Quantitative TRADING

The predictive power and economic effectiveness of trading strategies

BY MATEY GEROV

The moving average is considered to be an effective tool to reveal trends. It ostensibly removes noise and smoothes underlying data by averaging a chosen number of observations. It is employed in various market forecasting models and trading strategies which exploit well-documented low-order serial correlation and non-linearity of return series. Empirical evidence from a number of developed and emerging equity markets indicates that moving average models can predict reversals in price and yield statistically and economically significant excess returns over a buy-and-hold strategy, and that this cannot as yet be attributed to measurement errors arising from the use of non-synchronous trading data or the omission of dividend yields.[1] Surprisingly, genetic algorithms that simultaneously search for optimal trading strategies typically produce models similar to moving averages.[2]

This study evaluates the performance of ten double crossover moving average strategies with and without a trading filter band on the Toronto Stock Exchange S&P/TSX Composite Index.[3] While such a study is of obvious relevance to individual investors, it does have value for institutional investors who, regardless of whether they hold broad market-based investments, either benchmark against a market index or allocate capital to individual investments of varying correlation with the market.

The strategy is simple. A long position is held when the short-period average price is greater than the long-period average price, and a short position is held when the opposite is true. A filter band is a predefined fixed percentage around the long moving average that the short moving average must pass before the investment position is changed. This presumably filters out weak signals. These moving average strategies can be summarized by MA (p, q, r) model, where p is the length of the short period, q is the length of the long period, and r is the size of the filter band. So, for example, MA (5, 150, 0.01) uses five days to calculate the short moving average, 150 days for the long, and imposes a filter band of 1% above and below the long moving average.

To avoid selection biases, this study uses the same parameters as in the seminal study on the topic by Brock, Lakonishok and LeBaron (1992).[4] Ito (2003) has studied moving average strategies on TSE 300 data from 1977 to 1995. Here, the same sample is employed for a number of analyses of what will be referred to as the in-sample. Eight more years of data, 1996 to 2003, were collected and used as an out-of-sample period for comparison purposes. The predictive power of the strategies was evaluated by testing the trading signals for randomness because non-random signals have the potential to be exploited for excess returns. The economic performance of the strategies was then studied by comparing both simple returns and total returns, adjusted for risk and transaction costs, with the return on a buy-and-hold position. Finally, a Consistency Check of the strategies was performed to see whether they maintained their performance ranking over time. Did the best remain the best and the worst the worst?

The analysis reveals that the strategies generate non-random signals. Despite the large qualitative differences between sample periods, the average holding period of a position and the proportion of long to short positions

Matey Gerov, M.Sc., MSTA, graduated from the John Molson School of Business in 2005.

remain constant. All of the strategies yield excess returns over buy-and-hold, and none of them is riskier than the passive approach. The best performing models have negative betas. The results are largely consistent over various comparison criteria, and are robust to two-way transaction costs of 25 basis points. A rank order test of the models' risk-adjusted performance points to a degree of forecast reliability. There is also a positive and significant relationship between conditional returns and unconditional variance, which implies higher profits in more volatile periods.[5]

Previous Research

Brock, Lakonishok and LeBaron (1992) apply two of the most popular ad-hoc trading models, the moving average and the trading range breakout, to Dow Jones Industrial Average (DJIA) index daily data from 1897 to 1986. Their findings are noteworthy because every one of the models yields significant excess returns over the simple buy-and-hold strategy. The authors employed a bootstrap technique to deal with the non-normality of equity return distributions, comparing actual trading strategy returns to returns simulated from a number of generating processes, and found that actual returns are not consistent with a random walk, autoregressive, GARCH-M, or E-GARCH processes. Buy signals consistently produce higher returns than sell signals, and returns following buy signals are less volatile than those following sells. This asymmetry has been studied by Gencay (1998), who documents a non-linear predictability of stock market returns whereas moving averages experience at least a 10% forecast improvement in more volatile years.

Although Brock, Lakonishok and LeBaron do not find differences in trading model performance across sub-samples, replications of their work by others suggest weakening profits in more recent periods. Sullivan, Timmermann, and White (1999), for instance, document that the best model from 1897 to 1986 does not repeat its superior performance in the out-of-sample period 1987 to 1996. In fact, none of the models is able to maintain its profitability. They offer three possible explanations: that the out-of-sample period is not representative; that the trading models themselves are no longer representative; and that markets have become more efficient. They ultimately favour the third, citing low-cost computing power, lower transaction costs, and increased liquidity as possible causes of heightened market efficiency. Ready (2002) reports similar findings and speculates that arbitrage activities on Wall Street have minimized trading models, return opportunities. Employing data from 1962 to 1996 for NYSE and from 1973 to 1996 for NASDAQ, Kwon and Kish (2002) find weakening profits from trading models and suggest that markets are becoming more efficient in disseminating information to a wider range of investors.

Moving average strategies are, however, economically effective where trading costs are low. Isakov and Hollistein (1999) find that investors who face two-way transaction costs of less than 0.66% can benefit on the Swiss market (SBC Index, 1969-1997) and conclude that, for a large fraction of participants, the weak form of market efficiency cannot be rejected. Two separate studies employing large samples, 1935 to 1994, for the FT30 [Hudson, Dempsey and Keasey (1996)] and 1926 to 1991 for the DJIA [Bessembinder and Chan (1998)], estimate identical average break-even two-way transaction costs of 0.8% and 0.78%, respectively. In contrast, Ito (1999) reports an average break-even cost of 1.58% for Canada from 1980 to 1996. The break-even levels estimated in this study are much lower for roughly the same period, whereas the differences resulting from data samples and trading strategies cannot explain the gap. It appears that Ito did not distinguish between buy-sell differences and actual returns in the 1999 study, as he does in a later study.

TABLE 1 TRADING SIGNALS

	Panel A. 1977-1995			Panel B. 1996-2003		
MA (p, q, r)	Long	Short	Trade	Long	Short	Trade
MA (1, 50, 0)[S]	60.9%	39.1%	17.2	60.9%	39.1%	21.0
MA (1, 50, 0.01)[S]	48.3%	28.0%	21.4	53.8%	31.4%	20.8
MA (1, 150, 0)[S]	66.2%	33.8%	44.5	67.1%	32.9%	40.3
MA (1, 150, 0.01)[S]	61.5%	28.4%	37.3	62.9%	28.9%	35.4
MA (5, 150, 0)	66.1%	33.9%	72.8	67.3%	32.7%	84.0
MA (5, 150, 0.01)	61.1%	28.2%	71.7	63.2%	28.7%	77.5
MA (1, 200, 0)[S]	68.6%	31.4%	51.6	65.9%	34.1%	48.0
MA (1, 200, 0.01)[S]	63.8%	27.8%	46.3	63.0%	31.7%	50.4
MA (2, 200, 0)	68.5%	31.5%	64.6	66.3%	33.7%	59.3
MA (2, 200, 0.01)	64.0%	27.6%	58.1	62.8%	31.5%	61.1
Average	62.9%	31.0%	48.6	63.3%	32.5%	49.8

Data and Methodology

Since the omission of dividends may lead to overstated excess returns, the strategies are examined using both simple returns, which are based solely on daily price changes, and total returns, which are based on price changes and distributions (stocks, dividends, etc.). The sample consists of three daily time series: the S&P/TSX Composite Price Index, the S&P/TSX Composite Total Return Index, and the average yield on 91-day Government of Canada T-bills, the risk-free interest rate being used in Sharpe and Jensen risk adjustments. The sample runs from the first trading day of 1977, January 3rd, to the last trading day of 2003, December 31st, for a total of 6,803 observations. The first 200 trading days are set aside for calculation of the first observation of the moving averages. To test for consistency, the sample was split into two sub-samples: an in-sample, January 3rd, 1977 to December 29th, 1995, which is the same as Ito's, and an out-of-sample, January 2nd, 1996 to December 31st, 2003. All data were obtained from the Toronto Stock Exchange—Canadian Financial Markets Research Centre Summary Information Database (TSX-CFMRC), a product of the TSX Group.

Returns were calculated as differences of the natural logarithms of closing prices, and annual returns, where used, were estimated using continuous compounding for an average of 252 trading days. Long, short and out-of-market positions are initiated following buy, sell and neutral signals, respectively. A long position yields a return equal to the buy-and-hold; a short position yields a return equal to the negative of the buy-and-hold return; and an out-of-market position yields the risk-free rate. The cost of borrowing to sell short was implemented conservatively: it included the dividend yield and the risk-free rate. This method is the least biased when compared to other approaches, such as replacing short positions with out-of-market positions [Ready (2002)], which assumes equality between negative risk premia and the risk-free rate, or a double-or-out strategy [Bessembinder and Chan (1998)], which assumes equality between long and short position returns.

Recognizing the non-normality of equity returns, a set of distribution-free tests, the Chi-square, Mood's median and Wilcoxon's rank-sum test, were used to test hypotheses about the randomness of buy and sell signals. The premise is that the process of generating buy and sell signal return series is equivalent to a sampling procedure with the MA (p, q, r) model as the sample selection method and the buy-and-hold return series as the underlying population. As long as the selection method is random, buy and sell sampling distributions should approximate the population distribution; otherwise, the model possesses some predictive power. The economic performance was then examined by adjusting the moving average returns for risk and transaction costs. The notion of risk-adjusted returns is fundamental for the precise evaluation of trading strategies, because it demands comparison of returns at equal levels of risk. This study implemented it using the standard

TABLE 2 RISK-ADJUSTED PERFORMANCE

MA (p, q, r)	Panel A. 1977-1995 Sharpe	Alpha	Beta	Panel B. 1996-2003 Sharpe	Alpha	Beta
MA (1, 50, 0)[S]	2.92%	0.022%	-8.55%	4.25%	0.048%	-23.71%
MA (1, 50, 0.01)[S]	4.40%	0.031%	-9.32%	4.71%	0.049%	-18.28%
MA (1, 150, 0)	1.45%	0.011%	2.08%	2.80%	0.030%	3.50%
MA (1, 150, 0.01)	1.95%	0.015%	2.70%	2.63%	0.027%	3.20%
MA (5, 150, 0)	-0.39%	-0.003%	4.42%	2.17%	0.023%	4.41%
MA (5, 150, 0.01)	0.38%	0.003%	5.94%	2.33%	0.024%	4.43%
MA (1, 200, 0)	0.90%	0.007%	7.49%	3.13%	0.033%	1.34%
MA (1, 200, 0.01)	1.22%	0.009%	6.74%	3.42%	0.036%	2.08%
MA (2, 200, 0)	0.38%	0.003%	7.65%	2.74%	0.029%	1.77%
MA (2, 200, 0.01)	0.94%	0.007%	8.25%	2.85%	0.030%	2.47%

TABLE 3 ADJUSTMENT FOR TRANSACTION COSTS

MA (p, q, r)	Panel A. 1977-1995 ER	ER25	B/E	Panel B. 1996-2003 ER	ER25	B/E
MA (1, 50, 0)	7.23	3.06	0.44	9.38	5.91	0.70
MA (1, 50, 0.01)	9.90	6.47	0.74	9.83	6.31	0.72
MA (1, 150, 0)	4.00	2.42	0.64	4.95	3.21	0.72
MA (1, 150, 0.01)	4.91	3.01	0.66	4.19	2.24	0.54
MA (5, 150, 0)	0.03	-0.90	0.01	3.04	2.21	0.93
MA (5, 150, 0.01)	1.66	0.69	0.43	3.32	2.43	0.94
MA (1, 200, 0)	2.78	1.44	0.52	5.94	4.47	1.03
MA (1, 200, 0.01)	3.43	1.91	0.57	6.66	5.23	1.20
MA (2, 200, 0)	1.66	0.58	0.39	4.75	3.57	1.02
MA (2, 200, 0.01)	2.81	1.60	0.59	4.92	3.77	1.08
Average	3.84	2.03	0.50	5.70	3.93	0.89

Sharpe ratio (1966) and Jensen measure (1968).

Subtracting 25 basis points for every two-way transaction accounts for the trading costs presumably borne by institutional investors. One-way transaction costs were applied only for trading models with filter bands when there was a shift from a neutral to a long or short position or vice versa. Transaction costs were subtracted from the conditional returns on the day they were incurred. The analysis also includes an estimation of the break-even two-way transaction cost across all trading models. Finally, I asked whether there is a significant difference across moving average strategies and, most importantly, whether there is a degree of forecasting reliability. To do this, the trading models were first ranked according to their Sharpe ratios. Then a median exact test for differences across *k* independent samples of the returns was performed. To estimate the exact significance with 99% confidence, a Monte Carlo simulation was conducted on 10,000 sample "tables." Whenever the hypothesis that the trading models are not equally effective was rejected, the best and the worst strategies were excluded, and the procedure was repeated on the remaining models until all were tested. I refer to this as the Consistency Check.

TABLE 4

CONSISTENCY CHECK

Sharpe Ranking	Monte Carlo Exact Lower B	Monte Carlo Exact Upper B	Best Model	Worst Model
Panel A. 1977-1995				
MA Ranked 1-10	0.001	0.004	MA (1, 50, 0.01)*	MA (5, 150, 0)*
MA Ranked 2-9	0.003	0.007	MA (1, 50, 0)*	MA (5, 150, 0.01)*
MA Ranked 3-8	0.020	0.028	MA (1, 150, 0.01)	MA (2, 200, 0)
MA Ranked 4-7	0.010	0.016	MA (1, 150, 0)	MA (1, 200, 0)
MA Ranked 5-6	0.890	0.906	MA (1, 200, 0.01)	MA (2, 200, 0.01)
Panel B. 1996-2003				
MA Ranked 1-10	0.000	0.000	MA (1, 50, 0.01)*	MA (5, 150, 0)*
MA Ranked 2-9	0.000	0.000	MA (1, 50, 0)*	MA (5, 150, 0.01)*
MA Ranked 3-8	0.023	0.032	MA (1, 200, 0.01)	MA (1, 150, 0.01)
MA Ranked 4-7	0.159	0.179	MA (1, 200, 0)	MA (2, 200, 0)
MA Ranked 5-6	0.043	0.054	MA (2, 200, 0.01)	MA (1, 150, 0)

Empirical Results

A description of the trading signals is presented in Table 1. The columns Long and Short contain the number of days long and short positions were held relative to the sample's duration; Trade corresponds to the trading frequency, and it is given as the average holding period of a position in number of days. Moving average strategies result in a long position being held roughly two-thirds of the time and a short position for the remaining third with remarkable consistency between samples. Moreover, the holding periods are similar, in-sample and out-of-sample. These are striking findings considering how different market conditions were: from 1977 to 1995 daily price and total price returns were 0.0341% and 0.0484%, with a standard deviation of approximately 0.77%, and one-day serial correlation of 23.5%; from 1996 to 2003 those readings were 0.0276%, 0.0342%, 1.07%, and 8.6%, respectively. There is a strong positive relationship between the length of the averaging period and the trading frequency, or the length of time a position was held. A 50-day moving average, for example, resulted in a position being held for three weeks, whereas a 200-day strategy meant a position was held for two months. The effect of the filter bands was to increase the time out of the market to about 10% without a change in the relative duration of the long and short positions.

The distribution-free tests for the randomness of buy and sell signals reveal significant differences between the moving average and buy-and-hold distributions, implying some predictive power to the former. This predictive power is partly reflected in the observation that average returns following sell signals are all significantly positive, and this result is largely consistent with Brock, Lakonishok and LeBaron (1992) and Mills (1997), among others. A few of the tests of the two double crossover strategies, however, fail to reject the null hypotheses from 1996 to 2003 (see Table 1, models not marked with S in the first column). Interestingly, these are the models that have the lowest signal-generating frequency and thus the longest holding periods. Clearly these properties affect their conditional return distributions in a way that brings them closer to the buy-and-hold strategy.

The returns associated with the moving average strategies are less risky. Long positions are less volatile than short positions, and overall, none of the trading models is riskier than buy-and-hold, in- or out-of-sample. Table 2 highlights the risk properties of moving average strategies using the Sharpe ratio as well as Jensen's Alpha as a performance measure and Beta as the risk parameter. Alpha, as it is used here, is the average incremental risk-adjusted return due to the predictive power of the trading models, where positive and significant intercepts are indicative of superior performance; the beta is a measure of tendency of conditional returns to move in line with the market risk premia.

All of the moving average strategies outperform buy-and-hold under the Sharpe criterion with average ratios of 1.897% and 3.1029% for the two periods as compared to –0.4134% and 1.1309% for the passive approach. Nine in-sample and all ten out-of-sample alphas are positive, but for only two models, MA (1, 50, 0) and MA (1, 50, 0.01), are they highly significant (see Table 2, models marked with S in the first column). Moving average models outperformed the buy-and-hold strategy by an average of 0.0106% or 2.71% annually, and from 1996 to 2003, by 0.0328% or 8.62%. It is worth pointing out that the two models with the highest alphas have significantly negative betas. This is typical of strategies designed to counter the market, such as those employed by hedge funds. Lastly, while buy-and-hold returns are higher when based on total return series, moving average returns are still superior.

Table 3 summarizes the returns adjusted for transaction costs. Columns ER and ER25 present the annualized continuously compounded excess price return differences between moving average strategy and buy-and-hold returns before and after two-way transaction costs of 25 basis points, respectively. Column B/E shows the break-even or two-way transaction cost measured in basis points at which a trading model and buy-and-hold average returns are equal. Before transaction costs, all strategies yield higher mean daily returns than buy-and-hold, in-sample and out-of-sample. MA (1, 50, 0.01) generates the highest return of 0.0686% or 18.87% at an annual rate. This is 9.9% greater than buy-and-hold return. Overall, it appears that trading models with lower smoothing parameters tend to yield higher returns, and that the introduction of trading filter bands improves performance in every case. These findings are valid when total price returns are employed.

After adjusting returns with two-way transaction costs of 25 basis points, only the MA (5, 150, 0) model is unable to outperform the buy-and-hold in the first period. The other trading strategies require an average two-way transaction cost of less than 50 basis points in order to be profitable. Interestingly, despite its high trading frequency, the best performing model is again MA (1, 50, 0.01); it yields approximately 6.47% excess return per year and is profitable up to trading costs of 74 basis points. Accounting for dividend yields, the average break-even for a single crossover moving average model is 23 basis points. From 1996 to 2003 the break-even levels increase to 89 basis points for a two-way transaction, which suggests higher profits. This is attributable to the higher market volatility and its relationship to conditional returns. Overall, the break-even costs are strikingly similar to those reported on DJIA, FT30 and SBC.

One gauge of forecast reliability is to see whether the best and worst performing trading models in the in-sample keep their ranks in the out-of-sample. Assume that moving average strategies are not significantly different one from another. Hence, the probability of a trading model changing rank from best to worst or vice versa in a later period is equal across all models, and so combining them and using a weighted signal would be as good as any. If, on the other hand, trading models differ significantly in performance, then an optimal strategy will involve separating the best from the worst. That, in essence, is the Consistency Check, as described previously and presented here in Table 4. The column Sharpe Ranking shows the trading models ranked according to their Sharpe ratios, where 1-10 includes all models, 2-9 excludes the best and the worst after the first simulation, etc.; columns Lower B and Upper B are the lower and upper band of the 99% confidence interval calculated with Monte Carlo simulations based on 10,000 sample tables of the corresponding p-value. Best Model and Worst Model show the strategies that yield the highest and the lowest mean returns within those groups.

For 1977 to 1995, the null hypothesis of equality among moving average strategies is rejected at 1% level of significance as long as the two best and the two worst performing models are present. There is virtually no significant difference among trading models ranked from three to eight. The results for the second period are similar. Most important is that the two best performing, MA (1, 50, 0,01) and MA (1, 50, 0), and two worst performing, MA (5, 150, 0,01) and MA (5, 150, 0) strategies are the same, in-sample and out-of-sample (see Table 4, models marked with "*" in the last two columns).

Conclusion

Moving average strategies demonstrate striking efficiency and remarkable consistency in two completely different samples of Canadian equity market data. This is not attributable to chance, as the various comparison criteria used here provide strong support for a degree of forecasting reliability. Moreover, these strategies yield returns very similar to those obtained using DJIA, FT30 and SBC data, which implies broader effectiveness. The findings in this paper suggest avenues for future practitioner research. Moving averages can be easily incorporated in various equity pricing models, cloned in zero-cost portfolios by matching buy and sell signals for different stocks, or simply tested with lower smoothing parameters that promise higher profits. ■

Acknowledgments

The author is grateful to Professor Gregory Lypny, Ph.D., for his guidance and support.

Endnotes

1. Indices may contain stale information or exclude dividend distributions, leading to overstated predictive power of tested strategies.
2. Allen and Karjalainen (1999) conclude that with 0.25% transaction costs most of the genetically generated rules "are effectively similar to a 250-day moving average rule," page 262.
3. Three ad-hoc trading-range-breakout strategies have been studied as well. Results are given in the full version of the paper, which is available upon request.
4. Sullivan, Timmermann and White (1999) provide evidence that the selected models are robust to data snooping using a test that accounts for the bias inherent in a researcher's choice of trading models to apply.
5. To my knowledge, the full version of this paper is the first to document and study this relationship.

References

Allen, F., and R. Karjalainen, 1999, "Using Genetic Algorithms to Find Technical Trading Models," *Journal of Financial Economics* 51, 245-271.

Bessembinder, H., and K. Chan, 1998, "Market efficiency and the returns to technical analysis," *Financial Management* 27, 5-17.

Brock, W., J. Lakonishok, and B. LeBaron, 1992, "Simple Technical Trading Models and the Stochastic Properties of Stock Returns," *Journal of Finance* 47, 1731-1764.

Gencay, R., 1998, "The predictability of security returns with simple technical trading models," *Journal of Empirical Finance* 5, 347-359.

Hudson, R., M. Dempsey, and K. Keasey, 1996, "A note on the weak form efficiency of capital markets: The application of simple technical trading models to UK stock prices—1935 to 1994," *Journal of Banking & Finance* 20, 1121-1132.

Isakov, D., and M. Hollistein, 1999, "Application of simple technical trading models to Swiss stock prices. Is it profitable?," *Finanzmarkt und Portfolio Management* 13, 9-26.

Ito, A., 1999, "Profits on technical trading models and time-varying expected returns: Evidence from Pacific-Basin equity markets," *Pacific-Basin Finance Journal* 7, 283-330.

Ito, A., 2003, "Technical Trading Models and Nonlinear Time Series Models," *Asia-Pacific Journal of Financial Studies*.

Jensen, M., 1968, "Problems in Selection of Security Portfolios: The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance* 23, 398-416.

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

Mills, T., 1997, "Technical Analysis and the London Stock Exchange: Testing Trading Models Using the FT30," *International Journal of Financial Economics* 2, 319-331.

Ready, M., 2002, "Profits from technical trading models," *Financial Management* 31, 43-61.

Sharpe, W., 1966, "Mutual Fund Performance," *Journal of Business* 39, 119-138.

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