NORTH-HOLLAND

Financial Services Review 10 (2001) 173–186

FINANCIAL SERVICES REVIEW

Evaluating a stock market timing strategy: the case of RTE Asset Management

Ahmet Tezel*, Ginette McManus

Department of Finance, Haub School of Business, Saint Joseph's University, 5600 City Avenue, Philadelphia, PA 19131, USA

Received 3 April 2001; received in revised form 15 June 2001; accepted 12 September 2001

Abstract

Market timing is a popular active investment strategy that promises to beat the market. However, the evidence on the ability of timers to outperform the market is mixed. This paper provides strong supporting evidence of the timing ability of RTE Asset Management by investigating the implemented buy and sell recommendations derived from its proprietary computerized model over the 1979–1999 period and several subperiods. We use various performance-evaluation methodologies that investors can easily implement. The evidence obtained on market timing skills is essentially invariant to the evaluation method used if the analysis is performed over a long time period.

JEL classification: G14

Keywords: Market timing; Risk-adjusted performance measures; Nonparametric and parametric performance-evaluation models

1. Introduction

Investors are always yearning for investment strategies that promise to outperform the market. One popular strategy, referred to as asset class reallocation or market timing, involves active investment techniques that seek to enhance performance by shifting funds to asset classes that are expected to outperform other asset classes or produce above-average expected returns. Market timers employ quantitative models in assessing the relative attractiveness of asset classes, ranging from equity valuation techniques to cyclical indicators. Other market timers apply optimization models based on modern portfolio

* Corresponding author. Tel.: +1-610-660-1670; fax: +1-610-110-1986.
E-mail address: atezel@sju.edu (A. Tezel).

1057-0810/01/$ – see front matter
PII: S1057-0810(01)00082-8

theory, an approach requiring forecasts of standard deviations and correlations among asset classes, in addition to expected returns.

All market timing strategies are based on the hypothesis that stock market returns are not random but are, in fact, partly predictable (Brock, Lakonishok, & LeBaron, 1992; Lee, 1997; Sullivan, Timmermann, & White, 1999). Despite the predictability of stock returns, the evidence on the ability of market timers to outperform the market (or their benchmark) is mixed. In this paper we add to the debate by extensively testing the performance and timing ability of a professional market timer known as RTE Asset Management, a division of Rightime Econometrics, Inc. We use the actual implemented equity trade dates and weekly asset allocations to calculate the timed returns for the clients of Lincoln Investment Planning, Inc., a company providing investment advisory services that include the Rightime Funds and the Lincoln Timing Program. Rich and Reichenstein (1993), Larsen and Wozniak (1995), Prather and Bertin (1998), and Copeland and Copeland (1999) all proposed and tested a successful market timing strategy that could have been implemented by investors. In contrast, we investigate a market timing strategy that has actually been implemented by investors in the real-world. As such, this paper provides more direct evidence of market timing ability, if present.

The remainder of this paper is organized as follows. In Section 2, we provide a review of the relevant current literature. Section 3 describes the market timing strategy of RTE and the data used in the analysis of its performance and market timing ability. Section 4 contains several performance-evaluation methodologies classified in four categories: (1) performance-evaluation using total returns; (2) performance-evaluation using risk-adjusted returns; (3) performance-evaluation using a nonparametric test of market timing ability; and (4) performance-evaluation using a parametric test of market timing ability. We discuss the results in this section. A summary and conclusion is presented in Section 5.

2. Literature review

There have been many studies on the ability of professional investment managers to time the market since the early 1970s. Recent studies on the market timing skills of mutual funds by Chen, Lee, Rahman, and Chan (1992), Grinblatt and Titman (1994), Malkiel (1995), Daniel, Grinblatt, Titman, and Wermers (1997), Becker, Ferson, Myers, and Schill (1999), and Goetzmann, Ingersoll, and Ivkovic (2000) find little evidence of market timing ability. A similar conclusion is reached in Cumby and Glen (1990), Eun, Kolodny, and Resnick (1991), Droms and Walker (1994), and Kao, Cheng, and Chan (1998) who studied the timing and selectivity skills of managers of U.S.-based international mutual funds. In contrast, Lee and Rahman (1990), Weigel (1991), Ferson and Schadt (1996), as well as Bello and Janjigian (1997) find some evidence of market timing skills.

The lack of strong results in support of market timing ability may be explained by the fact that few mutual funds specialize in market timing. Market timing information is mostly delivered to investors via investment newsletters (Kane & Marks, 1990). Graham and Harvey (1996, 1997) evaluate the performance of newsletter strategies tracked by *Hubert Financial Digest* and find some evidence of market timing ability for newsletters that are on a hot streak.

More impressive results in support of market timing as a successful strategy is obtained by Wagner, Shellans, and Paul (1992) and Chance and Hemler (1998). Both studies use data from the MoniResearch Corporation, a newsletter publishing timing signals and monitoring the performance of professional market timers. Although the market timers evaluated in these studies are professionals who actually execute their buy and sell recommendations for clients, the timing signals provided by the MoniResearch newsletter do not necessarily correspond to the actual dates on which trades were implemented for the clients of the timers.

Similar to Chance and Hemler (1998), but unlike Wagner et al. (1992), we use more appropriate performance evaluation tests over a longer time period (Brocato & Chandy, 1994). If a market timing strategy is to work, it will likely occur primarily over a lengthy period of time (Beehover & Varikooty, 1991). In addition, our tests have more power to detect market timing ability since, unlike prior research, we use weekly instead of monthly data to measure returns (Goetzmann et al., 2000). Chance and Hemler (1998) are the first to use daily (or more frequent) observations and to demonstrate that the frequency with which timers' recommendations are observed is crucial in detecting significant market timing ability.

3. Data

3.1. RTE and its market timing strategy

RTE Asset Management is an independent registered investment advisor to its Rightime Family of Funds since 1979 and also markets the timing recommendations of its computerized econometric forecasting model, the RTE Market Model (RTMM®), through independent brokers/dealers and financial planners who implement them using mutual funds.

RTE follows a risk-averse approach to managing investment advisory accounts. More specifically, RTE believes that consistent performance and the preservation of capital (avoiding losses in down-markets) can be achieved with reduced risk by systematically reallocating client funds to the asset class with the most favorable prospects for superior future returns. According to RTE, above-average return can be achieved with less risk using a stock market timing strategy. RTE is a practitioner of market timing in its purest form, being always fully invested in one of two broad asset classes: stocks and cash.

The RTMM® is a weighted-average of over 60 individual indicators, which may be broadly grouped into four categories: fundamental, monetary, sentiment, and momentum. Fundamental indicators are generally related to inflation expectations and include the Consumer Price Index, the CRB Futures Index, the Dow Jones Futures Index, individual commodity prices, and the U.S. dollar. Monetary indicators are generally related to the monetary policy of the Federal Reserve Bank, interest rate movements, the yield curve and yield spreads as well as bond technical indicators such as volume, advances, declines, new highs and new lows. Indicators of investor sentiment in both the bond and stock markets are generally contrarian in nature and include various short interest ratios, call/put volume and premium ratios, and big block activity. Momentum indicators are related to actual market activity and do not include any price indices. RTE attaches the greatest weights to fundamental and monetary indicators and lesser weights to sentiment and momentum

indicators. Therefore, its forecasting model is economically based and quantitatively driven, without any subjective adjustment. The RTMM® has been under continuous development since 1969 and its asset allocation decisions actively used to manage investment advisory accounts since 1979. Trades are generally executed during the second or third business day following the generation of buy and sell signals and asset allocation decisions are reviewed weekly. Today, RTE manages approximately $2 billion in client assets.

3.2. Return calculations and composition

Over the 1979–1999 time period, RTMM® has given 48 timing decisions: 24 buys and 24 sells, averaging about two signals a year. Using RTMM® buy and sell recommendations and the specific dates these recommendations were implemented for the clients of Lincoln Investment Planning, Inc., we calculate total weekly timed returns, R_T, for a 20-year period from June 18, 1979 to February 12, 1999 (1,025 weeks). We use weekly returns because RTE timing recommendations are updated weekly.

Total timed returns are measured using the Russell 2000 (a total return index) and the S&P 500 (adjusted for dividends). The timed return figures do not include advisory or management fees of about 2.3% annually for accounts of $100,000 or less. Note that trade costs are small (probably less than 0.3%) since most of the capital of RTE and Lincoln clients is invested in mutual funds allowing switching (at least a few switches per year) to a money market account at no cost. The 3-month T-bill is used to proxy the risk-free rate, R_f, and the Russell 2000 as well as the S&P 500 indices are used for measuring total weekly market returns, R_m. While the S&P 500 Index is widely recognized as a benchmark for evaluating performance, it proxies large capitalization stocks. Since RTE uses the Lipper Growth Fund Index as a benchmark, we consider the Russell 2000 Index a more representative benchmark for the types of funds (growth and small capitalization stocks) that are timed by the investor. Therefore, results using the S&P 500 that are quantitatively similar to those using the Russell 2000 are not reported to conserve space.

To investigate whether the performance and market timing ability of RTE persist over time, the entire 20-year period is divided into four relatively equal time periods. The first subperiod ranges from June 18, 1979 to August 6, 1984 and the second one extends from August 6, 1984 to May 22, 1989. The third subperiod is from May 22, 1989 to November 30, 1993 while the fourth subperiod is from November 30, 1993 to February 12, 1999. The length of the subperiods is not exactly 5 years since they were created using the actual signal (trading) dates. The sensitivity of our results to the length of the subperiods was investigated by examining the performance and timing ability of RTE using six equal subperiods of 3 years starting in 1981. Since we did not observe any significant changes, results are reported for the entire 20-year time period as well as the four subperiods only.

4. Performance-evaluation: methodology and results

We employ several methodologies since there is a lack of consensus on a generally accepted performance-evaluation model. We first analyze the performance of RTE using

total returns. However, this performance measure can be highly misleading since it ignores risk. We then evaluate RTE's performance with various measures of risk-adjusted returns. Finally, because risk-adjusted measures do not involve any statistical testing, we perform a nonparametric and a parametric test of RTE's market timing ability.

4.1. Performance-evaluation using total returns

One conventional performance-evaluation method employs total returns. Table 1 reports average total weekly excess returns of RTE, $R_T - R_f$, and its benchmark, $R_m - R_f$, using the Russell 2000 Index as R_m, for the entire 20-year time period as well as the four subperiods. More importantly, Table 1 shows the composition of these total returns under the following four investment scenarios: RTE is in-the-market and the market is up (down) and RTE is out-of-the-market and the market is up (down). An up-market is defined as one where the market return equals or exceeds the risk-free rate, $R_m - R_f \geq 0$, and a down-market is characterized by a market return less than the risk-free rate, $R_m - R_f < 0$.

The results in Table 1 indicate that RTE's total weekly returns exceed those of its benchmark in all of the time periods tested. Over the 1979–1999 period, RTE was in-the-market 61.4% of the time and realized a total weekly excess return of 0.22%, which exceeds the market weekly excess return of 0.10%. More specifically, the timer was in-the-market in up-market conditions 36.6% of the time earning an average weekly excess return of 1.56% and in-the-market 24.8% of the time in down-market conditions earning an average weekly excess return of −1.42%.

While RTE appears to be successful at capturing a significant portion of up-market conditions, its value added as a market timer is also reflected in the losses avoided when out-of-the-market (Shilling, 1992; Wagner, 1997). Over the entire period tested, RTE avoided a large average weekly excess return of −2.23% despite foregoing a excess return of 1.48%. The difference between the average weekly excess return of −2.23% when RTE is out-of-the-market in down-market conditions (the loss avoided) and the average weekly excess return of 1.48% when RTE is out-of-the-market in up-market conditions (the forgone opportunity) is statistically significant at the 0.5% level or better ($t = 3.61$). This test result supports that it is profitable to be out-of-the-market in down-market conditions even if many of the up-market weeks are missed completely.

The only *t*-statistic in Table 1 not significant at the conventional 5% level or better is that of the second subperiod. RTE eliminated on April 6, 1987 its stock exposure of 2 years, thus successfully avoiding the crash of October 19, 1987. After the Federal Reserve quickly lowered interest rates in response to the stock market crash, RTE resumed its equity exposure on November 9, 1987. These allocation decisions contributed greatly to its superior returns against the benchmark. However, RTE produced several bad timing decisions in 1988–1989, therefore losing the value added before and during the crash.

The economic value added of RTE's market timing strategy is also investigated by calculating weekly excess returns net of management fees of 0.04% per week (2.3% annually/52 weeks). The weekly excess returns net of management fees are then compounded to show that the investor's terminal wealth, per $100 of initial investment, in the active market timing strategy of RTE is $2,134 which exceeds the terminal wealth of

Table 1
RTE performance on the basis of total returns and the Russell 2000 as the benchmark

Investment scenarios	Percentage of time	Average excess returns	Weighted-average benchmark, $R_m - R_f$	Excess returns RTE, $R_T - R_f$	t-statistic[a]
Panel A: Returns for the entire period (June 18, 1979 to February 12, 1999)					
RTE in; up-market	36.6%	1.56%	0.57%	0.57%	
RTE in; down-market	24.8%	−1.42%	−0.35%	−0.35%	
RTE out; up-market	19.9%	1.48%	0.29%	0	
RTE out; down-market	18.7%	−2.23%	−0.41%	0	3.61**
Total weekly excess returns			0.10%	0.22%	
Panel B: Returns for the first subperiod (June 18, 1979 to August 6, 1984)					
RTE in; up-market	28.8%	1.92%	0.55%	0.55%	
RTE in; down-market	16.5%	−1.62%	−0.27%	−0.27%	
RTE out; up-market	26.2%	1.64%	0.43%	0	
RTE out; down-market	28.5%	−2.27%	−0.64%	0	2.32*
Total weekly excess returns			0.07%	0.28%	
Panel C: Returns for the second subperiod (August 6, 1984 to May 22, 1989)					
RTE in; up-market	40.0%	1.46%	0.58%	0.58%	
RTE in; down-market	30.4%	−1.27%	−0.38%	−0.38%	
RTE out; up-market	16.8%	1.34%	0.22%	0	
RTE out; down-market	12.8%	−2.33%	−0.30%	0	1.23
Total weekly excess returns			0.12%	0.20%	
Panel D: Returns for the third subperiod (May 22, 1989 to November 30, 1993)					
RTE in; up-market	46.6%	1.50%	0.70%	0.70%	
RTE in; down-market	33.5%	−1.38%	−0.46%	−0.46%	
RTE out; up-market	8.9%	1.16%	0.10%	0	
RTE out; down-market	11.0%	−2.34%	−0.26%	0	2.74**
Total weekly excess returns			0.08%	0.24%	
Panel E: Returns for the fourth subperiod (November 30, 1993 to February 12, 1999)					
RTE in; up-market	32.4%	1.42%	0.46%	0.46%	
RTE in; down-market	20.2%	−1.51%	−0.31%	−0.31%	
RTE out; up-market	26.1%	1.49%	0.39%	0	
RTE out; down-market	21.3%	−2.05%	−0.44%	0	1.70*
Total weekly excess returns			0.10%	0.15%	

[a] t-statistic of the difference between the average returns in up- and down-markets when RTE is out-of-the-market.

*,** Indicate significance at the 0.5 and 5.0% levels or better (one-tail test), respectively.

$815 in the passive strategy for the entire 20-year period. The results are presented graphically in Fig. 1. The market timing strategy of RTE consists in trading T-bills and the Russell 2000 Index on the specific dates that the buy and sell recommendations of RTMM® are implemented. The passive strategy consists of buying-and-holding the Russell 2000 Index. Fig. 1 also reports that the weekly compounded annual return of RTE over the entire

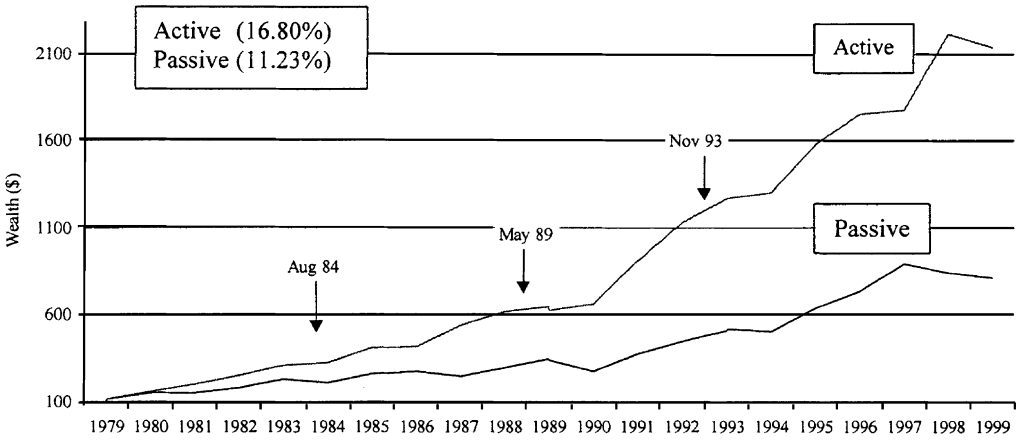

Fig. 1. Terminal wealth and compounded annual returns under several trading strategies.

20-year period is 16.80%, which is higher than the compounded annual return of the buy-and-hold strategy of 11.23%. Overall, the results over all time periods indicate that if the investor's horizon is sufficiently long and management fees (as well as trade costs) are sufficiently low, RTE's market timing strategy produces higher total returns than a passive (buy-and-hold) alternative.

4.2. Performance-evaluation using risk-adjusted return measures

We use four measures to estimate the risk-adjusted returns of RTE: the Sharpe measure, the Modigliani and Modigliani measure, and the Graham and Harvey measures.

The most common measure of risk-adjusted return is the Sharpe (S) ratio. RTE Sharpe Index, S, measures its total excess return per unit of total risk where risk is defined as the standard deviation:

$$S = \frac{R_T - R_f}{\sigma(R_T - R_f)} \tag{1}$$

where $R_T - R_f$ and $\sigma(R_T - R_f)$ are the mean and standard deviation of RTE excess return, respectively. RTE outperforms (underperforms) its benchmark on a risk-adjusted basis if its Sharpe ratio is higher (lower) than that of the benchmark.

Modigliani and Modigliani (MM, 1997) and Graham and Harvey (GH, 1997) propose risk-adjusted performance measures giving by how much, in basis points, RTE's return outperforms (underperforms) the return of its benchmark. The MM measure for RTE is obtained by calculating the mean weekly return of a portfolio with a constant proportion (σ_m/σ_T) earning R_T and the remaining proportion ($1 - \sigma_m/\sigma_T$) earning R_f and rearranging the terms as follows:

$$\mathrm{MM} = R_f + \left(\frac{\sigma_m}{\sigma_T}\right)(R_T - R_f) \tag{2}$$

where σ_T and σ_m are the standard deviations of the mean weekly returns of RTE, R_T, and the market index, R_m, respectively. The MM measure levers up the return of RTE, R_T, by borrowing at the risk-free rate so that it has the same level of risk or standard deviation as the benchmark return, $\sigma(R_T) = \sigma(R_m)$. The adjusted return of RTE, MM, is then compared directly with the return of the benchmark over the same time period. RTE outperforms (underperforms) its benchmark on a risk-adjusted basis if the difference between its adjusted return and that of the benchmark, $MM - R_m$, is positive (negative).

Graham and Harvey first measure, GH1, is obtained by subtracting from RTE's mean weekly return, R_T, the matched return of a portfolio with a constant proportion (σ_T/σ_m) earning R_m and the remaining proportion $(1 - \sigma_T/\sigma_m)$ earning R_f, and rearranging the terms as follows:

$$GH1 = R_T - \left[R_f + \left(\frac{\sigma_T}{\sigma_m}\right)(R_m - R_f)\right] \tag{3}$$

GH1 measures the difference between the mean weekly return of RTE and that of the benchmark, $R_T - R_m$, where the mean weekly return of the benchmark is levered up or down using the T-bill so that it has the same risk level or standard deviation as RTE's return. RTE outperforms (underperforms) its matched benchmark on a risk-adjusted basis if GH1 is positive (negative). Note that the GH1 measure divided by the risk or standard deviation of RTE's weekly returns is equal to the difference between the Sharpe ratio of RTE and that of the market. It is called ''ratio'' in Chance and Hemler (1998) who demonstrate that it is an excellent measure to rank the risk-adjusted performance of market timers.

The second measure introduced by Graham and Harvey, GH2, is the difference between the mean weekly return of RTE and that of the benchmark, $R_T - R_m$, where the return of RTE is levered up or down using the T-bill so that it has the same standard deviation as the return of the benchmark:

$$GH2 = \left[R_f + \left(\frac{\sigma_m}{\sigma_T}\right)(R_T - R_f)\right] - R_m \tag{4}$$

As with the first measure, RTE outperforms (underperforms) its benchmark on a risk-adjusted basis if GH2 is positive (negative). In addition, note that GH2 is essentially equal to $MM - R_m$ (Graham & Harvey, 1997).

The risk-adjusted measures of RTE's performance relative to the two benchmarks are presented in Table 2. Results in Table 2 confirm that RTE outperformed its benchmark, as measured by the Russell 2000 Index, on a risk-adjusted basis for the entire 20-year period and all subperiods. In Panel A, for the period 1979–1999, the weekly Sharpe ratio is 0.144 for RTE while that of the market is 0.042. Similarly, in Panel B, the risk-adjusted weekly return of RTE for the entire period, $MM = 0.47\%$, is well above that of the market benchmark, $R_m = 0.23\%$. Finally, both weekly adjusted returns GH1 (0.16%) and GH2 (0.24%) are positive for the entire period (and for all subperiods).

The risk-adjusted performance of RTE is less impressive when the S&P 500 is used as the market benchmark. In particular, its Sharpe ratio of 0.161 is lower than that of the market for the fourth subperiod. Similarly, the MM measure of RTE is less than that of the benchmark and the GH measures are negative in that subperiod. From RTE equity trades, we notice that

Table 2
Risk-adjusted measures of RTE performance

Time period	Sharpe Index–Russell 2000		Sharpe Index–S&P 500					
	RTE	Benchmark	RTE	Benchmark				
Panel A: Sharpe (S) measures								
Entire period[a]	0.144	0.042	0.126	0.089				
First subperiod[b]	0.187	0.028	0.108	0.028				
Second subperiod[c]	0.128	0.051	0.133	0.098				
Third subperiod[d]	0.146	0.043	0.106	0.075				
Fourth subperiod[e]	0.115	0.048	0.161	0.173				
	MM–Russell 2000		MM–S&P 500		GH–Russell 2000		GH–S&P 500	
	RTE	Benchmark	RTE	Benchmark	GH1	GH2	GH1	GH2
Panel B: Modigliani and Modigliani (MM) measures as well as Graham and Harvey (GH) measures								
Entire period[a]	0.47%	0.23%	0.42%	0.34%	0.16%	0.24%	0.06%	0.09%
First subperiod[b]	0.67%	0.28%	0.47%	0.28%	0.24%	0.39%	0.13%	0.20%
Second subperiod[c]	0.45%	0.26%	0.51%	0.41%	0.12%	0.19%	0.06%	0.10%
Third subperiod[d]	0.39%	0.19%	0.30%	0.24%	0.17%	0.21%	0.05%	0.06%
Fourth subperiod[e]	0.35%	0.20%	0.40%	0.42%	0.09%	0.15%	−0.02%	−0.02%

[a] Entire period: June 18, 1979 to February 12, 1999.
[b] First subperiod: June 18, 1979 to August 6, 1984.
[c] Second subperiod: August 6, 1984 to May 22, 1989.
[d] Third subperiod: May 22, 1989 to November 30, 1993.
[e] Fourth subperiod: November 30, 1993 to February 12, 1999.

it reversed its stock exposure of over 1.5 years on November 30, 1993 and it did not jump back into stocks before February 7, 1995. Meanwhile the S&P 500 Index was reaching new record highs! Similar bad timing decisions were made in 1997: RTE went out-of-the-market for the rest of the year on April 15, 1997 while the S&P 500 gained over 30% during 1997.

4.3. Performance-evaluation using a nonparametric test of market timing ability

We use a nonparametric procedure developed by Pesaran and Timmerman (1992) and utilized by Paláez (1998) for evaluating the market timing ability of RTE. Pesaran and Timmerman propose a test on the differences between the actual and expected percentage of correct predictions of market conditions. The test is referred to as the S-statistic, calculated as follows:

$$S = \frac{P - P^*}{[\sigma_P^2 - \sigma_{P^*}^2]^{1/2}} \tag{5}$$

where $P = [N_{\text{in;up}} + N_{\text{out;down}}]/N_{\text{tot}}$ is the percentage of the time that both up- and down-market conditions are correctly forecasted by the RTMM®. $N_{\text{in;up}}$ ($N_{\text{out;down}}$) is the number of weeks RTE is in-the-market (out-of-the-market) in up (down)-market conditions, and N_{tot} is the total number of weeks in the test period. On the other hand, $P^* = (P_{\text{in}})(P_{\text{tot}}) +$

$(1 - P_{in})(1 - P_{tot})$ is the predicted percentage of correct market forecasts under the null hypothesis of no market timing ability. $P_{in} = [N_{in;up} - N_{in;down}]/N_{tot}$ is the percentage of the time RTE is in-the-market and $P_{tot} = N_{tot;up}/N_{tot}$ is the percentage of the time the stock market is up. $\sigma_P^2 = (1/N_{tot})P^*(1 - P^*)$ is the variance of P. $\sigma_{P^*}^2 = (1/N_{tot})(2P_{tot} - 1)^2$-$(P_{in})(1 - P_{in}) + (1/N_{tot})(2P_{in} - 1)^2(P_{tot})(1 - P_{tot})$ is the variance of P^*. Only positive and statistically significant values of S provide evidence of market timing skills, since timing ability requires the percentage of actual correct predictions, P, to exceed the predicted proportion under the null hypothesis of no market timing ability, P^*.

As shown in Panel A of Table 3, the market, as measured by the Russell 2000 Index, was up 56.5% [$P_{tot} = 579/1{,}025$] of the time and down 43.5% of the time over the period 1979–1999. As mentioned before, RTE was in-the-market 61.4% [$P_{in} = 629/1{,}025$] of the time and out-of-the-market 38.6% of the time. More importantly, the predictions of the RTMM® were correct 55.3% [$P = (375 + 192)/1{,}025$] of the time. The expected percentage of

Table 3
Nonparametric test of RTE market timing ability

	Russell 2000					S&P 500				
	N_{in}	N_{tot}	P	P^*	S	N_{in}	N_{tot}	P	P^*	S
Panel A: *S*-statistic for the entire period (June 18, 1979 to February 12, 1999)										
Up-market	375	579				368	574			
Down-market	254	446				261	451			
Total	629	1,025	55.3%	51.5%	2.55**	629	1,025	54.4%	51.4%	2.04*
Panel B: *S*-statistic for the first subperiod (June 18, 1979 to August 6, 1984)										
Up-market	77	147				65	130			
Down-market	44	120				56	137			
Total	121	267	57.3%	49.5%	2.57**	121	267	54.7%	50.1%	1.50
Panel C: *S*-statistic for the second subperiod (August 6, 1984 to May 22, 1989)										
Up-market	100	142				97	144			
Down-market	76	108				79	106			
Total	176	250	52.8%	52.8%	0.01	176	250	49.6%	53.1%	−1.23
Panel D: *S*-statistic for the third subperiod (May 22, 1989 to November 30, 1993)										
Up-market	110	131				107	126			
Down-market	79	105				82	110			
Total	189	236	57.6%	53.3%	1.67*	189	236	57.2%	50.7%	1.99*
Panel E: *S*-statistic for the fourth subperiod (November 30, 1993 to February 12, 1999)										
Up-market	88	159				99	174			
Down-market	55	113				44	98			
Total	143	272	53.7%	50.4%	1.09	143	272	56.2%	50.7%	1.90*

N_{in}: number of weeks that RTE is in-the-market; N_{tot}: number of weeks in the test period; P: percentage of time that RTE correctly forecasted market conditions; P^*: expected proportion of correct forecasts under the null hypothesis of no forecasting ability. $S = (P - P^*)/[\sigma_P^2 - \sigma_{P^*}^2]^{1/2}$, where σ_P^2 and $\sigma_{P^*}^2$ are the variance of P and P^*, respectively.

*,** Indicate significance at the 0.5 and 5.0% levels or better (one-tail test), respectively.

correct predictions, P^*, is 51.5%. Overall, RTE was very successful with its market forecasts, since the null hypothesis of no forecasting ability can be rejected at the 0.5% level of significance of a one-tail test ($S = 2.55$). RTE was also very successful with its market forecasts during the first subperiod. The S-statistic of 2.57 for the 1979–1984 period is statistically significant at the 0.5% level. Less significant forecasting ability is implied for the third subperiod ($S = 1.67$, significant at the 5% level or better). However, the hypothesis of no forecasting ability cannot be rejected for the other subperiods at the 10% level of significance or better.

Table 3 also reports RTE's ability to forecast the direction of the market using the S&P 500 Index. While the S-statistic of 2.04 for the entire period is significant at the 5% level or better, we find stronger evidence of RTE's market timing ability when using the Russell 2000 to proxy the market. This result is consistent with those of Chance and Hemler (1998) and Kester (1990) who report that market timing ability appears the strongest relative to NASDAQ (small-firm stocks) and the weakest relative to the S&P 500 (large-firms stocks). This result is also persistent across the various time subperiods, except for the last one where the S-statistic using the S&P 500 is significant at the 5% level or better ($S = 1.90$) whereas the S-statistic using the Russell 2000 is not significant.

4.4. Performance-evaluation using a parametric test of market timing ability

We use a regression-based performance-evaluation model that clearly separates the market timing ability from nonmarket timing skills such as stock selection, hedging, and index arbitrage. The two-beta regression model is utilized by Kao et al. (1998) to investigate the selectivity and market timing ability of managers of international mutual funds and is expressed as follows:

$$R_{Tt} - R_{ft} = \alpha + \beta_1(R_{mt} - R_{ft}) + \beta_2 Y_t + e_t \tag{6}$$

where $R_{Tt} - R_{ft}$ = weekly excess return of RTE at time t; $R_{mt} - R_{ft}$ = weekly excess return of the market at time t; $Y_t = \max[0, -(R_{mt} - R_{ft})]$; α = intercept term; and e_t = random error term at time t. When $R_{mt} \geq R_{ft}$, Eq. (6) becomes $R_{Tt} - R_{ft} = \alpha + \beta_1(R_{mt} - R_{ft}) + e_t$, where β_1 is the up-market beta, β_u. When $R_{mt} < R_{ft}$, Eq. (6) becomes $R_{Tt} - R_{ft} = \alpha + (\beta_1 - \beta_2)(R_{mt} - R_{ft}) + e_t$ (since $R_{mt} - R_{ft} = -Y_t$), where $(\beta_1 - \beta_2)$ is the down-market beta, β_d. Therefore, $\beta_u - \beta_d = \beta_2$.

β_2 represents the change in the beta or systematic risk level of the market timer when market conditions change. Since the systematic risk of a market timer does not stay constant over time, market timing ability is measured as the change in beta from an up-to a down-market condition. For a skilled market timer, the up-market beta, β_u, should be greater than the down-market beta, β_d, that is β_2 must be positive and statistically significant. A positive and statistically significant t-statistic for β_2 provides evidence of successful market timing, since the expected value of β_2 is zero under the null hypothesis of no market timing ability. The intercept term α is a measure of the nonmarket timing skills of the market timer, after filtering out its market timing ability. Since RTE does not pick mutual funds but only provides a forecast of the market direction, α is likely to measure other types of activities than selectivity, namely hedging and index arbitrage (Weigel, 1991).

Table 4
Parametric test of RTE market timing ability

Dependent Variable	α	β_1	β_2[a]	R^2	N
Panel A: Regression coefficients using the entire period (June 18, 1979 to February 12, 1999)					
$R_{Tt} - R_{ft}$	−0.002 (−4.41)**	0.716 (25.19)	0.473 (11.22)	50.7	1,025
Panel B: Regression coefficients using the first subperiod (June 18, 1979 to August 6, 1984)					
$R_{Tt} - R_{ft}$	−0.002 (−2.19)*	0.680 (12.49)**	0.519 (6.00)**	47.6	267
Panel C: Regression coefficients using the second subperiod (August 6, 1984 to May 22, 1989)					
$R_{Tt} - R_{ft}$	−0.002 (−3.18)**	0.755 (13.80)**	0.568 (7.64)**	56.8	250
Panel D: Regression coefficients using the third subperiod (May 22, 1989 to November 30, 1993)					
$R_{Tt} - R_{ft}$	−0.003 (−3.12)**	1.012 (18.81)**	0.578 (6.87)**	75.2	236
Panel E: Regression coefficients using the fourth subperiod (November 30, 1993 to February 12, 1999)					
$R_{Tt} - R_{ft}$	0 (−0.23)	0.471 (8.06)**	0.168 (1.94)*	37.9	272

Results are from a two-beta regressions using RTE weekly excess returns, $R_{Tt} - R_{ft}$, as the dependent variable. The independent variables are market excess returns, $R_{mt} - R_{ft}$, using the Russell 2000 as R_m, and a variable capturing the down-market risk premium, $Y_t = \max[0, -(R_{mt} - R_{ft})]$. The numbers in parentheses are the t-statistics and their significance level.

*,** Indicate significance at the 0.5 and 5.0% levels or better (one-tail test), respectively.

[a] $\beta_2 = \beta_u - \beta_d$, where β_u is the up-market beta and β_d is the down-market beta.

We employ ordinary least squares method (OLS) and the Russell 2000 as the market index to estimate the parameters of the two-beta timing model. OLS parameter estimates are reported in Table 4. For the entire period, results support that RTE exhibits considerable skills at market timing. In Panel A of Table 4, β_2 is positive and statistically significant at the 0.5% level or better ($t = 11.22$). β_2 is also positive and statistically significant at the conventional 5% level or better for all subperiods. Finally, the statistically significant negative a implies that RTE pays for market timing skills by realizing negative returns to nonmarket timing activities. Negative intercept terms are also observed in several prior studies but there is no accepted single explanation for its cause (Bello & Janjigian, 1997; Weigel, 1991).

5. Summary and conclusion

We use explicit buy and sell recommendations derived by the RTMM®, a computerized model developed by a practitioner of market timing, and the specific dates these recommendations were implemented for the clients of Lincoln Investment Planning, Inc., to show how investors can use several performance-evaluation methodologies to investigate whether market timing adds value. We provide evidence of the superior performance and market timing skills of RTE Asset Management, especially over a 20-year time period (from 1979 to 1999) and using the Russell 2000 Index as a benchmark. Over the entire period, RTE outperformed its benchmark with all risk-adjusted measures. RTE also exhibited a strong ability to forecast up- and down-market conditions, the S-statistic being significant at the 0.5% level or better. In addition, the beta or systematic risk level of RTE in up-market

conditions is greater than its beta in down-market conditions since the timing coefficients, β_2, in the two-beta regressions are positive and significant at the 0.5% level or better.

As expected, the results are mixed when assessing RTE's market timing ability over several subperiods of approximately 5 years. The results for the 1979–1984 and the 1989–1993 subperiods are significant at the 0.5% level or better under all the methodologies. However, the results for the 1984–1989 and the 1993–1999 subperiods are not significant at the conventional level of 5% or better using the nonparametric and parametric tests of RTE's market timing ability. Overall, the findings provide some evidence of RTE's market timing skills despite that even a skillfull timer, at times, is going to perform poorly.

References

Becker, C., Ferson, W., Myers, D. H., & Schill, M. (1999). Conditional market timing with benchmark investors. *Journal of Financial Economics*, *52*, 119–148.

Beehover, G. L., & Varikooty, A. P. (1991). Measuring market timing strategies. *Financial Analysts Journal*, *47*, 78–84.

Bello, Z. Y., & Janjigian, V. (1997). A reexamination of the market-timing and security-selection performance of mutual funds. *Financial Analysts Journal*, *53*, 24–30.

Brocato, J., & Chandy, P. R. (1994). Does market timing really work in the real world? *Journal of Portfolio Management*, *20*, 39–44.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance*, *47*, 1731–1764.

Chen, C. R., Lee, C. F., Rahman, S., & Chan, A. (1992). A cross-sectional analysis of mutual funds' market timing and security selection skill. *Journal of Business Finance and Accounting*, *19*, 659–676.

Chance, D. M., & Hemler, M. L. (1998). The performance of professional market timers: Daily evidence from executed strategies. Paper presented at the 1998 Financial Management Association Annual Meeting.

Cumby, R. E., & Glen, J. D. (1990). Evaluating the performance of the international mutual funds. *Journal of Finance*, *45*, 497–521.

Copeland, M. M., & Copeland, T. E. (1999). Market timing: Style and size rotation using the VIX. *Financial Analysts Journal*, *55*, 73–81.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance*, *52*, 1035–1058.

Droms, W. G., & Walker, D. A. (1994). Investment performance of international mutual funds. *Journal of Financial Research*, *17*, 1–14.

Eun, C. S., Kolodny, R., & Resnick, B. G. (1991). U.S.-based international mutual funds: A performance evaluation. *Journal of Portfolio Management*, *17*, 88–94.

Ferson, W., & Schadt, R. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance*, *51*, 425–462.

Goetzmann, W. N., Ingersoll, J., Jr., & Ivkovic, Z. (2000). Monthly measurement of daily timers. *Journal of Financial and Quantitative Analysis*, *35*, 257–290.

Graham, J. R., & Harvey, C. R. (1996). Market timing ability and volatility implied in investment newsletters' asset allocation recommendations. *Journal of Financial Economics*, *42*, 397–421.

Graham, J. R., & Harvey, C. R. (1997). Grading the performance of market timing newsletters. *Financial Analysts Journal*, *53*, 54–66.

Grinblatt, M., & Titman, S. (1994). A study of mutual fund returns and performance evaluation techniques. *Journal of Financial and Quantitative Analysis*, *29*, 419–444.

Kane, A., & Marks, S. G. (1990). The delivery of market timing services: Newsletters versus market timing funds. *Journal of Financial Intermediation*, *1*, 150–166.

Kao, G., Cheng, L. T. W., & Chan, K. C. (1998). International mutual fund selectivity and market timing during up and down-market conditions. *Financial Review*, *33*, 127–144.

Kester, G. W. (1990). Market timing with small versus large-firm stocks: Potential gains and required predictive ability. *Financial Analysts Journal*, *46*, 63–69.

Larsen, G. A., Jr., & Wozniak, G. D. (1995). Market timing can work in the real world. *Journal of Portfolio Management*, *21*, 74–81.

Lee, W. (1997). Market timing and short-term interest rates. *Journal of Portfolio Management*, *23*, 35–46.

Lee, C., & Rahman, S. (1990). Market timing, selectivity, and mutual fund performance: An empirical examination. *Journal of Business*, *63*, 261–278.

Malkiel, B. G. (1995). Returns from investing in equity mutual funds: 1971–1991. *Journal of Finance*, *50*, 549–572.

Modigliani, F., & Modigliani, L. (1997). Risk-adjusted performance. *Journal of Portfolio Management*, *23*, 45–54.

Paláez, R. F. (1998). Economically significant stock market forecasts. *The Financial Review*, *33*, 65–76.

Pesaran, M. H., & Timmerman, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business and Economic Statistics*, *10*, 461–465.

Prather, L., & Bertin, W. J. (1998). The implication of discount rate changes for market timing. *Review of Financial Economics*, *7*, 21–33.

Rich, S. P., & Reichenstein, W. (1993). Market timing for the individual investor: Using the predictability of long-horizon stock returns to enhance portfolio performance. *Financial Services Review*, *3*, 29–43.

Shilling, A. G. (1992). Market timing: Better than a buy-and-hold strategy. *Financial Analysts Journal*, *48*, 46–50.

Sullivan, R., Timmermann, A., & White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *Journal of Finance*, *54*, 1647–1691.

Wagner, J. C. (1997). Why market timing works. *Journal of Investing*, *6*, 78–81.

Wagner, J., Shellans, S., & Paul, R. (1992). Market timing works where it matters most in the real world. *Journal of Portfolio Management*, *18*, 86–90.

Weigel, E. J. (1991). The performance of tactical asset allocation. *Financial Analysts Journal*, *47*, 63–71.