Available online at www.sciencedirect.com
ScienceDirect
Journal of Banking & Finance 32 (2008) 1810–1819
Journal of BANKING & FINANCE
www.elsevier.com/locate/jbf

Can commodity futures be profitably traded with quantitative market timing strategies? ☆

Ben R. Marshall *, Rochester H. Cahan, Jared M. Cahan

Department of Finance, Banking and Property, College of Business, Massey University, Private Bag 11222, Palmerston North, New Zealand

Received 25 June 2007; accepted 6 December 2007
Available online 23 December 2007

Abstract

Quantitative market timing strategies are not consistently profitable when applied to 15 major commodity futures series. We conduct the most comprehensive study of quantitative trading rules in this market setting to date. We consider over 7000 rules, employ two alternative bootstrapping methodologies, account for data-snooping bias, and consider different time periods. We cannot rule out the possibility that trading rules compliment some other trading strategy or that some traders may have success using a specific rule on its own, but we do conclusively show that none of these rules beat the market any more than expected given random data variation.

JEL classification: G12; G14

Keywords: Commodity; Futures; Technical analysis; Quantitative; Market timing

1. Introduction

We consider whether quantitative trading rules can be profitably applied to commodity futures trading. While this question has been considered in the past, we aim to provide the most comprehensive examination to date. We study a larger universe of technical trading rules, focus on more recent data and address the issue of data-snooping bias using robust statistical techniques.

Commodity futures have been traded for a long time. However, it is only recently that researchers have begun discussing the merits of including them in conventional portfolios. Gorton and Rouwenhorst (2006) show commodity futures can successfully provide diversification for both stock and bond portfolios. Vrugt et al. (2004) indicate that this diversification can be achieved while losing little or no return. There are several possible explanations for these strong diversification benefits. Gorton and Rouwenhorst (2006) suggest that the strong performance of commodities in periods of unexpected inflation offsets the weak performance of stocks and bonds in these periods, while Hiller et al. (2006) highlight that precious metal commodities may be seen as safe destinations for funds during times of high stock high market volatility.

Erb and Harvey (2006) suggest that out-performance is not assured by just adding commodity futures to a portfolio, which implies that consideration needs to be given to the value that can be added by active management. It is therefore interesting that recent studies have shown that commodity futures can be successfully traded with a variety of strategies. Basu et al. (2006) show that the Commitment of Traders Report published by the Commodity Futures Trading Commission (CFTC), which summaries the positions taken by different participants in the market, contains information that can be exploited by an active

☆ This paper was reviewed and accepted while Prof. Giorgio Szego was the Managing Editor of The Journal of Banking and Finance and by the past Editorial Board.
* Corresponding author. Tel.: +64 6 350 5799; fax: +64 6 350 5651.
E-mail address: B.Marshall@massey.ac.nz (B.R. Marshall).

0378-4266/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2007.12.011

manager. Vrugt et al. (2004) show variables related to the business cycle, monetary policy, and market sentiment can all be used to generate profitable trading signals for commodity futures. Miffre and Rallis (2007) show that Jegadeesh and Titman (1993) momentum strategies generate returns of over 9% a year when applied to commodity futures, while Wang and Yu (2004) find that short-term contrarian strategies, similar to those of Lehmann (1990) and Lo and MacKinlay (1990), produce abnormal returns on commodity futures.

Futures markets are more attractive for pursing active trading strategies than stock markets for several reasons. Any active trading strategy incurs higher transaction costs than a buy-and-hold approach. In fact the level of transactions costs incurred often determines whether the profits accruing to a trading strategy is economically significant or not. Bessembinder and Chan (1998) find that the profits documented by Brock et al. (1992) for technical trading rules on the Dow Jones industrial average are not higher than the transaction costs that would be incurred in implementing them. Transaction costs, which include spreads and commissions, are a lot lower in futures markets than stock markets. Locke and Venkatesh (1997) estimate futures markets transaction costs to be in the 0.0004–0.033% range, while Lesmond et al. (1999) suggest that transaction costs in US equity markets range from 1.2% for large firms to 10.3% for small firms.

The ability to short-sell is a key component of most active trading strategies. Lesmond et al. (2004) show that the assumed profits from short-sales of small illiquid stocks contribute a large portion of the profits attributed to equity market momentum trading strategies. Since selling these stocks short may be difficult, if not impossible, in reality Lesmond et al. (2004) suggest that the profits that have been documented for these strategies may be questionable. In contrast, short-selling is easily done in futures markets.

The first paper to consider technical trading rules on commodity futures data appears to be Donchian (1960) who considered channel trading rules on copper futures data. Since this paper, several authors have documented profitability that exceeds reasonable estimates of transaction costs. Irwin et al. (1997) find that a channel trading system generates statistically significant mean returns ranging 5.1–26.6% in soybeans, soybean oil, and soybean meal futures during the 1984–1988 period, while Lukac et al. (1988) find that several technical trading systems, such as moving average and channel break-out systems, yield statistically significant portfolio returns ranging from 3.8% to 5.6% in 12 futures markets (including agriculturals and metals) during the 1978–1984 period.[1] Surveys of journalists and market participants (e.g., Lui and Mole, 1998; Oberlechner, 2001) show these individuals rely on technical analysis a lot for shorter forecasting intervals.

We extend this literature by considering 7846 trading rule specifications from five rule families (filter rules, moving average rules support and resistance rules, channel breakouts, and on balance volume rules). We apply these rules to the 15 commodities considered by Wang and Yu (2004). The commodity series include cocoa, coffee, cotton, crude oil, feeder cattle, gold, heating oil, live cattle, oats, platinum, silver, soyabeans, soya oil, sugar, and wheat. Our data cover the 1/1/1984–31/12/2005 period. We study the entire series and two equal sub-periods. We focus on this later period separately because Olson (2004) shows that the profits to technical analysis in the currency market have been eroded over time.

Unlike the previous commodity futures technical analysis literature, we utilise a variety of tests to examine the statistically significance of the trading rules profits. We use the Brock et al. (1992) (hereafter BLL) approach which involves fitting null models to the data, generating random bootstrapped series and comparing the profits generated from running the rules on the original commodity series to the profits generated the random series. We also use the bootstrapping technique (Sullivan et al., 1999, hereafter STW) which adjusts for data-snooping bias.

We find that the best trading rule for each commodity series typically produces profits that are statistically significant at the 5% level. However, the trading rules we consider do not generate profits on 14 of the 15 commodity series after an adjustment is made for data-snooping bias. This underscores the importance of conducting a robust adjustment for data-snooping bias. A short-term moving average rule generates statistically significant profits (after data-snooping bias adjustment) for the oats series and this profitability appears to be in excess of reasonable estimates of transactions costs, however it is not robust to our sub-period analysis. Rather, the profitability disappears in the most recent sub-period. Our results do not imply that there are no profitable trading rules, but rather that the number of profitable rules is no different from what one would expect due to random variation. Indeed, traders may rationally apply technical trading rules that generate abnormal returns for them even though these rules may be the result of data-snooping. We are also unable to rule out the possibility that technical trading rules are popular with traders because they spend long periods of time out of the market and therefore provide traders with the opportunity to apply their equity to other uses at these times.[2] Finally, we cannot rule out the possibility that technical trading rules can compliment some other trading strategy. However, our results do conclusively demonstrate that the profits accruing to a large universe of technical trading rules are no larger than those expected due to random variation.

The rest of the paper is organized as follows. Section 2 contains a description of the technical trading rules we test. Our data and bootstrapping methodologies are described

[1] The interested reader should refer to Park and Irwin (2004) for an excellent review of these early technical analysis studies and many others.

[2] We thank an anonymous referee for pointing this out to us.

in Section 3. We present and discuss our results in Section 4, while Section 5 concludes the paper.

2. Technical trading rules tested

We consider the profitability of the 7846 technical trading rules adopted by STW (1999) on the US equity market. The rules come from five rule families: filter rules, moving average rules, support and resistance rules, channel breakouts, and on-balance volume rules. The interested reader should refer to the appendix of STW (1999) for a full description of each rule applied.

The simplest filter rules we consider involve buying (short-selling) after price increases (decreases) by x% and selling (buying) when price decreases (increases) by x% from a subsequent high (low). Following STW (1999), we consider two alternative definitions of subsequent highs and lows. The first is the highest (lowest) closing price achieved while holding a particular long (short) position. The second definition involves a most recent closing price that is less (greater) than the e previous closing prices. Rules that allow a neutral position are also considered. Under these rules a long (short) position is closed when price decreases (increases) y percent from the previous high (low). The final variation we consider involves holding a position for a prespecified number of periods, c, regardless of other signals generated during this time.

Moving average rules are mechanical trading rules that attempt to capture trends. These generate a buy (sell) signal when the price moves above (below) the longer moving average. Variations of moving average rules we consider include those that generate a buy (sell) signal when a short moving average (e.g., 10 days) moves above (below) a longer moving average (e.g., 200 days). In accordance with STW (1999), we also consider the impact of applying two filters. The first filter requires the shorter moving average to exceed the longer moving average by a fixed amount, b. The second requires a buy or sell signal to remain valid for a prespecified number of periods, d, before the signal is acted upon. We also consider holding a position for a prespecified number of periods, c.

Support and resistance or "trading range break" rules are the third rule family we consider. These rules aim to profit from the principle that trends typically begin when price breaks out of a fixed trading band. Support and resistance rules involve buying (short-selling) when the closing price rises above (falls below) the maximum (minimum) price over the previous n periods. The most recent closing price that is greater (less) than the e previous closing price can also be set as the extreme price level that triggers a buy or a sell. Positions can be held for prespecified number of periods, c, and we also impose a fixed percentage band filter, b, and a time delay filter, d.

Our fourth family of rules is channel breakouts. We follow STW (1999) in that the channel breakout rules we test involve buying (selling) when the closing price moves above (below) the channel. A channel is said to occur when the high over the previous n periods is within x percent of the low over the previous n periods. Positions are held for a fixed number of periods, c. We also investigate a sub-set of channel breakout rules which involve a fixed band, b, being applied to the channel as a filter.

Our last rule family is on-balance volume (OBV) averages. The OBV indicator involves adding (subtracting) to (from) the indicator the entire amount of daily volume when the closing price increases (decreases). In accordance with STW (1999), we apply a moving average of n periods to the OBV indicator and apply trading rules similar to the Moving Average rules, except the variable of interest is OBV rather than price. The interested reader should refer to the STW (1999) paper for a more in-depth description of the trading rules we apply.

3. Data and methodology

3.1. Data

We analyze daily data on settlement prices and trading volume for the 15 commodities considered by Wang and Yu (2004).[3] The commodity series include cocoa, coffee, cotton, crude oil, feeder cattle, gold, heating oil, live cattle, oats, platinum, silver, soyabeans, soya oil, sugar, and wheat. Wang and Yu (2004) choose this broad range of series due to their economic importance and market liquidity. Each commodity series covers the 1/1/1984–31/12/2005 interval with the exception of silver, which starts on 30/8/1988. In line with Wang and Yu (2004), we use Datastream continuous price series, which represent the price for the most actively traded contract.

Consistent with past research (e.g., Bessembinder, 1992; Miffre and Rallis, 2007) we measure daily returns as the log of the difference in price relatives. This effectively means that we are assuming that a trader is funding their position with 100% equity rather than using margin. We investigate the impact of assuming that margin of differing percentages is used and find that this does not impact our key result, which is that the profits accruing to technical trading strategies are less than those expected by chance.[4]

We present the summary statistics for each commodity series in Table 1. We examine the distribution characteristics using the following statistics: mean, standard deviation, skewness, kurtosis, and the autocorrelation characteristics using the Ljung–Box–Pierce (Q-stats) test at lags of 6, 12,

[3] We are unable to source a corn series for an extended period so we include an oats series instead.

[4] Both the bootstrapping methodologies we apply compare trading rule returns to a benchmark such as buy-and-hold returns. Therefore assuming say for example 50% margin is used results in the returns to the trading rule doubling, but the benchmark returns also doubling. This means that there is no change to the statistical significance of the trading rule. Hence, while the actual returns would be larger for a trader who uses margin the statistical significance of these returns would not be stronger. We thank an anonymous referee for suggesting we carefully consider the impact of margin assumptions on our results.

Table 1
Summary statistics

	N	Mean (%)	Std. (%)	Skew.	Kurt.	$\rho(1)$	$\rho(2)$	$\rho(3)$	$\rho(4)$	$Q(6)$	$Q(12)$	$Q(24)$
Cocoa	11,480	−0.0105	1.9040	0.1545**	6.2546**	0.0011	−0.0245	0.0075	−0.0074	0.4719	4.5600	17.5873
Coffee	11,480	−0.0045	2.4306	0.2137**	10.8884**	−0.0026	−0.0330*	0.0364**	0.0024**	0.4641**	22.1726**	31.3498**
Cotton	11,480	−0.0061	1.8993	−11.2831**	437.3499**	0.0172	−0.0281*	−0.0109	0.0167	0.4753	8.6335	12.7815
Crude oil	11,480	0.0126	2.3711	−1.1412**	22.7825**	−0.0099	−0.0517**	−0.0434**	0.0188	0.4658**	34.3410**	47.1361**
Feeder cattle	11,480	0.0089	0.8572	−0.1738**	7.1394**	0.0394**	0.0109	0.0205	−0.0027	0.4860**	27.7084**	58.4575**
Gold	11,480	0.0051	0.8929	0.0524*	12.0246**	−0.0325*	−0.0247	−0.0010	0.0007	0.4847	10.5015	14.5724
Heating oil	11,480	0.0125	2.4373	−1.6611**	24.4073**	−0.0154	−0.0080	−0.0400**	−0.0034	0.4666**	23.6366**	59.3482**
Live cattle	11,480	0.0061	1.0399	−0.6999**	10.9612**	0.0367**	−0.0001	0.0155	0.0175	0.4834**	20.4507**	40.2751**
Oats	11,480	0.0008	2.0151	−0.048*	9.4390**	0.0645**	−0.0216	−0.0130	−0.0173	0.4699**	36.6366**	52.7165**
Platinum	11,480	0.0159	1.4006	0.3228**	18.0310**	−0.0308**	−0.0192	−0.0329	−0.0225	0.4782**	19.4358**	26.7163**
Silver	9048	0.0068	1.4265	−0.3679**	8.2722**	−0.0073	−0.0028	−0.0060	−0.0078	0.4766	1.4817	15.0351
Soybeans	11,480	−0.0053	1.3695	−0.5001**	8.1828**	0.0012	0.0148	0.0124	0.0118	0.4789*	16.5106**	28.1074*
Soya oil	11,480	−0.0054	1.5075	0.0939**	4.9157**	0.0118	−0.0370**	−0.0147	0.0277	0.4775**	17.4790*	21.7579*
Sugar	11,480	0.0131	3.3189	10.7817**	460.4833**	−0.0117	−0.0551**	−0.0072	−0.0689**	0.4616**	46.6694**	76.5959**
Wheat	11,480	−0.0012	1.5319	−0.2539**	11.3228**	0.0313*	−0.0344**	−0.0223	−0.0095	0.4775**	29.2327**	41.1632**

This table contains summary statistics for each data series. The ρ columns are the autocorrelation results for the stated lags. Ljung–Box–Pierce Q autocorrelation results are also presented.
* Indicates statistical significance at the 5% level.
** Indicates statistical significance at the 1% level.

and 24 days, along with the estimated autocorrelation at lags of 1–4 days. Nine of the 15 commodity series have positive mean daily return. Platinum has the largest mean daily return while cocoa has the smallest. Sugar, heating oil, and coffee are the most volatile series. Statistically significant (at the 1% level) skewness is prevalent in each commodity series. However, there is an almost even split between positive (seven commodities) and negative skewness (eight commodities). Statistically significant (at the 1% level) kurtosis is present in returns in all commodity series, which indicates the presence of fat tails in each of the return distributions.

Turning to the time series properties of the samples, we observe that there is evidence of positive (negative) autocorrelation at one lag in four (two) of the series. Negative autocorrelation at two lags, three lags, and four lags is prevalent in six, two, and one of the series, respectively. However, the Ljung–Box test indicates that positive autocorrelation is more prevalent at long lags.

3.2. Methodology

Tests of the profitability of technical trading rules in commodity futures markets (e.g., Stevenson and Bear, 1970) usually rely on the assumption that returns are stationary, independent, and normally distributed. However, Lukac and Brorsen (1990) find that technical trading returns on commodities are positively skewed and leptokurtic so these tests may not be valid. In accordance with Marshall et al. (in press) we apply two more appropriate test procedures. The first is the BLL (1992) bootstrapping methodology, while the second is the Reality Check bootstrapping technique of STW (1999) that accounts for data-snooping bias. We describe each of these tests in more detail below.

We begin by applying the bootstrap methodology that BLL (1992) adopted. We fit a null model to the data and generate the parameters of this model. We then randomly resample the residuals 500 times. We use each series of re-sampled residuals and the model parameters to generate random price series which have the same times-series properties as the original series. Earlier work (e.g., BLL, 1992) shows that bootstrap results are invariant to the choice of null model so we follow the established precedent in the literature (e.g., Kwon and Kish, 2002) and focus on the GARCH-M null model.[5] The GARCH-M model we apply is presented in Eqs. (1)–(3) (see BLL, 1992 for a detailed description of this model):

$$r_t = \alpha + \gamma\sigma_t^2 + \beta\varepsilon_{t-1} + \varepsilon_t, \tag{1}$$
$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{2}$$
$$\varepsilon_t = \sigma_t z_t \qquad z_t \sim N(0,1). \tag{3}$$

The BLL (1992) bootstrap methodology involves comparing the conditional buy and sell returns generated by a trading rule on the original commodity series with the conditional buy or sell returns generated from the same trading rule on a random simulated series. We follow BLL (1992) and define the buy (sell) return as the mean return per period for all the periods where the rule is long (short). The difference between the two means is the buy–sell return. The proportion of times the buy–sell profit for the rule is greater on the 500 random series than the original series is the buy–sell p-value. If, for a given rule, 24 of the 500

[5] We verify that our results are not sensitive to the choice of null model by re-running them using a EGARCH model. These results are qualitatively identical so we do not report them. We thank an anonymous referee for highlighting the importance of this robustness test to us.

random series have a buy–sell profit greater than that on the original series the p-value will be 0.048.

Our second test of profitability is the so-called White Reality Check bootstrap, introduced by White (2000). This bootstrap-based test evaluates whether the profitability of the best trading rule is statistically significant after adjusting for data-snooping bias, which is introduced by selecting the rule from a wide universe of rules. When there is a large universe of rules some will be profitable due to randomness so explicitly adjusting for data-snooping is critical. The White Reality Check accounts for this by adjusting down the statistical significance of profitable trading rules if they are drawn from a large universe of unprofitable rules. This is in contrast to the BLL (1992) approach where each rule is evaluated in isolation.

Specifically, we follow STW (1999), and let $f_{k,t}$ ($k = 1,\ldots,M$) be the period t return from the kth trading rule (out of a universe of M rules), relative to the benchmark (which is the commodity return at time t). The performance statistic of interest is the mean period relative return from the kth rule, $\bar{f}_k = \sum_{t=1}^{T} f_{k,t}/T$, where T is the number of days in the sample.

Like STW (1999), our null hypothesis is that the performance of the best trading rule, drawn from the universe of M rules, is no better than the benchmark performance, i.e.,

$$H_0 : \max_{k=1,\ldots,M} \bar{f}_k \leqslant 0. \tag{4}$$

STW (1999) then use the stationary bootstrap of Politis and Romano (1994) on the M values of $\bar{f}_k$ to test the null hypothesis.[6] To do this, each time-series of relative returns, f_k ($k = 1,\ldots,M$), is resampled (with replacement) B times, i.e., for each of the M rules, we resample the time-series of relative returns B times. Note that for each of the M rules, the same B bootstrapped time-series are used. Following STW (1999), we set $B = 500$. For the kth rule, this generates B means, which we denote $\bar{f}^*_{k,b}(b = 1,\ldots,B)$, from the B resampled time-series, where

$$\bar{f}^*_{k,b} = \sum_{t=1}^{T} f^*_{k,t,b}/T \quad (b = 1,\ldots,B). \tag{5}$$

The test two statistics employed in the test are

$$\overline{V}_M = \max_{k=1,\ldots,M} \left[\sqrt{T}\bar{f}_k\right] \tag{6}$$

and

$$\overline{V}^*_{M,b} = \max_{k=1,\ldots,M} \left[\sqrt{T}(\bar{f}^*_{k,b} - \bar{f}_k)\right] \quad (b = 1,\ldots,B). \tag{7}$$

To generate the test statistic, $\overline{V}_M$ is compared to the quantiles of the $\overline{V}^*_{M,b}$ distribution, i.e., we compare the maximum mean relative return from the M rules run on the original series, with the maximum mean across the M rules from each of the 500 bootstraps. In this way, the test evaluates the performance of the best rule with reference to the performance of the whole universe. In the context of our analysis, the White Reality Check bootstrap test allows us to compute a data-snooping adjusted p-value for the best rule in each of the file rule families, in relation to the universe of 7846 rules from which they are drawn.

4. Results

Our results indicate there is evidence that certain rules generate profits, but the statistical significance of these profits disappears once data-snooping bias is accounted for. We do not interpret these results as indicating there are no profitable trading rules, but rather that the number of profitable rules is no different from what one would expect due to random variation.

Table 2 contains bootstrap results for the entire 1984–2005 period (with the exception of Silver which starts in 1988). The number of rules that are statistically significant out of the total universe of 7846 rules are documented in the p-value count columns. A rule is statistically significant at the 1% (5%) level if there are five (25) or less occasions where the rule generates more profit on bootstrapped series than the original series. Results relating to the STW (1999) bootstrap technique are presented in the remaining columns. The nominal p-value is the reality check p-value for the best rule, unadjusted for data-snooping. The STW (1999) p-value is the data-snooping adjusted p-value, after accounting for the fact the rule is drawn from a wider universe of 7846 rules. The remaining columns contain other results relating to the best trading rule for each commodity series.

It is clear that from the BLL (1992) results that there is at least one rule that generates statistically significant on each of the 15 commodity series. Coffee and cotton have the fewest rules generating statistically significant profits at the 5% level while live cattle has the most. The pre-data-snooping adjustment results for the STW (1999) bootstrapping procedure are similar to their BLL (1992) counterparts for eight commodities in that there is evidence of the best performing rule being statistically significant at the 5% level. For the remaining eight commodity series there is no evidence of even the best performing rule generating statistically significant profits at the 5% level. Despite these differences across the two alternative bootstrapping techniques prior to data-snooping, the results are very clear once data-snooping bias is accounted for.

After adjustment for data-snooping bias the statistical significance of the best performing rule on each of the series other than oats disappears. The difference between the nominal and STW (1999) p-values is considerable for each commodity (other than oats) which gives an indication of the size of the potential data-snooping problem. Anyone testing a few rules in isolation could incorrectly conclude that the profits to technical analysis are statistically significant, when it fact any profitability can be attributed to data-snooping bias.

In the White Reality Check methodology the data-snooping adjusted statistical significance of the best

[6] We refer the reader to Appendix C of STW (1999) for the details. As per STW (1999), we set the probability parameter to 0.1.

Table 2
Bootstrap results – full period

	BLL *p*-value count (1%)	BLL *p*-value count (5%)	Nominal *p*-value	STW *p*-value	Average daily return (%)	Average return per trade (%)	Total no. of trades	No. of winning trades	No. of losing trades	Average days per trade
Cocoa	13	111	0.032	0.628	0.047	1.597	170	101	69	26.06
Coffee	4	82	0.052	0.670	0.056	1.228	261	140	121	10.04
Cotton	26	129	0.054	0.532	0.037	0.070	3002	1358	1644	1.91
Crude oil	51	193	0.068	0.784	0.070	1.323	304	157	147	18.73
Feeder cattle	93	311	0.038	0.676	0.035	0.385	524	299	225	5.14
Gold	77	262	0.106	0.832	0.024	34.908	4	4	0	1401.25
Heating oil	27	137	0.038	0.740	0.079	0.178	2554	975	1579	2.25
Live cattle	325	573	0.006	0.428	0.047	3.668	74	53	21	51.35
Oats	77	253	0.000	0.010	0.141	0.403	2012	816	1196	2.85
Platinum	10	132	0.190	0.932	0.036	4.113	50	26	24	111.96
Silver	26	128	0.068	0.754	0.048	1.565	139	80	59	26.33
Soybeans	46	179	0.008	0.248	0.058	3.442	96	56	40	53.05
Soya oil	14	115	0.016	0.372	0.055	0.126	2473	926	1547	2.32
Sugar	13	100	0.132	0.806	0.068	4.037	96	54	42	52.96
Wheat	64	121	0.008	0.310	0.060	0.084	4131	1915	2216	1.38

This table contains bootstrap results for the entire 1984–2005 period (with the exception of silver which starts in 1988). The number of rules that are statistically significant out of the total universe of 7846 rules are reported in the *p*-value count columns. A rule is statistically significant at the 1% (5%) level, if there are five (25) or fewer instances of the rule generating more profit on bootstrapped series than the original series. The remaining columns contain results relating to the Sullivan et al. (1999) (STW) bootstrap technique. The nominal *p*-value is the reality check *p*-value for the best rule, unadjusted for data-snooping. The STW (1999) *p*-value is the data-snooping adjusted *p*-value, after accounting for the fact the rule is drawn from a wider universe of 7846 rules. The remaining columns contain results relating to the best trading rule for each commodity series.

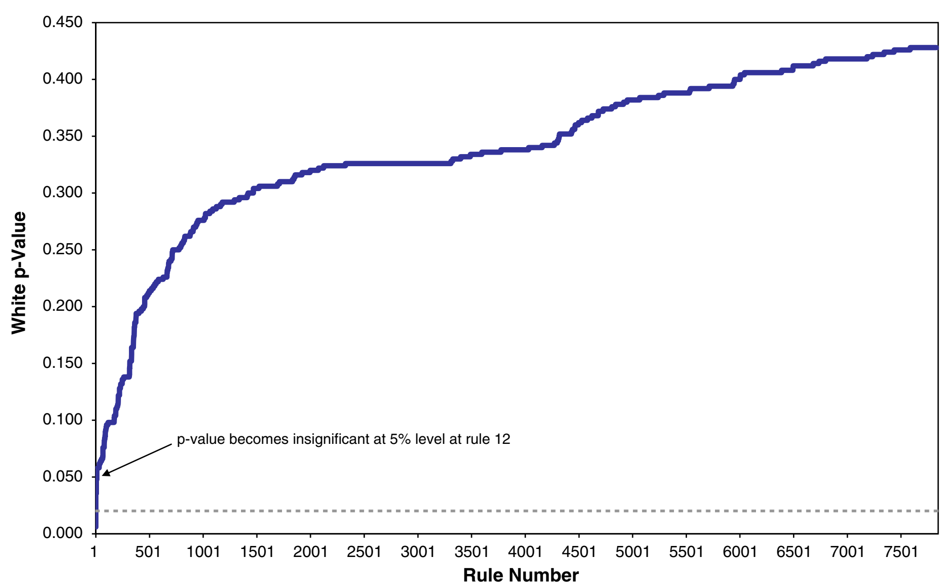

Fig. 1. Changes in white *p*-value for LCAT series as rule universe increases.

performing rule declines as more rules are added to the total universe of rules. It is therefore possible that a rule is unfairly penalized through comparison to a large number of unprofitable rules. We investigate the impact of this on our results by considering how the statistical significance of the best performing rule (chosen ex post after all 7846 rules are run) declines as more rules (starting with the most profitable first) are added to the universe. In Fig. 1 we display these results for live cattle. We choose this series because the best performing rule moves from being highly significant prior to the adjustment for data-snooping bias to highly insignificant after the adjustment, but the results are qualitatively similar for other commodity series. As the figure indicates, the best rule becomes insignificant at the 5% level when just 12 rules are included in the universe from which it was drawn. This suggests that our results are not being driven by the comparison to a large universe of rules.[7]

[7] We thank an anonymous referee for highlighting the importance of this analysis.

Table 3
Bootstrap results – first sub-period

	Nominal *p*-value	STW *p*-value	Average daily return (%)	Average return per trade (%)	Total no. of trades	No. of winning trades	No. of losing trades	Average days per trade
Cocoa	0.012	0.486	0.070	5.030	40	26	14	50.00
Coffee	0.052	0.720	0.080	7.676	30	19	11	50.27
Cotton	0.066	0.466	0.063	4.867	37	21	16	74.92
Crude oil	0.012	0.448	0.113	2.234	145	72	73	19.48
Feeder cattle	0.052	0.742	0.035	5.908	17	11	6	165.12
Gold	0.044	0.716	0.034	0.303	320	165	155	8.93
Heating oil	0.038	0.680	0.084	0.188	1282	488	794	2.24
Live cattle	0.032	0.610	0.044	0.507	249	156	93	5.08
Oats	0.000	0.016	0.172	0.473	1045	433	612	2.75
Platinum	0.066	0.728	0.053	3.720	41	24	17	68.24
Silver	0.032	0.674	0.003	1.415	4	3	1	493.75
Soybeans	0.014	0.320	0.054	5.337	29	20	9	92.76
Soya oil	0.026	0.680	0.055	26.076	6	5	1	477.67
Sugar	0.082	0.854	0.119	7.136	48	28	20	52.50
Wheat	0.022	0.456	0.078	0.189	1187	459	728	2.42

This table contains bootstrap results for the 1984–1994 period (except for silver which is 1988–1996). Each column contains results relating to the Sullivan et al. (1999) (STW) bootstrap technique. The nominal *p*-value is the reality check *p*-value for the best rule, unadjusted for data-snooping. The STW (1999) *p*-value is the data-snooping adjusted *p*-value, after accounting for the fact the rule is drawn from a wider universe of 7846 rules. The remaining columns contain results relating to the best trading rule for each commodity series.

The best performing trading rule on the oats series is the moving average rule involving price and a two-day moving average of price. This rule generates statistically significant (at the 1% level) profits after data-snooping bias has been accounted for. Even though the trading rule generates many trading signals (average days per trade is only 2.85), the average return per trade is 0.403%. This suggests that profits are available after transactions costs as estimated by Locke and Venkatesh (1997) (they estimate a range of 0.0004–0.033%). However, we leave it for the reader to decide whether these gross profits are sufficient to compensate for all costs incurred in exploiting them. The proportion of winning trades to total trades for oats (41%) indicates that technical trading rules can generate profits overall even if more losing than winning trades are generated.

Table 4
Bootstrap results – second sub-period

	Nominal *p*-value	STW *p*-value	Average daily return (%)	Average return per trade (%)	Total no. of trades	No. of winning trades	No. of losing trades	Average days per trade
Cocoa	0.060	0.774	0.065	1.211	153	81	72	10.32
Coffee	0.022	0.434	0.130	4.089	91	57	34	27.15
Cotton	0.022	0.444	0.079	2.622	86	53	33	27.03
Crude oil	0.118	0.970	0.079	0.822	276	102	174	10.38
Feeder cattle	0.036	0.782	0.025	7.068	10	4	6	266.80
Gold	0.084	0.772	0.041	1.325	88	56	32	27.68
Heating oil	0.186	0.932	0.082	33.498	7	6	1	373.71
Live cattle	0.028	0.674	0.053	4.506	34	29	5	50.00
Oats	0.020	0.528	0.114	0.151	2160	958	1202	1.31
Platinum	0.346	0.984	0.043	3.393	36	17	19	75.42
Silver	0.072	0.822	0.081	2.558	74	47	27	25.14
Soybeans	0.050	0.660	0.062	3.724	48	29	19	54.60
Soya oil	0.014	0.274	0.078	0.188	1179	472	707	2.43
Sugar	0.010	0.554	0.102	9.474	31	16	15	88.10
Wheat	0.000	0.212	0.108	0.143	2163	1058	1105	1.31

This table contains bootstrap results for the 1995–2005 period (except for silver which is 1997–2005). Each column contains results relating to the Sullivan et al. (1999) (STW) bootstrap technique. The nominal *p*-value is the reality check *p*-value for the best rule, unadjusted for data-snooping. The STW (1999) *p*-value is the data-snooping adjusted *p*-value, after accounting for the fact the rule is drawn from a wider universe of 7846 rules. The remaining columns contain results relating to the best trading rule for each commodity series.

No one rule performs best on each of the commodity series. Rather, rules form each of the rule families are represented across each of the 15 series. While the short-term moving average rule generates the largest profits for the oats series, a relatively long-term support and resistance rule generates the largest profits on the Gold series. This rule only signals 4 trades, or which all are profitable. It generates an average return per trade of 34.9% but this is not statistically significant (either before or after data-snooping adjustment) as the average daily return is only 0.024%.

We consider the robustness of these results by breaking each data series in half. Table 3 contains bootstrap results for the 1984–1994 period (except for silver which is 1998–1996). Given this consistency between the BLL (1992) and STW (1999) bootstrap results for the entire period we only present the more common STW (1999) results. The nominal p-value results indicate that, on average, the best performing rule on commodity series is more statistically significant in the early period than the entire period. Nine of the 15 nominal p-values are lower (more statistically significant) in the 1984–1994 period. Similarly, the average daily return is higher for the best rules on 10 of the 15 series in the early period.

Despite the evidence of more profitability to trading rules in the earlier period, the overall conclusions about their profitability made earlier still stand. There is strong evidence that the best performing trading rule generates statistically significant profits in the majority of series before data-snooping is accounted for, but this profitability disappears in all but the oats series once data-snooping bias is adjusted for. The Moving Average rule involving price and a two-day moving average of price is again the most profitable rule on the oats series. The proportion of all trades that turn out to be winning trades is similar (41%) to the entire series, while the average return per trade is slightly higher (0.473%).

Results for the second sub-period (1995–2005, except for silver which is 1989–2005) are presented in Table 4. Consistent with the entire period and first sub-period results, there is no evidence that the best performing trading rule

Table 5
Bootstrap results for long and short trades – entire period

	Average daily return (%)	Average return per trade (%)	Total no. of trades	No. of winning trades	No. of losing trades	Average days per trade
Panel A: Long trades						
Cocoa	0.06	1.620	88	53	35	26.20
Coffee	0.11	1.125	140	70	70	10.07
Cotton	0.03	0.059	1501	691	810	1.97
Crude oil	0.08	1.548	152	87	65	20.49
Feeder cattle	0.09	0.473	250	144	106	5.14
Gold	0.04	44.866	2	2	0	1192.50
Heating oil	0.09	0.207	1277	520	757	2.32
Live cattle	0.07	3.624	36	26	10	51.39
Oats	0.14	0.407	1006	397	609	2.90
Platinum	0.05	6.207	25	13	12	112.92
Silver	0.06	1.655	80	43	37	26.25
Soybeans	0.06	3.274	44	27	17	52.11
Soya oil	0.05	0.113	1236	468	768	2.31
Sugar	0.08	4.333	49	30	19	51.71
Wheat	0.06	0.084	2066	938	1128	1.39
Panel B: Short trades						
Cocoa	0.06	1.572	82	48	34	25.91
Coffee	0.13	1.347	121	70	51	10.00
Cotton	0.04	0.081	1501	667	834	1.84
Crude oil	0.06	1.097	152	70	82	16.97
Feeder cattle	0.06	0.305	274	155	119	5.15
Gold	0.02	24.950	2	2	0	1610.00
Heating oil	0.07	0.148	1277	455	822	2.17
Live cattle	0.07	3.711	38	27	11	51.32
Oats	0.14	0.398	1006	419	587	2.81
Platinum	0.02	2.020	25	13	12	111.00
Silver	0.05	1.442	59	37	22	26.44
Soybeans	0.07	3.584	52	29	23	53.85
Soya oil	0.06	0.139	1237	458	779	2.33
Sugar	0.07	3.728	47	24	23	54.26
Wheat	0.06	0.083	2065	977	1088	1.37

This table contains bootstrap results for long and short trades for the entire 1984–2005 period (with the exception of silver which starts in 1988). Each column contains results for to the best trading rule for each commodity series.

produces profits that are statistically significant once data-snooping bias is adjusted for in the majority of commodity series. This also applies to oats, which was previously able to be traded profitably using a short-term moving average rule. The fact that the best performing rule on the oats series in this period is a filter rule and even this is not profitable after data-snooping adjustment indicates that profitability of the moving average rule is not robust to different sub-periods.

Similar to gold in the entire period, the best performing rule on heating oil in the second sub-period generates profits in excess of 30% per trade. Despite the large size of these profits, they are not statistically significant before or after data-snooping adjustment due to the small number of trades generated (7) and the corresponding low average return per day. A comparison of the average return per trade figures for the first and second sub-periods indicates there is some evidence of a decline in profitability over time. Nine of the 15 commodity series have a best rule which yield lower profits in the second period.

We complete our analysis by considering whether there are major differences between the profits generated by the long and short signals generated by the best trading rule for the entire period.[8] It is possible that a rule generates particularly profitable long (short) signals but the lack of profitability in short (long) signals offsets this profitability. If this is the case then an investor may choose to act on the long (short) signals but ignore the short (long) signals.

The results presented in Table 5 indicate there is some evidence that the average daily return is higher for long trades than short trades. This is evident in 11 of the 15 commodity series. This is unsurprising as while commodity futures, as measured by the Reuters-CRB index, experienced considerable volatility over the 1984–2005 period we study, they did increase 27%. This indicates that, on average, there was more upward than downward movement. Although there is evidence of superior performance for long trades, the difference between the profits generated by long and short trades is generally small. The differences in average daily return per day are all 0.03% or less, which suggests an investor applying these technical trading rules is unlikely to be able to consistently add meaningful incremental profit by following the long signals of a rule and ignoring the short signals.

In summary, our results do not indicate that there are no profitable trading rules, but rather that the number of profitable rules is no different from what one would expect due to random variation. It is quite possible that traders who are unconcerned about data snooping bias apply technical trading rules that generate abnormal returns for them. Technical trading rules may also be popular with traders because they spend long periods of time out of the market and therefore earn the risk-free rate and provide traders with the opportunity to apply their equity to other uses at these times.[9] Finally, we cannot rule out the possibility that technical trading rules can compliment some other trading strategy. However, our results do conclusively demonstrate that the profits accruing to a large universe of technical trading rules are no larger than those expected due to random variation.

5. Conclusions

We reconsider whether quantitative trading rules can be profitably applied to commodity futures trading. Compared to previous work, we study a larger universe of technical trading rules, focus on more recent data and address the issue of data-snooping bias using robust statistical techniques. Commodity futures have been trading for long time, but it is only recently that debate has begun about the merits of including commodity futures in mainstream portfolios. Recent work has shown commodities can be very effective at providing diversification for both stock and bond portfolios, which may be due to their strong performance in periods of unexpected inflation or due to safe haven qualities of precious metal commodities.

Futures markets have several features that make them a more attractive market for active trading strategies than stock markets. In particular, transaction costs are lower and it is easier to short-sell. It is therefore interesting that recent studies have shown that commodity futures can be successfully traded with a variety of strategies, including using information on market positions from the Commodity Futures Trading Commission, medium-term momentum strategies and short-term contrarian strategies.

We find that the best trading rule for each commodity series typically produces profits that are statistically significant at the 5% level. However, the trading rules we consider do not generate statistically significant profits on 14 of the 15 commodity series after an adjustment is made for data-snooping bias. This underscores the importance of conducting a robust adjustment for data-snooping bias. A short-term moving average rule generates statistically significant profits (after data-snooping bias adjustment) for the oats series, however this profitability disappears in the most recent sub-period. We therefore conclude that technical trading rules are not profitable once data-snooping bias is taken into account. However, this does not rule out the possibility that individual traders can make abnormal profits by following these rules. We simply show that the profits they produce are less than those expected by chance. It is also possible that technical trading rules are popular with traders because they spend long periods of time out of the market and therefore provide traders with the opportunity to apply their equity to other uses at these times or that they compliment some other trading strategy.

[8] Equivalent results for each sub-period are very similar so are not reported in order to conserve space. The interested reader should contact the authors for these results.

[9] We thank an anonymous referee for pointing this out to us. The interested reader should refer to Kazemi and Li (2007) for more discussion on this.

Acknowledgement

We wish to thank Prof. Giorgio Szego and two anonymous referees for comments that improved our paper.

References

Basu, D., Oomen, R., Stremmer, A., 2006. How to time the commodity market. Working Paper. Warwick Business School.

Bessembinder, H., 1992. Systematic risk, hedging pressure and risk premiums in futures markets. Review of Financial Studies 5, 637–667.

Bessembinder, H., Chan, K., 1998. Market efficiency and the returns to technical analysis. Financial Management 272, 5–13.

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. Journal of Finance 485, 1731–1764.

Donchian, R.D., 1960. High finance in copper. Financial Analysts Journal (Nov/Dec), 133–142.

Erb, C., Harvey, C., 2006. The strategic and tactical value of commodity futures. Financial Analysts Journal 62 (2), 69–97.

Gorton, G., Rouwenhorst, K., 2006. Facts and fantasies about commodity futures. Financial Analysts Journal 622, 47–68.

Hiller, D., Draper, P., Faff, R., 2006. So precious metals shine? An investment perspective. Financial Analysts Journal 62 (2), 98–106.

Irwin, S.H., Zulauf, C.R., Gerlow, M.E., Tinker, J.N., 1997. A performance comparison of a technical trading system with ARIMA models for soybean complex prices. Advances in Investment Analysis and Portfolio Management 4, 193–203.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65–91.

Kazemi, H., Li, Y., 2007. Market timing of CTAs – An examination of systematic CTAs vs. Discretionary CTAs. University of Massachusetts. Working Paper.

Kwon, K.Y., Kish, R.J., 2002. A comparative study of technical trading strategies and return predictability: An extension of Brock, Lakonishok, and LeBaron 1992 using NYSE and NASDAQ indices. Quarterly Review of Economics and Finance 423, 611–631.

Lehmann, B., 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics 105, 1–28.

Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transactions costs. Review of Financial Studies 125, 1113–1141.

Lesmond, D.A., Schill, M.J., Zhou, C., 2004. The illusory nature of momentum profits. Journal of Financial Economics 71, 349–380.

Lo, A.W., MacKinlay, A.C., 1990. When are contrarian profits due to stock market overreaction? Review of Financial Studies 3, 175–208.

Locke, P.R., Venkatesh, P.C., 1997. Futures market transactions costs. Journal of Futures Markets 172, 229–245.

Lui, Y., Mole, D., 1998. The use of fundamental and technical analysis by foreign exchange dealers: Hong Kong evidence. Journal of International Money and Finance 17, 535–545.

Lukac, L.P., Brorsen, B.W., 1990. A comprehensive test of futures market disequilibrium. The Financial Review 244, 593–614.

Lukac, L.P., Brorsen, B.W., Irwin, S.H., 1988. A test of futures market disequilibrium using twelve different technical trading systems. Applied Economics 20, 623–640.

Marshall, B.R., Cahan, R.H., Cahan, J.M., in press. Does intraday technical analysis in the US equity market have value? Journal of Empirical Finance.

Miffre, J., Rallis, G., 2007. Momentum strategies in commodity futures markets. Journal of Banking and Finance 31 (6), 1863–1886.

Oberlechner, T., 2001. Importance of technical and fundamental analysis in the European foreign exchange market. International Journal of Finance and Economics 6, 81–93.

Olson, D., 2004. Have trading rule profits in the currency markets declined over time? Journal of Banking and Finance 28, 85–105.

Park, C.-H., Irwin, S.H., 2004. The profitability of technical analysis: A review. AgMAS Research Report.

Politis, D., Romano, J., 1994. The stationary bootstrap. Journal of American Statistical Association 89, 1303–1313.

Stevenson, R.A., Bear, R.M., 1970. Commodity futures: Trends or random walks? Journal of Finance 25, 65–81.

Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping, technical trading rule performance, and the bootstrap. Journal of Finance 245, 1647–1691.

Vrugt, Evert B., Bauer, R., Molenaar, R., Steenkamp, T., 2004. Dynamic commodity trading strategies. Working paper, July.

Wang, C., Yu, M., 2004. Trading activity and price reversals in futures markets. Journal of Banking and Finance 28, 1337–1361.

White, H., 2000. A reality check for data snooping. Economterica 685, 1097–1126.