Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 15 (2008) 199–210
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Does intraday technical analysis in the U.S. equity market have value? ☆

Ben R. Marshall [a,*], Rochester H. Cahan [b,1], Jared M. Cahan [a,c]

[a] *Department of Finance, Banking and Property, Massey University, Private Bag 11-222, Palmerston North, New Zealand*
[b] *Citi Investment Research, Sydney, Australia*
[c] *Macquarie Bank, Sydney, Australia*

Accepted 31 May 2006
Available online 23 May 2007

Abstract

This paper investigates whether intraday technical analysis is profitable in the U.S. equity market. Surveys of market participants indicate that they place more emphasis on technical analysis (and less on fundamental analysis) the shorter the time horizon; however, the technical analysis literature to date has focused on long-term technical trading rules. We find, using two bootstrap methodologies, that none of the 7846 popular technical trading rules we test are profitable after data snooping bias is taken into account. There is no evidence that the market is inefficient over this time horizon.

JEL classification: G12; G14
Keywords: High frequency data; Technical analysis; Bootstrapping

1. Introduction

The use of past price movements to predict future price movements (technical analysis) has been popular with the investment community for a considerable period of time. When the key word "technical analysis" is typed into the Internet search engine Google, 201,000,000 urls are located compared to only 71,300,000 urls for "fundamental analysis".[2] Despite this widespread acceptance and adoption by practitioners, academics have traditionally treated technical analysis with disdain. It has been described by Malkiel (1981) as an "anathema to the academic world" due to its conflict with market efficiency, one of the central pillars of academic finance.

☆ We wish to thank the Securities Industry Research Centre of Asia-Pacific (SIRCA) for supplying the data on behalf of Reuters, and Steven Cahan, Franz Palm and two anonymous referees for comments that have dramatically improved this paper.
* Corresponding author. Tel.: +64 6 350 5799x5402; fax: +64 6 350 5651.
E-mail address: B.Marshall@Massey.ac.nz (B.R. Marshall).
[1] Rochester Cahan is a Quantitative Associate in the Investment Research department of Citigroup. The views expressed are the author's own and are not necessarily shared by Citigroup or any of its affiliates.
[2] Both searches were conducted on September 22, 2005.

0927-5398/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2006.05.003

Surveys of market participants and journalists consistently find that these individuals place more emphasis on technical analysis (and less emphasis on fundamental analysis) the shorter the forecasting horizon (e.g., Carter and Van Auken, 1990; Allen and Taylor, 1992; Lui and Mole, 1998; Oberlechner, 2001). More specifically, respondents place approximately twice as much weight on technical analysis for intraday horizons as they do for one-year horizons.

Despite market participants ascribing the most value to short-term technical analysis, the academic literature has focused on testing the profitability of long-term technical trading rules.[3] Most studies find that technical analysis is not profitable once transaction costs are taken into account (e.g., Allen and Karjalainen, 1999; Bessembinder and Chan, 1998; Olson, 2004). However, Corrado and Lee (1992) and Lee, Chan, Faff, and Kalev (2003) point out that technical analysis may still have merit as a value-adding "overlay" strategy to assist investors such as fund managers in better timing the buying or selling of stocks as part of their normal trading activities. Under this scenario the stock trades would have occurred in the normal course of business so the transaction costs are already factored in.

This paper considers the value of equity market technical analysis on an intraday basis using 5-minute Standard and Poor's Depository Receipts (SPDR) data. In doing so, several contributions are made. First, to the best of our knowledge, this is the first paper to consider the profitability of intraday equity market technical analysis. This is important because it is heavily used by practitioners over this time horizon, and recent papers by Kavajecz and Odders-White (2004) and Osler (2003) find evidence of order clustering that is consistent with the propositions of technical analysis. Given that the price pressure from order clustering is a short-term phenomenon, it seems reasonable to expect this to lend more support to intraday technical analysis than daily or monthly technical analysis. In addition, the short-term nature of intraday technical analysis also means that any profitability is extremely unlikely to be driven by time varying risk premia.[4]

Second, the use of actual transactions data for the Standard and Poor's Depository Receipts (SPDRs), the exchange traded fund that replicates the S&P 500 index, by this paper has several advantages. Previous studies, such as Neely and Weller (2003) and Osler (2000), analyze the value of technical analysis on intraday foreign exchange market data, but the absence of foreign exchange market trade data necessitates that these papers estimate transaction prices based on bid and ask quotes.

In addition, the choice of SPDR data has several advantages over the index data that has been used by the majority of longer-term technical analysis papers. Indices are not tradable in their own right so any technical trading signals would therefore be unable to be implemented without purchasing each of the index components in the correct proportions. Moreover, as Day and Wang (2002) document, tests of technical trading rules on index data can be biased due to non-synchronous trading. Finally, technical analysts claim that technical analysis is most reliable on actively traded stocks (Morris, 1995). By the end of 1999 there was $19.8 billion invested in SPDRs, and in 1998 the daily dollar volume was the highest of any stock (Elton et al., 2002). We purposely study the January 1, 2002 to December 31, 2003 period to give us an insight into any difference between trading rule performance in bull and bear markets.[5] The S&P 500 declined 21.2% in 2002 and increased by 21.9% in 2003.

Third, the choice of 7846 trading rule specifications from five rule families (Filter Rules, Moving Average Rules, Support and Resistance Rules, Channel Breakouts, and On Balance Volume Rules), which were widely publicized prior to the start of this study, allows a fair test of market efficiency. Miller (1990) points out that the development of financial theories alters behavior so testing models with data from before the models were developed is less than adequate.

Finally, unlike the previous intraday technical analysis literature (on the foreign exchange market), which does not conduct robust statistical tests of the significance of profits they document, we apply a suite of tests. These are the Brock, Lakonishok and LeBaron (1992) (hereafter BLL) approach of fitting null models to the data, generating random series and comparing the results from running the rules on the original series to those from running on the randomly generated bootstrapped series, and the so-called White's Reality Check bootstrapping technique (Sullivan, Timmerman, and White (1999), hereafter STW) which adjusts for data snooping bias. To the best of our knowledge, this is the first paper to utilize both these techniques.

[3] These include tests of moving average rules that signal a buy (sell) when price moves above (below) a moving average of past prices (Brock et al., 1992), trading range break out rules that signal a buy (sell) when price moves above (below) local maxima (minima) (Brock et al., 1992), optimal combinations of moving average and trading range break out rules derived using genetic algorithms (Allen and Karjalainen, 1999), and chart patterns such as "head and shoulders" and "double top" formations (Lo et al., 2000).

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

[5] We wish to thank an anonymous referee for highlighting the importance of this.

Our results clearly demonstrate that intraday technical analysis is not profitable on the SPDR series over the 2002–2003 period. While there is evidence that a small number of rules are profitable prior to any formal adjustment for data snooping bias, none of the five rule families produced a statistically significant profitable rule once this was taken into account. Given these results, we conclude that there is no evidence of consistent inefficiency in the intraday SPDR data.

The remainder of the paper is organized as follows: Section 2 outlines the technical trading rules used to test market efficiency and the steps that were taken to minimize data snooping bias. Data and methodology are presented in Section 3. Section 4 contains the results and Section 5 concludes the paper.

2. Technical trading rules employed

It is clear that the application of new trading rules, or new specifications of existing trading rules, to historical data introduces the chance of data snooping bias. It is quite possible that the rules have been tailored to the data series in question and are only profitable because of this. If this is the case, there is nothing to suggest that the rules will be profitable out of sample, or that someone would have chosen those exact specifications ex ante to form a profitable trading rule.

To prevent data snooping bias, Pesaran and Timmerman (1995, p. 102) conclude that "as far as possible, rules for predicting stock returns should be formulated and estimated without the benefit of hindsight." Both Lo and MacKinlay (1990) and Lakonishok and Smidt (1988) maintain that this new data approach is the best protection against data snooping.

A second approach to minimizing data snooping bias involves adjusting the statistical significance of a particular trading rule by taking account of the universe of all trading rules from which it is drawn (e.g., STW, 1999). While it is not possible to quantify the entire universe of trading rules that one rule might have been chosen from (e.g., Ready, 2002), the inclusion of a wide range of different rules does significantly reduce the risk that any given rule's profitability is due to chance.

To minimize the chance of data snooping, we follow both approaches.[6] More specifically, we use the five broad types of rules that received wide publicity prior to the start of our data. These have been published in many different papers and were succinctly summarized in STW (1999).[7] Like STW, we also apply a data snooping adjustment technique.

Filter Rules are the first rule family we consider. Standard filter rules, which were first introduced by Alexander (1961), involve buying (short-selling) after price increases (decreases) by x% and selling (buying) when it decreases (increases) by x% from a subsequent high (low). We define subsequent highs and lows in two ways. The first definition is the highest (lowest) closing price achieved while holding a particular long (short) position. The second is the most recent closing price that is less (greater) than the e previous closing prices. We also investigate filter rules that permit a neutral position. This is accomplished by closing a long (short) position when price decreases (increases) y% percent from the previous high (low). We also consider holding a long or short position for a prespecified number of periods, c, effectively ignoring all other signals generated during this time.

The second rule family we consider are Moving Averages, which appear to have been developed by Gartley (1930). In their most basic form, a buy (sell) signal is generated when the price moves above (below) the longer moving average, because at this point a trend is considered to be initiated. We also investigate rules that generate a buy (sell) signal when a short moving average (e.g., 10 periods) moves above (below) a longer moving average (e.g., 200 periods). We investigate the impact of applying two filters. Firstly, we require the shorter moving average (or price) to exceed the longer moving average by a fixed multiplicative amount, b. Secondly, we require a buy or sell signal to remain valid for a prespecified number of periods, d, before action it taken. As with the filter rules, we also consider holding a position for a prespecified number of periods, c.

Support and Resistance or "Trading Range Break" rules were developed by Wyckoff (1910). In their most simple form, these involve buying (short-selling) when the closing price rises above (falls below) the maximum (minimum)

[6] A third data snooping treatment involves the assumption that agents trade recursively using rule specifications that are considered "best performing" based on information up to the previous day (Fong and Yong, 2005). However, with high frequency data most market participants would not have time to work out the most profitable rule and revise their trading approach at five minute intervals. For this reason, we do not use this approach.

[7] Rather than repeat their appendix here we refer the reader to their paper for more detail on the rule specifications.

price over the previous n periods. The extreme price level that triggers a buy or a sell can also be defined as the most recent closing price that is greater (less than) the e previous closing price. As with moving average rules, we also impose a fixed percentage band filter, b, and a time delay filter, d. Again, positions can be held for prespecified number of periods, c.

A fourth family of rules, Channel Breakouts, are similar to Support and Resistance Rules. A channel is said to occur when the high over the previous n periods is within x% of the low over the previous n periods, not including the current price. Following STW (1999), the channel breakout rules we test involve buying (selling) when the closing price moves above (below) the channel. Long and short positions are held for a fixed number of periods, c. Additionally, a fixed band, b, can be applied to the channel as a filter.

On-Balance Volume (OBV) Averages are the final rule family we consider. This indicator, which was popularized by Granville (1963), is calculated by keeping a running total of the indicator each period and adding (subtracting) the entire amount of daily volume when the closing price increases (decreases). We then apply a moving average of n periods to the OBV indicator, as per STW (1999). The OBV trading rules employed are the same as for the moving average rules, except the variable of interest is OBV rather than price.

3. Data and methodology

3.1. Data

The SPDR data we test is sourced from the Reuters database[8] for the period January 1, 2002 to December 31, 2003 and consists of close prices taken at 5-minute intervals over the trading day. Exchange Traded Funds (ETFs) trade from 9.30 am to 4.15 pm on the AMEX. Thus for a complete trading day, from 9.30 am to 4.15 pm, we have 81 5-minute periods. Following Hol and Koopman (2002), we define the 5-minute return as:

$$r_{t,d} = \ln P_{t,d} - \ln P_{t,d-1} \tag{1}$$

where $r_{t,d}$ is the return for intraday period d on trading day t. Similarly, we calculate the overnight return following each trading day as:

$$r_{t,d} = \ln P_{t,0} - \ln P_{t-1,D} \tag{2}$$

where $D=81$, $P_{t,0}$ is 9.30 am price on day t, and $P_{t-1,D}$ is the 4.15 pm price on day $t-1$. Thus for each trading day, we have 81 5-minute returns and one overnight return.

Because there are a number of trading days where the market closes early, e.g., Christmas Eve, not all trading days in the sample have 81 intraday periods. The final sample of 5-minute data consists of 38,816 returns from 488 distinct days.[9] Due to data errors, there are occasionally periods with zero prices.[10] In this case we replace the price with the price from the previous 5-minute period, with one exception: if the first n prices of a given day are zero we, again following Hol and Koopman (2002), replace these n rows with the first non-zero price that day. This ensures the overnight return is computed using the first available price on day t and the last price on day $t-1$.

Table 1 presents the summary statistics for our full data set, and also our two sub-periods. For each period, we present separate statistics for the intraday (5-minute) and overnight returns. We examine the distribution characteristics using the following statistics: mean, standard deviation, skewness, kurtosis, and the Kolmogorov–Smirnoff (D-stat) test for normality. We also examine the autocorrelation characteristics of the three periods using the Ljung–Box–Pierce (Q-stats) test at lags of 6, 12 and 24 days, along with the estimated autocorrelation at lags of 1 to 5 days.

As we would expect, in the sub-period January 1, 2002 to December 31, 2002, the mean for both the intraday and overnight returns are negative. Over this period, the S&P 500 lost 21.2% of its value. On the other hand, in the second sub-period (from January 1, 2003 to December 31, 2003) the S&P 500 gained 21.9% and the mean of both the intraday and overnight returns are positive.

[8] The data is supplied by the Securities Industry Research Centre of Asia-Pacific (SIRCA) on behalf of Reuters.
[9] Of these, 468 days have the full 81 returns, while 20 days have less than 81 returns.
[10] Out of our sample of 38,816 5-minute periods, there are 374 periods with zero prices.

Table 1
Statistical properties

2002–2003			2002		2003	
	Intraday (5-minute)	Overnight	Intraday (5-minute)	Overnight	Intraday (5-minute)	Overnight
N	38,328	488	18,987	244	19,341	244
Mean	1.6814×10^{-6}	−0.0002	-5.8571×10^{-6}	−0.0006	9.0820×10^{-6}	0.0002
Standard	0.0013	0.0071	0.0016	0.0086	0.0011	0.0052
Skew	−0.0176	0.5661**	−0.0037	0.6462**	−0.0332	0.3974*
Kurtosis	9.4319**	5.9176**	8.4043**	4.9724**	7.4298**	4.8889**
p(1)	−0.0594**	−0.1246**	−0.0613**	−0.1889**	−0.0553**	0.0392
p(2)	−0.0036	0.0220	−0.0141	0.0160	0.0192**	0.0154
p(3)	0.0099	−0.0970*	0.0174*	−0.1191	−0.0067	−0.0408
p(4)	−0.0013	−0.0516	−0.0094	−0.0786	0.0164*	−0.0038
p(5)	−0.0047	−0.0046	−0.0012	0.0169	−0.0120	−0.0664
Bartlett standard	0.0051	0.0453	0.0073	0.0640	0.0072	0.0640
D-stat	0.4967**	0.4898**	0.4964**	0.4893**	0.4974**	0.4944**
Q(6)	140.5525**	14.9881*	83.4637**	14.7189*	76.2794**	2.3013
Q(12)	151.1781**	31.6545**	94.5957**	31.6890**	77.1310**	4.0773
Q(24)	169.1107**	49.9211**	113.3156**	50.2105**	91.9635**	10.9612

Summary statistics for each data series. * indicates statistical significance at the 5% level, ** indicates statistical significance at the 1% level.

Statistically significant (at the 1% level) kurtosis is present in both intraday and overnight returns in the overall period, and also each of the sub-periods. This indicates the presence of fat tails in each of the return distributions. As a consequence of these characteristics, the Kolmogorov–Smirnoff (*D*-stat) test rejects normality in all periods.

Turning to the time-series properties of the samples, we observe that both the 5-minute data displays evidence of autocorrelation at one lag in all periods. As well, the overnight returns exhibit statistically significant (at the 1% level) autocorrelation in the overall sample and 2002 sub-period. Likewise, the Ljung–Box test is significant at the 1% level for all three lags tested in each of the 5-minute samples. For the overnight returns, the 2002 sub-period and the full sample both fail the Ljung–Box test. The exception is the 2003 overnight returns, where we cannot reject the null of no autocorrelation.

3.2. Methodology

Traditional tests of the profitability of technical trading rules were often based on a simple *t*-test methodology (e.g., Sweeney, 1988). However, as BLL (1992) point out, such tests rely on the assumption of normal, stationary and independent distributions of returns. As BLL (1992) find, even in conventional daily close return data, this is often not the case. In an intraday framework, this problem is further exacerbated as is evident in Table 1. Thus, we focus our testing on two more appropriate methodologies, namely the BLL (1992) bootstrapping methodology, and the Reality Check bootstrapping technique of STW (1999) that accounts for data snooping bias. Each of the two tests as described in detail below.

3.2.1. BLL Bootstrapping Methodology

Our first test is based on that employed by BLL (1992). The aim of the BLL (1992) methodology is to fit a null model to a data set such that random realizations of the model replicate the time-series properties of the original date series. This allows profitability statistics generated by a technical trading rule on the original series to be compared with the same statistics computed on a large number of random realizations of the series (based on a specific null model). From this, one can determine the probability that the trading profits generated by a rule are the results of a particular null model.

In their seminal paper, BLL (1992) select four null models: the random walk, AR(1), GARCH-M and E-GARCH. For our intraday analysis, we focus on the GARCH family of models. This family of models has become widely used in financial modeling because they take into account two important features of financial time-series: so-called fat-tails, i.e., significant kurtosis, and time-varying volatility clustering (Bollerslev et al., 1992). As Table 1 illustrates, these

Table 2
Model fit test statistics

Test	Lag	Raw returns	GARCH(1,1)	EGARCH	GARCH-M
LB	1	85.54**	0.03	0.03	0.03
	2	89.45 **	1.01	0.69	1.01
	3	89.62 **	1.80	0.72	1.80
	4	90.10 **	2.14	0.85	2.14
	5	91.38 **	2.25	2.02	2.25
	6	92.88 **	2.61	5.49	2.61
	7	93.74 **	2.68	7.22	2.68
	8	94.26 **	2.86	8.34	2.87
	9	94.46 **	3.01	8.34	3.01
	10	97.13 **	3.87	8.38	3.87
	11	97.20 **	3.93	8.41	3.93
	12	98.36 **	3.96	9.79	3.96
ML	1	46.74 **	2.43	6.03*	2.43
	2	71.98 **	2.91	7.30*	2.91
	3	99.48 **	2.96	8.81*	2.95
	4	138.01 **	2.98	15.26 **	2.98
	5	169.83 **	3.04	19.4 **	3.03
	6	219.02 **	3.07	35.34 **	3.07
	7	262.70 **	3.22	41.54 **	3.22
	8	300.48 **	3.43	44.15 **	3.43
	9	334.97 **	3.57	46.02 **	3.57
	10	350.97 **	3.59	46.19 **	3.58
	11	382.60 **	3.87	46.68 **	3.86
	12	415.96 **	4.39	46.75 **	4.39

Results from the Ljung–Box (LB) and McLeod–Li (ML) model fit tests. * indicates statistical significance at the 5% level, ** indicates statistical significance at the 1% level.

features are particularly important for modeling intraday data. Nonetheless, the use of GARCH-type models to describe intraday data is not yet firmly cemented in the literature (see Andersen and Bollerslev, 1998; Blair et al., 2001). With no widely accepted model to fall on as a default, we pay particular attention to developing a model suitable for our intraday data.

Specifically, we fit a GARCH(1,1), EGARCH and GARCH-M model to the data and perform post-estimation analysis on the residuals to determine the best null model to use in the bootstrapping procedure. An important step is the introduction of an overnight dummy variable into the GARCH models to account for any difference in the return level and conditional volatility between the overnight and intraday returns.[11] To choose between the three models, we use the Ljung and Box (1978) Q-test and the McLeod and Li, (1983) test. The LB test is a portmanteau test for autocorrelation in the first n lags of a time-series, while the ML test is identical to the LB test, except applied to the squared sequence. These tests are popular both to motivate the fitting of GARCH models as well as for post-estimation analysis (Chen, 2003).

Our methodology is to apply the two tests first to the raw return series, and then to the standardized residuals generated from each of the three null models. If our model fits are valid, we would expect to broadly observe the test statistics going from statistically significant in the raw returns, to insignificant in the residuals, i.e., the models should explain the autocorrelation in both the return and variance series. Table 2 presents the results for raw return series and the three GARCH models.

The 5-minute raw returns and squared raw returns both exhibit significant (at the 1% level) autocorrelation at lags of one to twelve periods. This suggests that a GARCH-type model is likely to be appropriate so we proceed to fit the GARCH(1,1), EGARCH and GARCH-M models. The EGARCH residuals show no significant autocorrelation, but the model fails to adequately explain autocorrelation in the squared residuals so it is rejected. In contrast, the GARCH(1,1) and GARCH-M residuals both pass the LB and ML tests at all lags. We also compute Akaike's Information Criterion

[11] We wish to thank an anonymous referee for highlighting the importance of this.

Table 3
GARCH-M (with dummy) model coefficients

	α	b	K	α_1	β	γ	ϕ	φ
Coefficient	6.0400×10^{-6}	0.0566	1.2100×10^{-7}	0.1370	0.7597	−0.2145	-7.8200×10^{-5}	1.1100×10^{-5}
t-statistic	0.70	9.40**	75.43**	51.74**	272.05**	−0.04	−0.76	86.35**

Coefficients and *t*-statistics for the GARCH-M model with overnight dummy. * indicates statistical significance at the 5%, *** indicates statistical significance at the 1% Table 4.

(AIC) and Schwarz's Bayesian Criterion (SBC) for each of the three models.[12] Both criteria, which measure goodness of fit while penalizing over-parameterization, are essentially the same for the GARCH(1,1) and GARCH-M models. Because neither model is clearly superior, in the interests of consistency with the literature (e.g., BLL, 1992; Kwon and Kish, 1992), we choose the GARCH-M model given by:

$$r_t = \alpha + \gamma h_t + b\varepsilon_{t-1} + \phi\delta_t + \varepsilon_t. \tag{3a}$$

$$h_t = \kappa + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1} + \phi\delta_t \tag{3b}$$

$$\varepsilon_t = h_t^{1/2} z_t \qquad z_t \sim N(0,1) \tag{3c}$$

where ε_t is an independent, identically distributed normal random variable[13], h_t is the conditional variance which is a linear function of the conditional variance of the previous period and the square of the previous period's error term, and δ_t is an overnight dummy such that $\delta_t = 1$ if r_t is an overnight return and 0 otherwise.

The overnight dummy in the conditional mean and variance equations is an important addition because it allows us to account for any difference in the mean level and/or conditional volatility between the overnight and intraday returns.[14] From a bootstrapping perspective, the overnight dummy is particularly important because we want each random realization of the null model to have the same time-series structure as the original series. The addition of the overnight dummy ensures that all replications of the null model have the same intraday/overnight split. Table 3 shows the estimated coefficients for the GARCH-M model with overnight dummy. The coefficient on the overnight dummy is not significant in the mean equation, but is significant at the 1% level in the variance equation. This suggests that there is no significant difference in the mean between the overnight and intraday returns while there is a difference in volatility. This is not surprising since we would expect that the overnight and intraday returns would average close to zero over time while, given the longer time period over which the overnight returns are calculated, the overnight volatility should be higher than the intraday volatility.

Having selected an appropriate null model, and obtained the model parameters, the bootstrapping process is straightforward. To generate a new realization of the null model we resample (with replacement) the standardized residuals (i.e., the residuals divided by the conditional standard deviation) from the model. Using these resampled residuals in conjunction with the model parameters, we iteratively construct a new time-series realization. The key point to note regarding this process is that because the residuals are resampled from the original series rather than randomly generated, no distribution is assumed for the error term. Instead, the distribution of standardized residuals for the new series will be the same as that in the original, and as a consequence the new realization will more closely match the time-series properties of the original series. Also, note that because we use dummies to flag overnight returns, the new realization will also exhibit the same structure of 81 intraday returns followed by an overnight return.

Having obtained a new realization of the null model, the bootstrap methodology is based on the comparison of conditional buy and sell returns (for a given rule) from original SPDR series with the conditional buy or sell returns

[12] Results are not reported but are available on request.
[13] We also fitted each of the three GARCH models using student-T innovations. In each case, the model failed to adequately remove autocorrelation from the residuals so these results are not reported.
[14] We wish to thank an anonymous referee for highlighting the importance of this.

generated from a simulated series using the null model. As per BLL (1992), the buy return is simply the mean return per period for all the periods where the rule is long, while the sell return is the mean return per period for all the periods in which the rule is short. The buy–sell return is simply the difference between the two means.

The buy–sell *p*-value is then just the proportion of 500 replications of the null model in which the buy–sell profit for the rule is greater than that of the original series. A *p*-value of zero indicates that none of the 500 replications of the null model have a buy–sell profit greater than that on the original series, for a given rule. So, for example, if the *p*-value for a given rule is less than 0.025, then we would reject the null that the profitability of the rule can be explained by the null model (at the 5% level).

3.2.2. STW's Reality Check bootstrapping methodology

STW (1999) use a formal test, the so-called White Reality Check bootstrap (the details of which are published in White (2000)), to test whether the profitability of the best trading rule, drawn from a wide universe of rules, is statistically significant, after adjusting for data snooping biases. Their data was daily returns for a range of stock market indices, including the Dow Jones Industrial Average and S&P 500. We apply the same methodology as our second test of the statistical significance of technical trading on our intraday SPDR data. The advantage of this second approach is that it explicitly adjusts for data snooping by considering the entire universe from which a rule was selected. If one rule is significantly profitable, but this is the only profitable rule in a vast universe of rules, then White's Reality Check will reduce the significance (since, on average, we expect some rules out of many will be profitable simply due to random variation). This is in contrast to the BLL approach, where each rule is evaluated independently.

Following STW (1999), we let $f_{k,t}$ ($k=1,\ldots,M$)be the period t return from the k-th trading rule (out of a universe of M rules), relative to the benchmark (which in this case is simply the SPDR return at time t). Then, the performance statistic we are interested in is the mean period relative return from the k-th rule, $\bar{f}_k=\sum_{t=1}^{T} f_{k,t}/T$, where T is the number of periods in the sample.

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

$$H_0 : \max_{k=1,..,M} \bar{f}_k \leq 0.$$

STW go on to illustrate how the stationary bootstrap of Politis and Romano (1994) can be used on the M values of $\bar{f}_k$ to evaluate the null hypothesis.[15] Essentially, 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. As per STW (1999), we set $B=500$. For the k-th rule, this yields B means, which we denote $\bar{f}^*_{k,b}(b=1,\ldots,B)$, from the B resampled time-series, where

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

The test then involves comparing the following two statistics:

$$\bar{V}_m = \max_{k=1,\ldots,M} [\sqrt{T}\bar{f}_k]$$

and

$$\bar{V}^*_{M,b} = \max_{k=1,\ldots,M} [\sqrt{T}(\bar{f}^*_{k,b} - \bar{f}_k)] \qquad , (\mathrm{b} = 1, \ldots, \mathrm{B})$$

We then compare $\bar{V}_M$ with the quantiles of the $\bar{V}^*_{M,b}$ 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. Thus, we are evaluating the performance of the best rule with reference to the performance of the whole universe.

[15] 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 4
Bootstrap results: full period

Rule type	Filter	Moving average	Support and resistance	Channel breakout	On-balance volume
Panel A: all rules					
Total number of rules	497	2049	1220	2040	2040
BLL *p*-value count (1%)	1	8	0	33	0
BLL *p*-value count (5%)	14	58	18	68	5
Panel B: best rule					
Nominal *p*-value	0.058	0.026	0.068	0.064	0.106
STW *p*-value	0.810	0.530	0.764	0.744	0.882
Average period return	0.000014	0.000020	0.000016	0.000016	0.000012
Average return per trade	0.134465	0.000356	0.001003	0.001056	0.000952
Total number of trades	4	2212	608	591	473
Number of winning trades	4	775	319	318	248
Number of losing trades	0	1437	289	273	225
Average periods per trade	7393.25	17.46	51.29	51.59	50.50
Average long return per trade	0.172375	0.000335	0.001114	0.001159	0.001108
Total number of long trades	2	1106	300	304	222
Number of winning long trades	2	396	155	163	121
Number of losing long trades	0	710	145	141	101
Average periods per long trade	12980.50	17.11	51.50	51.45	50.90
Average short return per trade	0.096556	0.000377	0.000895	0.000947	0.000813
Total number of short trades	2	1106	308	287	251
Number of winning short trades	2	379	164	155	127
Number of losing short trades	0	727	144	132	124
Average periods per short trade	1806	17.81	51.09	51.74	50.14

Technical trading rule results for the entire period of January 1, 2002 to December 31, 2003.

Using this approach allows us to compute a data snooping adjusted *p*-value for the best rule in each of the five rule families, relative to the universe of 7846 rules from which they are drawn.

4. Results

Our results comprehensively demonstrate that intraday technical analysis does not have value on the SPDR series in either the bear market of 2002 when the S&P 500 fell 21.2% or in the bull market of 2003 when the S&P 500 gained 21.9%. The most profitable rule out of each of the rule families we test (Filter, Moving Average, Support and Resistance, Channel Breakout, and On-Balance Volume) sometimes generate profits that are statistically significant prior to adjustment for data snooping bias, but once this is accounted for no statistical significance remains.

As far as we are aware, we are the first paper to compare and contrast these two popular bootstrapping techniques. The results displayed in Table 4 relate to the January 1, 2002 to December 31, 2003 period. Panel A contains the results from the BLL (1992) technique of fitting a null model (we use a GARCH-M model with dummy variables) to the original series and then generating 500 random series that have the same time-series properties as the original. The *p*-value count rows document the number of rules that a statistically significant. For rule to be statistically significant at the 1% (5%) level, there would have to be 5 (25) or fewer instances of the rule generating more profit on bootstrapped series than the original series. Panel B contains the results from the STW (1999) bootstrap technique. This approach evaluates the statistical significance of the returns generated by the best trading rule, in the context of the performance of the universe of rules from which the rule is selected. The nominal *p*-value is simply the Reality Check *p*-value for the best rule in each family, unadjusted for data snooping, i.e., it is the *p*-value assuming that the best rule is the only rule in the universe. The STW *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 difference between the two is an indication of the extent of data snooping bias.

Turning to the Panel A results, it is clear that very few rules are statistically significant based on the BLL (1992) technique. Of the total universe of 7846 rules, only 42 are statistically significant at the 1% level and only 163 are statistically significant at the 5% level. The Channel Breakout rule family is the most profitable while the On-Balance

Table 5
Bootstrap results: two sub-periods

Rule type	Filter	Moving average	Support and resistance	Channel breakout	On-balance volume
Panel A: 2002					
Nominal *p*-value	0.026	0.000	0.006	0.004	0.008
STW *p*-value	0.596	0.158	0.338	0.532	0.346
Average period return	0.000018	0.000037	0.000028	0.000022	0.000028
Average return per trade	0.000665	0.000965	0.001822	0.002344	0.001797
Total number of trades	534	736	300	177	302
Number of winning trades	209	270	170	108	149
Number of losing trades	325	466	130	69	153
Average periods per trade	35.89	25.86	51.35	25.00	62.88
Panel B: 2003					
Nominal *p*-value	0.208	0.264	0.476	0.938	0.312
STW *p*-value	0.988	0.984	0.994	0.986	0.986
Average period return	0.000017	0.000018	0.000012	0.000018	0.000017
Average return per trade	0.164277	0.001418	0.001099	0.00115	0.00183
Total number of trades	2	255	222	298	185
Number of winning trades	2	139	125	165	108
Number of losing trades	0	116	97	133	77
Average periods per trade	9349.00	51.43	50.23	50.97	50.00

Technical trading rule results for the first sup-period of entire period of January 1, 2002 to December 31, 2002 and the second sub-period of January 1, 2003 to December 31, 2003.

Volume rule family is the least profitable. The statistical tests used to generate the Panel A results do not account for data snooping bias, so we now focus on Panel B which contains results that do.

The results in Panel B relate the best rule for each rule family. The nominal *p*-values, which do not account for data snooping bias, indicate that the best Moving Average Rule, which involves a short moving average of one period, a long moving average of 75 periods and a two period delay before entry, has the strongest statistical significance (*p*-value of 0.026). None of the other rule families contain a rule that is statistically significant at the 5% level. This indicates that while the BLL (1992) and STW (1999) bootstrapping approach (prior to data snooping adjustment) produce the same overall result — the majority of technical trading rules have no value, there are some minor differences. The BLL (1992) technique leads to a minority of rule shows profits that are statistically significant at the 5% level within each rule family while the STW (1999) technique generates no statistically significant rules in any of the non-Moving Average Rule families.

Even though there is this minor discrepancy between the two techniques, the results are unambiguously clear once data snooping is accounted for. Data snooping bias is a big factor for all the rule families. When the statistical significance of the best rule from each rule family is adjusted for all other rules in our universe, the (STW) *p*-values decline dramatically to the point where none are even close to being statistically significant.

The Average Period Return is the average return over the 5-minute interval. This is related to the Average Return Per Trade by the Average Periods Per Trade. The best Filter Rule stands out from the best rules from the other four rule families in that it only generates four trades over the two-year period. This is because the best Filter Rule happens to be a rule that requires the price to move above (below) the highest (lowest) price over the last five periods by a fixed percentage. This means that any time there is a sustained movement of price in one direction over subsequent 5-minute periods (a common phenomena in our data) a relatively large move relative to the most recent 5-minute period price is required in the opposite direction for a trade to be closed. Hence this rule generates infrequent trades. The best On-Balance Volume Rule gives 473 trades over the 488 days in our sample. This equates to 0.97 round-trip trades or 1.9 trading signals (buy or sell) on average each day. The best Support and Resistance and Channel Breakout Rules give 2.4 and 2.5 trading signals per day respectively while the best Moving Average Rule gives an average of 9.0 trading signals per day.

Not surprisingly, given that it is the most statistically significant, the best Moving Average Rule is the most profitable, generating returns of 0.002% per period or 0.16% per day. Interestingly, the best Moving Average Rule generates a winning trade only 36% of the time, compared to the 50%+winning trade percentage for the other rule

families. This indicates that the rule is particularly effective at closing out losing trades before they generates big losses and/or not closing out winning trades too early. The best Moving Average Rule also stands out from the best rule in each of the other rule families, in that its short trades generate better returns than the long its long trades. However, the difference in performance is only minor (0.0377% is the Average Return be Short Trade versus 0.0335% for Long Trades) and each of the other rules perform better on the long side.

We now consider the performance of our rules during bull and bear markets. Splitting the data into two distinct years is an ideal way to achieve this as 2002 was a strong bear market during which the S&P 500 fell 21.2% while 2003 was a strong bull market where the S&P 500 gained 21.9%. We focus on the STW (1999) bootstrap results as our main interest is in determining whether there are any rules that produce statistically significant profits after accounting for data snooping bias.

The results in Table 5 indicate that none of the rule families have a best rule that is statistically significant in either bull or bear markets once data snooping bias is formally accounting for. However, all of the rule families contain a rule that performs dramatically better in the bear market of 2002, than the bull market of 2003. This is evident from both the nominal and STW p-values. All of the rule families have a best rule that is highly statistically significant before data snooping adjustment in 2002, but none are close to being statistically significant prior to this adjustment in 2003. The post-data snooping adjustment STW p-values are all insignificant in both years, but the 2002 p-values are all less than half their 2003 equivalents.

The best Filter Rule in 2002 has an Average Periods Per Trade of 35.88, which is similar to the best rules in the other rule families and dramatically less than that for the best Filter Rule over the entire sample. This is due to the nature of the Filter Rule. In 2002, the best Filter Rule is a rule that generates a buy (sell) when price moves above (below) the most recent price by a fixed percentage. This criterion is clearly more easily met when movement above (below) the highest (lowest) price over the past five periods as required by the best rule for all the data. The best Filter Rule in 2003 is similar to the best rule for the entire data set so the Average Periods Per Trade increases dramatically.

The best Moving Average Rule is the most profitable rule in both 2002 and 2003. In 2002, the best Moving Average Rule involves a short moving average of two periods, a long moving average of 75 periods, and a time delay of two periods between receiving a signal and entering a trade. This rule generates returns if 0.0037% per period or 0.30% per day, which is almost twice as much as that produced by the best Moving Average Rule over the entire period. In 2003 the best Moving Average Rule involves a short moving average of two periods, a long moving average of 50 periods, a time delay of one period between receiving a signal and entering a trade, and a holding period of 50 periods during which all other signals were ignored. This rule generates returns if 0.0018% per period or 0.15% per day, slightly less than those produced by the best moving average rule in the entire period.

In summary, there is no evidence that any of the 7846 technical trading rules we investigate generate statistically significant profits after data snooping bias is accounted for. This is made without consideration of transaction and market impact costs. Once these are taken into account, the performance of these rules would be even worse.

5. Conclusions

We investigate whether intraday technical analysis is profitable in the U.S. equity market. Surveys of market participants consistently find that technical analysis is more highly valued the shorter the time horizon, but prior work has focused on the profitability of long-term technical analysis. Studying intraday equity market technical analysis has several other aspects. Recent work has found evidence of order clustering which may provide theoretical support for technical analysis. Since this is a short-term phenomenon the impact of order clustering is likely to be more pronounced on intraday technical analysis.

We examine the profitability of 7846 rules from five major rule families (Filter, Moving Average, Support and Resistance, Channel Breakout, and On-Balance Volume Rules). These rules were well documented in the literature prior to the start of our sample so we propose that data snooping bias is minimised. As a further precaution against this, we formally adjust our test statistics to account for this.

Our results clearly demonstrate that intraday technical analysis is not profitable on the SPDR series over the 2002–2003 period. While the evidence that a small number of rules are profitable prior to any formal adjustment for data snooping bias, none of the five rule families produced a statistically significant profitable rule once this was taken into account. Given these results, we conclude that there is no evidence of consistent inefficiency in the intraday SPDR data.

References

Allen, F., Karjalainen, R., 1999. Using genetic algorithms to find technical trading rules. Journal of Financial Economics 51, 245–271.

Allen, H., Taylor, M.P., 1992. The use of technical analysis in the foreign exchange market. Journal of International Money and Finance 113, 301–314.

Alexander, S.S., 1961. Price movements in speculative markets: trends or random walks. Industrial Management Review 2, 7–26.

Andersen, T.G., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate Forecasts. International Economic Review 39, 885–905.

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

Blair, B.J., Poon, S., Taylor, S.J., 2001. Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high frequency returns. Journal of Econometrics 105, 5–26.

Bollerslev, T., Chou, R.Y., Kroner, K.F., 1992. ARCH modelling in finance: a review of the theory and empirical evidence. Journal of Econometrics 52 (1–2), 5–60.

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

Carter, R.B., Van Auken, H.E., 1990. Security analysis and portfolio management: a survey and analysis. Journal of Portfolio Management 81–85 Spring.

Chen, Yi-Ting, 2003. On the discrimination of competing GARCH type models for Taiwan stock index returns. Academia Economic Press Papers 31 (3), 369–405 September.

Corrado, C.J., Lee, S.H., 1992. Filter rule tests of the economic significance of serial dependence in daily stock returns. Journal of Financial Research 154, 369–387.

Day, T.E., Wang, P., 2002. Dividends, nonsynchronous prices, and the returns from trading the Dow Jones Industrial Average. Journal of Empirical Finance 9, 431–454.

Elton, E.J., Gruber, M.J., Comer, G., Li, K., 2002. Spiders: where are the bugs? Journal of Business 753, 453–472.

Fong, W., Yong, L., 2005. Chasing trends: recursive moving average trading rules and Internet stocks. Journal of Empirical Finance 12 (1), 43–76.

Gartley, H.M., 1935. Profits in the Stock Market, Lambert–Gann Publishing. Pomeroy, Washington.

Granville, J., 1963. Granville's New Key to Stock Market Profits. Prentice Hall, Englewood Cliffs, N.J.

Hol, E., Koopman, J., 2002. Stock index volatility forecasting with high frequency data. Tinbergen Institute Discussion Papers. 02-068/4.

Kavajecz, K.A., Odders-White, E.R., 2004. Technical analysis and liquidity provision. Review of Financial Studies 17 (4), 1043–1071.

Kwon, K-Y., Kish, R.J., 2002. A comparative study of technical trading strategies and return predictability: an extension of Brock, Lakonishok, and LeBaronk (1992) using NYSE and NASDAQ indices. Quarterly Review of Economics and Finance 42 (3), 611–631.

Lakonishok, J., Smidt, S., 1988. Are seasonal anomalies real? A ninety-year perspective. Review of Financial Studies 3, 431–468.

Lee, D.D., Chan, H., Faff, R.W., Kalev, 2003. Short-term contrarian investing — it is profitable — yes and no. Journal of Multinational Financial Management 13, 385–404.

Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–467.

Lo, A.W., Mamaysky, H., Wang, J., 2000. Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. Journal of Finance 55, 1705–1770.

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.

Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series models. Biometrika 65, 297–303.

McLeod, A.I., Li, W.K., 1983. Diagnostic checking ARMA time series models using squared-residual autocorrelations. Journal of Time Series Analysis 4, 269–273.

Malkiel, B., 1981. A Random Walk Down Wall Street, 2nd ed. Norton, New York.

Miller, E.M., 1990. Atomic bombs, the Depression and equilibrium. Journal of Portfolio Management 37–41.

Morris, G.L., 1995. Candlestick charting explained: timeless techniques for trading stocks and futures, 2nd ed. McGraw–Hill Trade, New York.

Neely, C.J., Weller, P.A., 2003. Intraday technical trading in the foreign exchange market. Journal of International Money and Finance 22, 223–237.

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

Osler, C., 2000. Support for resistance: technical analysis and intraday exchange rate. Federal Reserve Bank of New York Economic Policy Review, pp. 53–68.

Osler, C., 2003. Currency orders and exchange-rate dynamics: an explanation for the predictive success of technical analysis. Journal of Finance 585, 1791–1820.

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

Pesaran, M.H., Timmerman, A., 1995. Predictability of stock returns: robustness and economic significance. Journal of Finance 50, 1201–1228.

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

Ready, M.J., 2002. Profits from technical trading rules. Financial Management 313, 43–61.

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

Sweeney, R.J., 1988. Some new filter tests, methods and results. Journal of Financial and Quantitative Analysis 23, 285–300.

White, H., 2000. A Reality Check for data snooping. Econometrica 68 (5), 1097–1126.

Wycjoff, R., 1910. Studies in tape reading. Fraser Publishing Company, Burlington.