Decision Support Systems 50 (2010) 128–139

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Profitability of technical analysis in financial and commodity futures markets — A reality check

Stéphane Meng-Feng Yen [a,*], Ying-Lin Hsu [b]

[a] Department of Accountancy and Institute of Finance and Banking, National Cheng Kung University, 1 University Rd., East District, Tainan 701, Taiwan
[b] Department of Applied Mathematics and Institute of Statistics, National Chung Hsing University, 250 Kuo Kuang Rd., Taichung 402, Taiwan

ARTICLE INFO

Article history:
Received 5 August 2009
Received in revised form 16 July 2010
Accepted 19 July 2010
Available online 24 July 2010

Keywords:
Technical analysis
Data-snooping bias
SPA test
Sortino ratio
Market efficiency

ABSTRACT

Based on the SPA test (test for superior predictive ability), Sortino and reversed Sortino ratios, we examined the profitability of a universe of 8061 technical trading rules in ten futures markets including five financial and five commodity underlying assets. We tested whether the best performing rule really beats its buy-and-hold benchmark strategy in bullish and bearish markets, respectively, during the in-sample testing period. The best rules' performance relative to the benchmark is also tested during the one-year out-of-sample period for all ten sets of data. A novel set of multi-indicator rules, MFI–RSI, and four popular categories of single-indicator rules, filter rules, moving averages, on-balance volume averages and momentum strategy in volume, were employed to form our universe of trading rules. The results on the SPA test suggest market efficiency in nine of the ten futures markets, while the results on the Sortino and reversed Sortino ratios reveal persistent outperformance of the best 'downside' and 'upside' rules relative to the buy-and-hold benchmark across time in four and three futures markets, respectively.

1. Introduction

Technical analysis has been one of the popular trading techniques among others in the futures markets for years. An excerpt from a report in BusinessWeek on March 19, 2001 [11] emphasizes the prevalence of technical analysis in futures markets: '…Most of the futures managers trade on the basis of technical, rather than fundamental, analysis, looking at such measures as price movements and changes in trading volume. They have developed analytical models based on the behavior of different futures markets over the years. "It's a systematic approach, and the systems are designed to profit when the futures products move through certain designated levels," says Sol Waksman, president of Barclay Trading Group, which researches and tracks futures funds….'

Despite its importance in the futures markets, the profitability of technical analysis might be subject to the so-called data-snooping bias, a stylized fact that is common to research on repetitively discovering the best model to explain an economic or financial time series. When the same set of data is tested using a chunk of models, we will always find that one or two of them are able to explain the data to a satisfactory extent, but only by chance rather than by any specific ability of the models themselves. Such a search for the best model from an enormous union constitutes the so-called data-snooping bias. An implication of the data-snooping bias is that any previous evidence of the profitability of technical analysis might be subject to questions. Without taking into account the data-snooping bias, such a conclusion might be fairly due to the mere luck, which is not robust to variability in sample periods or in variables.

Sullivan, Timmermann and White [23] tested five categories[1] of simple trading rules each based on one single technical indicator, amounting to 7846 rules in total, for the S&P 500 index futures. STW's[2] sample period is from 1984 through 1996 for the S&P 500 futures. Using White's [25] BRC to control for the data-snooping bias, STW found that, though some of the trading rules are able to beat the benchmark model (i.e. holding cash and staying out of the market) during the sample period above, the best rule of them is not able to show statistically significant outperformance relative to the benchmark with a possibility of 90.8%. STW concluded that profitability of technical analysis in the S&P 500 futures market might be due to the mere luck.

In this article we aimed to extend STW's work by using more futures contracts, a more powerful test than the BRC to alleviate the data-snooping bias, adding a more sophisticated type of trading rules as well as improving on the way to calculate daily trading returns. In particular, we examined ten futures contracts consisting of five financials and five commodities. To alleviate the significant reduction

* Corresponding author. Tel.: +886 6 275 7575x53443; fax: +886 6 274 4104.
E-mail addresses: yenmf@mail.ncku.edu.tw (S.M.-F. Yen), ylhsu@amath.nchu.edu.tw (Y.-L. Hsu).

[1] Filter rule, moving average, support and resistance, channel breakout and on-balance volume average.

[2] Acronyms STW and HK refer to, two groups of authors: Sullivan, Timmermann and White, and Hsu and Kuan. BRC stands for *bootstrap reality check*, while SPA test is abbreviated from *test for superior predictive ability*.

0167-9236/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2010.07.008

in the rejection rate of the null hypothesis under the BRC due to too many poorly performing trading rules, we followed Hsu and Kuan [14] and employed the more powerful SPA test introduced by Hansen [13]. All five categories of trading rules in STW were based on a single indicator of different parameter sets, while we modified this simple universe by including a category of technical analysis based on a multi-indicator strategy, the *MFI–RSI* rules, in our universe of trading rules. Adding more complex trading rules to the set of rules tested would reduce the gap between previous academic studies and what the futures market practitioners have been gradually and really doing. STW and previous studies all ignored the issue of calculating the rate of return for futures trading in a more practical manner. On the contrary, we computed the return for a trading rule using both the futures margin account and the risk-free current account. Without taking into account the market practices in calculating returns for trading futures, the profits or losses will be underestimated which then leads to an underestimated standard deviation of the daily returns. The results of the SPA test, based on the studentized average daily return in our context, will also be biased. Moreover, STW and previous studies did not separate the trading rules' performance during downside markets from during upside markets. We responded to this issue by applying the SPA test to bull and bear markets separately and comparing the (reversed) Sortino ratio of the trading rules with their buy-and-hold benchmark counterpart. Last but not the least, we replaced the benchmark of 'null' system (i.e. always staying out of the market) in STW by the buy-and-hold strategy.[3]

The ten futures contracts are CME Euro/USD FX futures (EUR/$ FX, hereafter), LIFFE FTSE 100 Index futures (FT-100, hereafter), EUREX DJ Euro Stoxx 50 Index futures (Stoxx-50, hereafter), SIMEX MSCI Taiwan Index futures (TW, hereafter), CBOT/CME US 30-year T-Bond futures (T-Bond, hereafter), CME live cattle futures (Cattle, hereafter), COMEX gold futures (Gold, hereafter), NYMEX/CME New York light sweet crude oil futures (Crude oil, hereafter), NYBOT/ICE coffee futures (Coffee, hereafter) and CBOT/CME soybean futures (Soybean, hereafter).[4] The five categories of technical trading rules are, firstly, a hybrid of the money flow index and the relative strength index (hereafter MFI–RSI), followed by the filter rules (hereafter FR), the moving averages (hereafter MA), the on-balance volume averages (hereafter OBV) and the momentum strategies in volume (hereafter MSV). The MFI–RSI category is formed by coupling the money flow index with the relative strength index, a novel application in the academic literature. The MFI tracks the flow of money into and out of a market, which is often used to warn of trend weakness and likely reversal points. The RSI measures price strength by comparing upward and downward close-to-close movements, which is often used to indicate whether a security has been overbought or oversold and thus a likely reversal. Both indicators are formulated to fluctuate between 0 and 100, enabling prespecified overbought or oversold levels. Mixing these two indicators will efficiently reduce the number of noisy trading signals and increase the percentage of successful trades.

Based on the SPA test to control for the data-snooping bias, we found that, the *MFI–RSI* and *MSV* types of technical trading rules tend to be identified as the best rule which significantly outperforms the benchmark in as many as seven futures markets, four financial and three commodity contracts. Nine of the fifteen sub-testing periods leading to a significant, best rule are in a bearish market, suggesting that the elites from the universe of our 8061 rules tend to perform better in a downside market than in an upside market. However, these 'real' outperformers selected from the most recent in-sample testing period generally fail to consistently provide a relatively good performance during the out-of-sample period. What is worse is that most of them generate a negative cumulative excess return, an evidence of no performance persistence among these in-sample 'real' outperformers. The only exception appears in the case of live cattle futures, in which the best trading rule selected from the last testing period, an *MSV* rule, still significantly beats the benchmark and generates a relatively large cumulative excess return over the one-year out-of-sample period. These results suggest that the remaining nine futures markets conform to the weak form of efficient market hypothesis, a finding in accord with what STW found with the S&P 500 index futures market.

Our results based on the Sortino and reversed Sortino ratios calculated for the entire in-sample period tend to suggest performance superiority of the best 'downside' rule over the best 'upside' rule in that the Sortino ratio of the former is larger than the reversed Sortino ratio of the latter across all ten contracts. Unlike the results on the SPA test, the best 'downside' rule consistently outperforms the benchmark in more than half the eight futures markets where the best rule generated trading signals during the one-year out-of-sample period. In contrast, the best 'upside' rule keeps beating the benchmark in less than half the ten futures markets where the best rule generated trading signals in the out-of-sample period. However, most of these persistent 'downside' and 'upside' outperformers tend to give a cumulative excess return larger than what their benchmark brings about. Our results based on the Sortino and reversed Sortino ratios indicate an evidence of better profitability relative to the benchmark from using the best technical trading rule out of the universe of our 8061 rules in the Eur/$ FX, Stoxx-50, T-Bond and Crude oil futures markets.

The rest of the paper is organized as follows. Section 2 presents the material and methods on the reality check of technical analysis' profitability, involving the trading rules, the SPA test, the (reversed) Sortino ratio and the data used. Section 3 details the practical issues of transaction cost and return calculation. Section 4 discusses the results, while Section 5 concludes.

2. Material and methods

2.1. Trading rules

Since the word 'technical analysis' itself concerns an enormous variety of simple or complex rules and strategies, it is unlikely to cover all candidates that were or are being used. However, it is possible to discuss the effectiveness of technical analysis from the perspectives of academics. In particular, we reviewed all past studies in the literature of this regard and summarized all previously discussed rules and strategies. Although we attempted to include all important rules, we recognized that we cannot exhaust all possible rules. As a result, we focused on Brock et al. [5], STW [23] and HK [14] which nearly covered all past efforts in this regard and selected those relatively good rules identified by these studies. In addition, we included a type of double-indicator trading rules in our universe to test whether the use of multi-indicators would help increase profitability. In total, we tested five categories of trading rules: money flow index coupled with relative strength index, filter rules, moving averages, on-balance volume averages, and momentum strategy in volume. Since all five categories were constructed based upon simple arithmetic calculation of the historical prices and/or trading volumes, a trade-off between

[3] Since futures investors are able to go long and go short a futures contract recurrently, staying out of the market or the risk-free rate seems to be a natural benchmark. However, futures investors are not supposed to use the risk-free reward to benchmark their trading performance because speculative trading in high-leveraged derivatives itself aims to earn a higher rate of return than trading spot assets. If the buy-and-hold strategy is conventionally regarded as the performance benchmark for trading spot assets, speculative futures investors would surely go after something better than what the buy-and-hold strategy brings.

[4] To explore a wider sample period than the S&P 500 index futures in STW, we included the 30-year US T-bond, live cattle and soybean futures contracts in the ten contracts, which all started trading earlier than 1984.