Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Financial Markets 15 (2012) 47–80
Journal of FINANCIAL MARKETS
www.elsevier.com/locate/finmar

An improved test for statistical arbitrage☆

Robert Jarrow[a], Melvyn Teo[b], Yiu Kuen Tse[c], Mitch Warachka[b,]*

[a]*Johnson Graduate School of Management, Cornell University, Singapore*
[b]*L.K.C. School of Business, Singapore Management University, Singapore*
[c]*School of Economics, Singapore Management University, Singapore*

Available online 16 August 2011

Abstract

We improve upon the power of the statistical arbitrage test in Hogan, Jarrow, Teo, and Warachka (2004). Our methodology also allows for the evaluation of return anomalies under weaker assumptions. We then compare strategies based on their convergence rates to arbitrage and identify strategies whose probability of a loss declines to zero most rapidly. These strategies are preferred by investors with finite horizons or limited capital. After controlling for market frictions and examining convergence rates to arbitrage, we find that momentum and value strategies offer the most desirable trading opportunities.

Jel classification: C12; G12; G14

Keywords: Bootstrap; Momentum strategy; Statistical arbitrage; Value strategy

1. Introduction

Traditional tests of market efficiency calculate trading profits relative to an assumed model for risk-adjusted returns, such as the Fama and French (1993) three-factor model. The statistical significance of the alpha from such a model is then evaluated using the intercept's t-statistic. Therefore, this test of market efficiency requires the empirical specification of risk factors, hence an assumed model of market equilibrium. According to Fama (1998), this caveat weakens our conclusions regarding market efficiency.

☆We thank Steve Hogan and Olivier Ledoit for their helpful comments as well as Dharma Jayanto for his excellent research assistance. We also acknowledge the valuable assistance of the Center for Academic Computing at SMU. All remaining errors are our own.

*Corresponding author.
E-mail address: mitchell@smu.edu.sg (M. Warachka).

1386-4181/$ - see front matter
doi:10.1016/j.finmar.2011.08.003

This paper develops a test of market efficiency that does not require an underlying equilibrium asset pricing model. Our methodology extends the statistical arbitrage framework of Hogan, Jarrow, Teo, and Warachka (2004) (henceforth HJTW) that axiomatically defines the conditions for a statistical arbitrage. To test for statistical arbitrage, a parametric model is assumed for incremental trading profits and the null hypothesis of no statistical arbitrage is formulated as a union of sub-hypotheses involving the model parameters. Thus, statistical arbitrage replaces the joint-hypothesis involving a model of market equilibrium with another joint-hypothesis involving a parametric specification for trading profits whose appropriateness can be assessed empirically.

Under the assumption of no statistical arbitrage, the parameters of the statistical model for trading profits must *simultaneously* satisfy several restrictions. HJTW test these restrictions using the Bonferroni inequality. While they reject the null of no statistical arbitrage for many trading strategies using their simple model of trading profits, they are unable to reject this null for their more general model of trading profit dynamics. This paper improves upon HJTW's test procedure for statistical arbitrage by introducing a Min-t statistic defined from the parametric restrictions that define the null hypothesis of no statistical arbitrage. This statistic provides a significant improvement in power over HJTW's implementation. Furthermore, we allow trading profit residuals to be non-normal and serially correlated.

We apply the Min-t test to four classes of trading strategies, namely, momentum (Jegadeesh and Titman, 1993, 2001), industry momentum (Grinblatt and Moskowitz, 1999), value (Lakonishok, Shleifer, and Vishny, 1994), and liquidity[1] (Brennan, Chordia, and Subrahmanyam, 1998). As we select trading strategies from the prior anomalies literature, the data-snooping critique may apply. Since the seminal work of Brock, Lakonishok, and LeBaron (1992), the issue of data-snooping when testing the profitability of trading rules has attracted considerable research interest. As emphasized by Sullivan, Timmermann, and White (1999), data-snooping is not limited to a particular researcher's efforts but can reflect the aggregate experience of many researchers over time. However, the anomalies literature typically focuses on the profitability of trading strategies. In contrast, we examine several parametric restrictions on trading profits. These additional restrictions (beyond positive profits) partially mitigate the data-snooping critique. Indeed, our study re-examines four classes of existing trading strategies to determine whether their profits constitute a statistical arbitrage.

Based on the results of the Min-t test, we find that over 50% of the strategies in each of the four classes generate a statistical arbitrage. In addition, we compare and contrast strategies based on their rates of convergence to arbitrage. Specifically, we identify dominant strategies whose loss probabilities decline to zero most rapidly. This feature of statistical arbitrage is important to investors with finite horizons or limited capital who are concerned about incurring intermediate losses. Such investors include mutual fund managers who typically face the risk of retrenchment after a few years, or even a few quarters, of poor performance (see, Brown, Harlow, and Starks, 1996; Khorana, 1996; Shleifer and Vishny, 1997; Chevalier and Ellison, 1999). Stein (2005) argues that open-ended fund managers "will stick primarily to short horizon strategies" instead of attacking

[1]According to Brennan, Chordia, and Subrahmanyam (1998), stock trading volume provides incremental explanatory power on the cross-section of stock returns after adjusting for momentum, size, and book-to-market effects.

long run mispricings. We show that such investors will prefer momentum, industry momentum, and value strategies over liquidity strategies as the last class of strategies converges relatively slowly to arbitrage.

To obtain further insights into the economic relevance of the strategies, we examine their sensitivity to the round-trip transaction costs. These costs range from 1.45% to 2.45%, which correspond to Chan and Lakonishok's (1997) estimates. Despite these conservative estimates, most of the statistical arbitrage strategies (except industry momentum) survive the adjustments for transaction costs. By our estimates, at least $0.64 billion may be profitably invested in nine of the 16 momentum strategies after transaction costs. Our findings on momentum dovetail with those of Korajczyk and Sadka (2004) but differ from those of Lesmond, Schill, and Zhou (2004).[2] We also investigate whether the statistical arbitrage profits are compensation for illiquidity by removing the bottom 50% of stocks based on past trading volume on a rolling basis and performing the test on the remaining sample of liquid stocks. Except for the industry momentum strategies, the strategies are again largely robust to this adjustment for illiquidity. Finally, we find little evidence to suggest that the activities of market participants have reduced the overall profitability of these strategies, although important differences exist across strategy classes. Specifically, the profitability of value and industry momentum strategies has increased over the sample period, while that of liquidity strategies has waned. Overall, after controlling for market frictions and examining convergence rates to arbitrage, we conclude that momentum and value strategies offer the most desirable trading opportunities.

Overall, we contribute to the return anomalies literature by extending the empirical application of the HJTW framework of statistical arbitrage in four important directions. First, our Min-t test significantly improves on HJTW's Bonferroni test. The Bonferroni test is well-known to be consistent when the null hypothesis involves an *intersection* of sub-hypotheses, and thus rejects an incorrect null hypothesis with probability one when the sample size approaches infinity. However, as the null hypothesis of no statistical arbitrage is defined by a *union* of hypotheses, the Bonferroni test is not guaranteed to be consistent and suffers a severe loss of power in empirical applications. In contrast, the Min-t test detects many more statistical arbitrage opportunities, even for their general model in which no statistical arbitrage was detected by HJTW.[3] Second, unlike HJTW, our test allows for serial correlation and non-normality in trading profit dynamics. These properties exert a significant influence on the convergence rates to arbitrage. Third, our test is based on a modified definition of statistical arbitrage that avoids penalizing positive profit deviations from their expected values. The critique of the Sharpe ratio by Bernardo and Ledoit (2000) motivates this modification as they argue that investors benefit from positive (unexpected) deviations in their strategy's profitability. Fourth, HJTW apply their statistical arbitrage test to only two classes of anomalies: momentum and value. We extend the set of strategies to include liquidity and industry momentum.

The remainder of the paper is as follows. Section 2 reviews the theoretical underpinnings of statistical arbitrage. The modified definition of statistical arbitrage and the Min-t testing methodology are proposed in Sections 3 and 4, respectively. Section 5 describes the data while Section 6 reports the empirical findings. Section 7 concludes.

[2]One caveat is that, unlike Lesmond, Schill, and Zhou (2004), we do not estimate time-varying transaction costs as that is beyond the scope of the paper.

[3]As the HJTW approach involves a joint-hypothesis of trading profits, it is important to have a general trading profit process with few restrictive assumptions.

2. Review of statistical arbitrage

A statistical arbitrage opportunity is a zero cost, self-financing trading strategy that has positive expected cumulative trading profits with a declining time-averaged variance and a probability of loss that converges to zero.[4] The strategies evaluated in our empirical application are known to have a positive cumulative trading profit. Let $\{\upsilon_i\}$, for $i=1, \ldots, n$, be a sequence of discounted portfolio values generated by a self-financing trading strategy. We denote $\upsilon(n) = \sum_{i=1}^{n} \Delta\upsilon_i$ as the trading strategy's cumulative discounted trading profit, with its incremental components represented by $\Delta\upsilon_i$. The following definition is found in HJTW.

Definition 1. A statistical arbitrage is a zero initial cost, self-financing trading strategy with cumulative discounted trading profits $\upsilon(n)$ such that:

1. $\upsilon(0)=0$,
2. $\lim_{n\to\infty} E^P[\upsilon(n)]>0$,
3. $\lim_{n\to\infty} P(\upsilon(n)<0)=0$ and
4. $\lim_{n\to\infty} \frac{Var^P[\upsilon(n)]}{n}=0$ if $P(\upsilon(n)<0)>0 \quad \forall n<\infty$.

To test for statistical arbitrage, we begin by assuming the following process for incremental trading profits[5]:

$$\Delta\upsilon_i = \mu i^{\theta} + \sigma i^{\lambda} z_i, \tag{1}$$

where z_i are i.i.d. $N(0,1)$ random variables (the assumptions of normality and independence are subsequently relaxed). The initial quantities $z_0=0$ and $\Delta\upsilon_0$ are both zero by definition. The parameters σ and λ determine the volatility of incremental trading profits while the parameters μ and θ determine their corresponding expectations. In addition, observe that the process for incremental trading profits is non-stationary when θ or λ is nonzero. More importantly, Eq. (1) does not impose assumptions on investor preferences or utility since both are irrelevant to the definition of statistical arbitrage, which converges to arbitrage as n approaches infinity.

Appendix A contains more information on the θ and λ parameters by detailing the conversion of cross-sectional returns into trading profits. Over time, a statistical arbitrage opportunity converges to a riskless asset whose incremental trading profits, and their volatility, may be reduced as a consequence.

Under the assumption that trading profit innovations are uncorrelated and normally distributed, we implement two tests of statistical arbitrage. Eq. (1) represents the unconstrained mean (UM) model, which allows for time-varying expected trading profits. We also consider a more restrictive constrained mean (CM) model that assumes constant expected trading profits by setting θ equal to zero. Consequently, the CM version of

[4] A trading strategy that rapidly decreases its exposure to the risky long/short portfolio is inappropriate since the persistence of the underlying anomaly is not addressed.

[5] A geometric Brownian motion (lognormal distribution) that prevents negative values is inappropriate for modeling cumulative or incremental trading profits. Instead, an arithmetic Brownian motion is suitable for the difference between two portfolios (long minus short) over a Δ time interval. For a profitable strategy, the functions $i^{\theta}$and i^{λ} alter this arithmetic process to account for the increasing investment in the risk-free asset over time. Further justification for this process is found in HJTW.

statistical arbitrage has incremental trading profits evolving as:

$$\Delta v_i = \mu + \sigma i^{\lambda} z_i. \tag{2}$$

According to HJTW, statistical arbitrage opportunities exist under the UM model when the following sub-hypotheses hold jointly:

1. H1: $\mu > 0$,
2. H2: $\lambda < 0$,
3. H3: $\theta > \max\{\lambda - (1/2), -1\}$.

The first sub-hypothesis tests for positive expected profits while the second sub-hypothesis tests for a declining time-averaged variance. This sub-hypothesis is expected to be satisfied in our empirical implementation involving existing anomalies that are known to produce positive trading profits. Nonetheless, the first axiom of a statistical arbitrage opportunity cannot be ignored. The third sub-hypothesis ensures that any potential decline in expected trading profits does not prevent convergence to arbitrage. This restriction involves the trend in expected profits and the trend in volatility. For emphasis, statistical arbitrage requires the volatility of *incremental* (not cumulative) trading profits to decline. This property is consistent with profitable trading strategies placing more accumulated profit in the risk-free asset over time, as emphasized in Appendix A. Under the CM model, the third sub-hypothesis is not required for statistical arbitrage; a statistical arbitrage opportunity exists when the first two sub-hypotheses hold jointly. This special case addresses the loss of statistical power that occurs when estimating θ is unnecessary. Implementing both the UM and CM models allows us to investigate our methodology's improvement over the Bonferroni approach, which cannot examine the UM model due to a loss of power. Nonetheless, it is important to emphasize that no model selection is conducted before testing for statistical arbitrage. Instead, testing for statistical arbitrage involves multiple hypotheses.

Given the manner in which portfolios used to test financial anomalies are typically constructed, autocorrelation may be manifest in their incremental trading profits.[6] To address this issue, we allow the innovations of Eq. (1) to follow an MA(1) process given by:

$$z_i = \varepsilon_i + \phi \varepsilon_{i-1}, \tag{3}$$

where ε_i are i.i.d. $N(0,1)$ random variables. We label the model with time-varying expected profits and serially correlated innovations as the unconstrained mean with correlation (UMC) model. Analogously, we label the model with constant expected profits and serially correlated innovations as the constrained mean with correlation (CMC) model. The UMC model is a combination of Eqs. (1) and (3) while the CMC model is a combination of Eqs. (2) and (3). As proved in HJTW, the presence of an MA(1) process neither alters the conditions for statistical arbitrage nor increases the number of sub-hypotheses. However, including the additional parameter ϕ may improve the statistical efficiency of the remaining parameter estimates and avoid inappropriate standard errors. The probability of a trading strategy generating a loss after n periods is as follows:

$$\Pr\{\text{Loss after } n \text{ periods}\} = \Phi\left(\frac{-\mu \sum_{i=1}^{n} i^{\theta}}{\sigma(1+\phi)\sqrt{\sum_{i=1}^{n} i^{2\lambda}}}\right), \tag{4}$$

[6] This property may arise from negative serial correlation in stocks or cross-autocorrelation among stocks.

where $\Phi(\cdot)$ is the cumulative standard normal distribution function. This probability converges to zero at a rate that is faster than exponential, as shown in HJTW. This loss probability finds the dominate strategy within each class of return anomaly. The dominate strategy is defined as the permutation that converges to arbitrage most rapidly

3. A modified definition of statistical arbitrage

This section proposes a modified definition of statistical arbitrage. This is motivated by the overly conservative nature of the original statistical arbitrage definition in HJTW, which penalizes positive trading profit deviations from expected value. The problem stems from the asymmetry between desirable positive deviations and undesirable negative deviations, which compromises the ability of variance to properly account for risk. This shortcoming motivates the use of semi-variance instead of variance in the modified fourth axiom of the following definition for statistical arbitrage:

Definition 2. A statistical arbitrage is a zero initial cost, self-financing trading strategy with cumulative discounted trading profits $\upsilon(n)$ such that:

1. $\upsilon(0)=0$,
2. $\lim_{n\to\infty} E^P[\upsilon(n)]>0$,
3. $\lim_{n\to\infty} P(\upsilon(n)<0)=0$ and
4. $\lim_{n\to\infty} Var[\Delta\upsilon(n)|\Delta\upsilon(n)<0]=0$

Observe that only the fourth axiom,

$$\lim_{n\to\infty} Var[\Delta\upsilon(n)|\Delta\upsilon(n)<0]=0, \tag{5}$$

is altered; the first three axioms are identical to those in Definition 1. Under Definition 2, investors are only concerned about the variance of a potential "drawdown" in wealth. Provided the incremental trading profits are non-negative, their variability is not penalized. To facilitate empirical tests of statistical arbitrage under Definition 2, we turn to the following proposition:

Proposition 1. *Under the modified fourth axiom in Eq.* (5)*, a trading strategy generates statistical arbitrage if incremental trading profits satisfy the following conditions:*
1. H1: $\mu>0$,
2. H2: $\lambda<0$ or $\theta>\lambda$,
3. H3: $\theta>\max\{\lambda-(1/2),-1\}$.

Appendix B provides the details verifying the $\theta>\lambda$ condition in addition to H2 in HJTW. Observe that our proposed modification only applies to the UM and UMC models, and may detect additional statistical arbitrage opportunities when $\theta>0$. Conversely, a negative point estimate for θ implies that H2 reverts to the original hypothesis that $\lambda<0$. Intuitively, positive θ estimates are consistent with right-skewness in the incremental trading profits. Fig. 1 illustrates the modified fourth axiom by showing the boundary between no statistical arbitrage and statistical arbitrage under both definitions. Note that the upper half of the first quadrant (above the 45° line) is classified as a statistical arbitrage opportunity under the modified, but not under the original definition.

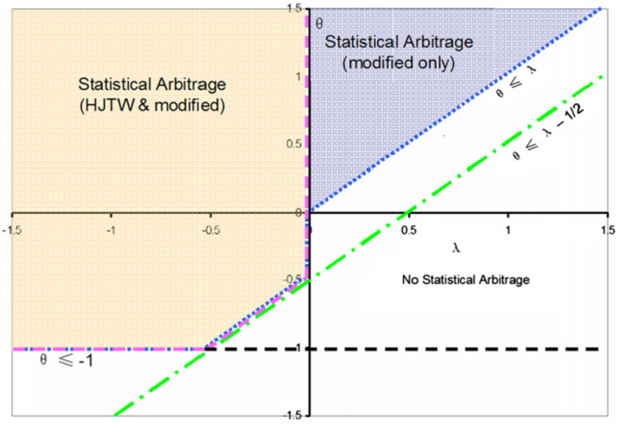

Fig. 1. Comparing the HJTW and our definition of statistical arbitrage. This figure illustrates the region where the null of no statistical arbitrage is accepted and the region where it is rejected. This is done for both the HJTW (Hogan, Jarrow, Teo, and Warachka, 2004) definition and our modified definition of statistical arbitrage. Unlike the HJTW version of statistical arbitrage, our modified version of statistical arbitrage does not penalize positive deviations in trading profits from the mean. The parameters θ and λ denote the rate of change of the expectation and volatility of the strategy's incremental profits, respectively.

As a final observation, the probability of a loss in Eq. (4) is unaltered by the modified fourth axiom. Furthermore, the economic consequences of a statistical arbitrage opportunity, in terms of its Sharpe ratio and contribution to expected utility, are preserved. Thus, under Definition 2, statistical arbitrage continues to contradict market efficiency.

4. Improved implementation of statistical arbitrage

In this section, we provide an improved methodology to test for statistical arbitrage. Although statistical tests may be conducted with either statistical arbitrage or no statistical arbitrage as the null, the accepted paradigm has market efficiency as the null.

The hypothesis of market inefficiency, namely the existence of statistical arbitrage, consists of joint restrictions on the parameters underlying the evolution of trading profits. For the UM model, the following restrictions have to be satisfied simultaneously for a statistical arbitrage opportunity to exist[7]:

1. R_1: $\mu > 0$,
2. R_2: $-\lambda > 0$ or $\theta - \lambda > 0$,
3. R_3 : $\theta - \lambda + (1/2) > 0$ and
4. R_4: $\theta + 1 > 0$.

Thus, statistical arbitrage is defined by an *intersection* of sub-hypotheses. Conversely, the no statistical arbitrage null hypothesis involves a *union* of sub-hypotheses (as a consequence of DeMorgan's Laws). In particular, the no statistical arbitrage null hypothesis

[7]A slight change of notation is adopted to separate H3 into two restrictions to facilitate the exposition of the proposed test.

is written as:

1. $R_1^c:\ \mu \leq 0$ or
2. $R_2^c:\ -\lambda \leq 0$ and $\theta - \lambda \leq 0$, or
3. $R_3^c:\ \theta - \lambda + (1/2) \leq 0$ or
4. $R_4^c:\ \theta + 1 \leq 0$.

Therefore, the null cannot be rejected provided a single sub-hypothesis R_t^c is satisfied. Statistically, the no statistical arbitrage null hypothesis presents a challenge as the Bonferroni procedure applies to an intersection, not union, of sub-hypotheses.[8] Appendix C examines the Bonferroni approach in HJTW and highlights its lack of power as the number of sub-hypotheses increase. For emphasis, each sub-hypothesis arises from (at least) one of the statistical arbitrage axioms. Although certain axioms may appear to be redundant or obvious (such as the constraint on μ), our statistical procedure must ensure that all four axioms are satisfied.

Given the limitations of the Bonferroni approach in testing for statistical arbitrage, this section proposes a new methodology centered on the Min-*t* statistic. We adopt Eq. (1) as the maintained parametric specification of the trading-profit process for testing the four axioms of statistical arbitrage. We first assume that the trading profit innovations are normally distributed and serially uncorrelated. The model is estimated using maximum likelihood estimation (MLE). Robust standard errors are also computed, from which we obtain the t ratios required for computing the Min-*t* statistic. The critical values of the Min-*t* statistic are estimated using a Monte Carlo procedure to facilitate inferences regarding the null hypothesis of no statistical arbitrage. We then allow the trading profit innovations to be non-normal, as well as serially dependent. In this case, there are nuisance parameters in the model and Monte Carlo estimates of the critical values are not obtainable. To overcome this problem, we estimate the *p*-values for the Min-*t* statistics using a bootstrap procedure.

Although other specifications (e.g., stochastic volatility) besides Eq. (1) are possible, additional enhancements would complicate the link with the statistical arbitrage axioms. Overall, there is a tradeoff between the statistical accuracy of any description for trading profits and the ease at which its specification represents these four axioms. This tradeoff highlights the dependence of the statistical arbitrage methodology on a parametric process for trading profits.

4.1. Monte Carlo procedure for uncorrelated normal errors

We estimate the parameters of the CM, UM, CMC, and UMC models using maximum likelihood estimation (MLE).[9] Our inference concerning the null hypothesis of no

[8]See Gourieroux and Monfort (1995), Chapter 19, for an exposition of testing joint hypotheses using the Bonferroni procedure.

[9]Phillips and Xu (2006) examine the asymptotic properties of the MLE for an autoregressive process with heteroscedastic errors $g(\cdot)z_i$, where $g(\cdot)$ is a deterministic or stochastic function with $g(\cdot)>0$. They prove the asymptotic normality of the MLE for the autoregressive coefficients and provide a consistent estimator for the integrated variance of the residuals. Our model has an error process that is nested within the Phillips and Xu (2006) model, with a deterministic mean. While a consistent estimator of the integrated volatility function $\int g(\cdot)^2$ can be obtained using the Phillips and Xu (2006, p. 296) method, σ cannot be estimated consistently if λ is negative. Furthermore, Xu and Phillips (2008) propose an improved methodology to estimate the autoregressive coefficients.

statistical arbitrage is based on the proposed Min-t statistic. In this subsection, we discuss the case when the errors of the models are independently and normally distributed, so that the *finite-sample* critical values of the test statistic can be estimated by Monte Carlo simulation. In the next subsection, we describe the bootstrap methodology for the estimation of the p-values when the errors may be serially correlated and non-normally distributed.

When each R_i is considered separately, the t-statistics $t(\hat{\mu})$, $t(-\hat{\lambda})$, $t(\hat{\theta}-\hat{\lambda})$, $t(\hat{\theta}-\hat{\lambda}+0.5)$, and $t(\hat{\theta}+1)$ may be used (the hats denote the MLE of the parameters). Since all four restrictions in Proposition 1 must be simultaneously satisfied to reject the null hypothesis of no statistical arbitrage, the minimum of their associated t-statistics serves as the rejection criterion. Thus, we consider the Min-t statistic defined as[10]:

$$\text{Min-}t = \text{Min}\{t(\hat{\mu}), t(\hat{\theta}-\hat{\lambda}+0.5), t(\hat{\theta}+1), \text{Max}[t(-\hat{\lambda}), t(\hat{\theta}-\hat{\lambda})]\}. \tag{6}$$

Intuitively, the Min-t statistic evaluates the ''weakest'' element in the union by focusing on the sub-hypothesis that is ''closest'' to being accepted. Thus, the null of no statistical arbitrage is rejected if Min-$t > t_c$, where the critical value t_c depends on the test's significance level α. This approach is equivalent to evaluating a Max-t statistic after reversing each of the inequalities that define the null hypothesis.

Under the statistical arbitrage axioms of Definition 1, the Min-t statistic in Eq. (6) is simplified to:

$$\text{Min}\{t(\hat{\mu}), t(\hat{\theta}-\hat{\lambda}+0.5), t(\hat{\theta}+1), t(-\hat{\lambda})\}. \tag{7}$$

Therefore, Eqs. (6) and (7) facilitate a direct comparison between the fourth axioms of Definitions 1 and 2.

For the CM models with $\theta = 0$, Eq. (6) becomes:

$$\text{Min-}t = \text{Min}\{t(\hat{\mu}), t(-\hat{\lambda})\}, \tag{8}$$

regardless of the fourth axiom examined. Indeed, as alluded to in the previous section, Definitions 1 and 2 have identical implementations in the CM and CMC models.

As the null of no statistical arbitrage is a composite hypothesis rather than a simple hypothesis, the probability of rejecting the null varies across different parameter values within the null. However, the probability of rejecting the null cannot exceedα. In other words, we require:

$$\Pr\{\text{Min-}t > t_c \,|\, \mu, \lambda, \theta, \sigma\} \leq \alpha \tag{9}$$

for all $(\mu,\lambda,\theta,\sigma)$ combinations satisfying the null. Thus, two issues have to be addressed. First, the derivation of the distribution of Min-t requires solving for the distribution of an order statistic of a multivariate distribution of correlated random variables. Thus, the theoretical distribution of Min-t is intractable. We propose to overcome this difficulty using Monte Carlo simulation. Second, to achieve a size-α test as in Eq. (9), we need to maximize t_c over the null's parameter space.

We first consider the CM model whose two statistical arbitrage sub-hypotheses are R_1 and R_2. Obviously, t_c is maximized when $(\mu,\lambda) = (0,0)$. Furthermore, as the t-statistics are scale free, we are able to select an arbitrary value of σ when estimating t_c. We assume $\sigma = 0.01$, which approximates its sample MLE estimate in our later empirical study.

[10]The original sub-hypotheses for Definition 1 may be tested using a special case of Eq. (6).

To estimate t_c, residuals z_i are obtained from a normal random number generator to form the incremental trading profits Δv_i in Eq. (2) based on assumed model parameters $(\mu,\lambda,\sigma)=(0,0,0.01)$. The estimated parameters, their individual t-statistics, and the corresponding Min-t statistic are then computed. This procedure is repeated 5,000 times, from which t_c is estimated as the $100(1-\alpha)$ percentile of the Min-t statistics.

Note that the distribution of Min-t is a function of the sample size n. As the series of trading profits used in our empirical study vary from 288 to 414 observations, sample sizes of 300 and 400 are examined. However, the results for both values of n are similar. Overall, the critical values of 0.4754, 0.7484, and 1.2694 at the 10%, 5%, and 1% significance levels are utilized in subsequent tests of the CM model. These critical values correspond to the largest estimates in the Monte Carlo simulations.

For the UM model, there are five inequality restrictions involving three parameters and not all the restrictions are necessarily binding. Thus, a model within the null family and on the boundary of all the inequality restrictions is not available. Nonetheless, as the t-statistics that comprise the Min-t statistic are monotonic in the underlying restrictions, it is appropriate to focus on their boundaries. Consequently, μ is set to -10^{-6} while the λ and θ parameters are varied along the boundary of the no statistical arbitrage/statistical arbitrage region, as depicted in Fig. 1.[11]

To control the probability of the Type I error at the stated nominal level, the maximum simulated critical values across different parameters are utilized in subsequent UM tests. These are 0.4034, 0.6004, and 0.9074 at the 10%, 5%, and 1% significance levels, respectively.[12]

4.2. Bootstrap procedure for correlated non-normal errors

The previous section assumes the innovations in incremental profits are normally distributed and serially uncorrelated. However, both assumptions have been shown to be dubious in empirical finance (e.g., Lo and MacKinlay, 1988; Affleck-Graves and McDonald, 1989). We now relax these assumptions by allowing the trading profit innovations to be non-normal and serially correlated. In particular, we assume that the innovations follow an MA(1) process, as defined in Eq. (3). Thus, there is an unspecified nuisance parameter ϕ in the model when testing the null hypothesis of no statistical arbitrage. As the influence of ϕ on the individual components of the Min-t statistic is unknown, searching for the maximum critical values for the test statistic using the Monte Carlo method is intractable. To overcome this difficulty, we employ a bootstrap procedure to estimate the p-values.

Brock, Lakonishok, and LeBaron (1992) introduce the bootstrap technique to the empirical finance literature to study the profitability of technical trading rules. Since then, this procedure has been adopted by many authors including Bessembinder and Chan (1998) and Sullivan, Timmermann, and White (1999). Ruiz and Pascual (2002) provide an excellent survey of the bootstrap method in empirical finance.

[11]Note that the influence of θ disappears when $\mu=0$. We also vary the values of μ from -0.01 to -0.0001 and obtain similar results.

[12]Since Monte Carlo simulation is employed to estimate the critical values of the Min-t statistic in finite samples, the nonstationarity of the UM model when $\theta\neq 0$ bears no consequence on our test procedure.

We adopt the parametric bootstrap procedure (Berkowitz and Kilian, 2000) to allow for time dependence in the residuals. The steps we employ for the UMC model are:

1. Estimate the parameters of the UM model with MA(1) errors using QMLE and calculate the residuals $\hat{z}_i$ using the following equation:

$$\hat{z}_i = \frac{\Delta \upsilon_i - \hat{\mu} i^{\hat{\theta}}}{\hat{\sigma} i^{\hat{\lambda}}} \tag{10}$$

   In addition, the Min-t statistic in Eq. (6) is calculated.
2. Sample with replacement blocks of $\hat{z}_i$ of random length, where the length of each block has a geometric distribution, to form a sample of residuals $\{z_1^*, \ldots, z_n^*\}$.
3. Generate a bootstrap sample of trading profits $\Delta \upsilon_i^*$ from $\{z_1^*, \ldots, z_n^*\}$ with the parameter values $(\mu, \lambda, \theta, \sigma) = (-10^{-6}, -0.5, -1, 0.01)$ so that:[13]

$$\Delta \upsilon_i^* = \mu i^{\theta} + \sigma i^{\lambda} z_i^*.$$

4. Calculate the MLE for the data $\Delta \upsilon_i^*$, $i = 1, \ldots, n$, and hence the Min-t statistic, denoted Min-$t^{\#}$.
5. Repeat Steps 2–4 a total of 1,000 times. The estimated p-value of the Min-t statistic is given by the empirical percentage of the bootstrapped Min-$t^{\#}$ values that are larger than Min-t calculated in Step 1.

A similar bootstrap procedure for the CMC model is implemented with $(\mu, \lambda) = (0, 0)$. Our procedure adheres to the guidelines in Hall and Wilson (1991) and Horowitz (2001). For the CM and UM models with uncorrelated errors, the bootstrap estimates of the p-values are also computed using the parametric bootstrap procedure. Thus, they provide robust estimates of the p-values even when the independent error assumption is mispecified.

Depending on the complexity of the model assumptions, the bootstrap procedure can be adjusted to improve the robustness of the estimates. Before reporting our results, we now discuss some possible alternative bootstrap methods that may be adopted. First, we use a parametric bootstrap method assuming the innovations follow an MA(1) process. Karolyi and Kho (2004) used a parametric bootstrap method assuming AR(1) errors. If no specific time series process is assumed for the innovation process, we may use nonparametric approaches, such as the stationary bootstrap method of Politis and Romano (1994) or the block bootstrap method of Künsch (1989). However, as pointed out by Karolyi and Kho (2004), the block bootstrap procedure has problems with unbalanced panel data due to differences in exit and entry. Second, re-sampling can be performed across stocks to account for their cross correlation. This is not performed in our study due to our problem's increased dimension. Third, to address data-snooping, the reality-check bootstrap method may be used, as in Sullivan, Timmermann, and White (1999). However, this method requires the computation of the profit processes over the universe of trading strategies under consideration. Thus, as we investigate four classes of trading strategies involving fundamental (accounting) variables, the dimension of the problem renders the use of reality-check bootstrap computationally very intensive in comparison to the Sullivan, Timmermann, and White (1999) study that evaluates technical trading strategies.

[13]Under the null of no statistical arbitrage with normally and serially uncorrelated errors, these parameter values provide the largest critical value t_c in the Monte Carlo simulation.

Finally, different bootstrap methods may have different rates of convergence to the unknown true values being estimated, if they converge. A rigorous asymptotic convergence theory for models with serially correlated errors is beyond the scope of this paper. However, Xu (2008) proposes the use of a recursive wild bootstrap method for an autoregressive process with non-stationary variance and proves its asymptotic validity.

5. Data and terminology

Our sample period starts in January 1965 and ends in December 2000.[14] Monthly equity returns data are derived from the Center for Research in Security Prices at the University of Chicago (CRSP). 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 U.S., and Americus Trust Components.

The stock characteristics underlying the trading strategies include book-to-market equity, cash flow-to-price ratio, earnings-to-price ratio, annual sales growth, and monthly trading volume. To calculate book-to-market equity, book value per share is taken from the CRSP/Compustat price, dividend, and earnings database. We treat all negative book values as missing. We take the sum of Compustat data item 123 [income before extraordinary items (SCF)] and data item 125 [depreciation and amortization (SCF)] as cash flow. Only data item 123 is used to calculate the 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)]. Share volume is the number of shares traded divided by the number of shares outstanding. All price and number of outstanding common shares information employed in the calculation of the ratios are computed at the end of the year.

To ensure that the accounting variables are known beforehand and to accommodate variation in fiscal year ends among firms, sorting on stock characteristics is performed in July of year t using the 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 the book equity is in the fiscal year ending in year $t-1$ and the 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 $t+1$, the stocks are sorted into deciles based on their cash flow-to-price, where cash flow is in the fiscal year ending in year $t-1$ and the 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 evaluation period.

The momentum strategies we implement follow Jegadeesh and Titman (1993). These strategies buy the top return decile and short the bottom return decile based on formation and holding period combinations of 3, 6, 9, and 12 months. The individual value strategies follow Lakonishok, Shleifer, and Vishny (1994) and buy 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, along with past sales growth over the past three years. These portfolios are then held for one, three, and five years. The individual liquidity strategies are based on volume (share volume over number of shares outstanding) and buy the bottom

[14]We choose the same sample period as HJTW to facilitate an easier comparison with their results.

trading volume decile and short the top trading volume decile of stocks in the spirit of Brennan, Chordia, and Subrahmanyam (1998). The industry momentum strategies follow Grinblatt and Moskowitz (1999). Stocks are first classified into 20 industries based on SIC Codes.[15] The industry momentum strategy buys the top three return industries and shorts the bottom three return industries as in Grinblatt and Moskowitz (1999). Like the momentum strategies, the liquidity and industry momentum strategies are based on formation and holding period combinations of 3, 6, 9, and 12 months. For all strategies, once the long and short portfolio returns are generated, a self-financing condition is enforced by investing (borrowing) trading profits (losses) at the risk-free rate. Risk-free rate data are obtained from Kenneth French's website.

Given the possible permutations of formation and holding periods, we investigate 16 momentum, 12 volume, 16 liquidity, and 16 industry momentum strategies. We adopt the notational convention of JTx_y for the momentum strategy, with a formation period of x months and a holding period of y months. The book-to-market, cash flow-to-price, earnings-to-price, and sales growth based value portfolios with a holding period of y years are denoted BMy, CPy, EPy, and SALEy respectively. The formation period for all the sales growth strategies is three years while that for the other value strategies is one year. The liquidity and industry momentum portfolios with a formation period of x months and a holding period of y months are abbreviated VOLx_y and INDx_y, respectively.

6. Empirical results

In this section, we apply the statistical arbitrage methodology developed in this paper to four broad classes of anomalies: momentum, value, liquidity, and industry momentum. In our empirical analysis, we implement the four trading profit models introduced in Section 2: the constrained mean model (CM), the unconstrained mean model (UM), the constrained mean model with correlation (CMC), and the unconstrained mean model with correlation (UMC). The UM and UMC models allow for time variation in expected trading profits while the CM and CMC models assume this expectation is constant. In addition, unlike the CM and UM models, the CMC and UMC models permit autocorrelation and non-normality in trading profits. The Schwartz information criterion and the Akaike information criterion determine the preferred trading profit model and streamline the presentation of our results.

For emphasis, our objective is not to find the parametric specification for each trading strategy that produces normally distributed error terms. Instead, obtaining reliable p-values for our joint-hypothesis is important. When estimating the critical values for the CM model, we operate on the boundary of the alternative hypothesis. This approach maximizes the critical value across different alternatives to ensure a size-α test. Likewise, in the bootstrap simulations, we also operate on the boundary.

In addition to testing anomaly-based trading strategies for statistical arbitrage, we leverage on our methodology's ability to explore their economic relevance. Concretely, we compute the convergence rates of the loss probabilities to zero and compare them across strategies. These probabilities are particularly relevant for investors like mutual fund and

[15] The 20 industries are mining, food, apparel, paper, chemical, petroleum, construction, primary metals, fabricated metals, machinery, electrical equipment, transport equipment, manufacturing, railroads, other transportation, utilities, department stores, retail, financial, and others. We refer the interested reader to Grinblatt and Moskowitz (1999) for further details.

hedge fund managers who have short (e.g., quarterly) evaluation horizons. We also test whether the statistical arbitrage profits are sensitive to transaction costs and market frictions, to removing the least liquid stocks, and to the sample period.

6.1. Basic empirical results

First, we make the assumption that the trading profit increments are normally distributed and serially uncorrelated, and test the CM and UM models. Table 1 displays the results from the CM and UM tests of statistical arbitrage. The statistical significance of the Min-t statistics (relative to critical values obtained from Monte Carlo simulations described in Section 4.1) indicates that for the CM model, 41 of the 60 strategies generate statistical arbitrage opportunities at the 5% level of significance. For the more general UM model, 30 of the 60 strategies generate statistical arbitrage opportunities at this significance level. Hence, at least half of the strategies under each trading profit model generate statistical arbitrage.

The results in Table 1 also indicate that the various anomaly classes differ markedly in their sensitivity to the trading profit model. Sales growth and volume strategies consistently test for statistical arbitrage under both models. In contrast, more industry momentum-based strategies test for statistical arbitrage under the UM model than under the CM model. The reverse holds for the momentum-based strategies. Of the 16 momentum strategies, 15 test for statistical arbitrage at the 5% level under the CM model while only one strategy tests for statistical arbitrage at the 5% level under the UM model. This disparity motives our later use of model specification tests.

We note that the statistical arbitrage test results for the value strategies illustrate the greater statistical power of our Min-t test approach vis à vis the HJTW Bonferroni test approach. Unlike the HJTW Bonferroni test, which cannot detect any UM statistical arbitrage opportunities amongst the value strategies, the Min-t test detects six UM statistical arbitrage opportunities within the same set of strategies. Indeed, without our procedure's enhanced statistical power, more complex trading profit formulations (e.g., the CMC and UMC models) cannot be reliably examined.

As discussed in Section 4.1, the critical Min-t values in Table 1 are estimated from a large scale Monte Carlo experiment, since the errors are normally distributed and uncorrelated in the CM and UM models. To check these assumptions, we test the residuals for normality and serial correlation using the Jacque-Bera test and the Q-statistic. The test statistics reveal that departures from normality and serial independence are detected at the 5% level of significance for many strategies. Hence, to test the robustness of the results with respect to the assumption of normality and to cross-validate the bootstrap methodology developed in Section 4.2 for the CMC and UMC models, we apply the methodology to the CM and UM models by constraining the autocorrelation coefficient ϕ to zero. With minor exceptions, the resulting bootstrapped p-value estimates for the CM and UM models reported in Table 1 agree with those from the Monte Carlo procedure (which assumes $\phi=0$ as well as normality). This reassuring result demonstrates convergence of the bootstrap procedure, and the robustness of the Min-t statistic with respect to the assumption of normality.

Next, we relax the assumption that the trading profit innovations are normally distributed and serially uncorrelated, and examine the CMC and UMC models. The statistical arbitrage results for the CMC model are reported in Table 2. In the interest of

Table 1
Tests of statistical arbitrage with CM (constrained mean) and UM (unconstrained mean) models.

Min-*t* statistics and bootstrapped *p*-values from tests of statistical arbitrage are presented for all four classes of trading strategies and for the UM and CM models. The sample period is from Jan 1965–Dec 2000. The UM model allows for time-varying expected trading profits. The CM model constrains expected trading profits to a constant. The JT portfolios are Jegadeesh and Titman (1993) stock momentum portfolios. The BM, CP, EP, and SALE portfolios are Lakonishok, Shleifer, and Vishny (1994) value portfolios based on book-to-market equity, cash flow-to-price, earnings-to-price, and sales growth, respectively. The VOL portfolios are liquidity-based portfolios. The IND portfolios are Grinblatt and Moskowitz (1999) industry momentum portfolios. JTx_y denotes a stock momentum portfolio with a formation period of x months and an evaluation period of y months. The VOLx_y and INDx_y portfolios are defined analogously. BMy, CPy, EPy, and SALEy denote value portfolios with an evaluation period of y years. The Min-t statistics are defined in Eqs. (6) and (8) for the UM and CM models respectively. Asterisks in parentheses denote statistical significance associated with bootstrapped *p*-values while those without parentheses denote statistical significance generated from Monte Carlo simulation. *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Portfolio	*Sample size*	CM model		UM model	
		Min-*t*	bstrap *p*-value	Min-*t*	bstrap *p*-value
Panel A: Momentum strategies					
JT3_3	*398*	−1.128	0.608	−0.439	0.468
JT3_6	*398*	2.294***	0.001 (***)	0.099	0.261
JT3_9	*398*	3.803***	0.000 (***)	0.119	0.231
JT3_12	*398*	2.508***	0.000 (***)	0.348	0.087 (*)
JT6_3	*398*	1.760***	0.005 (***)	0.796**	0.021 (**)
JT6_6	*398*	4.162***	0.000 (***)	0.119	0.245
JT6_9	*398*	3.049***	0.000 (***)	0.367	0.095 (*)
JT6_12	*398*	1.928***	0.000 (***)	0.375	0.078 (*)
JT9_3	*398*	3.153***	0.000 (***)	0.025	0.309
JT9_6	*398*	3.207***	0.000 (***)	0.371	0.091 (*)
JT9_9	*398*	2.239***	0.000 (***)	0.430*	0.061 (*)
JT9_12	*398*	1.491***	0.000 (***)	0.445*	0.042 (**)
JT12_3	*398*	2.924***	0.000 (***)	0.323	0.098 (*)
JT12_6	*398*	2.344***	0.000 (***)	0.405*	0.062 (*)
JT12_9	*398*	1.700***	0.000 (***)	0.485*	0.029 (**)
JT12_12	*398*	1.167**	0.006 (***)	0.404*	0.021 (**)
Panel B: Value strategies					
BM1	*414*	0.110	0.158	1.005***	0.001 (***)
BM3	*372*	0.233	0.169	0.993***	0.009 (***)
BM5	*324*	1.293***	0.025 (**)	1.157***	0.010 (***)
CP1	*414*	−0.266	0.466	−0.637	0.629
CP3	*372*	0.827**	0.075 (*)	0.451*	0.095 (*)
CP5	*324*	1.016**	0.020 (**)	0.371	0.128
EP1	*414*	−5.320	0.995	−0.662	0.866
EP3	*372*	−1.413	0.701	−0.140	0.439
EP5	*324*	−0.191	0.285	−1.371	0.761
SALE1	*378*	1.336***	0.007 (***)	1.198***	0.004 (***)
SALE3	*336*	1.530***	0.010 (***)	1.362***	0.003 (***)
SALE5	*288*	2.627***	0.001 (***)	1.073***	0.003 (***)

Table 1 (*continued*)

Portfolio	*Sample size*	CM model		UM model	
		Min-t	bstrap p-value	Min-t	bstrap p-value
Panel C: Volume strategies					
VOL3_3	398	0.897**	0.082 (*)	0.896**	0.019 (**)
VOL3_6	398	0.881**	0.070 (*)	0.882**	0.025 (**)
VOL3_9	398	0.973**	0.074 (*)	0.977***	0.019 (**)
VOL3_12	398	1.121**	0.050 (**)	1.126***	0.003 (***)
VOL6_3	398	0.974**	0.050 (**)	0.980***	0.023 (**)
VOL6_6	398	1.027**	0.060 (*)	1.034***	0.005 (***)
VOL6_9	398	1.109**	0.050 (**)	1.120***	0.003 (***)
VOL6_12	398	1.245**	0.041 (**)	1.234***	0.004 (***)
VOL9_3	398	1.035**	0.050 (**)	0.414*	0.047 (**)
VOL9_6	398	1.098**	0.041 (**)	1.094***	0.003 (***)
VOL9_9	398	1.219**	0.035 (**)	0.870**	0.025 (**)
VOL9_12	398	1.347***	0.038 (**)	1.244***	0.006 (***)
VOL12_3	398	1.166**	0.039 (**)	1.174***	0.005 (***)
VOL12_6	398	1.214**	0.034 (**)	0.960***	0.010 (***)
VOL12_9	398	1.331***	0.029 (**)	1.141***	0.003 (***)
VOL12_12	398	1.455***	0.024 (**)	1.362***	0.000 (***)
Panel D: Industry momentum strategies					
IND3_3	398	0.791**	0.042 (**)	0.808**	0.018 (**)
IND3_6	398	0.829**	0.058 (*)	0.836**	0.042 (**)
IND3_9	398	1.174**	0.009 (***)	0.934**	0.021 (**)
IND3_12	398	−0.065	0.221	0.852**	0.027 (**)
IND6_3	398	0.350	0.126	0.353	0.135
IND6_6	398	0.366	0.143	0.715**	0.047 (**)
IND6_9	398	−0.899	0.496	0.692**	0.040 (**)
IND6_12	398	−0.823	0.459	0.426*	0.093 (*)
IND9_3	398	1.870***	0.003 (***)	0.626**	0.054 (*)
IND9_6	398	−0.321	0.315	0.556**	0.065 (*)
IND9_9	398	−0.793	0.487	0.392	0.103
IND9_12	398	−0.561	0.367	0.269	0.164
IND12_3	398	−0.596	0.441	0.440*	0.100 (*)
IND12_6	398	−0.728	0.451	0.055	0.227
IND12_9	398	−0.518	0.378	−0.174	0.262
IND12_12	398	0.152	0.192	−0.091	0.281

brevity, we do not report the results for the UMC model[16] but note that the λ, σ and ϕ estimates are very similar across both models. A comparison of the statistical arbitrage results in Tables 1 and 2 reveals a strong congruence between the results for the CM model and those for the CMC model. Strategies that test positively for statistical arbitrage under the CM model usually do so for the CMC model as well. A similar relationship exists

[16]These results are available from the authors upon request.

Table 2
Tests of statistical arbitrage with CMC (constrained mean with correlation) model.

Results from tests of statistical arbitrage are presented for all four classes of trading strategies and for the CMC model. The sample period is from Jan 1965–Dec 2000. The CMC model features correlated trading profit innovations that follow an MA(1) process and constant expected trading profits as in Eqs. (2) and (3). The JT portfolios are Jegadeesh and Titman (1993) stock momentum portfolios. The BM, CP, EP, and SALE portfolios are Lakonishok, Shleifer, and Vishny (1994) value portfolios based on book-to-market equity, cash flow-to-price, earnings-to-price, and sales growth, respectively. The VOL portfolios are liquidity-based portfolios. The IND portfolios are Grinblatt and Moskowitz (1999) industry momentum portfolios. JTx_y denotes a stock momentum portfolio with a formation period of x months and an evaluation period of y months. The VOLx_y and INDx_y portfolios are defined analogously. BMy, CPy, EPy, and SALEy denote value portfolios with an evaluation period of y years. The Min-t statistic is defined in Eq. (8). *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Portfolio	Parameters (t-statistics)				Min-t	p-value
	Mean profit μ	Std dev σ	Growth rate of std dev λ	Auto-correlation ϕ		
Panel A: Momentum strategies						
JT3_3	−0.002 (−1.11)	0.086 (3.51)	−0.183 (−3.59)	0.022 (0.41)	−1.107	0.635
JT3_6	0.003 (2.37)	0.085 (3.50)	−0.208 (−3.96)	−0.031 (−0.50)	2.374	0.000 (***)
JT3_9	0.005 (4.16)	0.070 (3.65)	−0.191 (−3.83)	−0.043 (−0.64)	3.831	0.000 (***)
JT3_12	0.006 (5.32)	0.045 (3.84)	−0.122 (−2.53)	−0.066 (−0.81)	2.534	0.000 (***)
JT6_3	0.004 (1.81)	0.125 (3.49)	−0.220 (−4.21)	−0.026 (−0.45)	1.811	0.006 (***)
JT6_6	0.008 (4.37)	0.108 (3.66)	−0.209 (−4.19)	−0.043 (−0.68)	4.190	0.000 (***)
JT6_9	0.009 (5.59)	0.073 (3.86)	−0.147 (−3.07)	−0.053 (−0.71)	3.073	0.000 (***)
JT6_12	0.008 (5.04)	0.055 (3.56)	−0.101 (−1.96)	−0.069 (−0.80)	1.961	0.000 (***)
JT9_3	0.007 (3.30)	0.125 (3.63)	−0.203 (−4.11)	−0.040 (−0.63)	3.298	0.000 (***)
JT9_6	0.010 (5.20)	0.090 (3.87)	−0.153 (−3.22)	−0.053 (−0.71)	3.221	0.000 (***)
JT9_9	0.009 (4.95)	0.071 (3.59)	−0.116 (−2.27)	−0.067 (−0.78)	2.270	0.000 (***)
JT9_12	0.007 (4.09)	0.057 (3.35)	−0.084 (−1.54)	−0.074 (−0.78)	1.540	0.002 (***)
JT12_3	0.009 (4.12)	0.097 (3.71)	−0.144 (−2.93)	−0.060 (−0.74)	2.928	0.000 (***)
JT12_6	0.009 (4.35)	0.084 (3.49)	−0.124 (−2.37)	−0.069 (−0.79)	2.373	0.000 (***)
JT12_9	0.008 (3.82)	0.069 (3.37)	−0.095 (−1.75)	−0.071 (−0.76)	1.750	0.001 (***)
JT12_12	0.006 (2.95)	0.058 (3.14)	−0.070 (−1.22)	−0.074 (−0.77)	1.221	0.004 (***)
Panel B: Value strategies						
BM1	0.015 (5.51)	0.052 (2.78)	−0.004 (−0.06)	0.094 (1.41)	0.055	0.181
BM3	0.012 (5.38)	0.040 (2.70)	−0.017 (-0.23)	0.139 (2.38)	0.234	0.143
BM5	0.011 (5.36)	0.043 (3.55)	−0.069 (−1.23)	0.194 (3.06)	1.227	0.023 (**)
CP1	−0.001 (−0.25)	0.209 (3.74)	−0.200 (−3.40)	0.189 (2.33)	−0.251	0.396
CP3	0.002 (0.77)	0.111 (3.55)	−0.168 (−2.40)	0.077 (1.52)	0.768	0.100 (*)
CP5	0.002 (0.96)	0.125 (3.13)	−0.233 (−2.73)	0.080 (1.33)	0.956	0.068 (*)
EP1	−0.000 (−0.06)	0.013 (4.23)	0.291 (5.00)	0.205 (3.03)	−4.998	0.989
EP3	−0.001 (−0.19)	0.024 (2.76)	0.116 (1.36)	0.118 (2.09)	−1.359	0.673
EP5	−0.001 (−0.19)	0.044 (3.17)	−0.047 (−0.63)	0.173 (2.60)	−0.155	0.297
SALE1	0.009 (5.23)	0.050 (2.57)	−0.092 (−1.30)	0.066 (0.92)	1.301	0.009 (***)
SALE3	0.006 (3.94)	0.046 (2.44)	−0.127 (−1.60)	0.103 (1.47)	1.598	0.000 (***)
SALE5	0.005 (3.20)	0.049 (3.43)	−0.169 (−2.85)	0.148 (2.36)	2.846	0.000 (***)
Panel C: Volume strategies						
VOL3_3	0.005 (1.68)	0.065 (5.07)	−0.040 (−1.03)	0.111 (2.00)	1.025	0.050 (**)
VOL3_6	0.006 (2.00)	0.063 (5.12)	−0.039 (−1.01)	0.115 (2.09)	1.010	0.048 (**)

Table 2 (*continued*)

Portfolio	Parameters (t-statistics)				Min-t	p-value
	Mean profit μ	Std dev σ	Growth rate of std dev λ	Auto-correlation ϕ		
VOL3_9	0.006 (2.14)	0.063 (4.98)	−0.043 (−1.09)	0.123 (2.27)	1.094	0.039 (**)
VOL3_12	0.006 (2.23)	0.064 (4.87)	−0.049 (−1.23)	0.130 (2.41)	1.230	0.032 (**)
VOL6_3	0.006 (2.07)	0.067 (5.14)	−0.043 (−1.11)	0.112 (2.04)	1.112	0.035 (**)
VOL6_6	0.007 (2.26)	0.066 (4.98)	−0.045 (−1.15)	0.118 (2.16)	1.149	0.030 (**)
VOL6_9	0.007 (2.30)	0.067 (4.81)	−0.050 (−1.22)	0.127 (2.36)	1.223	0.034 (**)
VOL6_12	0.007 (2.30)	0.067 (4.73)	−0.056 (−1.35)	0.133 (2.47)	1.349	0.024 (**)
VOL9_3	0.007 (2.19)	0.068 (4.89)	−0.046 (−1.14)	0.110 (2.03)	1.136	0.046 (**)
VOL9_6	0.007 (2.28)	0.068 (4.76)	−0.049 (−1.19)	0.121 (2.24)	1.193	0.037 (**)
VOL9_9	0.007 (2.28)	0.069 (4.66)	−0.054 (−1.30)	0.129 (2.41)	1.303	0.023 (**)
VOL9_12	0.007 (2.31)	0.069 (4.64)	−0.060 (−1.43)	0.134 (2.47)	1.426	0.009 (***)
VOL12_3	0.007 (2.24)	0.070 (4.79)	−0.051 (−1.26)	0.121 (2.24)	1.126	0.034 (**)
VOL12_6	0.007 (2.25)	0.069 (4.70)	−0.053 (−1.28)	0.125 (2.31)	1.283	0.027 (**)
VOL12_9	0.007 (2.27)	0.070 (4.65)	−0.058 (−1.40)	0.130 (2.40)	1.398	0.024 (**)
VOL12_12	0.007 (2.29)	0.070 (4.63)	−0.064 (−1.52)	0.135 (2.46)	1.522	0.008 (***)
Panel D: Industry momentum strategies						
IND3_3	0.008 (4.41)	0.040 (3.74)	−0.049 (−0.97)	0.121 (2.19)	0.975	0.034 (**)
IND3_6	0.006 (3.66)	0.036 (3.62)	−0.058 (−1.11)	0.187 (3.50)	1.114	0.018 (**)
IND3_9	0.006 (4.35)	0.032 (4.10)	−0.063 (−1.31)	0.211 (3.89)	1.313	0.010 (***)
IND3_12	0.006 (4.43)	0.023 (3.90)	−0.012 (−0.23)	0.223 (3.79)	0.233	0.130
IND6_3	0.008 (3.66)	0.042 (4.17)	−0.031 (−0.67)	0.175 (3.34)	0.672	0.071 (*)
IND6_6	0.008 (3.94)	0.036 (4.51)	−0.023 (−0.53)	0.217 (4.44)	0.531	0.094 (*)
IND6_9	0.008 (4.14)	0.026 (4.09)	0.025 (0.51)	0.267 (5.14)	−0.514	0.408
IND6_12	0.006 (3.37)	0.025 (3.45)	0.026 (0.46)	0.256 (4.53)	−0.457	0.388
IND9_3	0.009 (3.86)	0.054 (4.46)	−0.080 (−1.87)	0.233 (4.76)	1.874	0.000 (***)
IND9_6	0.009 (3.84)	0.033 (4.54)	0.010 (0.23)	0.259 (5.23)	−0.255	0.296
IND9_9	0.007 (3.39)	0.028 (3.72)	0.033 (0.63)	0.253 (4.84)	−0.629	0.405
IND9_12	0.005 (2.53)	0.029 (3.84)	0.018 (0.36)	0.239 (4.38)	−0.356	0.335
IND12_3	0.010 (4.15)	0.036 (4.03)	0.010 (0.19)	0.261 (5.11)	−0.193	0.285
IND12_6	0.008 (3.49)	0.032 (3.81)	0.021 (0.41)	0.251 (4.91)	−0.409	0.358
IND12_9	0.006 (2.74)	0.033 (4.18)	0.010 (0.22)	0.252 (4.85)	−0.220	0.293
IND12_12	0.004 (1.90)	0.036 (4.00)	−0.020 (−0.43)	0.234 (4.49)	0.429	0.094 (*)

between the UM and the UMC results. Overall, non-normality and serial correlation in trading profit innovations do not alter our earlier conclusions regarding statistical arbitrage.

Table 3 reports the number of statistical arbitrage opportunities, identified using the bootstrap, for each strategy class and trading profit model. As noted earlier, the value and volume strategies display the greatest consistency in generating statistical arbitrage across trading profit models. The industry-based strategies generate statistical arbitrage more often when expected profits are allowed to be time-varying, while momentum strategies generate statistical arbitrage more frequently when they are assumed to be constant. More importantly, regardless of the assumed trading profit model, at least 50% of the trading strategies yield statistical arbitrage at the 5% level of significance. Therefore, Table 3

Table 3
Summary of statistical arbitrage opportunities.

The number of statistical arbitrage opportunities is reported for each strategy class and trading profit model. The sample period is from Jan 1965–Dec 2000. The trading profit models include CM (constrained mean), UM (unconstrained mean), CMC (constrained mean with correlation), and UMC (unconstrained mean with correlation). The UM and UMC models allow for time-varying expected trading profits. The CMC and UMC models allow for correlated trading profit increments that follow an MA(1) process as in Eq. (3). The classes of strategies include stock momentum, stock value, stock liquidity, and industry momentum. The stock momentum strategies buy the top decile and short the bottom decile of stocks based on past returns as in Jegadeesh and Titman (1993). The stock value strategies buy the top and short the bottom decile of stocks based on book-to-market, cash flow-to-price, and earnings-to-price, as in Lakonishok, Shleifer, and Vishny (1994). They also buy the bottom and short the top decile of stocks based on past sales growth. The stock liquidity strategies buy the bottom and short the top decile of stocks based on trading volume, in the spirit of Brennan, Chordia, and Subrahmanyam (1998). The industry momentum strategies buy the top three industries and short the bottom three industries based on past returns as in Grinblatt and Moskowitz (1999).

Class of trading strategy	Portfolios tested	Trading profit model			
		CM	UM	CMC	UMC
Panel A: Statistical arbitrage opportunities at the 10% level of significance					
Stock momentum (Jegadeesh and Titman, 1993)	16	15	11	15	10
Stock value (Lakonishok, Shleifer, and Vishny, 1994)	12	6	7	6	7
Stock liquidity (Brennan, Chordia, and Subrahmanyam, 1998)	16	16	16	16	15
Industry momentum (Grinblatt and Moskowitz, 1999)	16	4	10	7	11
Total	60	41	44	44	43
Panel B: Statistical arbitrage opportunities at the 5% level of significance					
Stock momentum (Jegadeesh and Titman, 1993)	16	15	4	15	5
Stock value (Lakonishok, Shleifer, and Vishny, 1994)	12	5	6	4	6
Stock liquidity (Brennan, Chordia, and Subrahmanyam, 1998)	16	12	16	16	15
Industry momentum (Grinblatt and Moskowitz, 1999)	16	3	6	4	5
Total	60	35	32	39	31
Panel C: Statistical arbitrage opportunities, using the Bonferroni approach of HJTW, at the 10% level of significance					
Stock momentum (Jegadeesh and Titman, 1993)	16	5	0	13	0
Stock value (Lakonishok, Shleifer, and Vishny, 1994)	12	1	0	1	0
Stock liquidity (Brennan, Chordia, and Subrahmanyam, 1998)	16	0	0	0	0
Industry momentum (Grinblatt and Moskowitz, 1999)	16	1	0	1	0
Total	60	7	0	15	0
Panel D: Statistical arbitrage opportunities, using the Bonferroni approach of HJTW, at the 5% level of significance					
Stock momentum (Jegadeesh and Titman, 1993)	16	3	0	11	0
Stock value (Lakonishok, Shleifer, and Vishny, 1994)	12	1	0	1	0
Stock liquidity (Brennan, Chordia, and Subrahmanyam, 1998)	16	0	0	0	0
Industry momentum (Grinblatt and Moskowitz, 1999)	16	0	0	0	0
Total	60	4	0	12	0

summarizes the improved power of our Min-t statistic when testing for statistical arbitrage. The number of statistical arbitrage opportunities cannot be attributed to the procedure's Type I error (even if the strategies are not independent).

The statistical arbitrage results are even stronger when we confine our attention to the trading profit models chosen by the Akaike Information Criterion (AIC) and the Schwartz Information Criterion (SC). The two leftmost columns in Table 4 report the preferred trading profit model for each trading strategy chosen using the AIC and the SC,

Table 4
Comparing loss probabilities across strategies.

The number of months till loss probability falls below 5% and 1% are reported for all strategies that test positively for statistical arbitrage at the 5% level. The sample period is from Jan 1965–Dec 2000. The JT portfolios are Jegadeesh and Titman (1993) stock momentum portfolios. The BM, CP, EP, and SALE portfolios are Lakonishok, Shleifer, and Vishny (1994) value portfolios based on book-to-market equity, cash flow-to-price, earnings-to-price, and sales growth, respectively. The VOL portfolios are liquidity-based portfolios. The IND portfolios are Grinblatt and Moskowitz (1999) industry momentum portfolios. JT*x*_*y* denotes a stock momentum portfolio with a formation period of *x* months and an evaluation period of *y* months. The VOL*x*_*y* and IND*x*_*y* portfolios are defined analogously. BM*y*, CP*y*, EP*y*, and SALE*y* denote value portfolios with an evaluation period of *y* years. For each strategy, we report the most appropriate model for describing the incremental trading profit dynamics, according to the Akaike information criterion (AIC) and the Schwartz information criterion (SC). The four trading profit models are CM, UM, CMC, and UMC. The UM and UMC models allow for time-varying expected trading profits. The CMC and UMC models allow for correlated trading profit increments that follow an MA(1) process as in Eq. (3).

Portfolio	Preferred model (AIC)	Preferred model (SC)	Months till loss probability $<5\%$				Months till loss probability $<1\%$			
			CM	UM	CMC	UMC	CM	UM	CMC	UMC
Panel A: Momentum strategies										
JT3_3	CM	CM	–	–	–	–	–	–	–	–
JT3_6	CM	CM	286	–	273	–	468	–	447	–
JT3_9	CM	CM	123	–	116	–	205	–	192	–
JT3_12	CM	CM	73	–	64	–	128	–	112	–
JT6_3	CM	CM	423	117	408	117	688	278	663	278
JT6_6	CM	CM	122	–	113	–	200	–	185	–
JT6_9	CM	CM	71	–	65	–	121	–	112	–
JT6_12	CM	CM	74	–	66	–	132	–	117	–
JT9_3	CM	CM	177	–	158	–	292	–	276	–
JT9_6	CM	CM	81	–	73	–	138	–	125	–
JT9_9	CM	CM	79	–	71	113	139	–	124	173
JT9_12	CM	CM	100	129	86	115	182	211	156	188
JT12_3	CM	CM	112	–	102	–	193	–	175	–
JT12_6	CM	CM	99	–	88	–	173	–	154	–
JT12_9	CM	CM	114	135	101	121	204	226	180	202
JT12_12	CM	CM	170	190	149	167	313	325	273	286

Panel B: Value strategies									
BM1	CMC	CM	–	60	–	69	–	101	–	116
BM3	CMC	CMC	–	53	–	63	–	88	–	106
BM5	CMC	CMC	31	43	41	55	58	73	76	93
CP1	CMC	CMC	–	–	–	–	–	–	–	–
CP3	CMC	CM	–	–	–	–	–	–	–	–
CP5	CMC	CM	703	–	–	–	1137	–	–	–
EP1	CMC	CMC	–	–	–	–	–	–	–	–
EP3	CMC	CMC	–	–	–	–	–	–	–	–
EP5	CMC	CMC	–	–	–	–	–	–	–	–
SALE1	CM	CM	50	49	55	55	90	89	99	99
SALE3	CMC	CM	78	40	92	49	137	85	160	106
SALE5	CMC	CMC	97	42	119	55	165	97	201	128
Panel C: Liquidity based strategies										
VOL3_3	CMC	CM	–	317	385	388	–	584	731	711
VOL3_6	CMC	CM	–	232	279	282	–	432	530	522
VOL3_9	CMC	CM	–	202	249	252	–	375	470	465
VOL3_12	CMC	CMC	185	187	230	234	350	347	432	428
VOL6_3	CMC	CM	213	217	265	264	405	407	502	493
VOL6_6	CMC	CM	–	183	224	227	–	342	422	422
VOL6_9	CMC	CMC	174	178	221	222	328	334	415	410
VOL6_12	CMC	CMC	174	176	222	225	326	328	414	414
VOL9_3	CMC	CM	193	61	239	236	366	115	451	440
VOL9_6	CMC	CM	177	182	223	224	334	342	418	416
VOL9_9	CMC	CMC	178	177	225	223	333	330	421	412
VOL9_12	CMC	CMC	173	176	219	222	322	325	408	407
VOL12_3	CMC	CM	184	187	231	229	346	348	433	424
VOL12_6	CMC	CMC	183	184	231	231	344	344	432	427
VOL12_9	CMC	CMC	179	184	226	–	334	341	420	–
VOL12_12	CMC	CMC	179	176	222	228	332	326	410	418

Table 4 (*continued*)

Portfolio	Preferred model (AIC)	Preferred model (SC)	Months till loss probability <5%				Months till loss probability <1%			
			CM	UM	CMC	UMC	CM	UM	CMC	UMC
Panel D: Industry momentum strategies										
IND3_3	CMC	CM	52	48	66	63	99	93	123	120
IND3_6	CMC	CMC	–	68	95	98	–	130	177	180
IND3_9	CMC	CMC	52	81	73	108	97	134	134	175
IND3_12	CMC	CMC	–	68	–	93	–	116	–	155
IND6_3	CMC	CMC	–	–	–	90	–	–	–	171
IND6_6	CMC	CMC	–	85	–	–	–	142	–	–
IND6_9	CMC	CMC	–	86	–	–	–	138	–	–
IND6_12	CMC	CMC	–	–	–	–	–	–	–	–
IND9_3	CMC	CMC	68	109	93	–	125	171	168	–
IND9_6	CMC	CMC	–	–	–	–	–	–	–	–
IND9_9	CMC	CMC	–	–	–	–	–	–	–	–
IND9_12	CMC	CMC	–	–	–	–	–	–	–	–
IND12_3	CMC	CMC	–	–	–	–	–	–	–	–
IND12_6	CMC	CMC	–	–	–	–	–	–	–	–
IND12_9	CMC	CMC	–	–	–	–	–	–	–	–
IND12_12	CMC	CMC	–	–	–	–	–	–	–	–

respectively. A few observations are noteworthy. First, the preferred models chosen by the AIC and SC usually coincide. Second, when differences exist, preference should be given to the SC given the AIC's lack of consistency. Third, only the CM and CMC models are chosen. This is consistent with the lack of significance for θ, which captures the rate of change in expected profits, in the UM and UMC models across all trading strategies. Fourth, for momentum-based strategies, the CM model is always preferred, while for industry momentum-based strategies, the CMC is usually the most appropriate specification. The last observation reinforces the statistically insignificant (significant) autocorrelation estimates, ϕ, for stock (industry) momentum strategies in Table 2.[17]

When we restrict the trading profit model for each trading strategy to the model chosen using the SC, the number of statistical arbitrage opportunities at the 5% level is 15, 5, 12, and 4 for momentum, value, volume, and industry momentum strategies, respectively. This implies that 60% of the trading strategies yield statistical arbitrage under the preferred model specification.

As a final note, for the strategies that we implement, the trading profits are not sufficiently right-skewed for the modified fourth axiom in Definition 2 to be consequential. When we compare the statistical arbitrage test results under Definition 1 [using the test statistic in Eq. (8)] to those under Definition 2, we find that the conclusions are the same for the SC preferred model for all strategies. Only for the UMC model do we find a few strategies that test positive for statistical arbitrage at the 10% level under Definition 2 but not under Definition 1. Recall that the test statistics are equivalent under the CM and CMC models as $\theta = 0$.

6.2. *Economic relevance of statistical arbitrage opportunities*

Our previous results have shown that over 50% of the simple anomaly-based trading strategies generate statistical arbitrage opportunities. However, in itself, this may not be of much relevance to investors if the trading profits are compensation for risk that has not been considered, if their profits take too long to materialize, if the strategies are not exploitable given market frictions and transaction costs, if the trading profits are compensation for liquidity, or if arbitrageurs are likely to reduce these profits going forward. In this section, we explore the economic significance of the anomaly-based trading profits along these lines.

6.2.1. *Relevance to short horizon investors*

Another concern is that the anomaly-based trading strategies may not be relevant to many investors if the profits take too long to materialize. Such investors include institutional fund managers who typically face the risk of retrenchment after a few years, or even a few quarters, of poor performance (see Brown, Harlow, and Starks, 1996; Khorana, 1996; Shleifer and Vishny, 1997; Chevalier and Ellison, 1999). Stein (2005) argues that, as a consequence, open-ended fund managers "will stick primarily to short horizon strategies" that pay off quickly instead of attacking long run mispricings in the market.

[17]The negative ϕ estimates for stock momentum in Table 2 may stem from negative serial correlation in stock returns over monthly horizons. In particular, the stocks held in the long and short portfolios remain in these respective portfolios over several months, which could induce negative serial correlation in their returns. Karolyi and Kho (2004) also find evidence of negative serial correlation in stock momentum returns.

By leveraging on the statistical arbitrage framework and computing the probability of loss for each trading strategy, as in Eq. (4), one can gain a unique perspective on the issue. Table 4 reports for each strategy that test for statistical arbitrage at the 5% level, the number of months before the probability of loss declines below 5% and 1%. The results reveal systematic differences in the rates of decline in loss probability across strategy classes. The volume-based strategies require on average 214 months before the probability of loss declines below 5%. In contrast, momentum, value, and industry momentum-based strategies require on average[18] only 140, 72, and 78 months, respectively, before the probability of loss declines below 5%. Some of these strategies are immensely relevant for investors with a strong preference for strategies that payoff in the short run. For example, the momentum strategy with a formation period of six months and a holding horizon of nine months needs just 71 months under its SC preferred model before its probability of loss declines below 5%. The sales growth-based value strategy with a holding horizon of one year and the industry momentum strategy with a formation and holding period of three months are even more impressive in this regard. They only need 50 and 52 months, respectively, to generate profits with high probability. Thus, investors who prefer to attack short run mispricings in the market will favor momentum, value, and industry momentum strategies.

6.2.2. Transaction costs and market frictions

Yet another concern is that the statistical arbitrage trading profits may not be sufficient to overcome transaction costs and market frictions. This concern is particularly relevant given that many of the statistical arbitrage opportunities (see Table 4) feature short (e.g., three months) formation and holding periods. To account for the effects of transaction costs, we adjust the trading profits downwards using conservative estimates of institutional investor round-trip transaction costs from Chan and Lakonishok (1997). Specifically, we use transaction costs of 1.45%, 1.71%, and 2.45%, which correspond to Chan and Lakonishok's (1997) estimates for NASDAQ trade packages with complexity (package size relative to outstanding equity) levels between the 75th and 90th percentiles, between the 90th and 95th percentiles, and above the 95th percentile, respectively. We use estimates for NASDAQ stocks as opposed to NYSE stocks because there are more NASDAQ stocks than NYSE stocks in the portfolios. Moreover, estimates of transaction costs for NASDAQ stocks tend to be more conservative than those for NYSE stocks. Next, we adjust the trading profits downwards for a short sales liquidity buffer (10%), a levy on the margin accounts (27.5 bps), a margin rate (50%), and a borrowing rate that is 2% higher than the lending rate, as in Jacobs and Levy (1995) and Alexander (2000). Then, we re-test the resultant post-transaction costs portfolios for statistical arbitrage. For brevity, we confine our analysis to strategies that test for statistical arbitrage (under the model chosen by the SC) at the 5% level of significance and to strategies that require less than 400 months before their probability of loss falls below 5%. Since we find that volume-based strategies in general may not be particularly relevant to short-term investors, given their relatively slow rate of decline in loss probability, we limit the analysis to volume strategies with a nine-month formation period. This yields a total of 26 strategies on which we apply the transaction costs adjustment.[19]

[18]The average number of months is calculated using the preferred model determined by the Schwartz Information Criterion. In the calculation of the stock value-based strategy average, we exclude the cash flow-to-price strategy with a five-year holding horizon so as to avoid the effect of this outlier.

[19]The rest of our analysis focuses on this set of strategies.

Table 5 reports the annual turnover for the 26 strategies, as well as their statistical arbitrage test results under the various transaction costs regimes. Of the 26 strategies, 23 test for statistical arbitrage at the 5% level with the 1.45% round-trip transaction costs adjustment. Even with the very conservative estimate of 2.45% per round-trip transaction, we find that 17 of the 26 strategies still test for statistical arbitrage at the 5% level of significance. Value and volume strategies are highly robust to the transaction costs adjustments. Out of the four value and four volume strategies tested, all four value strategies and three volume strategies test for statistical arbitrage at the 5% level with the 2.45% estimate of transaction costs. Momentum strategies are also robust to transaction costs. Of the 14 momentum strategi es tested, 9 remain statistical arbitrage opportunities at the 5% level with the 2.45% estimate of transaction costs. In contrast, the industry momentum strategies display the most sensitivity to transaction costs with only one of the four strategies surviving the final adjustment for transaction costs. This is not surprising given the short formation periods (e.g., three months), and consequently high turnover, of the industry momentum strategies tested. On the other hand, the low turnover of the annually rebalanced value strategies render them almost impervious to the effects of transaction costs.

In addition, we can leverage on the results in Table 5 and estimates from Chan and Lakonishok (1997) to determine the amount that may be invested under the various transaction costs regimes. The rightmost column of Panel B, Table III in Chan and Lakonishok (1997) provides estimates of the number of trade packages and the dollar value traded for NASDAQ stocks with market capitalization below the 30th NYSE percentile, and for each complexity group. From those estimates we can infer the average size of the trade packages for the smallest stocks in the NASDAQ. This together with the average number of stocks in the long/short momentum portfolios (e.g., 412) allows us to derive an estimate of the amount of funds that may be profitably invested in momentum strategies under each transaction costs regime. For example, the average trade package size is \$1.56m, \$0.81m, and \$0.40m under the 2.45%, 1.71%, and 1.45% transaction cost regimes, respectively, for the smallest stocks. Hence, the statistical arbitrage results in Table 5 imply that at least \$0.64bn ($=$\$1.56m $\times$ 412) may be profitably invested in each of the 17 strategies that test for statistical arbitrage with the 2.45% round-trip estimate of transaction costs. This number is somewhat more optimistic than those from Korajczyk and Sadka (2004).[20]

6.2.3. *Illiquidity*

Transaction costs aside, it will be important to also check the sensitivity of the statistical arbitrage trading profits to illiquidity. If the trading profits are driven by highly illiquid stocks, it would be hard for large institutional investors to take advantage of the market inefficiencies uncovered. To gauge the effects of illiquidity, we constrain the stock sample to the top 70% and top 50% of stocks in terms of volume (share volume/number of shares outstanding) averaged over the evaluation period and re-test for statistical arbitrage. The results reported in Table 6 indicate that the trading profits from momentum and value strategies are not driven by illiquid stocks. In fact, the rightmost column in Table 6 reveals that the loss probabilities often decline faster without the 50% most illiquid stocks than with those stocks in the sample. Indeed, the probability of loss declines more rapidly for 10 of the 14 the momentum and three of the four value strategies without the illiquid stocks. One view is that by removing illiquid stocks, we reduce the overall volatility of the risky

[20] See Figs. 4a, 5a, and 6a in Korajczyk and Sadka (2004).

Table 5

Tests of statistical arbitrage with adjustments for market frictions and transaction costs.

The sample period is from Jan 1965–Dec 2000. The strategies and models are as per defined in the previous tables. The 1.34%, 1.71%, and 2.45% round-trip cost corresponds to the estimated cost for the third, fourth, and highest complexity (trade size over stock size) groups of NASDAQ portfolios for institutional investors, respectively, as in Chan and Lakonishok (1997). Incremental profits are also adjusted downwards for a short sales liquidity buffer (10%), a levy on the margin accounts (27.5 bps), a margin rate (50%) and a borrowing rate that is 2% higher than the lending rate, as in Alexander (2000) and Jacobs and Levy (1995). Only strategies that test for statistical arbitrage at the 5% level, before transaction costs, are included in the analysis. The preferred model is chosen using the Schwartz information criterion. *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Portfolio	Preferred model	Annual turnover	Round-trip transaction costs = 1.34%				Round-trip transaction costs = 1.71%				Round-trip transaction costs = 2.45%			
			Mean μ	Min-t	p-value	Months till loss prob < 5%	Mean μ	Min-t	p-value	Months till loss prob < 5%	Mean μ	Min-t	p-value	Months till loss prob < 5%
JT3_6	CM	1.72	−0.001	−1.026	0.533	–	−0.003	−2.087	0.794	–	−0.006	−4.212	0.970	–
JT3_9	CM	1.16	0.002	1.769	0.002(***)	402	0.001	1.190	0.013(**)	728	−0.000	−0.278	0.258	–
JT3_12	CM	0.89	0.003	2.498	0.000(***)	153	0.003	2.495	0.000(***)	198	0.002	1.742	0.000(***)	391
JT6_6	CM	1.72	0.003	1.959	0.001(***)	350	0.002	1.364	0.004(***)	585	−0.000	−0.046	0.192	–
JT6_9	CM	1.15	0.006	3.047	0.000(***)	124	0.005	3.047	0.000(***)	152	0.004	2.404	0.000(***)	243
JT6_12	CM	0.89	0.005	1.922	0.000(***)	131	0.004	1.920	0.000(***)	157	0.004	1.918	0.001(***)	231
JT9_3	CM	2.11	0.002	0.871	0.034(**)	–	0.000	0.124	0.152	–	−0.003	−1.620	0.708	–
JT9_6	CM	1.45	0.006	3.205	0.000(***)	154	0.005	2.793	0.000(***)	191	0.004	1.930	0.000(***)	337
JT9_9	CM	1.15	0.006	2.236	0.000(***)	142	0.005	2.236	0.000(***)	176	0.004	0.086	0.000(***)	288
JT9_12	CM	0.88	0.004	1.485	0.001(***)	189	0.004	1.484	0.002(***)	233	0.003	1.480	0.001(***)	388
JT12_3	CM	1.84	0.004	1.997	0.000(***)	319	0.003	1.512	0.002(***)	488	0.001	0.511	0.064(*)	–
JT12_6	CM	1.29	0.005	2.348	0.000(***)	200	0.005	2.284	0.000(***)	251	0.003	1.577	0.000(***)	450
JT12_9	CM	1.04	0.005	1.701	0.002(***)	230	0.004	1.702	0.001(***)	292	0.003	1.429	0.000(***)	512
JT12_12	CM	0.89	0.003	1.165	0.007(***)	418	0.003	1.164	0.008(***)	576	0.002	0.814	0.014(**)	1414
BM5	CMC	0.17	0.010	1.074	0.031(**)	39	0.009	1.072	0.024(**)	40	0.009	1.067	0.029(**)	41
SALE1	CM	0.49	0.007	1.263	0.014(**)	54	0.007	1.283	0.006(***)	58	0.006	1.316	0.005(***)	70
SALE3	CM	0.29	0.005	0.979	0.039(**)	73	0.004	0.969	0.049(**)	79	0.004	0.943	0.039(**)	90
SALE5	CMC	0.18	0.004	2.566	0.001(***)	115	0.004	2.551	0.001(***)	121	0.004	2.519	0.000(***)	131
VOL9_3	CM	0.81	0.004	1.043	0.061(*)	383	0.004	1.044	0.056(*)	484	0.003	1.047	0.058(*)	808
VOL9_6	CM	0.67	0.005	1.101	0.046(**)	313	0.004	1.102	0.048(**)	354	0.004	1.103	0.039(**)	515
VOL9_9	CMC	0.58	0.005	1.302	0.027(**)	358	0.004	1.301	0.023(**)	404	0.004	1.300	0.029(**)	549
VOL9_12	CMC	0.49	0.005	1.427	0.011(**)	323	0.005	1.426	0.017(**)	362	0.004	1.409	0.015(**)	443
IND3_3	CM	3.05	0.001	0.522	0.100(*)	–	−0.001	−0.793	0.491	–	−0.006	−3.977	0.989	–
IND3_6	CMC	1.58	0.002	1.112	0.023(**)	622	0.001	0.699	0.051(*)	–	−0.001	−0.673	0.455	–
IND3_9	CMC	1.04	0.003	1.278	0.012(**)	182	0.003	1.267	0.016(**)	261	0.002	1.153	0.014(**)	744
IND9_3	CMC	1.78	0.004	1.893	0.000(***)	288	0.003	1.484	0.010(***)	496	0.001	0.459	0.102	–

Table 6
Tests of statistical arbitrage with controls for stock liquidity.

The sample period is from Jan 1965–Dec 2000. The strategies and models are as per defined in the previous tables. The strategies are based on the top 70% and top 50% of stocks in terms of past volume. Only strategies that test for statistical arbitrage at the 5% level, on the full sample, are included in the analysis. The preferred model is chosen using the Schwartz information criterion. *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Portfolio	Preferred model	Top 70% of stocks based on volume				Top 50% of stocks based on volume			
		Mean (μ)	Min-t	p-value	Months till loss prob < 5%	Mean (μ)	Min-t	p-value	Months till loss prob < 5%
JT3_6	CM	0.006	3.151	0.000(***)	166	0.008	2..839	0.000(***)	106
JT3_9	CM	0.007	3.333	0.000(***)	104	0.009	3.117	0.000(***)	77
JT3_12	CM	0.007	2.532	0.000(***)	68	0.009	2.384	0.000(***)	54
JT6_6	CM	0.009	3.598	0.000(***)	109	0.011	3.209	0.000(***)	82
JT6_9	CM	0.010	2.861	0.000(***)	71	0.012	2.607	0.000(***)	58
JT6_12	CM	0.008	1.997	0.000(***)	81	0.010	1.908	0.001(***)	68
JT9_3	CM	0.009	3.557	0.000(***)	138	0.011	3.295	0.000(***)	106
JT9_6	CM	0.011	2.983	0.000(***)	79	0.013	2.740	0.000(***)	67
JT9_9	CM	0.010	2.195	0.000(***)	85	0.011	2.012	0.000(***)	132
JT9_12	CM	0.007	1.628	0.001(***)	122	0.008	1.482	0.000(***)	108
JT12_3	CM	0.010	2.792	0.000(***)	106	0.012	2.664	0.000(***)	87
JT12_6	CM	0.009	2.187	0.000(***)	113	0.010	2.150	0.000(***)	96
JT12_9	CM	0.007	1.599	0.002(***)	151	0.008	1.518	0.001(***)	129
JT12_12	CM	0.005	1.207	0.002(***)	276	0.005	1.156	0.003(***)	226
BM5	CMC	0.010	1.424	0.004(***)	110	0.010	1.317	0.017(**)	48
SALE1	CM	0.009	0.797	0.026(**)	39	0.010	0.467	0.090(*)	37
SALE3	CM	0.006	1.085	0.024(**)	74	0.006	1.086	0.034(**)	70
SALE5	CMC	0.005	1.763	0.004(***)	110	0.005	1.541	0.012(**)	89
VOL9_3	CM	0.009	−0.368	0.379	–	0.008	1.678	0.016(**)	84
VOL9_6	CM	0.008	−0.709	0.458	–	0.008	1.544	0.030(**)	82
VOL9_9	CMC	0.008	−0.485	0.416	–	0.008	1.831	0.010(***)	99
VOL9_12	CMC	0.008	−0.438	0.430	–	0.008	1.743	0.007(***)	99
IND3_3	CM	0.010	−2.062	0.867	–	0.011	−2.377	0.877	–
IND3_6	CMC	0.008	−1.166	0.601	–	0.008	−1.341	0.662	–
IND3_9	CMC	0.007	−0.916	0.568	–	0.008	−1.187	0.628	–
IND9_3	CMC	0.011	0.093	0.180	–	0.013	−1.752	0.772	–

Table 7
Tests of statistical arbitrage by subperiod.

The sample period is from Jan 1965–Dec 2000. The strategies and models are as per defined in the previous tables. The statistical arbitrage test is applied separately to the first half and the second half of the sample period. Only strategies that test for statistical arbitrage at the 5% level, on the entire sample period, are included in the analysis. The preferred model is chosen using the Schwartz information criterion for each subperiod. *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Portfolio	1st half of sample period					2nd half of sample period				
	Preferred model	Mean(μ)	Min-t	p-value	Months till loss prob < 5%	Preferred model	Mean(μ)	Min-t	p-value	Months till loss prob < 5%
JT3_6	CM	0.001	0.128	0.210	–	CM	0.004	−0.151	0.276	–
JT3_9	CM	0.002	0.688	0.059(*)	658	CM	0.007	0.030	0.227	–
JT3_12	CM	0.004	1.077	0.019(**)	178	CM	0.007	0.179	0.156	–
JT6_6	CM	0.004	0.094	0.029(**)	409	CM	0.009	−0.181	0.294	–
JT6_9	CM	0.006	1.271	0.004(***)	148	CM	0.010	0.223	0.137	–
JT6_12	CM	0.005	0.442	0.094(*)	155	CM	0.009	0.226	0.149	–
JT9_3	CM	0.003	0.591	0.076(*)	881	CM	0.009	0.311	0.141	–
JT9_6	CM	0.007	1.248	0.005(***)	151	CM	0.011	0.361	0.119	–
JT9_9	CM	0.006	0.516	0.064(*)	149	CM	0.010	0.490	0.084(*)	38
JT9_12	CM	0.005	−0.007	0.204	–	CM	0.008	0.537	0.067(*)	58
JT12_3	CM	0.007	0.963	0.016(**)	200	CM	0.010	0.382	0.086(*)	51
JT12_6	CM	0.007	0.513	0.075(*)	177	CM	0.010	0.422	0.116	–
JT12_9	CM	0.006	0.092	0.166	–	CM	0.008	0.634	0.048(**)	68
JT12_12	CM	0.004	−0.426	0.404	–	CM	0.006	0.679	0.048(**)	111
BM5	CMC	0.009	−1.666	0.838	–	CMC	0.011	0.144	0.196	–
SALE1	CM	0.007	−0.102	0.247	–	CM	0.010	0.878	0.055(*)	27
SALE3	CM	0.007	−1.983	0.927	–	CM	0.005	1.405	0.021(**)	60
SALE5	CMC	0.009	0.222	0.105	–	CMC	0.005	1.950	0.004(***)	95
VOL9_3	CM	0.009	2.003	0.001(***)	123	CM	0.006	−1.682	0.761	–
VOL9_6	CM	0.009	2.084	0.000(***)	118	CM	0.007	−1.722	0.753	–
VOL9_9	CMC	0.009	1.981	0.000(***)	129	CMC	0.006	−1.406	0.716	–
VOL9_12	CMC	0.008	1.938	0.000(***)	136	CMC	0.007	−1.468	0.754	–
IND3_3	CM	0.009	0.513	0.069(*)	41	CM	0.007	2.264	0.001(***)	72
IND3_6	CMC	0.006	0.413	0.131	–	CMC	0.005	1.780	0.002(***)	109
IND3_9	CMC	0.005	−0.029	0.220	–	CMC	0.006	3.170	0.000(***)	76
IND9_3	CMC	0.005	0.021	0.205	–	CMC	0.009	3.453	0.000(***)	77

portfolio and allow for faster convergence to arbitrage. It is not surprising that the volume-based strategies display some sensitivity to the removal of the illiquid stocks. However, it is intriguing to note that the volume strategies without the 50% most illiquid stocks actually outperform those without the 30% most illiquid stocks. Moreover, the strategies in the former group also outperform their counterparts based on the full sample of stocks (see Table 4) in terms of the loss probability metric. This suggests that the profitability of volume strategies is not simply a function of the cross-sectional variation in volume within the sample nor is it solely dependent on the presence of extremely illiquid stocks in the sample. In contrast to the momentum and value strategies, the industry momentum strategies do not test for statistical arbitrage when we remove the most illiquid stocks from the sample. This together with the transaction costs results suggest that it would be difficult for institutional investors to take advantage of the industry momentum profits.

6.2.4. *Rational arbitrageurs and statistical arbitrage*

Finally, there are also concerns that, going forward, market participants will quickly arbitrage away most of the profits from the statistical arbitrage opportunities, rendering them irrelevant in the near future. While we cannot directly test this assertion, we can conduct subperiod tests to see whether investors are arbitraging away the trading profits over the 1965–2000 sample period. To this end, we split the sample period into two and conduct statistical arbitrage tests on each subperiod. If arbitrageurs are reducing the trading profits from the statistical arbitrage opportunities, then we would expect the statistical arbitrage profits to diminish significantly in the second half of the sample period relative to the first half of the sample period.

The results from the subperiod analysis reported in Table 7 demonstrate that value and industry momentum strategies perform better, while the volume and momentum strategies fare worse, in the second half than in the first half of the sample period. Of the original set of statistical arbitrage strategies, 15 and 12 test for statistical arbitrage at the 10% level[21] in the first and second subperiods, respectively. Overall, market participants have only been able to reduce the number of statistical arbitrage opportunities by 20%. Moreover, the average number of months until a loss probability declines below 5% has actually decreased for those strategies that test for statistical arbitrage in the second period versus the first. The average is 70.2 months for the statistical arbitrage opportunities in the second subperiod and 243.5 months for the statistical arbitrage opportunities in the first subperiod. Hence, while the number of statistical arbitrage opportunities declines somewhat over time, the opportunities that survive are much more relevant to investors.

7. Conclusion

The traditional intercept test of market efficiency is a widely-used metric in empirical asset pricing studies. However, the uncertainty surrounding equilibrium risk factors limit this metric's conclusions. This paper develops an alternative approach to the traditional intercept test that replaces an assumed model of market equilibrium with an assumed statistical model for trading profits. Our test improves on the power and consistency of the test for statistical arbitrage in Hogan, Jarrow, Teo, and Warachka (2004) and permits the

[21] Note that the power of the test falls significantly with the reduction in sample size.

examination of trading profits under fewer assumptions. When applying our test to a wide range of financial anomalies, we find that over 50% of the strategies are inconsistent with market efficiency, conditional on our assumed trading profit process.

Moreover, we provide several new insights into the economic relevance of these trading strategies. We show that short-horizon investors will prefer momentum, value, and industry momentum strategies over liquidity strategies. Furthermore, while the trading profits from the momentum, value, and liquidity strategies are fairly robust to adjustments for transaction costs and market frictions, the industry momentum strategies are not as they tend to be driven by illiquid stocks. Finally, investors do not, in general, appear to be arbitraging away the financial anomaly profits over our 35-year sample period.

Appendix A. Conversion of returns into trading profits

Define R^L and R^S as the long and short portfolio returns respectively from a zero-cost trading strategy (return of past winners and return of past losers in the case of momentum). The conversion of these portfolio returns into dollar-denominated profits places the cumulative profit of the underlying trading strategy into the risk-free asset (money market account) after each period. To properly capture an anomaly's persistence, a \$1 buy/sell position in the risky long/short portfolios is maintained across time.[22] This conversion process yields cumulative trading profits equaling:

$$V(j) = \exp\{r\} V(j-1) + \$1[\exp\{R_j^L\} - \exp\{R_j^S\}], \tag{11}$$

whose incremental trading profits (losses) are harvested each month, with cumulative trading profits being constructed recursively as a consequence.

Observe that time-varying moments are induced by allocating wealth between the risk-free asset and the risky portfolio. In particular, \$1 is exposed to the risky long minus short position while the cumulative trading profit is deposited into the risk-free asset (or borrowed from in the event of a cumulative loss).

Overall, the cumulative trading profit may be decomposed as:

$$[1-\pi(j)]\exp\{r\} + \pi(j)[\exp\{R_j^L\} - \exp\{R_j^S\}] \tag{12}$$

over a single time increment. Thus, $1-\pi(j)$ is invested in the risk-free asset while the remaining $\pi(j)$ fraction is maintained in the risky portfolio. A statistical arbitrage has $\pi(j) \to 0$. Intuitively, this conversion eventually creates a "risk-free arbitrage" (with a zero investment). Note that the interest rate is not assumed to be constant nor is stationarity imposed on the long/short portfolio returns. More importantly, the time-varying nature of $\pi(j)$ implies that θ and λ are not necessarily zero since incremental trading profits reflect differences in the amount invested in the long/short portfolios underlying a particular cross-sectional return anomaly. Eq. (11) may be altered by investing $x(j)$ in the risky portfolios rather than \$1:

$$V_x(j) = \exp\{r\} V_x(j-1) + \$x(j)[\exp\{R_j^L\} - \exp\{RR_j^S\}]. \tag{13}$$

[22]The standard buy- and hold-strategy yields trading profits that are very sensitive to the start date and is inappropriate for frequent (e.g., monthly) realizations of intermediate gains and losses.

For example, having $x(j)=B(j)$ gradually increases the statistical arbitrage's exposure to the risky portfolios over time (in nominal but not real terms) while maintaining its self-financing property.

Recall that the purpose of statistical arbitrage is to test whether the persistence of cross-sectional return anomalies generate arbitrage profits in the long run. This methodology is not intended to test whether one can manipulate the returns of an anomaly so as to reject market efficiency. Consequently, a rapidly decreasing $x(j)$ function is not valid since the investment in the underlying anomaly diminishes over time. Indeed, this conversion eventually puts a negligible weight on the anomaly's returns, and therefore fails to address its persistence. Similarly, investing the cumulative trading profit $x(j)=V_x(j-1)$ from a return anomaly into the risky long/short positions would not evaluate this anomaly's persistence. Instead, this conversion would allow the investor to lose their entire profit over a single time interval.

Overall, economic considerations dictate that the conversion of returns into trading profits be accomplished using a simple transformation. The θ and λ lambda parameters guard against inappropriate $x(j)$ functions that cause incremental trading profits to rapidly decrease or their volatility to rapidly increase, respectively.

Appendix B. Verification of semi-variance sub-hypotheses

The quantity $Var[\Delta \upsilon(t)|\Delta \upsilon(t)<0]$ is computed from the distribution of $\Delta \upsilon(t)$, which equals $N(\mu t^{\theta},\sigma^2 t^{2\lambda})$. The conditional variance is expressed as:

$$Var\left[\Delta \upsilon(t)\middle|\Delta \upsilon(t)<0\right] = \frac{1}{\sqrt{2\pi\sigma^2 t^{2\lambda}}}\int_{-\infty}^{0} (x-\mu t^{\theta})^2 e^{-(x-\mu t^{\theta})^2/2\sigma^2 t^{2\lambda}}\, dx, \tag{14}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\mu t^{\theta}/\sigma t^{\lambda}} (\sigma t^{\lambda} y)^2 e^{-y^2/2}\, dy, \tag{15}$$

$$= \frac{\sigma^2 t^{2\lambda}}{\sqrt{2\pi}}\int_{-\infty}^{-\mu t^{\theta}/\sigma t^{\lambda}} y^2 e^{-y^2/2} dy, \tag{16}$$

$$\leq \sigma^2 t^{2\lambda}, \tag{17}$$

after a change of variables $y=(x-\mu t^{\theta})/\sigma t^{\lambda}$ which implies $\sigma t^{\lambda} dy=dx$. The inequality in Eq. (17) stems from;

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\mu t^{\theta}/\sigma t^{\lambda}} y^2 e^{-y^2/2}\, dy \leq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-y^2/2}\, dy = 1, \tag{18}$$

since the second term equals the second moment (or variance) of a standard normal random variable. Thus, the constraint $\lambda<0$ is a sufficient condition for the fourth axiom to hold. However, the integral:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(-\mu/\sigma)t^{\theta-\lambda}} y^2 e^{-y^2/2}\, dy \tag{19}$$

also converges to zero provided $\theta > \lambda$. Indeed, if $\theta > \lambda$, then $t^{\theta-\lambda} \to \infty$ as $t \to \infty$, which implies the range of integration declines to zero. Thus, a weaker version of the fourth axiom implies statistical arbitrage occurs if either $\lambda < 0$ or $\theta > \lambda$.

To provide an alternative perspective and confirm the above result, observe that the integral in Eq. (16) equals:

$$\frac{\mu t^{\theta-\lambda}}{\sqrt{2\pi\sigma^2}} e^{-\mu^2 t^{2(\theta-\lambda)}/2\sigma^2} + N\left(\frac{-\mu t^{\theta-\lambda}}{\sigma}\right). \tag{20}$$

Although there is no closed form solution for the standard normal cdf, a polynomial approximation (for $x < 0$) is available in Hull (2000) as:

$$N(x) = N'(x)\left(a_1 \frac{1}{1+\gamma x} + a_2 \frac{1}{(1+\gamma x)^2} + a_3 \frac{1}{(1+\gamma x)^3} + \text{h.o.t.}\right), \tag{21}$$

where a_1, a_2, a_3, and γ are constants. Ignoring the constants a_1, σ, and μ, as well as the contribution of $1/(1+\gamma x)$ implies the relevant terms of Eq. (19) are of the order:

$$t^{\theta-\lambda} e^{-t^{2(\theta-\lambda)}} + e^{-t^{2(\theta-\lambda)}}. \tag{22}$$

The product $t^{2\lambda}$ from Eq. (16) or (17) results in the above expression becoming:

$$t^{\theta+\lambda} e^{-t^{2(\theta-\lambda)}} + t^{2\lambda} e^{-t^{2(\theta-\lambda)}}. \tag{23}$$

Since the exponential function converges to zero for $\theta-\lambda$ faster than the power function increases towards ∞, the conditional semi-variance becomes zero in the limit as $t \to \infty$.

Appendix C. Bonferroni approach for multiple hypotheses

This appendix discusses the Bonferroni approach for testing sub-hypotheses, with particular reference to testing for statistical arbitrage.

Let H_0 be the null hypothesis consisting of K sub-hypotheses $h_1, \ldots, h_K$, all of which are required to hold under H_0. Thus, the rejection of even one sub-hypothesis rejects the null H_0. As a consequence, H_0 is the *intersection* of sub-hypotheses given by:

$$H_0 : \bigcap_{i=1}^{K} h_i.$$

In the Bonferroni procedure, each sub-hypothesis h_i is tested at a given level of significance α_i with a critical region denoted C_i so that $\Pr(C_i | H_0) = \alpha_i$. The critical region of the null hypothesis H_0 is the union $\bigcup_{i=1}^{K} C_i$. Let C_i^C be the complement of C_i. The null hypothesis H_0 cannot be rejected if all the sub-hypotheses are accepted. Suppressing the conditioning notation, the probability of failing to reject H_0 equals $\Pr(\bigcap_{i=1}^{K} C_i^C)$.

The Bonferroni inequality states that:

$$\Pr\left(\bigcap_{i=1}^{K} C_i^C\right) \geq 1 - \sum_{i=1}^{K} \Pr(C_i) = 1 - \sum_{i=1}^{K} \alpha_i, \tag{24}$$

from which we obtain:

$$\sum_{i=1}^{K} \alpha_i \geq 1 - \Pr\left(\bigcap_{i=1}^{K} C_i^C\right). \tag{25}$$

Therefore, $\sum_{i=1}^{K} \alpha_i$ is an upper bound on the size of the statistical test, that is, the probability of committing a Type I error. If H_0 is not satisfied, then at least one sub-hypothesis, say h_j, is not satisfied. As:

$$\Pr\left(\bigcup_{i=1}^{K} C_i\right) \geq \Pr(C_j), \tag{26}$$

we observe that if all the sub-tests reject their sub-hypothesis with probability one as the sample size tends to infinity, $\Pr(C_j) \to 1$, then $\Pr(\bigcap_{i=1}^{K} C_i) \to 1$. As a result, the Bonferroni test is consistent.

However, in the statistical arbitrage test conducted by HJTW, the null hypothesis of no statistical arbitrage is a *union* of sub-hypotheses. This statement is a consequence of the fact that to reject no statistical arbitrage, all the sub-hypotheses must be rejected. Rejecting one sub-hypothesis is not sufficient to reject no statistical arbitrage. Thus, the null hypothesis is defined as:

$$H_0^* : \bigcup_{i=1}^{K} h_i, \tag{27}$$

and the probability of being unable to reject H_0^* is $\Pr(\bigcap_{i=1}^{K} C_i^C)$. As the probability of a union is greater than its corresponding intersection, we have:

$$\Pr\left(\bigcup_{i=1}^{K} C_i^C\right) \geq \Pr\left(\bigcap_{i=1}^{K} C_i^C\right), \tag{28}$$

which, when combined with Eq. (25), yields the relationship:

$$\sum_{i=1}^{K} \alpha_i \geq 1 - \Pr\left(\bigcap_{i=1}^{K} C_i^C\right) \geq 1 - \Pr\left(\bigcup_{i=1}^{K} C_i^C\right). \tag{29}$$

Thus, we conclude that $\sum_{i=1}^{K} \alpha_i$ is also an upper bound on the size of the test for the null hypothesis H_0^* defined in terms of a union. However, Eq. (29) implies the Bonferroni inequality is a weaker bound for H_0^* than for H_0. Furthermore, the Bonferroni test is generally not consistent for H_0^*, in contrast to H_0. Indeed, when K is large, the actual size of the Bonferroni test for H_0^* may be far below $\sum_{i=1}^{K} \alpha_i$, resulting in a test with low power. Conversely, the Min-t test has the correct nominal size. To the extent that searching for the maximum probability of rejection over the parameter space H_0 results in the true maximum, the power of the test is also enhanced.

References

Affleck-Graves, J., McDonald, B., 1989. Nonnormalities and tests of asset pricing theories. Journal of Finance 44, 889–908.

Alexander, G., 2000. On back-testing 'zero-investment' strategies. Journal of Business 73, 255–278.

Berkowitz, J., Kilian, L., 2000. Recent developments in bootstrapping time series. Econometric Reviews 19, 1–48.

Bernardo, A., Ledoit, O., 2000. Gain, loss, and asset pricing. Journal of Political Economy 108, 144–172.

Bessembinder, H., Chan, K., 1998. Market efficiency and the returns to technical analysis. Financial Management 21, 5–17.

Brennan, M., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics and the cross-section of expected stock returns. Journal of Financial Economics 49, 345–373.

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. Journal of Finance 47, 1731–1764.
Brown, K., Harlow, W., Starks, L., 1996. Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry. Journal of Finance 51, 85–110.
Chan, L., Lakonishok, J., 1997. Institutional equity trading costs: NYSE versus NASDAQ. Journal of Finance 52, 713–735.
Chevalier, J., Ellison, G., 1999. Career concerns of mutual fund managers. Quarterly Journal of Economics 114, 389–432.
Fama, E., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49, 283–306.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.
Gourieroux, C., Monfort, A., 1995. Statistics and Econometric Models, Vol 2. Cambridge University Press, Cambridge, UK.
Grinblatt, M., Moskowitz, T., 1999. Do industries explain momentum? Journal of Finance 54, 1249–1290.
Hall, P., Wilson, S., 1991. Two guidelines for bootstrap hypothesis testing. Biometrics 47, 757–762.
Hull, J., 2000. In: Options, Futures, and Other Derivative Securities, 4th ed. Prentice-Hall, New Jersey, USA.
Hogan, S., Jarrow, R., Teo, M., Warachka, M., 2004. Testing market efficiency using statistical arbitrage with applications to momentum and value trading strategies. Journal of Financial Economics 73, 525–565.
Horowitz, J., 2001. The bootstrap and hypothesis tests in econometrics. Journal of Econometrics 100, 37–40.
Jacobs, B., Levy, K., 1995. More on long-short strategies. Financial Analysts Journal 51, 88–90.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.
Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: an evaluation of alternative explanations. Journal of Finance 56, 699–720.
Karolyi, A., Kho, B., 2004. Momentum strategies: some bootstrap tests. Journal of Empirical Finance 11, 509–536.
Khorana, A., 1996. Top management turnover: an empirical investigation of mutual fund managers. Journal of Political Economy 98, 225–264.
Korajczyk, R., Sadka, R., 2004. Are momentum profits robust to trading costs? Journal of Finance 59, 1039–1082.
Künsch, H., 1989. The jackknife and the bootstrap for general stationary observations. Annals of Statistics 17, 1217–1241.
Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.
Lesmond, D., Schill, M., Zhou, C., 2004. The illusory nature of momentum profits. Journal of Financial Economics 71, 349–380.
Lo, A., MacKinlay, C., 1988. Stock market do not follow random walks: evidence from a simple specification test. Review of Financial Studies 1, 41–66.
Phillips, P., Xu, K., 2006. Inference in autoregression under heteroskedasticity. Journal of Time Series Analysis 27, 289–308.
Politis, D., Romano, J., 1994. The stationary bootstrap. Journal of the American Statistical Association 89, 1303–1313.
Ruiz, E., Pascual, L., 2002. Bootstrapping financial time series. Journal of Economic Surveys 16, 271–300.
Shleifer, A., Vishny, R., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.
Stein, J., 2005. Why are most mutual funds open-end? Competition and the limits of arbitrage. Quarterly Journal of Economics 120, 247–272.
Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping, technical trading rule performance and the bootstrap. Journal of Finance 54, 1647–1691.
Xu, K., 2008. Bootstrapping autoregression under non-stationary volatility. Econometrics Journal 11, 1–26.
Xu, K., Phillips, P., 2008. Adaptive estimation of autoregressive models with time-varying variances. Journal of Econometrics 142, 265–280.