on the  $\mu$  (mean incremental profit) and  $\lambda$  (growth rate of standard deviation of incremental profits) parameters of all trading strategies. This additional robustness check involves bootstrapping  $t$ -statistics to ascertain their distribution and reaffirms our conclusions regarding statistical arbitrage. Our emphasis is on reporting the results for  $\lambda$ , the parameter most often responsible for determining whether a trading strategy constitutes statistical arbitrage. In most cases, when  $\mu$  is not significantly positive,  $\lambda$  is not significantly negative (with possible exceptions being the cash flow-to-price-based value strategies CP3 and CP5, whose results are reported below).

We implement the following four-step test procedure to verify whether the estimates for  $\lambda$  (as well as  $\mu$ ) are statistically negative (positive):

1. Compute residuals based on the sample estimates of the  $\mu$ ,  $\lambda$ , and  $\sigma$  (standard deviation of incremental profits) values. This sequence of residuals forms the basis for the bootstrap procedure in the next step. The bootstrap alters the sequence of residuals, hence returns, according to a set of uniformly distributed integers based on the sample size of each individual strategy.
2. Bootstrap 10,000 sequences of trading profits using the residuals from the original sequence under the assumption that  $\lambda = 0$ , while  $\mu$  and  $\sigma$  are set equal to their sample estimates.
3. For each of the 10,000 trials, compute the corresponding  $t$ -statistic for  $\lambda$ . Because we are testing whether  $\lambda$  is negative, we are interested in the 5th percentile of these simulated  $t$ -statistics as smaller values are indicative of greater significance. Denote the value at the 5th percentile as the critical  $t$ -statistic,  $t_{\text{crit}}$ .
4. Compare the sample  $t$ -statistic with  $t_{\text{crit}}$  for each trading strategy. If the sample  $t$ -statistic is lower than  $t_{\text{crit}}$ , reject the hypothesis that  $\lambda$  is non-negative and conclude that the trading strategy is a statistical arbitrage.

We find that the values of  $t_{\text{crit}}$  are similar for the various trading strategies, with most larger than  $-1.70$ . This result supports our earlier conclusions reported in Tables 2 and 3. For example, the MOM 9/12 strategy has a  $t_{\text{crit}}$  of  $-1.655$ , while its sample  $t$ -statistic is  $-3.031$ , allowing us to confidently reject the null hypothesis and

conclude that  $\lambda$  is negative, as required for statistical arbitrage. Similarly, for the SALES 5 strategy,  $t_{\text{crit}}$  equals  $-1.641$ , while the sample  $t$ -statistic is  $-4.176$ .

Furthermore, we also confirm that our rejections of statistical arbitrage are justified. For example, the MOM 12/12 strategy has a critical value of  $-1.652$ , smaller than the sample  $t$ -statistic of  $0.012$ . This outcome is consistent with accepting the null hypothesis that  $\lambda$  is nonnegative, reinforcing the conclusion that MOM 12/12 is not a statistical arbitrage.

When testing whether  $\mu$  is statistically positive, we are interested in the 95th percentile of the bootstrapped  $t$ -statistics to determine  $t_{\text{crit}}$ . For CP3 and CP5, the  $t_{\text{crit}}$  statistics for  $\mu$  from the bootstrap procedure are  $1.781$  and  $1.713$ , respectively, both of which are below the sample  $t$ -statistics. Thus, we are able to confidently reject the hypothesis that  $\mu$  is nonpositive and conclude that CP3 and CP5 remain statistical arbitrage opportunities.

In summary, the bootstrap procedure confirms that a trading strategy's ability to test for (or not test for) statistical arbitrage is not unduly influenced by the observed sequence of returns.

#### 6.6. Final comments on robustness

It is important to emphasize the conservative nature of our statistical test. First, despite the accuracy of our estimation procedure, the Bonferroni inequality requires the sum of the  $p$ -values to be below the  $\alpha$  level. Second, deviations from our assumed process bias the test toward accepting the null hypothesis of no statistical arbitrage. Hence, evidence of statistical arbitrage is unlikely the consequence of a misspecified trading profit process.

For completeness and in view of the potential applications to other trading strategies in future research, a brief summary of statistical tests that account for autocorrelated trading profits is provided in Appendix D. These tests do not lead to different conclusions regarding market efficiency for the trading strategies considered in this paper. Furthermore, we consider the maximum losses incurred while pursuing our statistical arbitrage opportunities, which equal the minimum values of the cumulative trading profits. These values are reported in Table 6 and proxy for the amount of capital required to engage in trading strategies that offer statistical arbitrage. As shown in Table 6, these values are not prohibitive, especially in light of the potential rewards.

### 7. An application: the size effect

It is often believed that the size effect of Banz (1981) disappeared after the mid-1980s. Hence, the size effect offers an interesting test case for our proposed statistical arbitrage methodology. In this section, we investigate whether the strategy of buying small stocks and shorting large stocks constitutes a statistical arbitrage. Using the Fama-French Small Minus Big (SMB) factor as a proxy for the size effect, we consider the incremental and cumulative trading profits based on a trading strategy

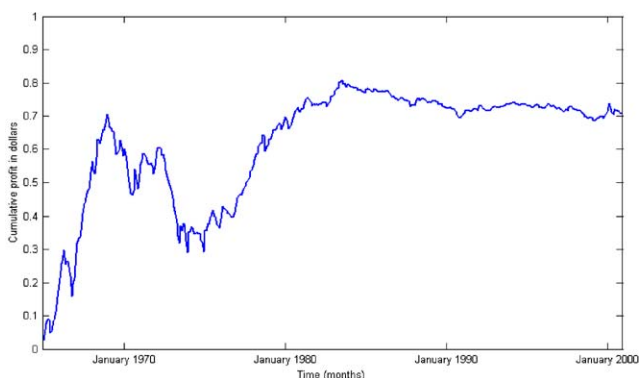


Fig. 8. Cumulative trading profits from \$1 invested in the Fama and French (1992) Small Minus Big (SMB) portfolio. Sample period is from January 1965 to December 2000. The SMB portfolio longs stocks with market equity (*ME*) below the 30th NYSE *ME* percentile and shorts stocks above the 70th NYSE *ME* percentile. Any profits (losses) are reinvested in (financed with) the risk-free asset.

that buys stocks with market capitalizations below the 30th NYSE percentile and shorts stocks with market capitalizations above the 70th NYSE percentile. The simple *t*-statistic on the mean profit is 2.11, consistent with positive expected profits. However this does not imply that the SMB-based trading strategy violates market efficiency.

Applying the statistical arbitrage methodology yields parameter estimates for  $\mu$ ,  $\sigma$ , and  $\lambda$  (mean, standard deviation, and growth rate of standard deviation of incremental profits) of 0.0002, 0.6404, and  $-0.7961$ , respectively, while the *p*-value for statistical arbitrage equals 0.294. Hence, the size effect does not constitute a statistical arbitrage. A plot of the cumulative profits from the SMB-based strategy in Fig. 8 reveals that the SMB trading profits dissipated around the mid-1980s. In fact, strategy's mean incremental profit after 1985 is negative, implying that statistical arbitrage portrays the market efficiency implications of the size effect more accurately than a *t*-ratio test. In summary, our findings corroborate the popular belief that the size effect disappeared around the mid-1980s.

## 8. Conclusion

This paper introduces the concept of statistical arbitrage, which allows one to test for market efficiency without having to specify an equilibrium model. In the limit, statistical arbitrage converges to arbitrage. Consequently, the joint hypothesis dilemma is avoided by appealing to a long horizon trading strategy. Methodologies that test for the existence of statistical arbitrage are provided. Simulations reveal the robustness and the conservative nature of our statistical test. It is biased toward accepting the null hypothesis of no statistical arbitrage in the presence of deviations from our assumed trading profit process, such as autocorrelation, jumps, and nonstationary parameters.

We test the cumulative discounted trading profits of momentum (Jegadeesh and Titman, 1993) and value trading strategies (Lakonishok et al., 1994) for statistical arbitrage. In contrast to the  $t$ -test of positive risk-adjusted expected returns, the existence of statistical arbitrage is sufficient to reject market efficiency for any model of market returns.

Roughly half of the momentum and value strategies we evaluate test positively for statistical arbitrage. This implies that several trading opportunities that converge to riskless arbitrages with decreasing time-averaged variances exist, a result difficult to reconcile with market efficiency. Our test results are robust to the exclusion of small stocks and to the incorporation of several market frictions such as transactions costs, margin accounts, short-selling buffers, and higher borrowing rates.

Promising avenues for future research include testing other anomalies, such as the abnormal returns from dividends or earnings announcements, for statistical arbitrage. Also, the Bonferroni inequality could be replaced by a more computationally intensive Monte Carlo procedure to simulate the critical values underlying our joint hypothesis test.

#### Appendix A. Importance of fourth condition

Consider the standard Black-Scholes economy with

$$S_t = S_0 e^{\alpha t - \sigma^2 t/2 + \sigma W_t}, \quad (\text{A.1})$$

and assume  $r = 0$  implying  $B_t = 1$  for all  $t$ . This economy is in equilibrium but would be rejected by a definition of statistical arbitrage that does not include the fourth condition. To prove this, first consider a simpler economy with stock dynamics

$$\bar{s}_t = \bar{s}_0 + \alpha t - \sigma^2 t/2 + \sigma W_t \quad (\text{A.2})$$

and a buy and hold strategy similar to Example 2 whose value at time  $t$  equals  $\bar{v}(t) = 1 \cdot \bar{s}_t - \bar{s}_0 \cdot 1$ . For  $r = 0$ , cumulative trading profits are identical to discounted cumulative trading profits. Cumulative trading profit dynamics are represented as

$$\begin{aligned} \bar{v}(t) &= \bar{s}_0 + \alpha t - \sigma^2 t/2 + \sigma W_t - \bar{s}_0 \\ &= [\alpha - \sigma^2/2]t + \sigma W_t. \end{aligned} \quad (\text{A.3})$$

If  $\alpha - \sigma^2/2 > 0$ , then  $P(\bar{v}(t) < 0)$  converges to zero. This claim follows from Appendix C with  $\lambda$  and  $\theta$  both equaling zero and  $\mu = \alpha - \sigma^2/2$ . This portfolio would also satisfy the first and second conditions of statistical arbitrage. Returning to the Black-Scholes economy with  $S_0 = 1$ , consider the same buy and hold trading strategy with cumulative trading profits

$$v(t) = 1 \cdot S_t - S_0 \cdot 1 = S_t - 1 = e^{\bar{s}_t} - 1. \quad (\text{A.4})$$

The following sequence illustrates that  $P(v(t) < 0)$  also converges to zero:

$$P(v(t) < 0) = P(e^{\bar{s}_t} - 1 < 0) < P(\bar{s}_t - 1 < 0) = P(\bar{v}(t) < 0), \quad (\text{A.5})$$

where the second relation employs the inequality  $S_t = e^{\bar{s}_t} > \bar{s}_t$ . Thus, without the fourth condition, the definition of statistical arbitrage would apply to a buy and hold strategy in the Black-Scholes economy, which is known to be in equilibrium.

**Appendix B. Doubling strategies**

The canonical discrete time doubling strategy is a fair coin toss with an investor winning \$1 if the coin toss is heads and doubling the bet if the coin toss is tails. This strategy continues until the first heads is attained, at which time the strategy is stopped. LeRoy (2002) also considers such a strategy [Eq. (15) of Section 5]. To demonstrate that our empirical findings are not the result of a doubling strategy, we consider their cumulative trading profits  $v(t_i)$ , which equal  $-1, -3, -7, -15, \dots$  and so on for  $t_i = 1, 2, 3, 4, \dots$  until the stopping time  $t^*$  when  $v(t^*) = 1$ .<sup>19</sup> Thus, an investor collects \$1, after accounting for the series of increasingly large previous losses, when the first heads is attained.

Until the stopping time is reached, all incremental trading profits are negative with the investor losing  $2^{t_i-1}$  dollars at  $t_i = 1, 2, 3, 4, \dots$  before  $t^*$ . The doubling strategy has cumulative trading profits equal to

$$v(t_i) = \begin{cases} 1 & \text{with probability } 1 - (\frac{1}{2})^{t_i}, \\ -2^{t_i} + 1 & \text{with probability } (\frac{1}{2})^{t_i} \end{cases} \tag{B.1}$$

and converges almost surely to a riskless arbitrage although  $v(t_0 = 0) = 0$ . However, the doubling strategy is not a statistical arbitrage opportunity because the expectation and variance of  $v(t_i)$  are zero (instead of positive) and  $2^{t_i} - 1$ , respectively. Therefore, the time-averaged variance does not decline as  $(2^{t_i} - 1)/t_i \rightarrow \infty$ . In summary, doubling strategies violate the second and fourth properties of statistical arbitrage in Definition 1.

**Appendix C. Zero probability of loss in limit**

Eq. (13) states that  $v(t_n)$  is distributed

$$N\left(\mu \sum_{i=1}^n i^\theta, \sigma^2 \sum_{i=1}^n i^{2\lambda}\right). \tag{C.1}$$

Convergence of the tail probability,  $P(v(t_n) < 0)$ , to zero is determined by the cumulative normal distribution

$$P(v(t_n) < 0) = \frac{1}{\sqrt{2\pi\sigma^2 \sum_{i=1}^n i^{2\lambda}}} \int_{-\infty}^0 e^{-\frac{1}{2\sigma^2 \sum_{i=1}^n i^{2\lambda}} (x - \mu \sum_{i=1}^n i^\theta)^2} dx. \tag{C.2}$$

<sup>19</sup>Without loss of generality, we assume the risk-free interest rate is zero.

The above probability can be expressed as

$$P(v(t_n) < 0) = P\left(Z < \frac{-\mu \sum_{i=1}^n i^\theta}{\sigma \sqrt{\sum_{i=1}^n i^{2\lambda}}}\right) = N\left(\frac{-\mu \sum_{i=1}^n i^\theta}{\sigma \sqrt{\sum_{i=1}^n i^{2\lambda}}}\right), \tag{C.3}$$

where  $Z$  is a standard normal and  $N(x)$  is the standard normal cumulative distribution function. Although there is no closed form solution for the standard normal cumulative density function, a polynomial approximation (for  $x < 0$ ) is available (see Hull, 1997, p. 243).

$$N(x) = N'(x) \left( a_1 \frac{1}{1 + \gamma x} + a_2 \frac{1}{(1 + \gamma x)^2} + a_3 \frac{1}{(1 + \gamma x)^3} + h.o.t. \right), \tag{C.4}$$

where  $a_1 = 0.4361836$ ,  $a_2 = -0.1201676$ ,  $a_3 = 0.9372980$ , and  $\gamma = 0.33267$  are constants. The leading term governs the rate of convergence implying

$$P(v(t_n) < 0) = O\left( \frac{a_1}{\sqrt{2\pi}} \exp\left\{ -\frac{\mu^2 (\sum_{i=1}^n i^\theta)^2}{2\sigma^2 \sum_{i=1}^n i^{2\lambda}} \right\} \frac{1}{1 + \gamma \frac{\mu \sum_{i=1}^n i^\theta}{\sigma \sqrt{\sum_{i=1}^n i^{2\lambda}}}} \right), \tag{C.5}$$

which converges to zero as  $n$  increases under the assumption of statistical arbitrage. In particular, the probability converges faster than rate

$$O\left( \exp\left\{ -\frac{\mu^2 (\sum_{i=1}^n i^\theta)^2}{2\sigma^2 \sum_{i=1}^n i^{2\lambda}} \right\} \right) \tag{C.6}$$

as a function of  $n$ .

**Appendix D. Correlated trading profit increments**

Trading strategies could produce autocorrelated trading profits. For example, the momentum trading strategies have profits derived from overlapping periods. Consider an  $MA(1)$  process to instill correlation between consecutive trading profit increments according to the parameter  $\phi$ .<sup>20</sup> Specifically, the correlation between  $\Delta v_i$  and its previous value  $\Delta v_{i-1}$  is approximately  $\phi/(1 - \phi^2)$  under the following assumption.

**Assumption D.1.** Let the discounted incremental trading profits satisfy

$$\Delta v_i = \mu i^\theta + i^\lambda z_i + \phi(i - 1)^\lambda z_{i-1} \tag{D.1}$$

for  $i = 1, 2, \dots, n$ , where  $z_i$  are i.i.d.  $N(0, \sigma^2)$  random variables with  $z_0 = 0$ .

For parameter estimation, several iterative procedures exist as detailed in Box and Jenkins (1976, Chapter 7). The distribution of  $v(t_n)$  is required to construct a test for

<sup>20</sup> Generalizing the process to an  $MA(q)$  for  $q \geq 2$  follows in a straightforward manner.

statistical arbitrage and determine the convergence rate to arbitrage. The distribution of the cumulative discounted trading profits is the topic of Lemma D.1.

**Lemma D.1.** *The distribution of  $v(t_n)$  with correlated increments equals*

$$v(t_n) = N\left(\mu \sum_{i=1}^n i^\theta, \sigma^2 \left[ (1 + \phi)^2 \sum_{i=1}^{n-1} i^{2\lambda} + n^{2\lambda} \right]\right). \tag{D.2}$$

The distribution of  $v(t_n)$  is solved by recursion

$$\begin{aligned} v(t_i) &= v(t_{i-1}) + \mu i^\theta + i^\lambda z_i + \phi(i-1)^\lambda z_{i-1} \\ &= v(t_{i-2}) + \mu(i-1)^\theta + (i-1)^\lambda z_{i-1} + \phi(i-2)^\lambda z_{i-2} \\ &\quad + \mu i^\theta + i^\lambda z_i + \phi(i-1)^\lambda z_{i-1} \\ &= v(t_{i-3}) + \mu(i-2)^\theta + (i-2)^\lambda z_{i-2} + \phi(i-3)^\lambda z_{i-3} \\ &\quad + \mu(i-1)^\theta + (i-1)^\lambda z_{i-1} + \phi(i-2)^\lambda z_{i-2} \\ &\quad + \mu i^\theta + i^\lambda z_i + \phi(i-1)^\lambda z_{i-1} \end{aligned} \tag{D.3}$$

and so forth generating the following solution for  $v(t_n)$ :

$$\begin{aligned} v(t_n) &= \mu \sum_{i=1}^n i^\theta + (1 + \phi) \sum_{i=1}^{n-1} i^\lambda z_i + n^\lambda z_n \\ &= N\left(\mu \sum_{i=1}^n i^\theta, \sigma^2 \left[ (1 + \phi)^2 \sum_{i=1}^{n-1} i^{2\lambda} + n^{2\lambda} \right]\right). \end{aligned} \tag{D.4}$$

As expected, Eq. (D.2) reduces to Eq. (13) when  $\phi = 0$ .

**Theorem D.1.** *Under the previous assumption, a trading strategy generates a statistical arbitrage with  $1 - \alpha$  percent confidence if the following conditions are satisfied:*

$$H1 : \hat{\mu} > 0, \tag{D.5}$$

$$H2 : \hat{\lambda} < 0, \text{ and} \tag{D.6}$$

$$H3 : \hat{\theta} > \max\{\hat{\lambda} - \frac{1}{2}, -1\}, \tag{D.7}$$

with the sum of the individual  $p$ -values forming an upper bound for the Type I error.

The variance in Eq. (D.4) can be modified to equal  $\sigma^2(1 + \phi)^2 \sum_{i=1}^n i^{2\lambda}$  by adding the term  $\sigma^2(2\phi + \phi^2)n^{2\lambda}$ . The marginal contribution of this additional term

$$\frac{\sigma^2((1 + \phi)^2 \sum_{i=1}^{n-1} i^{2\lambda} + n^{2\lambda}) + \sigma^2(2\phi + \phi^2)n^{2\lambda}}{\sigma^2((1 + \phi)^2 \sum_{i=1}^{n-1} i^{2\lambda} + n^{2\lambda})} \rightarrow 0 \tag{D.8}$$

as  $n$  increases, ensuring asymptotic inferences remain valid. Consequently, for large  $n$ , the approximation  $v(t_n) \sim^d N(\mu \sum_{i=1}^n i^\theta, \sigma^2(1 + \phi)^2 \sum_{i=1}^n i^{2\lambda})$  is analyzed. Therefore, testing for statistical arbitrage involves the same exact joint hypotheses as Theorem 1 but with a modified variance that accounts for  $\phi$ .

The next theorem follows immediately from Theorem 2.

**Theorem D.2.** *The rate of convergence from statistical arbitrage to arbitrage is exponential with respect to  $n$ . Specifically, the convergence rate is faster than*

$$O\left(\exp\left\{-\frac{\mu^2(\sum_{i=1}^n i^\theta)^2}{2\sigma^2(1+\phi)^2\sum_{i=1}^n i^{2\lambda}}\right\}\right). \quad (\text{D.9})$$

Theorems D.1 and D.2 illustrate that, conditional on parameters  $\mu, \sigma^2, \theta$ , and  $\lambda$  being equal, positive correlation implies slower convergence (higher variance) with  $(1 + \phi)^2 > 1$ , while negative correlation implies faster convergence (lower variance) with  $(1 + \phi)^2 < 1$ .

## References

- Alexander, G.J., 2000. On back-testing “zero-investment” strategies. *Journal of Business* 73, 255–278.
- Banz, R.W., 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9, 3–18.
- Bernardo, A.E., Ledoit, O., 2000. Gain, loss, and asset pricing. *Journal of the Political Economy* 108, 144–172.
- Bodurtha Jr., J.N., Courtadon, G.R., 1986. Efficiency tests of the foreign currency options market. *Journal of Finance* 41, 151–162.
- Box, G.E.P., Jenkins, G.M., 1976. *Time Series Analysis: Forecasting and Control*, 3rd Edition. Prentice-Hall, Englewood Cliffs, NJ.
- Carhart, M., 1997. On persistence of mutual fund performance. *Journal of Finance* 52, 57–82.
- Carr, P., Geman, H., Madan, D., 2001. Pricing and hedging in incomplete markets. *Journal of Financial Economics* 62, 131–167.
- Casella, G., Berger, R.L., 1990. *Statistical Inference*. Duxbury Press, California.
- Chan, L.K.C., Lakonishok, J., 1997. Institutional equity trading costs: NYSE versus Nasdaq. *Journal of Finance* 52, 713–735.
- Chan, L.K.C., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. *Journal of Finance* 51, 1681–1713.
- Cochrane, J.H., Saá-Requejo, J., 2000. Beyond arbitrage: good-deal asset price bounds in incomplete markets. *Journal of the Political Economy* 108, 79–119.
- Duffie, D., 2001. *Dynamic Asset Pricing Theory*, 3rd Edition. Princeton University Press, Princeton, NJ.
- Dybvig, P.H., Huang, C., 1988. Nonnegative wealth, absence of arbitrage, and feasible consumption plans. *Review of Financial Studies* 1, 377–401.
- Fama, E.F., 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49, 283–306.
- Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. *Journal of Finance* 47, 427–465.
- Fama, E.F., French, K.R., 1993. Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics* 33, 3–56.
- Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55, 265–295.
- Hull, J.C., 1997. *Options, Futures, and Other Derivatives*, 3rd Edition. Prentice-Hall, Englewood Cliffs, NJ.
- Jacobs, B.I., Levy, K.L., 1995. More on long-short strategies. *Financial Analysts Journal* 51, 88–90.
- Jacobs, B.I., Levy, K.L., 1997. The long and short of long-short. *Journal of Investing* 6, 73–86.



- Jarrow, R., 1988. *Finance Theory*. Prentice-Hall, Englewood Cliffs, NJ.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48, 65–91.
- Jensen, M.C., 1978. Some anomalous evidence regarding market efficiency. *Journal of Financial Economics* 6, 95–101.
- Kamara, A., Miller Jr., T.W., 1995. Daily and intradaily tests of European put-call parity. *Journal of Financial and Quantitative Analysis* 30, 519–539.
- Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49, 1541–1578.
- LeRoy, S.F., 2002. Infinite portfolio strategies. Unpublished Working Paper. University of California, Santa Barbara.
- Mitchell, M., Stafford, E., 2000. Managerial decisions and long-term stock price performance. *Journal of Business* 73, 287–329.
- Ross, S., 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13, 341–360.
- Shleifer, A., 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford University Press, Oxford, UK.
- Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. *Journal of Finance* 52, 35–55.