NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

A Reality Check on Technical Trading Rule Profits in US Futures Markets

by
Cheol-Ho Park and Scott H. Irwin

Suggested citation format:

Park, C.-H., and S. H. Irwin. 2005. "A Reality Check on Technical Trading Rule Profits in US Futures Markets." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.uiuc.edu/nccc134].

A Reality Check on Technical Trading Rule Profits in US Futures Markets

Cheol-Ho Park
and
Scott H. Irwin*

Paper presented at the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 18-19, 2005

Copyright 2005 by Cheol-Ho Park and Scott H. Irwin. All rights reserved.
Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

* Cheol-Ho Park (cpark7@uiuc.edu) is a Graduate Research Assistant and Scott H. Irwin (sirwin@uiuc.edu) is the Laurence J. Norton Professor of Agricultural Marketing in the Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign.

A Reality Check on Technical Trading Rule Profits in US Futures Markets

Practitioners Abstract

This paper investigates the profitability of technical trading rules in US futures markets over the 1985-2004 period. To account for data snooping biases, we evaluate statistical significance of performance across technical trading rules using White's Bootstrap Reality Check test and Hansen's Superior Predictive Ability test. These methods directly quantify the effect of data snooping by testing the performance of the best rule in the context of the full universe of technical trading rules. Results show that the best rules generate statistically significant economic profits only for two of 17 futures contracts traded in the US. This evidence indicates that technical trading rules generally have not been profitable in US futures markets after correcting for data snooping biases.

Keywords: Technical Analysis, Data Snooping, Reality Check, Futures Markets

Introduction

Technical analysis has been widely used among market participants in a variety of asset markets. Surveys show that the use of technical analysis has become more prevalent in recent years (e.g., Cheung and Chinn; Gehrig and Menkhoff). In their survey of foreign exchange professionals, for example, Gehrig and Menkhoff conclude that "… technical analysis dominates foreign exchange and most foreign exchange traders seem to be chartists now" (p. 3). In addition, recent behavioral finance models such as feedback models show that noise traders who buy when prices rise and sell when prices fall (trend chasers) on average can earn higher returns than rational investors in the short run because of "noise trader risk," and they can survive and even dominate the market in the long-run (De Long et al. 1990, 1991; Slezak).

Numerous empirical studies have investigated the profitability of technical trading rules and many find evidence of positive trading profits (Park and Irwin, 2004). As an example, Lukac, Brorsen, and Irwin reported that four technical trading systems, including the dual moving average and the price channel systems, produced statistically significant monthly portfolio net returns of 1.89%-2.78% over 1978-1984. However, proponents of the efficient markets hypothesis, where prices fully incorporate all available information, argue that many of the predictable patterns that have been identified in financial markets may be simply due to "chance" as a result of data snooping (Fama; Malkiel). Malkiel states that, "Given enough time and massaging of data series, it is possible to tease almost any pattern out of most data sets" (p. 72).

A fairly blatant form of data snooping in the technical trading literature is an ex post and "in-sample" search for profitable trading rules. More subtle forms of data snooping occur when a set of data is repeatedly used to search for profitable "families" of trading systems, markets, in-sample estimation periods, out-of-sample periods, and trading model assumptions including performance criteria and transaction costs. For example, a researcher may investigate a number of in-sample optimization periods (or methods) on the same dataset and select one that provides the most successful result. Even if a researcher selects a single in-sample period in an ad-hoc fashion, it is likely to be strongly affected by previous research. Moreover, if there are many researchers who choose one in-sample optimization method on the same dataset, then they are collectively snooping the data. Collective data snooping is potentially even more dangerous because it is not easily recognized by each individual researcher. In the presence of such data snooping, conventional statistical tests may mislead researchers with exaggerated significance levels (Denton; Lo and MacKinlay; Sullivan, Timmermann, and White 1999).

Data snooping problems have not been completely ignored in previous studies of technical analysis. In the financial economics literature, a number of authors suggest replicating models used in a previous study on a new set of data as a method to avoid data snooping problems (e.g., Lovell; Lo and MacKinlay; Schwert; Sullivan, Timmermann, and White 2003). This procedure is robust to data snooping problems because all choice variables (e.g., trading systems, in- and out-of-sample periods, and markets) in the original trading model are 'written in stone' and thus 'true' out-of-sample verification can be conducted on new data. To date, however, only two technical trading studies have followed this suggestion. For the stock market, Sullivan, Timmermann, and White (1999) replicated Brock, Lakonishok, and LeBaron's trading rules on the subsequent 10 years (1987-1996) of Dow Jones Industrial Average (DJIA) data, and found that the best of their trading rules generated a statistically insignificant mean return of 8.63% per year before transaction costs. For futures markets, Park and Irwin (2005) confirmed Lukac, Brorsen, and Irwin's (1988) original findings and then replicated the same trading rules on new futures price data over the 1985-2003 period. They found that the earlier successful performance of Lukac, Brorsen, and Irwin's technical trading systems did not persist in the subsequent sample period.

Despite their usefulness, replication studies are by definition limited to the trading systems and markets analyzed in the original study. Timmerman and Granger argue that such a fixed approach is unlikely to uncover profitable models in dynamic markets. Instead, they suggest testing a broad set of models in a large set of markets to uncover "hot spots of forecastability." Examining more trading systems, parameters, and/or contracts, however, may result in data snooping biases unless dependencies across all trading rules tested are taken into account.

White proposes a procedure, termed the Bootstrap Reality Check methodology, which can directly quantify the effect of data snooping by evaluating the performance of the best rule in the context of the full universe of technical trading rules. As mentioned above, conventional hypothesis tests are invalid in the presence of data snooping because they do not consider dependence in performance across all trading rules tested. White's procedure avoids this problem by testing the null hypothesis that the performance of the best model in the full universe of models is no better than the performance of a benchmark model. In his approach, the best rule is obtained by applying a performance statistic to the full set of trading rules, and then a desired p-value can be obtained from comparing the performance of the best trading rule with approximations to the asymptotic distribution of the performance statistic across all the trading rules. Hansen (2005) improves White' procedure. He argues that White's procedure may reduce rejection probabilities of the superior predictive ability test under the null hypothesis by the inclusion of poor and irrelevant alternative models because it does not satisfy a relevant similarity condition that is necessary for a test to be unbiased. In evaluating the performance of technical trading rules, poor-performing trading rules are unavoidably included because there is no theoretical guidance regarding the proper selection of parameters (i.e., individual trading rules). Hansen shows that his new test, termed the Superior Predictive Ability (SPA) test, can improve the power of the test by adopting a studentized test statistic and a data dependent null distribution.

The purpose of this study is to determine whether technical trading rules have been profitable in US futures markets over 1985-2004 after explicitly accounting for the effect of data snooping. To achieve the purpose, this study expands the number of technical trading rules and futures contracts analyzed by Lukac, Brorsen, and Irwin (1988) and Park and Irwin (2005) and tests statistical significance of technical trading profits using White's and Hansen's tests. We construct a universe of technical trading rules with more than 9,000 rules from 14 trading systems. This is so far the largest universe of trading rules in studies of technical analysis on the futures market. In addition to the 12 futures contracts that were investigated by Lukac, Brorsen, and Irwin, five highly traded contracts also are investigated to ensure a more general test of the profitability of technical trading rules. The contracts represent each major group of futures contracts, i.e., grains, meats, metals, energies, softs, currencies, equity indices, and interest rates. Hence, this study improves upon previous studies of technical analysis in futures markets by incorporating more trading rules and contracts and conducting superior statistical tests.

Data and the Trading Model

Daily futures price data for the 1985-2004 period are used to evaluate the performance of technical trading rules. The data set consists of the 12 futures contracts analyzed by Lukac, Brorsen, and Irwin and 5 additional contracts. These actively traded contracts are selected from each major category of futures contracts, i.e., grains (corn, soybeans, and wheat), meats (live cattle and pork bellies), metals (silver and copper), energy (crude oil), equity index (S&P 500), interest rates (treasury-bills and Eurodollar), currencies (pound, mark, and yen), and softs (cocoa, sugar, and lumber). The price data for the 17 futures markets is provided by the Commodity Research Bureau, Inc. Table 1 presents descriptions of each futures contract, including exchange, contract size, value of one tick, daily price limits, and contract months used.

Because Lukac, Brorsen, and Irwin's investigated the 1975-1984 period, this study considers the 1985-2004 period as the full sample period, with the exception of two financials: Deutsche mark (1985-1998) and treasury-bills (1985-1996). The full sample period is divided into two 10-year subperiods, 1985-1994 and 1995-2004. Out-of-sample performance of the best trading rule identified in the 1985-1994 period, therefore, is evaluated as well as in-sample performance in each sub-period and the full sample period.

The basic trading model is similar to that used in Park and Irwin (2005). The trading model typically consists of input data, technical trading systems, performance measures, an optimization method, and other relevant assumptions. As input data, this study uses daily futures price series. To obtain a price series that reflects the most important market characteristics, we use dominant contracts that have the highest open interest (Dale and Workman). The current dominant contract rolls over to the new dominant contract on the second Tuesday of the month preceding its delivery month.

Performance criteria adopted are the mean net return and the Sharpe ratio. Although defining a rate of return may be problematical because there is no initial investment except for a margin deposit in the futures market, recent technical analysis studies about futures markets tend to use the continuously compounded (log) return per unit (Kho; Szakmary and Mathur; Sullivan, Timmermann, and White 1999). Kho argues that, "… it provides a sufficient statistic for testing the profitability of trading rules because there exists a one-to-one correspondence between a daily price change and dollar gains (p. 252)." The continuously compounded daily gross return on a technical trading rule *k* at time *t* can be calculated by:

$$r^{g}_{k,t+1} = [\ln(P_{t+1}) - \ln(P_t)]S_{k,t}, \tag{1}$$

where P_{t+1} and P_t are futures prices at time $t+1$ and *t*, respectively, and $S_{k,t}$ is an indicator variable that takes one of three values: +1 for a long position, 0 for a neutral position (i.e., out of the market), and -1 for a short position.[1] Measuring trading returns on a daily basis is consistent with the process of the daily settlement (marking-to-market) in the futures market. The daily net trading return is then given by:

$$r_{k,t+1} = r^{g}_{k,t+1} + d_{t+1}\left(\frac{n_k}{N^{in}_k}\right)\ln(1-c), \tag{2}$$

where n is the number of round-turn trades for a contract, N^{in} is the number of days "in" the market (e.g., $N^{in} = N - N^{out}$, where N^{out} is the number of days "out" of the market), d_{t+1} is an indicator variable having a value of 1 for in-days and 0 for out-days, and *c* is round-turn proportional transaction costs. As in Lukac, Brorsen, and Irwin, this study assumes transaction costs of \$100 per round-turn trade, which are quite conservative when compared with the sum of the bid-ask spread and commissions estimated in other studies of technical trading rules (see Park and Irwin for more details).

The Sharpe ratio measures the excess return per unit of total risk. In futures markets, because traders can deposit treasury-bills for margin requirement, there is no need to sacrifice the risk-free return in order to participate in an alternative investment. The Sharpe ratio (SR_k) of a trading rule *k* can then be calculated by:

$$SR_k = \bar{r}_k / \hat{\sigma}_k, \tag{3}$$

where $\bar{r}_k$ denotes the annualized mean net return during a sample period and $\hat{\sigma}_k$ denotes the standard deviation. Based on both performance criteria, technical trading rules are optimized. That is, for a given sample period a trading rule showing the highest mean net return or Sharpe ratio among the full set of trading rules is chosen.

It is important to incorporate accurate daily price limits into the trading model because for certain futures contracts price movements are occasionally locked at the daily allowable limits. Trend-following trading rules typically generate buy (sell) signals in up (down) trends. Thus, if

---

[1] P_t may differ depending on the execution price of a trade, e.g., today's closing price, tomorrow's open price, or a daily stop.

a trading signal is triggered on the day a price limit move occurs, then a buy (sell) trade may be executed at a higher (lower) price than that at which the trading signal was triggered. This may result in seriously overstated trading returns if trades are assumed to be executed at the limit 'locked' price levels. For contracts having daily trading limits, therefore, no position is taken or closed out if the high price equals the low price and both of these equal the closing price (lock-limit days), or if the contract's opening price is up or down the daily allowable limit. The history of daily price limits for each contract is obtained from exchange statistical yearbooks and the annual *Reference Guide to Futures/Options Markets* and *Source Book* issues of *Futures* magazine. Other assumptions included in the trading model are: (1) all trading is on a one contract basis, i.e., only one contract is used for each transaction; (2) no pyramiding of positions or reinvestments of profits is allowed; and (3) sufficient funds are assumed available to meet the margin requirement that may occur due to trading losses.

Universe of Technical Trading Rules

"Reality Check" tests such as White's and Hansen's tests evaluate the performance of the best rule in terms of the full universe of technical trading rules. Thus, it is critical to the analysis to construct an appropriate universe of technical trading rules, because it directly influences test results. For example, considering only currently popular trading rules, which are likely to have subtle 'survivorship biases' over a long time period, may bring about spurious results. Sullivan, Timmermann, and White (1999, p. 1649) argued that "If enough trading rules are considered over time, some rules are bound by pure luck, even in a very large sample, to produce superior performance even if they do not genuinely possess predictive power over asset returns. Of course, inference based on the subset of surviving trading rules may be misleading in this context because it does not account for the full set of initial trading rules, most of which are likely to have underperformed."

On the other hand, a potential limitation of a replication study on technical analysis arises because its search space is limited to the original trading systems, which may lack other important trading systems. For example, Lukac, Brorsen, and Irwin's 12 technical trading systems do not include several key systems such as the Relative Strength Index (RSI) and the Moving Average Convergence-Divergence (MACD) systems that have been prominently featured in well-known books on technical analysis. Schwager states that "Of all the momentum oscillators currently in wide use, RSI responds the best to basic technical analysis methods such as trend lines, chart patterns, and support and resistance. Applying these methods to RSI in conjunction with overbought/oversold levels and divergences can provide very valuable insight into market behavior" (p. 542). He further states that the MACD system is "one of the most interesting and dependable technical indicators" (p. 538). Moreover, some of Lukac, Brorsen, and Irwin's 12 trading systems are quite simple and may not respond adequately to a wide variety of possible market situations. These simple trading systems require more parameters, such as a confirmation device (e.g., a price band or time delay) that has long been used by actual traders. In this sense, replication results by Park and Irwin (2005) may not fully represent performance of technical trading rules in US futures markets during recent years.

Lukac, Brorsen, and Irwin's 12 trading systems cover the major groups of technical trading systems such as moving averages, channels, momentum oscillators, filters, and combinations. This study, therefore, approximates the full "universe" of technical trading rules based on their trading systems and three additional important trading systems. A total of 9,385 trading rules are drawn from those trading systems, all of which were available to investors before the beginning of the full sample period. The availability of technical trading systems during the sample period is of particular importance in testing market efficiency. It would be inappropriate to apply relatively new searching techniques such as a genetic algorithm or an artificial neural network to the period before the discovery of such techniques in order to reveal evidence on the profitability of technical trading rules (Sullivan, Timmermann, and White 1999; Cooper and Gulen; Timmermann and Granger).

Among Lukac, Brorsen, and Irwin's 12 technical trading systems, seven trading systems (L-S-O Price Channel, Directional Indicator, Range Quotient, Reference Deviation, Directional Movement, Parabolic Time/Price, and Directional Parabolic) remain the same as in their study, with five trading systems (Outside Price Channel, M-II Price Channel, Simple Moving Average with a Percentage Price Band, Dual Moving Average Crossover, and Alexander's Filter Rule) slightly modified to include an additional parameter (See Lukac, Brorsen, and Irwin (1990) for details of the 12 trading systems). For the Outside Price Channel and the M-II Price Channel systems, a percentage price band is incorporated as a new parameter. According to Schwager (p. 613), such confirmation devices are, "An important modification that can be made to a basic trend-following system" to reduce potential whipsaw losses. If prices frequently move far enough for a trading rule to trigger a trading signal and then reverse direction, then the rule may generate a number of false signals. To avoid such false signals, a variety of confirmation devices such as bands and time delay can be used, although they tend to produce delayed entry points on correct signals and thus reduce trading profits. The inclusion of confirmation devices will not introduce selection biases, because they have long been used for many technical trading systems. Two moving average systems, the Simple Moving Average with a Percentage Price Band and the Dual Moving Average Crossover, are integrated into the Moving Average Crossover system, so that the previous Dual Moving Average Crossover system has an additional parameter of a percentage price band. Alexander's Filter Rule system is also expanded by taking a different filter depending on whether a position is initiated or liquidated. Three additional trading systems are the Exponential Moving Average Crossover, the Moving Average Convergence-Divergence, and the Relative Strength Index (RSI) systems. These three systems are selected because as noted above, they have been prominently featured in well-known books on technical analysis, such as Schwager, Murphy, and Kaufman. These modified trading systems and additional trading systems are fully described in the Appendix. Hence, the full set of technical trading rules is represented by 9,385 trading rules parameterized from 14 technical trading systems.

Statistical Testing Procedure

In the literature on technical trading rules, researchers generally search for profitable trading rules by applying a large number of trading rules to past price and volume data. As a result, some of the trading rules may work by chance rather than their inherent forecasting ability. White describes the data snooping problem as follows:

> Whenever a 'good' forecasting model is obtained by an extensive specification search, there is always the danger that the observed good performance results not from actual forecasting ability, but is instead just luck. Even when no exploitable forecasting relation exists, looking long enough and hard enough at a given set of data will often reveal one or more forecasting models that look good, but are in fact useless. (p. 1097)

Moreover, such data snooping practices inevitably overstate significance levels of conventional hypothesis tests because they do not take account of the dependence between performance across all trading rules tested, therefore misleading researchers.

Based on this idea, White proposed the Reality Check methodology to account for data snooping problems. Unlike conventional tests, White's approach provides a comprehensive test across all trading rules considered. More specifically, White's procedure directly quantifies the effect of data snooping by testing the null hypothesis that the performance of the best trading rule is no better than the performance of the benchmark. The best rule is searched by applying a performance statistic to the full set of trading rules, and then a desired *p*-value is obtained from comparing the performance of the best rule to approximations to the asymptotic distribution of the performance statistic, which is derived from the performance of the full set of trading rules.

According to Hansen (2005), however, White's test may reduce rejection probabilities of the test under the null by the inclusion of poor and irrelevant alternative models (in our case, trading rules) because it does not satisfy a relevant similarity condition that is necessary for a test to be unbiased. In a study of technical trading rules, poor-performing trading rules are inevitably included because there is no theoretical guidance regarding the proper selection of parameters (i.e., trading rules). Hansen shows that his new test can improve the power of the superior predictive ability test by adopting a studentized test statistic and a data dependent null distribution and thus reducing the influence of poor-performing models. Hence, in this study statistical tests on significance of technical trading returns are conducted by employing White's and Hansen's procedures, both of which are implemented by utilizing the stationary bootstrap. Previous studies that applied the standard bootstrap or recursive bootstrap methods construct bootstrap samples by resampling raw prices and then applying a trading rule to the resampled price series, whereas White's and Hansen's bootstrap procedures allow researchers to obtain bootstrap samples by directly resampling observations of a test statistic (e.g., differences between returns of a technical trading rule and returns of a benchmark rule). White's and Hansen's procedures are outlined next.

White's Bootstrap Reality Check Procedure

Basic Framework

White's testing procedure is based on the following $m \times 1$ performance measure: for a one-step ahead forecasting horizon,

$$\overline{\mathbf{Y}} = N^{-1} \sum_{t=R}^{T} \hat{Y}_{k,t+1}, \qquad (4)$$

where $k = 1, \ldots, m,$ m is the number of technical trading rules, N is the number of prediction periods indexed from R to $T,$ so that $T = R + N - 1,$ and $\hat{\Upsilon}_{k,t+1} = Y(\Lambda_{k,t}, \hat{\theta}_{k,t})$ is the observed performance measure at $t+1.$ In general, $\mathbf{\Lambda}$ consists of a vector of dependent variables and $\hat{\boldsymbol{\theta}}_t$ is an estimate of an unknown parameter vector $\boldsymbol{\theta}^* \equiv p\lim \hat{\boldsymbol{\theta}}_T$ that relies on data from period t and earlier.[2] In our case, however, parameters (i.e., technical trading rules) are not estimated but pre-specified. Each pre-specified parameter or set of parameters represents individual technical trading rules $(\theta_k, k = 1, \ldots, m)$, and this study considers a universe of 9,385 technical trading rules that result from parameterizations of 14 trading systems.

To measure the performance of trading rules, two performance statistics are used: the mean net return and the Sharpe ratio. First, to evaluate whether technical trading rules (each rule is indexed by a subscript "k") generate a mean net return superior to that of a benchmark strategy (indexed by a subscript "0"), the following form of a performance measure is constructed:

$$\mathbf{Y}_{k,t+1} = \left\{ [\ln(1 + X_{t+1}) S_k(\boldsymbol{\psi}_t, \theta_k)] + d_{k,t+1} \left(\frac{n_k \ln(1-c)}{N_k^{in}} \right) \right\}$$
$$- \left\{ [\ln(1 + X_{t+1}) S_0(\boldsymbol{\psi}_t, \theta_0)] + d_{0,t+1} \left(\frac{n_0 \ln(1-c)}{N_0^{in}} \right) \right\}, \qquad (5)$$

where $X_{t+1} = (P_{t+1} - P_t) / P_t,$

$P_t =$ original prices of a futures contract,

$c =$ round-turn proportional transaction costs,

$\boldsymbol{\psi}_t = \{P_i\}_{i=1}^t,$ a sequence of past futures prices,

$n_0, n_k =$ the number of round-turn trades during a sample period,

$N_0^{in}, N_k^{in} =$ the number of days 'in' the market (e.g., $N_k^{in} = N - N_k^{out},$ where N_k^{out} is the number of days 'out' of the market),

$d_0, d_k =$ indicator variables having a value of 1 for in-days and 0 for out-days,

$S_0, S_k =$ functions that indicate trading signals generated by simulating trading rules $(\theta_k$ and $\theta_0)$ on past prices $(\boldsymbol{\psi}_t)$. The trading signals have one of three values: +1 for a long position, 0 for neutral, and –1 for short.

The second term on the right hand side in equation (5) becomes a zero vector because this study considers "zero mean profits" as the benchmark strategy. Then, we can define the null hypothesis as follows: provided that $\mu_k \equiv E(Y_k)$ is well-defined,

$$H_0 : \max_{k=1,\ldots,m} \mu_k \leq 0, \qquad (6)$$

which states that the performance of the best technical trading rule is no better than the performance of the benchmark strategy in terms of the mean net return. Of course, the alternative hypothesis $H_1 : \max_{k=1,\ldots,m} \mu_k > 0$ suggests that the best trading rule performs better

[2] Notation and usage are similar to those in White and Sullivan, Timmermann, and White (1999).

than the benchmark. Note that the null hypothesis is a multiple hypothesis that represents the intersection of the one-sided individual hypotheses $\mu_k \le 0,\ \ k = 1,\ldots,m.$

Similarly, for the Sharpe ratio criterion the following null hypothesis is tested:

$$H_0 : \max_{k=1,\ldots,m} \{f(E(g_k))\} \le f(E(g_0)), \tag{7}$$

where g is a 2×1 vector with two components given by:

$$g^1_{k,t+1} = [\ln(1+X_{t+1})S_k(\boldsymbol{\psi}_t,\theta_k)] + d_k\left(\frac{n_k \ln(1-c)}{N_k^{in}}\right), \tag{8}$$

$$g^2_{k,t+1} = \left\{[\ln(1+X_{t+1})S_k(\boldsymbol{\psi}_t,\theta_k)] + d_k\left(\frac{n_k \ln(1-c)}{N_k^{in}}\right)\right\}^2, \tag{9}$$

and the form of $f(\cdot)$ can be expressed as:

$$f(E(g_{k,t+1})) = \frac{E(g^1_{k,t+1})}{\sqrt{E(g^2_{k,t+1}) - [E(g^1_{k,t+1})]^2}}. \tag{10}$$

The expectations are assessed with arithmetic averages. Performance statistics are then given by:

$$\bar{Y}_k = f(\bar{g}_k) - f(\bar{g}_0), \tag{11}$$

where $\bar{g}_k$ and $\bar{g}_0$ are averages calculated over a sample period for the *k*th technical trading rule and the benchmark, respectively. For example, $\bar{g}_k$ is given by:

$$\bar{g}_k = N^{-1}\sum_{t=R}^{T} g_{k,t+1}, \qquad k = 0,\ldots,m. \tag{12}$$

As a first step to test the null hypothesis, $H_0 : \max_{k=1,\ldots,m} \mu_k \le 0,$ White applies West's Theorem 4.1 (a) that provides the asymptotic distribution of $\bar{\mathbf{Y}}$: if $\lim_{T\to\infty}(N/R) = 0$ or $\mathbf{F} \equiv E[(\partial \mathbf{Y}_t/\partial\boldsymbol{\theta})(\boldsymbol{\Lambda},\boldsymbol{\theta}^*)] = \mathbf{0}$,[3] then

$$N^{1/2}(\bar{\mathbf{Y}} - \boldsymbol{\mu}) \xrightarrow{d} N(\mathbf{0},\boldsymbol{\Omega}), \tag{13}$$

where $\xrightarrow{d}$ indicates convergence in distribution as $T\to\infty,$ and a variance-covariance matrix $\boldsymbol{\Omega} = \lim_{T\to\infty} \text{var}[N^{-1/2}\sum_{t=R}^{T}\mathbf{Y}(\boldsymbol{\Lambda}_{t+1},\boldsymbol{\theta}^*)]$. This theorem states that the random vector $N^{1/2}(\bar{\mathbf{Y}} - \boldsymbol{\mu})$ has an asymptotic multivariate normal distribution with a mean vector $\mathbf{0}$ and a variance-covariance matrix $\boldsymbol{\Omega}$. Note that the null hypothesis tested is a composite hypothesis whose asymptotic distribution typically depends on nuisance parameters and thus is not uniquely known. This result leads White to employ the least favorable configuration (LFC), the points of the null least favorable to the alternative. In our application, the points under the LFC are $\mu_k = 0$ for all *k*, where $k = 1,\ldots,m,$ which assumes that all alternative models perform as well as the benchmark. Then, the asymptotic behavior of the test statistic, which is given by:

$$T_m^{RC} \equiv \max_{k=1,\ldots,m} N^{1/2}\bar{Y}_k, \tag{14}$$

[3] These conditions allow asymptotic analysis (For more details, see West (p. 1073)).

is known, and an asymptotic p-value for the test of the null hypothesis can be obtained based on the value of T_m^{RC}. White calls any method for obtaining such a p-value a "Reality Check."

There are two possible methods to obtain the unknown distribution of the test statistic T_m^{RC}: Monte Carlo simulation and the bootstrap. An approach based on the bootstrap that uses a resampled version of $\overline{\mathbf{Y}}$ delivers the "Bootstrap Reality Check" p-value for testing the null hypothesis. A resampling procedure is directly related to selecting random indexes $\eta(t)$ for $t = R,\ldots,T$. Once the random indexes have been selected, the resampled performance statistic can be computed as

$$\overline{\mathbf{Y}}^* = N^{-1}\sum_{t=R}^{T}\hat{\mathbf{Y}}_{t+1}^*, \qquad \hat{\mathbf{Y}}_{t+1}^* \equiv \mathbf{Y}(\mathbf{\Lambda}_{\eta(t)+1}, \hat{\mathbf{\theta}}_{\eta(t)}), \qquad t = R,\ldots,T. \tag{15}$$

To generate the bootstrap resamples, White employs Politis and Romano's stationary bootstrap that can be applicable to a stationary and weakly dependent time series. The stationary bootstrap resamples blocks of random length from the original data, where the block length follows the geometric distribution, with mean block length b. White (p. 1104) applies the following resampling algorithm of the stationary bootstrap to obtain the random indexes $\eta(t)$:

(1) Start by selecting a smoothing parameter $q = 1/b = q_N$, $0 < q_N \le 1$, $q_N \to 0$, $Nq_N \to \infty$ as $N \to \infty$.
(2) Set $t = R$. Draw $\eta(R)$ at random, independently and uniformly from $\{R,\ldots,T\}$.
(3) Increment t. If $t > T$, stop. Otherwise, draw a standard uniform random variable U (supported on [0,1]) independently of all other random variables.
    A. If $U < q$, draw $\eta(t)$ at random, independently and uniformly from $\{R,\ldots,T\}$.
    B. If $U \ge q$, set $\eta(t) = \eta(t-1)+1$; if $\eta(t) > T$, reset to $\eta(t) = R$.
(4) Repeat (3).

This method ensures that the block lengths are randomly distributed according to the geometric distribution with mean block length $b = 1/q$.

When θ^* is known, as noted by Diebold and Mariano (1995), Politis and Romano's (1994a) Theorem 2 can be immediately applied to establish that under appropriate conditions,[4] the distribution of $N^{1/2}(\overline{\mathbf{Y}}^* - \overline{\mathbf{Y}})$ conditional on $\{\mathbf{\Lambda}_{R+1},\ldots,\mathbf{\Lambda}_{T+1}\}$ converges to the distribution of $N^{1/2}(\overline{\mathbf{Y}} - \boldsymbol{\mu})$ as N increases. Thus, we can construct an estimate of the desired distribution of $N(\mathbf{0},\mathbf{\Omega})$ by repeatedly drawing realizations of $N^{1/2}(\overline{\mathbf{Y}}^* - \overline{\mathbf{Y}})$, and obtain the Bootstrap Reality Check p-value for testing $H_0 : \max_{k=1,\ldots,m} \mu_k \le 0$ by comparing T_m^{RC} to the quantiles of

$$T_m^{RC*} \equiv \max_{k=1,\ldots,m} N^{1/2}(\overline{Y}_k^* - \overline{Y}_k). \tag{16}$$

White shows that when $\hat{\mathbf{\theta}}_{\eta(t)}$ appears in $\overline{\mathbf{Y}}^*$, careful argument under additional conditions yields the same conclusion (p. 1104).

[4] This implies that the conditions of Politis and Romano's (1994a) Theorem 2 hold for each element of $\mathbf{Y}_t^*$.

The above procedure can also be extended to the Sharpe ratio. Relevant bootstrapped values of the Sharpe ratio are given by:

$$\bar{Y}_k^* = f(\bar{g}_k^*) - f(\bar{g}_0^*), \;\; k = 0,\ldots,m, \tag{17}$$

where $\bar{g}_k^* = N^{-1}\sum_{t=R}^{T} \hat{g}_{k,t+1}^*$ and $\hat{g}_{k,t+1}^* \equiv g_k(\mathbf{\Lambda}_{\eta(t)+1}, \hat{\mathbf{\theta}}_{\eta(t)})$, $t = R,\ldots,T$. We can then obtain the Bootstrap Reality Check *p*-value for the Sharpe ratio criterion in a similar manner to that in the case of the mean net return criterion (for more details, see White).

Bootstrap Implementation

Given *N* prediction observations and the set of 9,385 technical trading rules, implementation of the Bootstrap Reality Check for in-sample periods begins with determining the smoothing parameter, $q = q_N$, and the number of resamples, *B*. In practice, it is inevitable that the smoothing parameter *q* is chosen arbitrarily based on the data used. The parameter *q* is inversely related to the block length. A larger value of *q*, therefore, is relevant for data with little dependence and a smaller value of *q* for data with more dependence. As a special case, $q = 1$ implies independent bootstrap resampling. Sullivan, Timmermann, and White (1999), however, found that *p*-values for their several combinations of different samples and performance criteria were insensitive to the choice of the smoothing parameter *q*. This study sets $q = 0.1$, which is the same value of *q* as in Sullivan, Timmermann, and White. The value of the smoothing parameter gives a mean block length of 10. The number of bootstrap samples *B* should be a sufficiently large number, because it may significantly influence the accuracy of *p*-values estimated. Brock, Lakonishock, and LeBaron and Kho, however, demonstrated that their estimated bootstrap *p*-values were insensitive to the replication size, *B*, once it was extended beyond 500. Sullivan, Timmermann, and White (1999) also simulated 500 replications. This study, therefore, sets $B = 500$. Next, the stationary bootstrap is applied to generate *B* sets of random observation indexes of length *N*, $\{\eta_i(t), t = R,\ldots,T\}$, $i = 1,\ldots,B$. These indexes are drawn once and for all at the beginning of the analysis.

The specification search and statistical inference can be done in a recursive manner. In case of the mean net return criterion, first, we compute daily net returns for the benchmark strategy from the second term in equation (2) and then compute those for the first technical trading rule from the first term. With the return values, we obtain $Y_{1,t+1}$ and $\bar{Y}_1 = N^{-1}\sum_{t=R}^{T} Y_{1,t+1}$. Using the random indexes generated by the stationary bootstrap, we can also construct $\bar{Y}_{1,i}^* = N^{-1}\sum_{t=R}^{T} Y_{1,\eta_i(t)+1}^*$, $i = 1,\ldots,B$. Now, set $T_1^{RC} = N^{1/2}\bar{Y}_1$ and $T_{1,i}^{RC*} = N^{1/2}(\bar{Y}_{1,i}^* - \bar{Y}_1)$, $i = 1,\ldots,B$. The *p*-value for the null hypothesis that the first technical trading rule has no predictive superiority over the benchmark strategy is obtained by comparing the value of T_1^{RC} to the percentiles of $T_{1,i}^{RC*}$. To do this, we form the order statistics $T_{1,(1)}^{RC*}, T_{1,(2)}^{RC*},\ldots,T_{1,(B)}^{RC*}$ by sorting the values of $T_{1,i}^{RC*}$ from smallest to largest, and then determine *L* such that $T_{1,(L)}^{RC*} \le T_1^{RC} \le T_{1,(L+1)}^{RC*}$. The *p*-value (p_1) for the null hypothesis that the first trading rule performed no better than the benchmark is given by:

$$p_1 = 1 - (L/B). \tag{18}$$

This procedure alone provides a "nominal" p-value for the first rule.

We proceed further to obtain the Bootstrap Reality Check p-value. For the second trading rule, we compute the daily net returns and form $Y_{2,t+1}$, $\overline{Y}_2 = N^{-1}\sum_{t=R}^{T} Y_{2,t+1}$, and $\overline{Y}_{2,i}^* = N^{-1}\sum_{t=R}^{T} Y_{2,\eta_i(t)+1}^*$, $i = 1,\ldots,B$. Then, we set

$$T_2^{RC} = \max\{N^{1/2}\overline{Y}_2, T_1^{RC}\}, \text{ and}$$

$$T_{2,i}^{RC*} = \max\{N^{1/2}(\overline{Y}_{2,i}^* - \overline{Y}_2), T_{1,i}^{RC*}\}, \; i = 1,\ldots,B.$$

To test whether the better of the two trading rules outperforms the benchmark, we compare the sample value of T_2^{RC} with the percentiles of $T_{2,i}^{RC*}$. By proceeding recursively in this manner for $k = 3,\ldots,m,$ we can test whether the best of the k technical trading rules so far outperforms the benchmark. The test can be done by comparing the sample value of

$$T_k^{RC} = \max\{N^{1/2}\overline{Y}_k, T_{k-1}^{RC}\}$$

to the quantiles of

$$T_{k,i}^{RC*} = \max\{N^{1/2}(\overline{Y}_{k,i}^* - \overline{Y}_k), T_{k-1,i}^{RC*}\}, \qquad i = 1,\ldots,B.$$

Specifically, we form the order statistics $T_{m,(1)}^{RC*}, T_{m,(2)}^{RC*}, \ldots, T_{m,(B)}^{RC*}$ by sorting the values of $T_{m,i}^{RC*}$ in ascending order, and then find L such that $T_{m,(L)}^{RC*} \le T_m^{RC} \le T_{m,(L+1)}^{RC*}$. The Bootstrap Reality Check p-value (p_{RC}) for the null hypothesis that the best trading rule performs no better than the benchmark is then given by:

$$p_{RC} = 1 - (L/B). \tag{19}$$

The Bootstrap Reality Check p-value for the Sharpe ratio performance criterion can also be obtained from implementing a similar procedure to that described above.

Hansen's Superior Predictive Ability (SPA) Test

Basic Framework

Based on the same framework as that of White, Hansen (2005) proposes another testing procedure, named the Superior Predictive Ability (SPA) Test. Hansen (p. 11) shows that the power of the Bootstrap Reality Check test may be reduced by deliberately adding poor-performing alternatives to the full set of models. To avoid the problem, Hansen makes two modifications to White's procedure. The first is to studentize a test statistic and the second is to construct a data-dependent null distribution. The studentized test statistic is given by:

$$T_m^{SPA} \equiv \max\left\{\left(\max_{k=1,\ldots,m} \frac{N^{1/2}\overline{Y}_k}{\hat{\sigma}_k}\right), 0\right\}, \tag{20}$$

where $\hat{\sigma}_k^2$ is a consistent estimator of $\sigma_k^2 \equiv \mathrm{var}(N^{1/2}\overline{Y}_k)$. T_m^{SPA}, therefore, delivers the largest t-statistic of relative performance. The studentization of the individual statistics typically improves the power of the test by avoiding a comparison between performance of alternative models that are measured in different units of standard deviation.

In his early study, Hansen (2003) showed that LFC-based tests such as the Reality Check test do not satisfy an asymptotic similarity condition that is necessary for a test to be unbiased, because $\boldsymbol{\mu} \equiv E(\mathbf{Y})$ is on the boundary of the null hypothesis. As a result, these tests are sensitive to the inclusion of poor and irrelevant forecasting models. A proper test should reduce the influence of the poor alternatives while preserving the influence of the alternatives with $\mu_k = 0,$ so that Hansen (2005) proposes the following estimator, $\hat{\mu}_k^c$:

$$\hat{\mu}_k^c = \overline{Y}_k 1_{\{N^{1/2}\overline{Y}_k / \hat{\sigma}_k \geq -\sqrt{2\log\log N}\}}, \qquad k = 1,\ldots,m, \tag{21}$$

where $1_{\{\cdot\}}$ is an indicator function that takes the value one if the expression in $\{\cdot\}$ is true and the value zero otherwise. The threshold rate, $\sqrt{2\log\log N},$ used to calculate $\hat{\boldsymbol{\mu}}^c$ is the slowest rate that captures all alternatives with $\mu_k = 0.$ Because different threshold rates generate different *p*-values in finite samples, Hansen introduces two additional estimators, $\hat{\mu}_k^l \equiv \min(\overline{Y}_k, 0)$ and $\hat{\mu}_k^u = 0,$ $k = 1,\ldots,m,$ which provide a lower bound and an upper bound of the *p*-value, respectively. Of course, $\hat{\boldsymbol{\mu}}^u$ will lead to the LFC-based test and thus $\hat{\boldsymbol{\mu}}^l \leq \hat{\boldsymbol{\mu}}^c \leq \hat{\boldsymbol{\mu}}^u$. Hansen further shows that $\hat{\boldsymbol{\mu}}^c$ is a consistent estimate of the asymptotic distribution of the test statistic and that the consistent estimate of the null distribution can improve the power of the test (See Hansen (2005, p. 41) for more details).

Bootstrap Implementation

A bootstrap implementation of the SPA tests is similar to that of White. Hansen also utilizes the stationary bootstrap of Politis and Romano, in which the pseudo time-series, $\{\mathbf{Y}_{b,t}^*\} \equiv \{\mathbf{Y}_{\eta_{b,t}}\},$ $b = 1,\ldots,B,$ are bootstrap resamples of $\mathbf{Y}_t,$ where $\{\eta_{b,1},\ldots,\eta_{b,N}\}$ is constructed by combining blocks of $\{1,\ldots,N\}$ with random lengths. As described in earlier sections, the block length has the geometric distribution with parameter $q \in (0,1].$ Using the pseudo time-series, the bootstrap sample averages $\overline{\mathbf{Y}}_b^* \equiv N^{-1}\sum_{t=1}^N \mathbf{Y}_{b,t}^*,$ $b = 1,\ldots,B,$ can be easily calculated. In this study, we use the same values of *B* and *q* as those applied to White's Reality Check bootstrap procedure. That is, $B = 500$ and $q = 0.1.$

The test statistic, $T_m^{SPA} \equiv \max\{(\max_{k=1,\ldots,m} N^{1/2}\overline{Y}_k / \hat{\sigma}_k), 0\},$ requires estimates $\sigma_k^2 \equiv \mathrm{var}(N^{1/2}\overline{Y}_k),$ $k = 1,\ldots,m.$ In estimating $\sigma_k^2,$ Hansen recommends using the bootstrap-population value directly, which is given by Politis and Romano's (1994a) Lemma 1. That is,

$$\hat{\sigma}_k^2 \equiv \hat{\gamma}_{0,k} + 2\sum_{i=1}^{N-1} b_N(i)\hat{\gamma}_{i,k}, \tag{22}$$

where $\hat{\gamma}_{i,k} \equiv N^{-1}\sum_{j=1}^{N-i}[(Y_{k,j} - \overline{Y}_k)(Y_{k,j+i} - \overline{Y}_k)],$ $i = 0,1,\ldots,N-1,$ are the empirical covariances, and the kernel weights under the stationary bootstrap are given by:

$$b_N(i) = \left(1 - \frac{i}{N}\right)(1-q)^i + \frac{i}{N}(1-q)^{N-i}. \qquad (23)$$

Define

$$Z^*_{k,b,t} \equiv d^*_{k,b,t} - g_i(\bar{d}_k), \quad i = l, c, u, \quad b = 1, \ldots, B, \quad t = 1, \ldots, N,$$

where $g_l(x) = \max(0, x)$, $g_c(x) = x \cdot 1_{\{x \geq -\sqrt{(\hat{\sigma}_k^2 / N) 2 \log \log N}\}}$, and $g_u(x) = x$. The distribution of the test statistic under the null hypothesis can then be approximated by the empirical distribution obtained from the bootstrap resamples $\mathbf{Z}^*_{b,t}$, $t = 1, \ldots, N$.

By calculating $T^{SPA*}_{b,m} \equiv \max\{(\max_{k=1,\ldots,m} N^{1/2} \bar{Z}^*_{k,b} / \hat{\sigma}_k), 0\}$ for $b = 1, \ldots, B$, we can obtain the *p*-values of the three tests for SPA, which are given by:

$$p_{SPA} \equiv \sum_{b=1}^{B} \frac{1_{\{T^{SPA*}_{b,m} > T^{SPA}_m\}}}{B}, \qquad (24)$$

where $1_{\{\cdot\}}$ denotes an indicator function. The null hypothesis is rejected at the $\alpha\%$ significance level if $p_{SPA} \leq (\alpha / 100)$.

Empirical Results

Results for the Mean Net Return Criterion

Tables 2-4 report the performance results of the best technical trading rule under the mean net return criterion for each of the sample periods. Each table includes White's Reality Check bootstrap *p*-value (p_{RC}), Hansen's Lower, Consistent, and Upper SPA bootstrap *p*-values (p_l, p_c, and p_u), and White's nominal *p*-value $(p_{N,W})$ and Hansen's nominal *p*-value $(p_{N,H})$. As discussed above, the Consistent SPA *p*-value (p_c) is consistent for the true *p*-value, whereas the Lower SPA *p*-value (p_l) and the Upper SPA *p*-value (p_l) provide upper and lower bounds for the true *p*-value. White's and Hansen's nominal *p*-values are obtained by applying each bootstrap procedure to the best trading rule, thereby ignoring the effect of data snooping.

Table 2 shows that over 1985-1994 the best technical trading rules identified under the mean net return criterion generate positive annual mean net returns for each of 17 contracts. The mean net returns range from a low of 1.06% for treasury-bills to a high of 20.18% for crude oil and are statistically significant in all but corn and soybeans at the 10% level when the effect of data snooping is ignored. Even at the 5% significance level, the statistical significance holds for 12 of the 17 contracts as White's nominal *p*-values indicate. Hansen's nominal *p*-values provide similar results, showing statistical significance for 13 contracts at the 10% level. In contrast, White's Reality Check bootstrap *p*-values that account for data snooping biases indicate that technical trading returns are statistically significant only in two of the 17 contracts: Eurodollar (*p*-value of 0.03) and the yen (*p*-value of 0.10). For the rest of contracts, the null hypothesis that the mean net return of the best technical trading rule is not greater than that of the benchmark (in our case, mean zero profits) cannot be rejected at the 10% significance level, although financial contracts have much lower *p*-values than commodity contracts. These results are consistent with the findings of Silber, who documented that moving average rules generated substantial profits for several major currency futures (e.g., the mark, yen, and Swiss franc) and Eurodollar futures but negative profits for commodity futures such as silver and gold over 1980-1991. Hansen's Consistent SPA *p*-values also indicate that the best rule generates a statistically significant return only for Eurodollar (*p*-value of 0.02). Even considering his Lower SPA *p*-values that result from removing the effects of poor-performing trading rules does not alter the result. These results suggest that the conventional statistical test, which does not account for the dependence in performance across technical trading rules, may mislead researchers in interpreting the predictability of technical trading rules. An interesting thing to note is that Hansen's tests do not necessarily produce lower *p*-values than those of White's tests. Consistent SPA *p*-values are equal to or higher than Reality Check *p*-values in 7 of the 17 contracts.

Out-of-sample performance of the best rules chosen over 1985-1994 is disappointing. During the out-of-sample period 1995-2004, as shown in table 2, the best rules generate positive mean net returns for six contracts, with statistically significant returns only for silver and Eurodollar. For the rest of 11 contracts, annual mean net returns are negative. Moreover, the size of positive returns is substantially reduced compared with that of in-sample returns in most cases (e.g., 10.46% to 2.04% for the mark; 10.90% to 4.33% for the yen; 1.76% to 0.53% for Eurodollar; and 14.15% to 1.49% for the S&P 500 Index). To investigate whether the poor out-of-sample results were caused by higher transaction costs of $100 per round turn, we also apply lower transaction costs of $50 to the same best rules.[5] Such lower transaction costs may be possible because commissions through discount brokers are around $12.50 per round turn (Lukac, Brorsen, and Irwin; Lukac and Brorsen), and even lower for both high volume traders and electronic trades introduced in the early 1990s. Results show that annual mean net returns are still negative for nine of 17 contracts and are statistically significant only for three contracts (silver, the yen, and Eurodollar), although they increased by 0.05%-6.15% across all markets. Technical trading rules, therefore, generally fail to produce economically and statistically significant profits in out-of-sample tests.

Table 3 shows in-sample performance of the best technical trading rules over 1995-2004. During the sample period, annual mean net returns of the best technical rules are positive across all contracts and statistically significant for all but two contracts (sugar #11 and the pound) in terms of the nominal *p*-value. Annual mean net returns range from 0.89% for treasury-bills to 13.69% for crude oil. When the effect of data snooping is considered, however, technical trading returns are statistically significant only for Eurodollar, with a Reality Check *p*-value of 0.03 and a Consistent SPA *p*-value of 0.10.

Table 4 shows in-sample performance of the best technical trading rules over the full sample period, 1985-2004. As expected by the previous two tables, when the effect of data snooping is not considered, annual mean net returns of the best rules are all positive and statistically significant for all but two contracts (cocoa and sugar #11). White's Reality Check tests and Hansen's SPA tests, however, reveal that technical trading rule profits are statistically significant

[5] Neely, Weller, and Dittmar showed that applying higher transaction costs to in-sample periods and lower transaction costs to out-of-sample periods may reduce the problem of overfitting in-sample associated with high trading frequency.

only for the yen and Eurodollar when the effects of data snooping is properly accounted for. The *p*-values from both tests indicate that an annual mean net return of 1.21% for Eurodollar is statistically significant at the 1% level and the annual mean net return of 9.10% for the yen is statistically significant at the 5% level in White's test and at the 10% level in Hansen's test. In general, data-snooping-adjusted *p*-values are much lower in financial contracts than in commodity contracts, suggesting that technical trading rules have performed better in financial futures markets than in commodity futures markets.

Results for the Sharp Ratio Criterion

Tables 5-7 report the performance results of the best technical trading rules under the Sharpe ratio criterion for each sample period. As the Sharpe ratio measures the mean excess net return per unit of total risk (standard deviation), technical trading rules are now formally optimized based on the performance after adjustment for transaction costs and risk. In addition, the Sharpe ratio makes it possible to compare risk-adjusted returns from different contracts by standardizing returns of each contract by their standard deviations.

Over 1985-1994, in-sample performance of the best rules under the Sharpe ratio criterion is similar to that under the mean net return criterion, as shown in table 5. Sharpe ratios of the best rules range from 0.46 (corn) to 1.32 (Eurodollar), and are generally higher in financial contracts (0.57 to 1.32) than in commodity contracts (0.46 to 0.70). For each contract, Sharpe ratios are statistically significant as far as nominal *p*-values are concerned. In sharp contrast, the Reality Check tests and the SPA tests indicate that the Sharpe ratios are statistically insignificant for all but Eurodollar when the effect of data snooping is considered. For Eurodollar, the best rule from the Directional Parabolic system generates a Reality Check *p*-value of 0.05 and a Consistent SPA *p*-value of 0.02. Previously, the best rule under the mean net return criterion generated a statistically significant Reality Check *p*-value of 0.10 for the yen, whereas the best rule under the Sharp ratio criterion generates a Reality Check *p*-value of 0.32. Therefore, the null hypothesis that the Sharpe ratio of the best technical trading rule is not greater than zero cannot be rejected for all but Eurodollar. Note that for 13 of the 17 contracts the best rules differ from those identified under the mean net return criterion, although 10 contracts have the best rules from the same trading systems for both criteria. Another interesting thing to note is that Hansen's SPA *p*-values under the Sharpe ratio criterion are identical to those under the mean net return criterion in each of the sample periods even if the best trading rules chosen under each performance criterion differ. The reason is that Hansen's test statistic is standardized by its standard deviation.

Out-of-sample results over 1995-2004 are also similar to those for the mean net return criterion. Only for Eurodollar among the 17 contracts, the best rule produces a statistically significant Sharpe ratio of 1.23. Even with lower transaction costs of $50, the best rules produce statistically significant Sharpe ratios only for Eurodollar (1.30) and the yen (0.40).

Table 6 presents in-sample performance of the best rules identified under the Sharpe ratio criterion over 1995-2004. Sharpe ratios range from 0.42 (cocoa) to 1.81 (treasury-bills) and are statistically significant for all but the pound in terms of the nominal *p*-value. None of Reality Check *p*-values, however, are significant, and Consistent SPA *p*-values are significant only for

Eurodollar, with a *p*-value of 0.10. These results suggest that technical trading rules fail to produce statistically significant performance over 1995-2004, when the effect of data snooping is considered. In fact, during the sample period, the best rules identified under the Sharp ratio criterion are identical to those under the mean net return criterion in 10 contracts.

Table 7 provides the performance of the best technical trading rules over the full sample period, 1985-2004. During this sample period, the best rules generate Sharpe ratios ranging from 0.29 (corn) to 1.21 (Eurodollar), with significant nominal *p*-values. Sharpe ratios for financial contracts such as major currencies and short-term interest rates are higher (0.41-1.21) than those (0.29-0.54) of commodity contracts. The best trading rules identified under the Sharpe ratio criterion, however, generate statistically significant Reality Check *p*-values and Consistent SPA *p*-values only for Eurodollar and the yen. The Reality Check *p*-values are zero for Eurodollar and 0.10 for the yen, whereas the Consistent SPA *p*-values are zero for Eurodollar and 0.09 for the yen. Hence, technical trading rules do not generate successful results when transaction costs, risk, and the effect of data snooping are taken into account.

Finally, table 8 shows the performance of individual technical trading systems measured by the number of the best trading rules they produce. Overall, the Moving Average Crossover (MAC), the Relative Strength Index (RSI), the Directional Indicator (DRI), Alexander's Filter Rule (ALX), and the Directional Parabolic (DRP) systems performed the best across the performance criteria and sample periods. In particular, for each sample period the RSI system generates the best rule across various contracts, and the DRP system generates statistically significant profits for Eurodollar even after adjustment for transaction costs, risk, and data snooping biases. This result implies that future studies on technical analysis should include the RSI and the DRP systems in analysis as well as the popular MAC and ALX systems.

Summary and Conclusions

Proponents of the efficient markets hypothesis argue that various market anomalies, such as technical trading profits, found in financial markets may be simply due to "chance" as a result of data snooping (Fama; Malkiel). In the technical trading literature, a fairly blatant form of data snooping is an ex post and "in-sample" search for profitable trading rules. More subtle forms of data snooping occur when a set of data is repeatedly used to search for profitable "families" of trading systems, markets, in-sample estimation periods, out-of-sample periods, and trading model assumptions including performance criteria and transaction costs. In the presence of such data snooping, conventional statistical tests may mislead researchers with exaggerated significance levels (Denton; Lo and MacKinlay; Sullivan, Timmermann, and White 1999).

Data snooping problems have not been completely ignored in previous studies of technical analysis. In the financial economics literature, a number of authors have suggested replicating models used in a previous study on a new set of data as a method to avoid data snooping problems (e.g., Lovell; Lo and MacKinlay; Schwert; Sullivan, Timmermann, and White 2003). This procedure is robust to data snooping problems because all choice variables (e.g., trading systems, in- and out-of-sample periods, and markets) in the original trading model are 'written in stone' and thus 'true' out-of-sample verification can be conducted on new data. Despite their usefulness, however, replication studies are by definition limited to the trading systems and markets analyzed in the original study. Timmerman and Granger argue that such a fixed approach is unlikely to uncover profitable models in dynamic markets. Instead, they suggest testing a broad set of models in a large set of markets to uncover "hot spots of forecastability." Examining more trading systems, parameters, and contracts, however, may result in data snooping biases unless dependencies across all trading rules tested are taken into account.

The purpose of this study was to determine whether technical trading rules have been profitable in US futures markets for the last 20 years after explicitly accounting for the effect of data snooping. To achieve the purpose, this study expanded the number of technical trading rules and futures contracts analyzed in Lukac, Brorsen and Irwin (1988) and Park and Irwin (2005) and tested statistical significance of technical trading profits using White' Bootstrap Reality Check methodology and Hansen's Superior Predictive Ability (SPA) tests, which can quantify data snooping biases. To better approximate the full set of rules, this study considered a set of 9,385 technical trading rules parameterized from 14 technical trading systems, which were drawn from the major categories of technical trading systems, such as moving averages, channels (trading range breakouts), momentum oscillators, filters, and combinations. In addition to the 12 futures contracts that were investigated by Lukac, Brorsen, and Irwin, five highly traded contracts also were investigated to ensure a more general test of the profitability of technical trading rules. These contracts represent each major group of futures contracts, i.e., grains, meats, metals, energies, softs, currencies, equity indices, and interest rates. This study, therefore, improves upon previous studies of technical analysis in futures markets by incorporating more trading rules and contracts and conducting superior statistical tests.

The basic trading model used was similar to that introduced by Lukac, Brorsen, and Irwin. Two performance criteria, the mean net return and the Sharpe ratio, were used to measure the performance of technical trading rules. Based on these criteria, technical trading rules were optimized over a full sample period (1985-2004) and two subperiods (1985-1994 and 1995-2004). Net returns of technical trading rules were calculated by applying transaction costs of $100 per round-turn trade. We also applied lower transaction costs of $50 in out-of-sample tests to investigate the effect of transaction costs on the performance of technical trading rules.

The empirical evidence showed that the best trading rules identified under the mean net return criterion generated positive annual mean net returns in all 17 contracts over each of the sample periods. The trading profits were statistically significant in most contracts at a moderate significance level. When taken at face value, the performance of technical trading rules seemed to be promising. Once the effect of data snooping was considered, however, the technical trading profits were statistically significant only in two contracts, i.e., Eurodollar and the yen. These results strongly suggest that researchers may be misled by conventional statistical inference that does not account for the dependence across all the technical trading rules tested. Moreover, the superior profitability of the best technical trading rules identified over the in-sample period 1985-1994 did not continue for the out-of-sample period, 1995-2004. The best rules yielded positive net returns only in 7 of 17 contracts, with statistically significant net returns in two contracts (silver and Eurodollar). Similar results were observed under the Sharpe ratio criterion that considers transaction costs and risk. Among the 14 technical trading systems, the Moving Average Crossover (MAC), the Relative Strength Index (RSI), the Directional

Indicator (DRI), Alexander's Filter Rule (ALX), and the Directional Parabolic (DRP) systems perform the best across the performance criteria and sample periods. In particular, for each sample period the RSI system generated the best trading rule across various contracts, and the DRP system generated statistically significant profits for the Eurodollar contract even after adjustment for transaction costs, risk, and data snooping biases.

In conclusion, the evidence indicates that technical trading rules in general have not been profitable in US futures markets after correcting for transaction costs, risk, and data snooping biases during the 1985-2004 period. This finding confirms Park and Irwin's (2005) results in which the successful performance of Lukac, Brorsen, and Irwin's technical trading rules in the 1978-1984 period does not persist in the l985-2003 period. There are three possible explanations for the disappearance of technical trading profits in the 1985-2004 period: (1) data snooping biases (or selection bias) in previous studies, (2) structural changes in futures markets, and (3) the inherently self-destructive nature of technical trading strategies. To begin, the results of this study showed that over a relatively long time period U.S. futures markets were informationally efficient at least with respect to past prices. Lukac, Brorsen, and Irwin's (1988) successful finding, therefore, might result from examination of a relatively short and profitable sample period by chance. As noted previously, data snooping problems can occur by searching for profitable in- and out-of-sample periods, trading systems, and trading model assumptions, as well as profitable trading rules. As another explanation, Kidd and Brorsen (2004) report that returns to managed futures funds and commodity trading advisors (CTAs), which predominantly use technical analysis, declined dramatically in the 1990s. The decrease in technical trading profits could have been caused by structural changes in markets, such as reduced price volatility and increased kurtosis of daily price returns occurring while markets are closed. Since technical trading strategies make profits by the process of a market shifting to a new equilibrium, there may be fewer opportunities for profitable trading if prices are not volatile. Finally, forecasting methods are likely to be self-destructive (Malkiel 2003; Schwert 2003; Timmermann and Granger 2004). New forecasting models may produce economic profits when first introduced. Once these models become popular in the industry, however, their information is likely to be impounded in prices, and thus their initial profitability may disappear. Schwert (2003) finds that a wide variety of market anomalies in the stock market, such as the size effect and value effect, tend to have disappeared after the academic papers that made them famous were published.

References

Alexander, S.S. "Price Movements in Speculative Markets: Trends or Random Walks." *Industrial Management Review*, 2(1961):7-26.

Brock, W., J. Lakonishock, and B. LeBaron. "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns." *Journal of Finance*, 47(1992):1731-1764.

Brorsen, B.W., and S.H. Irwin. "Futures Funds and Price Volatility." *The Review of Futures Markets*, 6(1987):118-135.

Cheung, Y.W., and M.D. Chinn. "Currency Traders and Exchange Rate Dynamics: A Survey of the US Market." *Journal of International Money and Finance*, 20(2001):439-471.

Cooper, M., and H. Gulen. "Is Time-Series Based Predictability Evident in Real-Time?" *Journal of Business*, 79(2006):forthcoming.

Dale, C., and R. Workman. "Measuring Patterns of Price Movements in the Treasury Bill Futures Market." *Journal of Economics and Business*, 33(1981):81-87.

De Long, J.B., A. Shleifer, L.H. Summers, and R.J. Waldmann. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, 45(1990):379-395.

De Long, J.B., A. Shleifer, L.H. Summers, and R.J. Waldmann. "The Survival of Noise Traders in Financial Markets." *Journal of Business*, 64(1991):1-19.

Denton, F.T. "Data Mining as an Industry." *Review of Economics and Statistics*, 67(1985):124-127.

Fama, E.F. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, 25(1970):383-417.

Fama, E.F. "Market Efficiency, Long-Term Returns, and Behavioral Finance." *Journal of Financial Economics*, 49(1998):283-306.

Gehrig, T., and L. Menkhoff. "Extended Evidence on the Use of Technical Analysis in Foreign Exchange." *International Journal of Finance and Economics*, forthcoming.

Hansen, P.R. "Asymptotic Tests of Composite Hypotheses." Working Paper, Brown University, 2003.

Hansen, P.R. "A Test for Superior Predictive Ability." *Journal of Business and Economic Statistics*, forthcoming.

Kavajecz, K.A., and E.R. Odders-White. "Technical Analysis and Liquidity Provision." *Review of Financial Studies*, 17(2004):1043-1071.

Kaufman, P.J. *Trading Systems and Methods*, New York, NY: John Wiley & Sons, 1998.

Kho, B. "Time-Varying Risk Premia, Volatility, and Technical Trading Rule Profits: Evidence from Foreign Currency Futures Markets." *Journal of Financial Economics*, 41(1996): 249-290.

Kidd, W.V., and B.W. Brorsen. "Why Have the Returns to Technical Analysis Decreased?" *Journal of Economics and Business*, 56(2004):159-176.

Lo, A., and A.C. MacKinlay. "Data Snooping Biases in Tests of Financial Asset Pricing Models." *Review of Financial Studies*, 3(1990):431-467.

Logue, D.E., and R.J. Sweeney. "'White-Noise' in Imperfect Markets: The Case of the Franc/Dollar Exchange Rate." *Journal of Finance*, 32(1977):761-768.

Lovell, M.C. "Data Mining." *Review of Economics and Statistics*, 65(1983):1-12.

Lukac, L.P., and B.W. Brorsen. "A Comprehensive Test of Futures Market Disequilibrium." *Financial Review*, 25(1990):593-622.

Lukac, L.P., B.W. Brorsen, and S.H. Irwin. "A Test of Futures Market Disequilibrium Using Twelve Different Technical Trading Systems." *Applied Economics*, 20(1988):623-639.

Lukac, L.P., B.W. Brorsen, and S.H. Irwin. *A Comparison of Twelve Technical Trading Systems.* Greenville, SC: Traders Press, Inc, 1990.

Malkiel, B.G. "The Efficient Market Hypothesis and Its Critics." *Journal of Economic Perspectives*, 17(2003):59-82.

Murphy, J.J. *Technical Analysis of the Futures Markets*. New York, NY: New York Institute of Finance, 1986.

Neely, C., P. Weller, and R. Dittmar. "Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach." *Journal of Financial and Quantitative Analysis*, 32(1997):405-426.

Osler, C.L. "Currency Orders and Exchange Rate Dynamics: An Explanation for the Predictive Success of Technical Analysis." *Journal of Finance*, 58(2003):1791-1819.

Park, C.H., and S.H. Irwin. "The Profitability of Technical Analysis: A Review." AgMAS Project Research Report No. 2004-04, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, October 2004. http://www.farmdoc.uiuc.edu/agmas/

Park, C.H., and S.H. Irwin. "The Profitability of Technical Trading Rules in US Futures Markets: A Data Snooping Free Test." AgMAS Project Research Report No. 2005-04, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, May 2005. http://www.farmdoc.uiuc.edu/agmas/

Politis, D.N., and J.P. Romano. "The Stationary Bootstrap." *Journal of the American Statistical Association*, 89(1994a):1303-1313.

Politis, D.N., and J.P. Romano. "Limit Theorems for Weakly Dependent Hilbert Space Valued Random Variables with Application to the Stationary Bootstrap." *Statistica Sinica*, 4 (1994b):461-476.

Schwager, J.D. *Schwager on Futures*: *Technical Analysis*. New York, NY: John Wiley & Sons, 1996.

Schwert, G.W. "Anomalies and Market Efficiency." In G. Constantinides, M. Harris and R. Stulz (ed.) *Handbook of the Economics of Finance* (pp. 937-972). North-Holland, 2003.

Silber, W.L. "Technical Trading: When It Works and When It Doesn't." *Journal of Derivatives*, 1(1994):39-44.

Slezak, S.L. "On the Impossibility of Weak-Form Efficient Markets." *Journal of Financial and Quantitative Analysis*, 38(2003):523-554.

Sullivan, R., A. Timmermann, and H. White. "Data Snooping, Technical Trading Rule Performance, and the Bootstrap." *Journal of Finance*, 54(1999):1647-1691.

Sullivan, R., A. Timmermann, and H. White. "Forecast Evaluation with Shared Data Sets." *International Journal of Forecasting*, 19(2003):217-227.

Szakmary, A.C., and I. Mathur. "Central Bank Intervention and Trading Rule Profits in Foreign Exchange Markets." *Journal of International Money and Finance*, 16(1997):513-535.

Timmermann, A., and C.W.J. Granger. "Efficient Market Hypothesis and Forecasting." *International Journal of Forecasting*, 20(2004):15-27.

West, K.D. "Asymptotic Inference about Predictive Ability." *Econometrica*, 64(1996):1067-1084.

White, H. "A Reality Check for Data Snooping." *Econometrica*, 68(2000):1097-1126.

Wilder, J.W. *New Concepts in Technical Trading Systems*. Trend Research, Greensboro, NC: Hunter Publishing Company, 1978.

Table 1. Description of Futures Data, 1985-2004

Commodity	Exchange[a]	Contract Size	Value of One Tick (per contract)	Daily Price Limit	Contract Months Used
Corn (C)	CBOT	5,000 bu.	$12.50	10¢/15¢ per bu., expandable limits, before 7/15/93; 12¢/18¢ per bu., expandable limits, after 7/15/93; 20¢ (no more expanded limits) after 8/27/00.	Mar, May, Jul, Dec
Soybeans (S)	CBOT	5,000 bu.	$12.50	30¢/45¢ per bu., expandable limits, after 10/18/76; 50¢ (no more expanded limits) after 8/27/00.	Jan, Mar, May, Jul, Nov
Wheat (W)	CBOT	5,000 bu.	$12.50	20¢ per bushel above and below the previous day's settlement price, expandable limits; 30¢ on 8/27/2000. No limit in the spot month.	Mar, May, Jul, Sep, Dec
Cattle-Live (LC)	CME	40,000 lbs	$10.00	1.5¢/lb beginning with Dec. 1974 contract; 1.5¢ /3¢/5¢ per pound, expandable limits, effective 10/15/03.	Feb, Apr, Jun, Aug, Oct, Dec
Pork Bellies (PB)	CME	38,000 lbs, effective with Feb. 1979 contract; 40,000 lbs from Feb. 87 contract onwards.	$9.50; $10.00	2¢/lb before 10/1/96; 2¢/3¢/4.5¢ per pound, expandable limits, after 10/1/96.	Feb, Mar, May, Jul, Aug
Lumber (LB)	CME	130,000 before 3/16/88; 150,000 before 9/6/91; 160,000 before March 1996 contract; 80,000 before 8/11/00; 110,000 thereafter.	$13.00; $15.00; $16.00; $8.00; $11.00	$5.00/thousand board feet; No limit in spot month beginning with November 1984 contract; $10.00/$15.00, expandable limits, effective 2/3/93.	Jan, Mar, May, Jul, Sep, Nov
British Pound (BP)	CME	25,000 pounds; 62,500 pounds from Sep. 88 contract onwards	$12.50; $6.25	0.0500, expandable limits, after 3/18/74; No limit after 2/22/85.	Mar, Jun, Sep, Dec
Deutsch Mark (DM)	CME	125,000 marks	$12.50	0.0100, expandable limits, after 9/5/1978; No limit after 2/22/85.	Mar, Jun, Sep, Dec
Japanese Yen (JY)	CME	12,500,000 Japanese yen	$12.50	100 basis points, expandable limits; No limit after 2/22/85.	Mar, Jun, Sep, Dec
13-week Treasury Bills (TB)	CME	$1,000,000	$25.00	0.60, expandable limits, after 6/19/80; No limit after 12/19/85.	Mar, Jun, Sep, Dec
Eurodollar (ED)	CME	$1,000,000	$25.00	No limit.	Mar, Jun, Sep, Dec
S&P 500 (SP)	CME	$500 per index point; Reduced to $250 per index point on 11/3/1997.	$25.00	Price Limits corresponding to a 5.0%, 10.0%, 15.0% and 20.0% decline below the Settlement Price of the preceding Regular Trading Hours (RTH) session.	Mar, Jun, Sep, Dec

[a] CBOT: Chicago Board of Trade.
CME: Chicago Mercantile Exchange.

Table 1 continued.

Commodity	Exchange[a]	Contract Size	Value of One Tick (per contract)	Daily Price Limit	Contract Months Used
Cocoa (CC)	CSCE	10 Metric Tons for contract months beginning with 12/80.	$10 for Metric contract	$88, expandable to $132. For both contracts, limits are removed from the spot month on and after the first Notice day. No limit after 1998.	Mar, May, Jul, Sep, Dec
Sugar #11 (SB)	CSCE	112,000 lbs.	$11.20	1/2¢, expandable in increments of 1/2¢ to a maximum of 2¢. Limits do not apply to the nearest two months. No limit after 1998.	Mar, May, Jul, Oct
Crude Oil (CL)	NYMEX	1,000 U.S. barrels (42,000 gallons).	$10.00	$1.5/$3.0 per barrel until 1990 August contract; $7.5/$15.0 per barrel from 1990 September contract; $10.0/$20.0 effective on 9/05/2001.	All 12 month contracts
Copper (HG)	NYMEX	25,000 lbs.	$12.50	5¢, expandable limits, before 5/29/87; 20¢, expandable limits, after 6/1/87.	Mar, May, Jul, Sep, Dec
Silver (SI)	NYMEX	5,000 Troy oz.	$5.00 $25.00	50¢, expandable limits, before 5/29/87; $1.50, expandable limits, after 6/1/87.	Mar, May, Jul, Sep, Dec

[a] CSCE: Coffee, Sugar and Cocoa Exchange in the New York Board of Trade.
NYMEX: New York Mercantile Exchange.

Table 2. Performance of the Best Technical Trading Rules under the Mean Net Return Criterion, 1985-1994

		In-Sample Performance (1985-1994)							Out-of-Sample Performance (1995-2004)[a]			
Contracts	Best Rules[b]	AR_{100}	p_{RC}	$p_{N,W}$	p_l	p_c	p_u	$p_{N,H}$	AR_{100}	$p_{N,W}$	AR_{50}	$p_{N,W}$
Commodities												
Corn (CBOT)	DRP (26, 0.001)	5.25	1.00	0.13	0.87	0.97	1.00	0.16	4.85	0.20	6.98	0.11
Soybeans (CBOT)	RSI (13, 40)	6.31	1.00	0.16	0.68	0.92	0.94	0.15	-9.42	0.91	-7.62	0.86
Wheat (CBOT)	RSI (7, 12)	6.46	1.00	0.01	0.76	0.96	0.98	0.03	-4.30	0.84	-3.90	0.82
Live Cattle (CME)	ALX (0.03, 0.20)	7.19	0.95	0.06	0.86	0.97	0.98	0.05	-2.39	0.74	-1.26	0.62
Pork Bellies (CME)	RSI (9, 38)	19.03	0.89	0.05	0.78	0.88	0.91	0.04	-14.24	0.91	-11.64	0.86
Lumber (CME)	MACD (18, 26, 7, 0)	16.52	0.78	0.02	0.67	0.83	0.85	0.03	-0.96	0.55	3.11	0.37
Cocoa (CSCE)	RSI (8, 18)	12.06	0.99	0.02	0.73	0.89	0.94	0.02	-2.89	0.62	-1.47	0.55
Sugar #11 (CSCE)	DRP (6, 0.004)	18.17	0.96	0.05	0.86	0.96	0.98	0.05	-15.94	0.97	-9.93	0.88
Copper (COMEX)	MII (45, 0.001)	11.51	0.97	0.07	0.63	0.83	0.87	0.07	-2.95	0.66	-0.08	0.50
Silver (COMEX)	RSI (18, 38)	10.33	0.99	0.09	0.93	0.99	1.00	0.07	9.89	0.10	11.59	0.08
Crude Oil (NYMEX)	LSO (20, 1)	20.18	0.85	0.04	0.62	0.84	0.90	0.02	-4.43	0.67	1.72	0.39
Financials												
Mark (CME)	MAC (25, 55, 0)	10.46	0.27	0.00	0.28	0.38	0.40	0.00	2.04	0.33	2.54	0.30
Pound (CME)	DRP (10, 0.009)	9.01	0.59	0.01	0.43	0.52	0.55	0.01	-1.44	0.71	-0.72	0.62
Yen (CME)	MAC (7, 65, 0)	10.90	0.10	0.00	0.11	0.15	0.15	0.00	4.33	0.14	4.77	0.10
Eurodollar (CME)	DRP (2, 0.004)	1.76	0.03	0.00	0.01	0.02	0.02	0.00	0.53	0.02	0.58	0.01
T-Bills (CME)	PAR (0.026)	1.06	0.49	0.01	0.36	0.44	0.44	0.01	-0.16	0.58	-0.04	0.51
S&P 500 (CME)	RSI (6, 42)	14.15	0.30	0.05	0.54	0.69	0.70	0.01	1.49	0.38	1.92	0.34

Note: AR_{100} and AR_{50} denote annual mean net returns (%) after adjustment for transaction costs of $100 and $50, respectively. p_{RC} denotes White's Reality Check *p*-value. p_l, p_c, and p_u denote Hansen's Lower, Consistent, and Upper SPA *p*-values, respectively. $p_{N,W}$ and $p_{N,H}$ denote White's and Hansen's nominal *p*-values, respectively, which are obtained from applying their testing procedures only to the best rule or a single rule, thereby ignoring the effect of data snooping.

[a] Out-of-sample periods for the mark and T-bills are 1995-1998 and 1995-1996, respectively.

[b] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

Table 3. In-Sample Performance of the Best Technical Trading Rules under the Mean Net Return Criterion, 1995-2004[a]

Contracts	Best Rules[b]	AR_{100}	p_{RC}	$p_{N,W}$	p_l	p_c	p_u	$p_{N,H}$
Commodities								
Corn (CBOT)	MAC (25, 50, 0.005)	13.15	0.82	0.02	0.59	0.84	0.90	0.02
Soybeans (CBOT)	DRP (38, 0.002)	12.20	0.86	0.01	0.60	0.79	0.81	0.01
Wheat (CBOT)	ALX (0.07, 0.20)	15.33	0.74	0.03	0.43	0.86	0.91	0.02
Live Cattle (CME)	MAC (15, 35, 0.015)	8.52	0.87	0.00	0.26	0.36	0.42	0.01
Pork Bellies (CME)	RSI (14, 20)	14.75	0.99	0.04	0.81	0.92	0.95	0.04
Lumber (CME)	DRM (4)	15.84	0.95	0.05	0.84	0.98	0.98	0.04
Cocoa (CSCE)	RSI (10, 30)	12.59	0.98	0.08	0.85	0.99	1.00	0.06
Sugar #11 (CSCE)	CHL (50, 0.03)	5.07	1.00	0.14	0.59	0.86	0.93	0.17
Copper (COMEX)	RSI (9, 44)	10.74	0.96	0.04	0.89	0.98	0.99	0.05
Silver (COMEX)	RSI (7, 22)	14.30	0.73	0.03	0.75	0.92	0.94	0.02
Crude Oil (NYMEX)	EMC (5, 65, 0.001)	15.84	0.94	0.03	0.87	0.98	0.99	0.03
Financials								
Mark (CME)	MACD (14, 30, 17, 0.30)	7.87	0.91	0.01	0.67	0.82	0.84	0.01
Pound (CME)	RSI (16, 36)	3.09	1.00	0.12	0.99	1.00	1.00	0.11
Yen (CME)	ALX (0.05, 0.10)	10.40	0.19	0.00	0.29	0.35	0.36	0.00
Eurodollar (CME)	ALX (0.001, 0.0035)	0.82	0.03	0.00	0.09	0.10	0.11	0.00
T-Bills (CME)	RSI (10, 16)	0.94	0.91	0.02	0.53	0.65	0.66	0.01
S&P 500 (CME)	RSI (4, 38)	13.79	0.38	0.00	0.44	0.60	0.61	0.01

Note: AR_{100} denotes the annual mean net return (%) after adjustment for transaction costs of $100. p_{RC} denotes White's Reality Check p-value. p_l, p_c, and p_u denote Hansen's Lower, Consistent, and Upper SPA p-values, respectively. $p_{N,W}$ and $p_{N,H}$ denote White's and Hansen's nominal p-values, respectively, which are obtained from applying their testing procedures only to the best rule or a single rule, thereby ignoring the effect of data snooping.

[a] Sample periods for the mark and T-bills are 1995-1998 and 1995-1996, respectively.

[b] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

Table 4. In-Sample Performance of the Best Technical Trading Rules under the Mean Net Return Criterion, 1985-2004[a]

Contracts	Best Rules[b]	AR_{100}	p_{RC}	$p_{N,W}$	p_l	p_c	p_u	$p_{N,H}$
Commodities								
Corn (CBOT)	MAC (20, 65, 0)	5.81	1.00	0.10	0.94	0.99	1.00	0.07
Soybeans (CBOT)	DRP (34, 0.001)	5.88	1.00	0.07	0.92	0.99	1.00	0.06
Wheat (CBOT)	ALX (0.06, 0.20)	7.43	0.98	0.07	0.71	0.96	1.00	0.04
Live Cattle (CME)	MAC (15, 35, 0.015)	5.47	0.94	0.00	0.56	0.72	0.80	0.01
Pork Bellies (CME)	DRP (40, 0.001)	11.04	0.98	0.04	0.70	0.87	0.90	0.05
Lumber (CME)	DRP (6, 0.022)	9.53	0.98	0.05	0.84	0.97	0.98	0.05
Cocoa (CSCE)	RSI (7, 16)	5.38	1.00	0.12	0.94	1.00	1.00	0.15
Sugar #11 (CSCE)	RSI (16, 22)	5.17	1.00	0.16	0.91	1.00	1.00	0.12
Copper (COMEX)	ALX (0.02, 0.20)	8.98	0.92	0.04	0.75	0.93	0.95	0.05
Silver (COMEX)	RSI (18, 38)	10.01	0.85	0.03	0.65	0.91	0.95	0.03
Crude Oil (NYMEX)	MII (15, 0)	13.69	0.88	0.06	0.67	0.90	0.94	0.05
Financials								
Mark (CME)	MAC (25, 55, 0)	8.10	0.30	0.00	0.41	0.54	0.57	0.00
Pound (CME)	DRP (4, 0.010)	4.39	0.86	0.03	0.71	0.86	0.90	0.01
Yen (CME)	DRI (60, 3)	9.10	0.02	0.00	0.09	0.09	0.10	0.00
Eurodollar (CME)	DRP (6, 0.004)	1.21	0.00	0.00	0.00	0.00	0.00	0.00
T-Bills (CME)	DRP (2, 0.017)	0.89	0.53	0.01	0.34	0.43	0.44	0.01
S&P 500 (CME)	ALX (0.015, 0.100)	10.04	0.36	0.01	0.27	0.38	0.39	0.00

Note: AR_{100} denotes the annual mean net return (%) after adjustment for transaction costs of $100. p_{RC} denotes White's Reality Check *p*-value. p_l, p_c, and p_u denote Hansen's Lower, Consistent, and Upper SPA *p*-values, respectively. $p_{N,W}$ and $p_{N,H}$ denote White's and Hansen's nominal *p*-values, respectively, which are obtained from applying their testing procedures only to the best rule or a single rule, thereby ignoring the effect of data snooping.

[a] Sample periods for the mark and T-bills are 1985-1998 and 1985-1996, respectively.

[b] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

Table 5. Performance of the Best Technical Trading Rules under the Sharpe Ratio Criterion, 1985-1994

		In-Sample Performance (1985-1994)							Out-of-Sample Performance (1995-2004)[a]			
Contracts	Best Rules[b]	AR_{100}	p_{RC}	$p_{N,W}$	p_l	p_c	p_u	$p_{N,H}$	AR_{100}	$p_{N,W}$	AR_{50}	$p_{N,W}$
Commodities												
Corn (CBOT)	MAC (7, 10, 0.02)	0.46	1.00	0.08	0.87	0.97	1.00	0.09	-0.57	0.97	-0.49	0.94
Soybeans (CBOT)	RSI (7, 6)	0.53	0.99	0.04	0.68	0.92	0.94	0.02	-0.07	0.60	-0.06	0.58
Wheat (CBOT)	RSI (8, 6)	0.55	0.98	0.07	0.76	0.96	0.98	0.07	0.06	0.41	0.08	0.40
Live Cattle (CME)	DRI (40, 57)	0.60	0.98	0.06	0.86	0.97	0.98	0.07	-0.24	0.87	-0.22	0.86
Pork Bellies (CME)	DRP (40, 0.001)	0.66	0.88	0.03	0.78	0.88	0.91	0.02	0.11	0.35	0.14	0.31
Lumber (CME)	MACD (18, 26, 7, 0)	0.70	0.95	0.02	0.67	0.83	0.85	0.03	-0.03	0.55	0.11	0.37
Cocoa (CSCE)	RSI (7, 6)	0.65	0.90	0.03	0.73	0.89	0.94	0.03	-0.15	0.71	-0.13	0.70
Sugar #11 (CSCE)	DRP (6, 0.004)	0.47	0.99	0.05	0.86	0.96	0.98	0.05	-0.55	0.97	-0.34	0.88
Copper (COMEX)	DRI (25, 48)	0.66	0.79	0.01	0.63	0.83	0.87	0.02	-0.21	0.75	-0.16	0.66
Silver (COMEX)	MAC (15, 20, 0.04)	0.48	0.99	0.10	0.93	0.99	1.00	0.15	0.01	0.46	0.02	0.45
Crude Oil (NYMEX)	DRP (38, 0.001)	0.63	0.82	0.03	0.62	0.84	0.90	0.01	0.22	0.21	0.31	0.14
Financials												
Mark (CME)	MAC (25, 55, 0)	0.86	0.43	0.00	0.28	0.38	0.40	0.00	0.21	0.33	0.26	0.30
Pound (CME)	DRP (32, 0.009)	0.79	0.67	0.01	0.43	0.52	0.55	0.01	-0.21	0.73	-0.12	0.64
Yen (CME)	MAC (7, 65, 0)	1.00	0.32	0.00	0.11	0.15	0.15	0.00	0.37	0.12	0.40	0.10
Eurodollar (CME)	DRP (38, 0.007)	1.32	0.05	0.00	0.01	0.02	0.02	0.00	1.23	0.00	1.30	0.00
T-Bills (CME)	DRP (8, 0.026)	0.90	0.53	0.00	0.35	0.43	0.43	0.01	0.49	0.26	0.61	0.21
S&P 500 (CME)	RSI (8, 44)	0.57	0.78	0.02	0.54	0.69	0.69	0.02	0.15	0.38	0.17	0.34

Note: SR_{100} and SR_{50} denote Sharpe ratios after adjustment for transaction costs of $100 and $50, respectively. p_{RC} denotes White's Reality Check *p*-value. p_l, p_c, and p_u denote Hansen's Lower, Consistent, and Upper SPA *p*-values, respectively. $p_{N,W}$ and $p_{N,H}$ denote White's and Hansen's nominal *p*-values, respectively, which are obtained from applying their testing procedures only to the best rule or a single rule, thereby ignoring the effect of data snooping.

[a] Out-of-sample periods for the mark and T-bills are 1995-1998 and 1995-1996, respectively.

[b] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

Table 6. In-Sample Performance of the Best Technical Trading Rules under the Sharpe Ratio Criterion, 1995-2004[a]

Contracts	Best Rules[b]	SR_{100}	p_{RC}	$p_{N,W}$	p_l	p_c	p_u	$p_{N,H}$
Commodities								
Corn (CBOT)	MAC (25, 50, 0.005)	0.65	0.91	0.02	0.59	0.84	0.90	0.02
Soybeans (CBOT)	DRP (38, 0.002)	0.65	0.87	0.01	0.60	0.79	0.81	0.01
Wheat (CBOT)	ALX (0.07, 0.20)	0.63	0.91	0.03	0.43	0.86	0.91	0.02
Live Cattle (CME)	DRI (45, 36)	0.90	0.50	0.01	0.26	0.36	0.42	0.01
Pork Bellies (CME)	DRI (20, 63)	0.66	0.89	0.03	0.81	0.92	0.95	0.04
Lumber (CME)	DRM (4)	0.58	0.99	0.05	0.84	0.98	0.98	0.04
Cocoa (CSCE)	RSI (10, 30)	0.42	1.00	0.08	0.85	0.99	1.00	0.06
Sugar #11 (CSCE)	DRI (35, 51)	0.60	0.96	0.04	0.59	0.86	0.93	0.02
Copper (COMEX)	REF (50, 90)	0.51	0.98	0.04	0.89	0.98	0.99	0.02
Silver (COMEX)	RSI (7, 22)	0.61	0.88	0.03	0.75	0.92	0.94	0.02
Crude Oil (NYMEX)	EMC (5, 65, 0.001)	0.47	1.00	0.03	0.87	0.98	0.99	0.03
Financials								
Mark (CME)	MACD (14, 30, 17, 0.30)	0.95	0.89	0.01	0.67	0.82	0.84	0.01
Pound (CME)	ALX (0.15, 0.005)	0.52	0.98	0.14	1.00	1.00	1.00	0.15
Yen (CME)	ALX (0.05, 0.10)	0.88	0.30	0.00	0.29	0.35	0.36	0.00
Eurodollar (CME)	DRP (36, 0.007)	1.26	0.13	0.00	0.09	0.10	0.11	0.00
T-Bills (CME)	RSI (12, 16)	1.81	0.51	0.02	0.53	0.65	0.66	0.01
S&P 500 (CME)	RSI (4, 38)	0.73	0.61	0.00	0.44	0.60	0.61	0.01

Note: SR_{100} denotes the Sharpe ratio (%) after adjustment for transaction costs of $100. p_{RC} denotes White's Reality Check *p*-value. p_l, p_c, and p_u denote Hansen's Lower, Consistent, and Upper SPA *p*-values, respectively. $p_{N,W}$ and $p_{N,H}$ denote White's and Hansen's nominal *p*-values, respectively, which are obtained from applying their testing procedures only to the best rule or a single rule, thereby ignoring the effect of data snooping.

[a] Sample periods for the mark and T-bills are 1995-1998 and 1995-1996, respectively.

[b] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

Table 7. In-Sample Performance of the Best Technical Trading Rules under the Sharpe Ratio Criterion, 1985-2004[a]

Contracts	Best Rules[b]	SR[c]	p_{RC}	$p_{N,W}$	p_l	p_c	p_u	$p_{N,H}$
Commodities								
Corn (CBOT)	MAC (20, 65, 0)	0.29	1.00	0.10	0.94	0.99	1.00	0.07
Soybeans (CBOT)	DRP (34, 0.001)	0.33	1.00	0.05	0.92	0.99	1.00	0.06
Wheat (CBOT)	ALX (0.06, 0.20)	0.35	1.00	0.07	0.71	0.96	1.00	0.04
Live Cattle (CME)	MAC (15, 35, 0.015)	0.54	0.80	0.00	0.56	0.72	0.80	0.01
Pork Bellies (CME)	RSI (15, 18)	0.46	0.91	0.01	0.70	0.87	0.90	0.01
Lumber (CME)	RSI (16, 12)	0.45	0.99	0.05	0.84	0.97	0.98	0.06
Cocoa (CSCE)	RSI (7, 16)	0.32	1.00	0.05	0.94	1.00	1.00	0.15
Sugar #11 (CSCE)	RSI (15, 8)	0.31	1.00	0.10	0.91	1.00	1.00	0.08
Copper (COMEX)	REF (50, 85)	0.41	0.95	0.04	0.75	0.93	0.95	0.03
Silver (COMEX)	RSI (18, 38)	0.42	0.94	0.03	0.65	0.91	0.95	0.03
Crude Oil (NYMEX)	DRP (38, 0.001)	0.43	0.93	0.04	0.67	0.90	0.94	0.03
Financials								
Mark (CME)	MAC (25, 55, 0)	0.71	0.46	0.00	0.41	0.54	0.57	0.00
Pound (CME)	MAC (25, 30, 0.02)	0.55	0.78	0.10	0.71	0.86	0.90	0.09
Yen (CME)	DRI (60, 3)	0.81	0.10	0.00	0.09	0.09	0.10	0.00
Eurodollar (CME)	DRP (38, 0.007)	1.21	0.00	0.00	0.00	0.00	0.00	0.00
T-Bills (CME)	DRP (8, 0.026)	0.85	0.45	0.00	0.34	0.43	0.44	0.00
S&P 500 (CME)	ALX (0.015, 0.100)	0.52	0.49	0.01	0.27	0.38	0.39	0.00

Note: SR_{100} denotes the Sharpe ratio (%) after adjustment for transaction costs of \$100. p_{RC} denotes White's Reality Check *p*-value. p_l, p_c, and p_u denote Hansen's Lower, Consistent, and Upper SPA *p*-values, respectively. $p_{N,W}$ and $p_{N,H}$ denote White's and Hansen's nominal *p*-values, respectively, which are obtained from applying their testing procedures only to the best rule or a single rule, thereby ignoring the effect of data snooping.

[a] Sample periods for the mark and T-bills are 1985-1998 and 1985-1996, respectively.

[b] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

Table 8. The Performance of Technical Trading Systems by the Number of the Best Trading Rules

		Technical Trading Systems[a]													
Sample Period[b]	Performance Criterion	MAC	EMC	MACD	CHL	LSO	MII	RSI	DRI	REF	RNQ	DRM	ALX	PAR	DRP
1985-1994	Mean Net Return	2[c] (1)[d]	0	1	0	1	1	6	0	0	0	0	1	1	4 (1)
	Sharpe Ratio	4	0	1	0	0	0	4	2	0	0	0	0	0	6 (1)
1995-2004	Mean Net Return	2	1	1	1	0	0	7	0	0	0	1	3 (1)	0	1
	Sharpe Ratio	1	1	1	0	0	0	4	3	1	0	1	3	0	2 (1)
1985-2004	Mean Net Return	3	0	0	0	0	1	3	1 (1)	0	0	0	3	0	6 (1)
	Sharpe Ratio	4	0	0	0	0	0	5	1 (1)	1	0	0	2	0	4 (1)

[a] MAC: Moving Average Crossover; EMC: Exponential Moving Average Crossover; MACD: Moving Average Convergence-Divergence; CHL: Outside Price Channel; LSO: L-S-O Price Channel; MII: M-II Price Channel; RSI: Relative Strength Index; DRI: Directional Indicator; RNQ: Range Quotient; REF: Reference Deviation; DRM: Directional Movement; ALX: Alexander's Filter Rule; PAR: Parabolic Time/Price; DRP: Directional Parabolic

[b] The second subsample periods for the mark and T-bills are 1995-1996 and 1995-1998, respectively, and the full sample periods are 1985-1998 and 1985-1996, respectively.

[c] This figure denotes the number of contracts for which the technical trading system generates the best rule.

[d] Figures in parentheses denote the number of contracts for which the best rule from the technical trading system has a significant White's *p*-value or Hansen's *p*-value or both, showing a statistically significant performance after accounting for the effect of data snooping.

Appendix

Outside Price Channel (CHL)

The Outside Price Channel system is modified to construct a band around support and resistance levels. This system has been widely used among academics since Brock, Lakonishock, and LeBaron tested it. Osler showed that the CHL system may be profitable because down-trends (up-trends) tend to reverse course at predictable support (resistance) levels in foreign exchange markets. Kavajecz and Odders-White also found that support and resistance levels tend to identify clusters of orders (high depth) already in place on the limit order book in the NYSE.

Specifications of the system are as follows:

A. Definitions and abbreviations

1. Price channel = a time interval including today, n days in length.

2. The Highest High $(HH_t) = \max\{P_{t-1}^h, \ldots, P_{t-n+1}^h\}$, where P_{t-1}^h is the high at time $t-1$.

3. The Lowest Low $(LL_t) = \min\{P_{t-1}^l, \ldots, P_{t-n+1}^l\}$, where P_{t-1}^l is the low at time $t-1$.

4. Upper Band Limit $(UB_t) = HH_t + (b)HH_t$, where b is the fixed band multiplicative value.

5. Lower Band Limit $(LB_t) = LL_t - (b)LL_t$.

B. Trading rules

1. Go long at P_t^c if $P_t^c > UB_t$, where P_t^c is the close at time t. Offset the current (long) position at P_t^c and then go neutral (out-of-market) if $LB_t \leq P_t^c < LL_t$. Offset the current (long) position and simultaneously take a short position at P_t^c if $P_t^c < LB_t$.

2. Go short at P_t^c if $P_t^c < LB_t$. Offset the current (short) position at P_t^c and then go neutral (out-of-market) if $UB_t \geq P_t^c > HH_t$. Offset the current (short) position and simultaneously take a long position at P_t^c if $P_t^c > UB_t$.

C. Parameters

1. n = 2, 3, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 (16 values).

2. b = 0, 0.001, 0.005, 0.01, 0.015, 0.02, 0.03 (7 values); for interest rate contracts (Eurodollar and treasury-bills) b = 0, 0.0002, 0.0005, 0.001, 0.0015, 0.002, 0.003 (7 values).

M-II Price Channel (MII)

The M-II Price Channel system generates trading signals by comparing today's close with the theoretical high or low of the first day of the price channel. This system is also modified to form a band around the Reference Day Theoretical High (RDTH) and the Reference Day Theoretical Low (RDTL).

Specifications of the system are as follows:

A. Definitions and abbreviations

1. Price channel = n consecutive days price action including today.
2. Reference Day (RD) = the first day of the price channel.
3. Reference Day Theoretical High $(RDTH_t) = \max\{P_{RD}^h, P_{RD-1}^c\}$, where P_{RD}^h is the high of the RD day and P_{RD-1}^c is the close of the $RD-1$ day.
4. Reference Day Theoretical Low $(RDTL_t) = \min\{P_{RD}^l, P_{RD-1}^c\}$, where P_{RD}^l is the low of the RD day.
5. Upper Band Limit $(UB_t) = RDTH_t + (b)RDTH_t$, where b is the fixed band multiplicative value.
6. Lower Band Limit $(LB_t) = RDTL_t - (b)RDTL_t$.

B. Trading rules

1. Go long at P_t^c if $P_t^c > UB_t$, where P_t^c is the close at time t. Offset the current (long) position at P_t^c and then go neutral (out-of-market) if $LB_t \le P_t^c < RDTL_t$. Offset the current (long) position and simultaneously take a short position at P_t^c if $P_t^c < LB_t$.
2. Go short at P_t^c if $P_t^c < LB_t$. Offset the current (short) position at P_t^c and then go neutral (out-of-market) if $UB_t \ge P_t^c > RDTH_t$. Offset the current (short) position and simultaneously take a long position at P_t^c if $P_t^c > UB_t$.

C. Parameters

1. n = 2, 3, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 (16 values).

2. b = 0, 0.001, 0.005, 0.01, 0.015, 0.02, 0.03 (7 values); for interest rate contracts (Eurodollar and treasury-bills) b = 0, 0.0002, 0.0005, 0.001, 0.0015, 0.002, 0.003 (7 values).

Moving Average Crossover (MAC)

The moving average crossover system integrates two moving average systems introduced by Lukac, Brorsen, and Irwin, i.e., the simple moving average with a percentage price band (MAB) and the dual moving average crossover (DMC). No change is made for the MAB system, but the DMC system can now be tested with and without various percentage price bands. The MAC system has been the most popular trading system among academics since Brock, Lakonishock, and LeBaron found its profitability on the DJIA data.

Specifications of the system are as follows:

A. Definitions and abbreviations

1. Shorter Moving Average over s days at time t $(SMA_t) = \sum_{i=1}^{s} P_{t-i+1}^{c} / s$, where P_t^c is the closing price at time t and $s < t$.

2. Longer Moving Average over l days at time t $(LMA_t) = \sum_{i=1}^{l} P_{t-i+1}^{c} / l$, where $s < l \leq t$.

3. Upper Band Limit $(UB_t) = LMA_t + (b)LMA_t$, where b is the fixed band multiplicative value.

4. Lower Band Limit $(LB_t) = LMA_t - (b)LMA_t$.

B. Trading rules

1. Go long at P_{t+1}^{o} if $SMA_t > UB_t$, where P_{t+1}^{o} is the open at time $t+1$. Offset the current (long) position at P_{t+1}^{o} and then go neutral (out-of-market) if $LB_t \leq SMA_t < LMA_t$. Offset the current (long) position and simultaneously take a short position at P_{t+1}^{o} if $SMA_t < LB_t$.

2. Go short at P_{t+1}^{o} if $SMA_t < LB_t$. Offset the current (short) position at P_{t+1}^{o} and then go neutral (out-of-market) if $UB_t \geq SMA_t > LMA_t$. Offset the current (short) position and simultaneously take a long position at P_{t+1}^{o} if $SMA_t > UB_t$.

C. Parameters

1. s = 1, 2, 3, 4, 5, 7, 10, 15, 20, 25 (10 values).

2. l = 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 (13 values).

3. b = 0, 0.001, 0.005, 0.01, 0.015, 0.02, 0.03, 0.04, 0.05 (9 values); for interest rate contracts (Eurodollar and treasury-bills) b = 0, 0.0002, 0.0005, 0.001, 0.0015, 0.002, 0.003, 0.004, 0.005 (9 values).

Exponential Moving Average Crossover (EMC)

The exponential moving average is another form of a weighted moving average. Whereas the simple moving average weights all prices equally, the exponential moving average gives more weight to recent prices relative to older prices. The importance of older prices decreases rapidly, but it never fully disappears. Although most academic research has focused on simple moving averages, exponential moving averages may be useful because it responds faster to recent price movements than simple moving averages.

Specifications of the system are as follows:

A. Definitions and abbreviations

1. Shorter Moving Average (SMA) over s days at time t:

$$SMA_t = \left(\frac{2}{s+1}\right)P_t^c + \left(1 - \frac{2}{s+1}\right)SMA_{t-1},$$

where P_t^c is the closing price at time t and $s < t$. SMAs are calculated over 71 days before the first day of actual trade with the first SMA being equal to the closing price at 71 days before.

2. Longer Moving Average over l days at time t:

$$LMA_t = \left(\frac{2}{l+1}\right)P_t^c + \left(1 - \frac{2}{l+1}\right)LMA_{t-1},$$

where $s < l \leq t$. LMAs are calculated over 71 days before the first day of actual trade with the first LMA being equal to the closing price at 71 days before.

3. Upper Band Limit $(UB_t) = LMA_t + (b)LMA_t$, where b is the fixed band multiplicative value.

4. Lower Band Limit $(LB_t) = LMA_t - (b)LMA_t$.

B. Trading rules

1. Go long at P^o_{t+1} if $SMA_t > UB_t$, where P^o_{t+1} is the open at time $t+1$. Offset the current (long) position at P^o_{t+1} and then go neutral (out-of-market) if $LB_t \leq SMA_t < LMA_t$. Offset the current (long) position and simultaneously take a short position at P^o_{t+1} if $SMA_t < LB_t$.

2. Go short at P^o_{t+1} if $SMA_t < LB_t$. Offset the current (short) position at P^o_{t+1} and then go neutral (out-of-market) if $UB_t \geq SMA_t > LMA_t$. Offset the current (short) position and simultaneously take a long position at P^o_{t+1} if $SMA_t > UB_t$.

C. Parameters

1. s = 1, 2, 3, 4, 5, 7, 10, 15, 20, 25 (10 values).

2. l = 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 (13 values).

2. b = 0, 0.001, 0.005, 0.01, 0.015, 0.02, 0.03, 0.04, 0.05 (9 values); for interest rate contracts (Eurodollar and treasury-bills) b = 0, 0.0002, 0.0005, 0.001, 0.0015, 0.002, 0.003, 0.004, 0.005 (9 values).

Moving Average Convergence-Divergence (MACD)

The moving average convergence-divergence system is a technique developed by Gerald Appel. According to Schwager (p. 538), this system is one of the most interesting and dependable technical indicators in that it integrates positive features of both momentum oscillators and trend-following indicators. In its original form, the MACD system consists of two trendlines derived from three exponential moving averages (EMAs). The first is the MACD line that indicates the difference between a 12-day EMA and a 26-day EMA. The second is the signal line that indicates a 9-day EMA of the MACD line. Trading signals are generated when the MACD line penetrates above or below the signal line. A band around the signal line may be used to reduce losses from frequent whipsaws. The zero line can also be used to produce a trading signal, which is generated when the MACD line crosses above or below the zero line. Another popular variant is the MACD-histogram that is obtained by subtracting the signal line from the MACD line. In this study, however, the MACD line and the signal line are used with a band around the signal line. Four parameters that consist of a band and three EMAs are optimized by applying various combinations of parameter values.

Specifications of the system are as follows:

A. Definitions and abbreviations

1. Shorter Moving Average (SMA) over *s* days at time *t*:

$$SMA_t = \left(\frac{2}{s+1}\right)P_t^c + \left(1 - \frac{2}{s+1}\right)SMA_{t-1},$$

where P_t^c is the closing price at time *t* and $s < t$. SMAs are calculated over 71 days before the first day of actual trade with the first SMA being equal to the closing price at 71 days before.

2. Longer Moving Average over *l* days at time *t*:

$$LMA_t = \left(\frac{2}{l+1}\right)P_t^c + \left(1 - \frac{2}{l+1}\right)LMA_{t-1},$$

where $s < l \leq t$. LMAs are calculated over 71 days before the first day of actual trade with the first LMA being equal to the closing price at 71 days before.

3. Moving Average Convergence-Divergence $(MACD_t) = SMA_t - LMA_t$.

4. Signal Line over *n* days at time *t* $(SL_t) = \left(\frac{2}{n+1}\right)MACD_t + \left(1 - \frac{2}{n+1}\right)SL_{t-1}$.

SLs are calculated over 70 days before the first day of actual trade with the first SL being equal to the MACD at 70 days before.

5. Upper Band Limit $(UB_t) = SL_t + (b)SL_t$, where *b* is the fixed band multiplicative value.

6. Lower Band Limit $(LB_t) = SL_t - (b)SL_t$.

B. Trading rules

1. Go long at P_{t+1}^o if $MACD_t > UB_t$, where P_{t+1}^o is the open at time $t+1$. Offset the current (long) position at P_{t+1}^o and then go neutral (out-of-market) if $LB_t \leq MACD_t < SL_t$. Offset the current (long) position and simultaneously take a short position at P_{t+1}^o if $MACD_t < LB_t$.

2. Go short at P_{t+1}^o if $MACD_t < LB_t$. Offset the current (short) position at P_{t+1}^o and then go neutral (out-of-market) if $UB_t \geq MACD_t > SL_t$. Offset the current (short) position and simultaneously take a long position at P_{t+1}^o if $MACD_t > UB_t$.

C. Parameters

1. s = 1, 2, 4, 6, 8, 10, 12, 14, 16, 18 (10 values).

2. l = 5, 10, 15, 20, 26, 30, 40, 50, 60 (9 values).

3. n = 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 (12 values).

4. b = 0, 0.05, 0.10, 0.20, 0.30, 0.50 (9 values).

Relative Strength Index (RSI)

The Relative Strength Index, introduced by Wilder, is one of the most well-known momentum oscillator systems. Momentum oscillator techniques derive their name from the fact that trading signals are obtained from values which "oscillate" above and below a neutral point, usually given a zero value. In a simple form, the momentum oscillator compares today's price with the price of n-days ago. Wilder (p. 63) explains the momentum oscillator as follows:

> The momentum oscillator measures the velocity of directional price movement. When the price moves up very rapidly, as some point it is considered to be overbought; when it moves down very rapidly, at some point it is considered to be oversold. In either case, a reaction or reversal is imminent.

Momentum values are similar to standard moving averages, in that they can be regarded as smoothed price movements. Momentum oscillators, however, may identify a change in trend in advance because the momentum values generally decrease before a reverse in trend has taken place. The Relative Strength Index was designed to overcome two problems encountered in developing meaningful momentum oscillators: (1) erroneous erratic movement, and (2) the need for an objective scale for the amplitude of oscillators.[6]

Specifications of the system are as follows:

A. Definitions and abbreviations

1. Up Closes at time t $(UC_t) = P_t^c - P_{t-1}^c$, if $P_t^c > P_{t-1}^c$. P_t^c is the close at time t.

2. Down Closes at time t $(DC_t) = -(P_t^c - P_{t-1}^c)$, if $P_t^c < P_{t-1}^c$.

3. Average Up Closes over n days at time t, $t+1$, $t+2$, … :
$AUC_t = \sum_{i=1}^{n} UC_{t-i+1} / n$, $AUC_{t+1} = (AUC_t \times (n-1) + UC_{t+1}) / n$,
$AUC_{t+2} = (AUC_{t+1} \times (n-1) + UC_{t+2}) / n$, …. The first AUC is calculated at 51 days before the first day of actual trade.

[6] See Wilder for detailed discussion.

4. Average Down Closes over *n* days at time t, $t+1$, $t+2$, … :
   $ADC_t = \sum_{i=1}^{n} DC_{t-i+1} / n$, $ADC_{t+1} = (ADC_t \times (n-1) + DC_{t+1}) / n$,
   $ADC_{t+2} = (ADC_{t+1} \times (n-1) + DC_{t+2}) / n$, …. The first ADC is calculated at 51 days before the first day of actual trade.

5. Relative Strength at time t $(RS_t) = AUC_t / ADC_t$.

6. Relative Strength Index at time t $(RSI_t) = 100 - (100/(1 + RS_t))$.

7. Entry Thresholds $(ET, 100 - ET)$: RSI values beyond which buy or sell signals are generated.

B. Trading rules

1. Go long when RSI falls below *ET* and rises back above it.

2. Go short when RSI rises above $100 - ET$ and falls back below it.

3. Trading simulation begins 51 days before the first day (rollover date) of actual trade.

C. Parameters

1. n = 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 (16 values).

2. ET = 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 (20 values).[7]

Alexander's Filter Rule (ALX)

Alexander's original filter rule generates a buy (sell) signal when today's closing price rises (falls) by x% above (below) its most recent low (high). In this modified version, a long (short) position is initiated when today's closing price rises (falls) by x% and the existing long (short) position is liquidated when today's closing price rises (falls) by y%. A similar trading system was used by Logue and Sweeney and Sullivan, Timmermann, and White (1999).

Specifications of the system are as follows:

A. Definitions and abbreviations

[7] Wilder originally set the parameter values at n = 14 and ET = 30.

1. High Extreme Point (HEP) = the highest closing price obtained while in a long trade.

2. Low Extreme Point (LEP) = the lowest closing price obtained while in a short trade.

3. *x* = change in futures price required to initiate a position.

4. *y* = change in futures price required to liquidate a position.

B. Trading rules

1. Initiate a long (short) position if today's close rises (falls) *x*% above (below) the LEP (HEP).

2. Liquidate a long (short) position if today's close falls (rises) *y*% below (above) the HEP (LEP).

3. The system allows to go neutral (out-of-market).

4. Trading simulation begins 51 days before the first day (rollover date) of actual trade.

C. Parameter

1. *x* = 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.045, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.12, 0.15, 0.2 (18 values); for interest rate contracts (Eurodollar and treasury-bills), *x* =0.0005, 0.001, 0.0015, 0.002, 0.0025, 0.003, 0.0035, 0.004, 0.0045, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.012, 0.015, 0.02 (18 values).

2. y = 0.005, 0.01, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.045, 0.05, 0.06, 0.08, 0.1, 0.15, 0.2 (15 values); for interest rate contracts (Eurodollar and treasury-bills), y = 0.0005, 0.001, 0.0015, 0.002, 0.0025, 0.003, 0.0035, 0.004, 0.0045, 0.005, 0.006, 0.008, 0.01, 0.015, 0.02 (15 values).