Risk-Adjusted, Ex Ante, Optimal, Technical Trading Rules in Equity Markets,

Authors	Christopher J. Neely
Working Paper Number	1999-015D
Citable Link	https://doi.org/10.20955/wp.1999.015
Suggested Citation	Neely, C.J.; Risk-Adjusted, Ex Ante, Optimal, Technical Trading Rules in Equity Markets,, Federal Reserve Bank of St. Louis Working Paper 1999-015. URL https://doi.org/10.20955/wp.1999.015

Published In	International Review of Economics & Finance
Publisher Link	https://doi.org/10.1016/S0261-5606(02)00101-8

Risk-Adjusted, Ex Ante, Optimal Technical Trading Rules in Equity Markets

Christopher J. Neely

August 8, 2001

Senior Economist, Research Department
Federal Reserve Bank of St. Louis
St. Louis, MO 63011
(314) 444-8568 (o), (314) 444-8731 (f), neely@stls.frb.org

Primary Subject Code: G0 - Financial Economics
Secondary Subject Code: G14 - Information and Market Efficiency
Keywords: technical analysis, genetic programming, trading rule, stock price, equity

---

The views expressed are those of the author and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, or the Federal Reserve System. The author thanks Kent Koch for research and programming assistance, Chuck Whiteman, Paul Weller and three anonymous referees for comments, Ramazan Gençay for private correspondence on the implementation of the X_{eff} measure and Franklin Allen and Risto Karjalainen both for making their programs available and for correspondence on the programs. Any errors are my own.

Risk-Adjusted, Ex Ante, Optimal Technical Trading Rules in Equity Markets

Abstract: This paper uses genetic programming to construct risk-adjusted, ex ante, optimal, trading rules for the S&P 500 Index and then characterizes the predictive content of these rules. These results extend previous results by using risk-adjustment selection criteria to generate ex ante rules with improved performance. There is, however, no evidence that the rules significantly outperform the buy-and-hold strategy on a risk-adjusted basis. Therefore, the results are consistent with market efficiency. Nevertheless, risk-adjustment techniques should be seriously considered when evaluating trading strategies.

Risk-Adjusted, Ex Ante, Optimal Technical Trading Rules in Equity Markets

The use of technical trading rules—trading rules based on past price behavior—has been common in equity markets since the turn-of-the-century analysis of *Wall Street Journal* editor Charles Dow. Because excess returns generated from publicly available information would seem to contradict the efficient markets hypothesis, a number of authors have studied the usefulness of technical analysis in equity markets (Brock, Lakonishok and Lebaron (1992), Bessembinder and Chan (1995, 1998), Allen and Karjalainen (1999), Lo, Mamaysky and Wang (2000)). Such studies have generally evaluated raw excess returns rather than explicitly risk-adjusted returns, leaving unclear the implications of their work for the efficient markets hypothesis. Risk-adjusting the returns is necessary because the technical trading strategies spend time out of the market and therefore have less volatile returns than the buy-and-hold rule. Therefore, simply comparing the returns to each strategy is insufficient to compare the usefulness of the rules. An exception to the failure to adjust for risk is the work of Brown, Goetzmann and Kumar (1998) that found value in the risk-adjusted returns generated by the market signals of William Peter Hamilton.[1]

Another common problem in the trading rule literature is that rules are evaluated precisely because they are widely used by technical traders (see Brock, Lakonishok and Lebaron (1992)). Ready (1998) argues that testing such rules is a form of data snooping. This practice is likely to produce spurious evidence of technical trading profits; the rules are widely used precisely because they would have been profitable on past data.

[1] Another exception is the work of Bessembinder and Chan (1998), which uses varying leverage with its technical trading rule to produce a rule with approximately the same risk as a buy-and-hold strategy.

Rather than evaluating widely used rules, one might search for optimal, ex ante rules with a nonlinear search procedure such as genetic programming (Koza (1992)). Allen and Karjalainen (1999)—hereafter AK—used genetic programming to generate optimal, ex ante, technical trading rules on daily S&P 500 data over the period 1929 through 1995.[2] They found that the transactions cost-adjusted returns to these rules failed to exceed the returns to a buy-and-hold strategy—despite the exclusion of dividends from the stock return—and that the market was efficient in this sense.[3] There was, however, some evidence of predictability in returns as the rules tended to be in the market during periods of high returns and out of the market during periods of low returns. Although AK attributed this predictability to low-order serial correlation in the stock index, they speculated that the rules might be useful on a risk-adjusted basis despite their lower returns. *"Even though the rules do not lead to higher absolute returns than a buy-and-hold strategy, the reduced volatility might still make them attractive to some investors on a risk-adjusted basis." (Allen and Karjalainen (1999), p. 261)* AK did not, however, use any risk-adjustment techniques in their work.

This paper extends the literature by investigating whether ex ante, optimal technical trading rules are useful on a risk-adjusted basis in equity markets. It is not sufficient merely to examine the results from previous genetic-programming rules with common methods of risk adjustment. To fairly evaluate risk-adjusted returns, new sets of rules that maximize risk-adjusted measures like the Sharpe ratio (Sharpe (1966)), the X* statistic (Sweeney and Lee (1990)) and the X_{eff} measure (Dacorogna et al. (2001)) are generated. In addition, this paper

[2] Neely, Weller and Dittmar (1997) and Neely and Weller (2000, 1999) have applied genetic programming to find trading rules in the dollar foreign exchange market and the European Monetary System.

[3] The return to a dynamic strategy—moving in and out of the market—will be reduced less by the exclusion of dividends than will the return to a buy-and-hold strategy.

more fully characterizes the predictability found by these rules and conducts formal tests of market timing (Cumby and Modest (1987)). A central point of this study is that risk adjustment is not a secondary issue; it is absolutely essential both for evaluating the usefulness of trading rules and for measuring the consistency of results with market efficiency (Sharpe (1966), Jensen (1968), Kho (1996), Brown, Goetzmann, and Kumar (1998), Ready (1998), Dowd (2000)).

The rules fail to consistently and significantly outperform the buy-and-hold strategy by any risk-adjusted measure. Thus, this exercise extends previous results to find that risk-adjusted, ex ante, optimal rule returns are consistent with market efficiency. The facts that the market indices used here exclude dividends and that some predictability may be due to spurious autocorrelation only reinforce the negative results.

2. METHODOLOGY

A. Genetic Programming

Genetic programming is a nonlinear search procedure for problems in which the solution may be represented as a computer program or decision tree (Koza (1992)). Like its cousin, the genetic algorithm (Holland (1975)), genetic programming uses the principles of parallel search and natural selection to search for candidate solutions to problems of interest.[4] Essentially, a computer randomly generates a population of candidate solutions—expressible as decision trees—to a problem of interest. The rules are required only to be well defined and to produce output appropriate to the problem of interest—a buy/sell decision in the present case. Of course, most of these random solutions will be quite poor, but some, purely by chance, will "fit" the in-sample data reasonably well, generating excess returns. The computer then allows the population to "evolve" using reproduction and mutation operators. Reproduction mixes subtrees of the population while mutation replaces subtrees with new, randomly generated subtrees. Members of the population that are more fit (profitable) have a greater chance to reproduce whereas less fit members have a greater chance of being replaced. In this way the genetic program searches promising areas of the solution space by evolving a population of rules that tends to become more adept at solving the problem in successive generations.[5]

B. Data Sets

Ten overlapping in-sample estimation periods (1929-35, 1934-40, 1939-45… 1974-80) of daily S&P500 data from 1929 to 1995 are input to the genetic program to construct ten sets of ex ante trading rules. There were ten independent rules from each in-sample period. Each in-sample period of seven years was broken down into a training period (five years) and a selection period (two years) to alleviate the problem of overfitting the data.[6] Rules with positive excess returns over the buy-and-hold strategy in the training period were saved for out-of-sample testing over the remainder of the data (1936-95, 1941-95…1981-95).[7]

---

[4] Genetic algorithms require the solution to the problem to be encoded as fixed-length character strings rather than as decision trees or computer programs as in genetic programming.

[5] An alternative (but not equivalent) nonparametric procedure would be to use neural networks to produce either forecasts (Gençay (1998b), Gençay and Stengos (1997)) or trading rule signals (Gençay (1998a)) on daily equity market data.

[6] Overfitting—the finding of spurious patterns in the data—often results from applying a flexible statistical technique to one sample of data. Breaking the in-sample period into two subsamples helps alleviate this by requiring the patterns to be in both samples.

[7] To check for robustness, all the exercises were repeated with 15-year in-sample periods, with and without truncated out-of-sample periods. Except where otherwise noted, all the results in this paper proved robust to longer in-sample periods, with and without shorter out-of-sample periods. Full results are available from the author.

C. Trading Procedures and Return Calculations

This paper uses modified versions of AK's programs and similar procedures both because the procedures are sensible and to maintain maximum comparability to the unadjusted results.[8] One difference between AK's procedures and those used here should be noted: Interest rates are treated differently. AK's code attributes one day's (1/365) interest rate to the rules during each business day—not calendar day—they are out of the market. This practice understates the returns to the genetic programming rules by 0.5 percent or less. In this paper, rules earn interest on calendar days—not business days—they are out of the market. Table 1 summarizes some of the important parameters of interest. AK provide more information on the program and parameters.

[Place Table 1 about here.]

Each day, the trading rules generated by the genetic program observe prices and generate a buy or sell signal indicating the position to take (the same day). The buy and sell signals are used along with stock prices and 30-day T-bill interest rates to compute the continuously compounded excess return of the rule over the return to a buy-and-hold strategy in the stock market. This daily excess return—ignoring any transactions costs—over the buy-and-hold strategy at time t is given by:

$$xsr_t = (z_t - 1)\left[\ln\left(\frac{P_{t+1}}{P_t}\right) - \ln(1 + i_t)\right] \tag{1}$$

where z_t is an indicator variable taking the value 1 if the rule is in the market or 0 if the rule is in T-bills, P_t is the stock index and i_t is the interest rate on the 30-day Treasury bill earned from business day t to business day $t+1$. The cumulative excess return—also called the "fitness"—for a trading rule from time zero to time *T* is the sum of the daily excess returns less a proportional transactions cost. AK considered transactions costs of 0.1 percent, 0.25 percent and 0.5 percent. For brevity's sake, this paper concentrates on transactions costs of 0.25 percent. Because it is likely that trading costs have decreased over time, the use of 25 basis point transactions costs might underestimate the true value in earlier periods but overestimate them recently.

D. Risk Adjustment Techniques

The criterion of judging the rules to be useful only if they generate a return that exceeds the buy-and-hold return is neither necessary nor sufficient to conclude that the rules do not violate the efficient markets hypothesis (EMH).[9] The EMH is usually interpreted to mean that asset prices reflect information to the point where the potential risk-adjusted excess returns do not exceed the transactions costs of acting (trading) on that information (Jensen (1978)). Risk adjustment is potentially important because dynamic strategies, such as those found by the genetic program, are often out of the market and therefore may bear much less risk than the buy-and-hold strategy. Although there is no universally accepted method of adjusting returns for risk, this paper will employ four techniques: the Sharpe ratio, the X* measure, Jensen's α and the X_{eff} measure of Dacorogna et al. (2001).[10]

---

[8] Genetic programs written by other authors produced results similar to those generated by the AK programs, suggesting that genetic programming is robust to small change in procedures.

[9] Brown, Goetzmann and Kumar (1998) find that risk adjustment is crucial in evaluating Dow Theory recommendations.

[10] There is some danger of data snooping in looking at multiple measures of risk. Sullivan, Timmerman and White (1999) propose a method to counter data snooping in the testing of multiple types of rules. This procedure is not appropriate for this paper as this paper does not test the "best" rule out of a group, but rather a portfolio rule.

The Sharpe ratio—the expected excess return per unit of risk for a zero-investment strategy (Campbell, Lo and MacKinlay (1997))—is usually expressed in annual terms as the annual excess return over the riskless rate—net of transactions costs—to a portfolio over that portfolio's annual standard deviation. The daily excess return over the riskless rate—ignoring transactions costs—to the rules at time *t* is given by:

$$r_t = z_t \left[\ln\left(\frac{P_{t+1}}{P_t} \right) - \ln(1 + i_t) \right], \tag{2}$$

where the variables are as defined in equation (1). Although the rules may have lower returns than the buy-and-hold strategy, lower volatility may permit the returns to be leveraged up to exceed the buy-and-hold return with similar risk. As an example, assume the excess return to the trading rule was only half that of the buy-and-hold strategy, but the trading rule's Sharpe ratio was higher. In this case, the trading rule could take leveraged positions in the market—buying with only a 50 percent margin—to obtain equal returns with lower risk.[11] Buying with a slightly lower margin would enable the rule to obtain higher expected returns for the same risk.

Sweeney and Lee (1990) developed another risk-adjustment strategy, the X* measure, in the context of the foreign exchange market that may be even more appropriate for equity

[11] Because dynamic strategies are at an inherent disadvantage, as the market return will, on average, exceed the riskless return, Bessembinder and Chan (1995, 1998) pursue another strategy to compare trading rules to a market return. They permit rules to use double leverage during periods in which they are in the market. Ready (1998) has questioned whether the strategy of leveraging returns is implementable, as the investor would have to know—or predict—the ex post moments to compute the proper amount of leverage.

markets.[12] They show that, in the presence of a constant risk premium, an equilibrium daily risk-adjusted return to a trading rule would be given by:

$$X^* = \frac{1}{T}\sum_{t=0}^{T-1}\left[z_t \ln\left(\frac{P_{t+1}}{P_t}\right) + (1-z_t)\ln(1+i_t)\right] + \frac{n}{2T}\ln\left(\frac{1-c}{1+c}\right) - \left[\frac{p_1}{T}\sum_{t=0}^{T-1}\ln\left(\frac{P_{t+1}}{P_t}\right) + \frac{p_2}{T}\sum_{t=0}^{T-1}\ln(1+i_t)\right] \tag{3}$$

where z_t, P_t, and i_t are defined as before, T is the number of observations, n is the number of one-way trades, c is the proportional transactions cost, p_1 is the proportion of the time spent in the market and p_2 is the proportion of the time spent in T-bills ($p_1 + p_2 = 1$). Note that the sum of the third and fourth terms estimates the expected return to a zero transactions-cost strategy that randomly is in the market on a fraction p_1 of the days, earning the market premium, and in T-bills otherwise. The risk-adjusted return—under the null of no timing ability—is the actual return less the expected return. Positive X* statistics are interpreted as evidence of superior risk-adjusted returns.

The third risk-adjustment measure considered is the X_{eff} measures advocated by Dacorogna et al. (2001). X_{eff} measures the utility that the trading strategy provides to a constant absolute risk averse individual over a weighted average of return horizons. The measure is:

$$X_{eff} = \frac{252 \cdot 100}{T}\left(\sum_{t=1}^{T} r_t - \frac{n}{2}\ln\left(\frac{1+c}{1-c}\right)\right) - \frac{\gamma}{2}\frac{\sum_{i=1}^{n}\tilde{w}_i \sigma_i^2 \left(1year/\Delta t_i\right)}{\sum_{i=1}^{n}\tilde{w}_i}, \tag{4}$$

[12] Sweeney (1988) uses the X* measure in the equity market. Ready (1998) constructs a statistic similar to Sweeney and Lee's (1990) X*. In turn, the test statistic of X*, proposed by Sweeney and Lee (1990), is virtually equivalent to the test statistic of the coefficient β_1 in the Cumby-Modest test of market timing if transactions costs are omitted from the X* calculation.

where $\frac{252 \cdot 100}{T}\left(\sum_{t=1}^{T} r_t - \frac{n}{2}\ln\left(\frac{1+c}{1-c}\right)\right)$ is the annualized excess return to the rule in percentage terms, net of transactions costs; r_t is defined in equation (2); σ_i^2 is the variance of nonoverlapping returns of length Δt_i days; (1 year/Δt_i days) is the number of returns of length Δt_i days in one year; and γ is a risk aversion parameter. Dacorogna et al. (2001) recommend values of γ between 0.08 and 0.15; this paper sets γ equal to 0.12. The sequence of weights $\{\tilde{w}_i\}$ takes on a maximum value at a holding period of 90 days:

$$\tilde{w}_i = \frac{1}{2 + \left(\ln\left(\frac{\Delta t_i}{90 days}\right)\right)^2} . \qquad (5)$$

The return horizons Δt_i are given by a geometric sequence {1,2,4,8, 16, etc}, whose maximum value is less than or equal to one quarter the number of days in the sample. Interested readers should consult Dacorogna et al. (2001) for additional details on this risk measure.

Finally, we also consider the performance of the rules according to Jensen's (1968) α, the return in excess of the riskless rate that is uncorrelated with the excess return to the market:

$$z_t\left[\ln(P_{t+1}/P_t) - \ln(1+i_t)\right] - \frac{n}{2T}\ln\left(\frac{1+c}{1-c}\right) = \alpha + \beta_M\left[\ln(P_{t+1}/P_t) - \ln(1+i_t)\right] + \varepsilon_t . \qquad (6)$$

If the intercept in equation (6)—α—is positive and significant, then the trading rule produces excess returns that cannot be explained by correlation with the market.

3. RESULTS

A. In-sample Results

Table 2 shows the in-sample results from implementing a uniformly weighted portfolio strategy using the excess return as the fitness criteria. The rules are assessed a 0.25 percent transactions cost for changing positions, and information on day *t* is used to trade the same day. As might be expected, the rules did very well in the sample on which they were trained and selected. The in-sample, annual excess return over the buy-and-hold strategy, averaged over the ten samples, was just over 5 percent. The analogous Sharpe ratio was 0.75, the risk-adjusted, annualized X* return was 4.34 percent and Jensen's α was 7.12 percent. Rules selected using the other fitness criteria (Sharpe ratios, X*, X_{eff}) also did well in-sample, producing especially good results by the metric for which they were constructed.

[Place Table 2 about here.]

It is worth noting that, although these rules are optimal in the sense of being the best rules that could be found within the limitations of the search procedure, they are not optimal in the sense of being the best possible rules one could find on the in-sample data. With sufficient computational power, a nonparametric search procedure like the genetic program could find an extremely good in-sample fit that would be unlikely to be informative about out-of-sample performance. That is, it would "overfit" the data. The use of both the training and selection periods is one method to help guard against such overfitting.

B. Comparison with Previous Results

Table 3 shows the out-of-sample results from implementing a uniformly weighted portfolio based on all the good rules found in-sample.[13] As in AK (compare with Table 2, Panel A in AK), the rules generally failed to produce positive excess returns over the buy-and-hold strategy in the sample. With the exception of the period 1949-55, for which no good in-sample

[13] Results for median portfolio rules are broadly similar to—slightly better than—those of the uniform portfolio rules. For the sake of brevity, they will not be reported separately. The median portfolio rule goes into the market if most of the N rules are in the market; otherwise it stays out of the market.

rules were found, the out-of-sample performance was similar to that found by AK.[14] While AK found only one period in which the mean excess return over the buy-and-hold strategy was positive, the current exercise found two such periods. The rules were long in the market about 51 percent of the time and traded 7.3 times a year, on average, though the figures varied widely with the in-sample period. The 1974-80 period produced uninteresting rules that stayed out of the market almost all the time.

[Place Table 3 about here.]

Column 5 of Table 3 shows the mean annual return to the market when the rules are in the market less the mean annual market return when the rules are out of the market (r_b-r_s). Although there is no measure of statistical significance, positive numbers favor the proposition that the rules have some market timing ability. AK found that rules from seven of ten in-sample periods had market timing ability by this measure. The results in this paper are similar, showing that six of nine have positive r_b-r_s. Because the rules' buy/sell decisions could be closely replicated by moving average rules, AK concluded that the genetic programming rules were taking advantage of low-order serial correlation. AK speculated that the rules might be of use to a risk-averse speculator, but did not seriously explore that possibility. As noted previously, risk adjustment is necessary to evaluate the consistency of the rules with the EMH. The next section reports the results of such procedures.

[14] There are two reasons why the results will not exactly replicate those found by AK: 1) Genetic programming is inherently stochastic, generating and recombining populations probabilistically; and 2) interest rate returns were treated differently in this analysis.

C. Risk-Adjusted Results

The average Sharpe ratio of the transactions cost-adjusted genetic programming rules is about 0.02, lower than the average 0.13 Sharpe ratio—the index doesn't include dividends—for the buy-and-hold strategy over the ten subsamples.[15]

As Dowd (2000) points out, the Sharpe ratio only provides enough information to choose between two investments that are uncorrelated with the return to the institution's portfolio. This is clearly not the case with these trading rules. Dowd (2000) proposes that an investment should be undertaken if it improves the Sharpe ratio of the portfolio, taken as a whole. To implement this strategy in the present situation, portfolio shares must be chosen. Three methods are considered: 1) to arbitrarily set the portfolio weights to ½ on the trading rule and ½ on the buy-and-hold strategy; 2) to choose the weights to maximize the in-sample Sharpe ratio; 3) to choose the weights ex post to maximize the out-of-sample Sharpe ratio. While the third method clearly constitutes data mining, it provides an upper bound on the benefit from using the rule.

Columns 7 through 9 of Table 3 show the improvement of the combined portfolio's out-of-sample Sharpe ratio over that of the buy-and-hold strategy, using each of the three strategies for choosing portfolio weights. None of the three portfolio strategies significantly improves upon the Sharpe ratio of the buy-and-hold strategy.

The second risk adjustment procedure considered is the X* procedure advocated by Sweeney and Lee (1990). Most of the annualized X* statistics—net of transactions costs—in

---

[15] Jorion and Goetzmann (1999) estimate that dividends made up much of the total return to U.S. equities over the period 1921 to 1995. The average Sharpe ratio for the buy-and-hold strategy over the ten overlapping out-of-sample subsamples is 0.13; whereas the Sharpe ratio from 1929 through 1995 is 0.06 for the index ex dividends.

Table 3 are negative, indicating that the rules would not have been useful, even by this risk-adjusted measure. Almost all the X* statistics would have been positive though, if transactions costs had not been netted out. This supports the evidence of predictability suggested by the r_b-r_s statistics.

The mean of the third risk-adjusted return measure, X_{eff}, is negative and lower than the X_{eff} generated by either the buy-and-hold strategy or the riskless strategy ($X_{eff} = 0$). That is, the genetic programming rules would not help an investor who has constant risk averse utility over the weighted average of returns at various horizons.

Recall the last measure of risk-adjusted returns—Jensen's α—estimates the mean return that is uncorrelated with the return to the market portfolio. To measure Jensen's α, returns to the market and to the trading rules were aggregated over nonoverlapping 30-day periods and regression (6) was performed by OLS using annualized returns. These results are shown in columns 14 and 15 of Table 3. The alphas are never significantly positive in any out-of-sample period, and the mean among all the out-of-sample periods is negative.

D. Rules to Maximize Risk-Adjusted Measures

Of course, the rules trained on an excess return criterion may not be the best risk-adjusted rules. To determine whether technical trading rules can produce better risk-adjusted returns than the buy-and-hold strategy, ideally we must train rules using risk-adjustment criteria.

The results of such an exercise using the Sharpe ratio as the in-sample fitness criterion are shown in Table 4. In several cases, the rules trained on Sharpe ratios turned out to be trivial— almost always long or short—in the out-of-sample period. The non-trivial rules failed to produce higher mean out-of-sample Sharpe ratios, but they did show marginally greater predictive ability by the standard of the r_b-r_s and X* statistics. They also spent less time in the market (27 percent long). Using the rules trained to maximize the Sharpe ratio in a portfolio with the buy-and-hold strategy, as suggested by Dowd (2000), failed to produce significant improvement over the buy-and-hold Sharpe ratio (see columns 7 through 9 of Table 4).

[Place Table 4 about here.]

Table 5 shows the analogous out-of-sample results from rules trained to maximize X* as the in-sample fitness criteria. There are no trivial (non-trading) X* portfolio rules, and the rules are very even handed. That is, there are no cases in which the rules are always in or always out of the market. The results are generally better to those of the rules trained on excess returns. The annualized excess return over the buy-and-hold is about 60 basis points greater than in the benchmark case, and the average Sharpe ratio is about the same as the average buy-and-hold Sharpe ratio over all sample periods (0.11 vs. 0.13). The mean annualized X* statistic is about zero, which is better than the mean X* statistics from the rules trained with excess returns or the Sharpe ratio as fitness criteria. However, it should be noted that even positive X* results may be consistent with the EMH in the presence of a time-varying risk premium. The rules' mean X_{eff} measure is slightly inferior to that of the buy-and-hold strategy. Jensen's alpha is never statistically significant in any sample, and the mean alpha is slightly positive but statistically and economically insignificant.

[Place Table 5 about here.]

Finally, we consider the results from training rules to maximize the X_{eff} statistics. The results are very similar to those from the other fitness criteria in Tables 3 through 5. Table 6 shows that the ex ante uniform portfolio rule chosen to maximize the X_{eff} statistic never outperforms the buy-and-hold strategy in a statistically and economically significant way by any of the risk-adjusted return metrics.

[Place Table 6 about here.]

4. CHARACTERIZING LOW-ORDER SERIAL CORRELATION

After finding that moving average rules could closely approximate the GP rules' buy/sell behavior, AK attributed the predictability found by their GP trading rules to "low-order serial correlation" in the returns (Campbell, Lo and MacKinlay (1997)). One might speculate that a simple time series model of returns could produce better decisions than the GP. To test this prediction and to attempt to better characterize the nature of the predictability found by AK, a variety of ARMA models were fit to the in-sample excess returns and the best in-sample models and parameters were chosen by the Akaike, Schwarz and excess return criteria. The best models were used to generate trading signals—as with the genetic programs—during the out-of-sample periods. Table 7 shows that the non-trivial ARIMA models are even less successful than the rules constructed by genetic programming. If low-order serial correlation generates the predictability, the genetic rules are apparently more successful at estimating it than are standard ARIMA models.

[Place Table 7 about here.]

Finally, Cumby-Modest tests of market timing ability are used to more formally determine whether the rules have predictive content. The statistical significance of the coefficient (β_1) in the regression of excess returns on signals from the trading rule summarizes the rules' one-day-ahead timing ability:

$$250 \cdot 100 \cdot \left[\ln\left(\frac{P_{t+1}}{P_{t+1}} \right) - \ln(1+i_t) \right] = \beta_0 + \beta_1 z_t + \varepsilon_t . \tag{7}$$

Table 8 presents strong evidence that the rules do possess predictive ability: 29 of the 32 available β_l coefficients are positive and 20 of those are significant at the 5 percent level. Of the three fitness criteria, the X* criteria seems to have produced the rules with the most predictive content. These timing tests illustrate the well-known result that profitability is not necessary for a rule to have predictive content.

[Place Table 8 about here.]

5. CONCLUSION

This paper has investigated whether ex ante, optimal trading rules created by genetic programming are useful on a risk-adjusted basis. Although risk-adjustment improves the relative attractiveness of the rules, neither Sharpe ratios nor Sweeney and Lee's X* statistic, nor Gençay's X_{eff} measure, nor Jensen's α provide evidence that rules developed by genetic programming would have been useful even to risk-averse speculators, contrary to reasonable speculation. Rules trained on the X* measure had the best risk-adjusted performance by all the measures, approximately equaling the buy-and-hold return performance. Not too much should be made of this result, however. With 15-year in-sample periods—results omitted for brevity—the Sharpe ratio became the best out-of-sample performer, again with performance approximately equal to the buy-and-hold strategy. Of course, risk is difficult to measure and any risk adjustment is subject to criticism. Nevertheless, this paper argues that trading rule results must be carefully interpreted in light of risk adjustment.

It is likely that the inclusion of dividends in the stock index, the removal of spurious autocorrelation from the index returns or accounting for price slippage would only strengthen the negative results of this exercise.

REFERENCES

Allen, F., Karjalainen, R., (1999) Using Genetic Algorithms to Find Technical Trading Rules, *Journal of Financial Economics*, 51, 245-271.

Bessembinder, H., Chan, K., (1995) The Profitability of technical Trading rules in the Asian stock markets, *Pacific-Basin Finance Journal*, 3, 257-284.

Bessembinder, H., Chan, K., (1998). Market Efficiency and the Returns to Technical Analysis. *Financial Management*, 27, 5-17.

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

Brown, S.J., Goetzmann, W.N., Kumar, A., (1998). The Dow Theory: William Peter Hamilton's Track Record Reconsidered, *Journal of Finance*, 53, 1311-1333.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., (1997). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Cumby, R.E., Modest, D.M., (1987). Testing for Market Timing Ability: A Framework for Forecast Evaluation, *Journal of Financial Economics*, 19, 169-189.

Dacorogna, M.M., R. Gençay, U.A. Müller, O.V. Pictet, (2001). Effective Return, Risk Aversion and Drawdowns, *Physica A*, 289, 229-248.

Dowd, K., (2000). Adjusting for Risk: An Improved Sharpe Ratio, *International Review of Economics and Finance*, 9, 209-222.

Gençay, R., (1998a). Optimization of Technical Trading Strategies and The Profitability in Security Markets, *Economics Letters*, 59, 249-254.

Gençay, R., (1998b). The Predictability of Security Returns with Simple Technical Trading Rules, *Journal of Empirical Finance*, 5, 347-359.

Gençay, R., Stengos, T., (1997). Technical Trading Rules and the Size of the Risk Premium in Security Returns, *Studies in Nonlinear Dynamics* 2(2), 23-34.

Holland, J., (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Jensen, M.C., (1968). Problems in Selection of Security Portfolios: The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance*, 23, 389-416.

Jensen, M.C., (1978). Some Anomalous Evidence Regarding Market Efficiency, *Journal of Financial Economics*, 6, 95-101.

Jorion, P., Goetzmann, W.N., (1999). Global Stock Markets in the Twentieth Century*, Journal of Finance*, 54, 953-980.

Kho, B.C., (1996). Time-Varying Risk Premia, Volatility, and Technical Trading Rule Profits: Evidence from Foreign Currency Futures Markets, *Journal of Financial Economics,* 41, 249-290.

Koza, J.R., (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA.

Lo, A.W., Mamaysky, H., Wang, J., (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, *The Journal of Finance*, 55, 1705 – 1770.

Neely, C.J., Weller, P.A., (2000). Technical Analysis and Central Bank Intervention, FRB St. Louis WP 97-002C. Forthcoming in the *Journal of International Money and Finance*.

Neely, C.J., Weller, P.A., (1999). Technical Trading Rules in the European Monetary System, *Journal of International Money and Finance*, 18, 429-458.

Neely, C.J., Weller, P.A., Dittmar, R.D., (1997). Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach, *Journal of Financial and Quantitative Analysis*, 32, 405-426.

Ready, M.J., (1998). Profits from Technical Trading Rules. Unpublished manuscript. University of Wisconsin, Madison.

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

Sullivan, R., Timmermann, A., White, H., (1999). Data-Snooping, Technical Trading Rule Performance, and the Bootstrap, *Journal of Finance*, 54, 1647-91.

Sweeney, R.J., (1988). Some New Filter Tests, Methods and Results. *Journal of Financial and Quantitative Analysis*, 23, 285-300.

Sweeney, R.J., Lee, E.J.Q., (1990). In: Aggarwal R., and Lee C.F. (Eds.), *International Dimensions of Securities and Currency Markets. Advances in Financial Planning and Forecasting Series Vol. 4, Part A*. JAI Press, Greenwich, CT, 55-79.

Table 1: Genetic programming parameters of interest for AK's implementation

Parameter	AK's Choice
Size of a generation	500
Termination criterion	50 generations or no improvement for 25 generations
Probability of selection for reproduction with rules ranked from 1 (worst) to 500 (best).	$\frac{2 \cdot \text{rank in population - 1}}{(\text{size of population})^2}$
arithmetic functions	+, -, *, /, norm, constant between (0,2)
Boolean operators	"if-then", "and", "or", "<", ">", "not", "true", "false"
functions of the data	"moving average", "local maximum", "local minimum", "lag of stock index", "current stock index"

Table 2: In-sample uniform portfolio results from the benchmark case

		# Rules	Excess Return	R_b-R_s	Sharpe	X*	(s.e.)	X_{eff}	X_{eff} B&H	alpha	(s.e.)	Trades per year	% long	B&H in-samp	B&H out-samp
1929	1935	10	18.26	310.58	0.63	10.50	(4.97)	6.43	-15.87	14.74	(4.21)	4.3	22.4	-8.53	6.34
1934	1940	10	6.38	41.11	0.52	6.69	(5.15)	5.08	-3.32	9.68	(4.14)	3.0	56.1	0.85	7.35
1939	1945	10	0.92	23.42	0.53	3.39	(3.34)	3.30	0.87	4.57	(3.09)	4.8	34.1	3.98	7.10
1944	1950	10	2.31	49.07	0.79	3.40	(4.59)	7.83	5.15	6.41	(2.45)	4.4	85.0	8.01	7.52
1949	1955	10	-2.83	10.30	1.25	0.30	(3.43)	9.92	12.39	1.99	(2.65)	4.1	78.0	15.63	6.47
1954	1960	10	1.48	159.19	1.14	2.14	(3.77)	9.46	7.87	3.08	(1.38)	2.7	93.4	12.06	6.70
1959	1965	10	2.42	70.81	0.98	3.17	(2.60)	5.65	2.60	6.42	(2.68)	10.2	83.0	7.31	6.28
1964	1970	10	5.85	49.96	0.63	5.04	(2.29)	2.80	-3.96	6.88	(2.39)	19.8	59.1	2.96	7.52
1969	1975	10	11.77	48.63	0.47	7.35	(3.04)	2.54	-10.72	10.27	(2.90)	7.3	44.3	-2.00	9.50
1974	1980	10	4.06	330.07	0.58	1.39	(0.90)	0.82	-5.51	NA	NA	1.0	0.8	4.67	9.97
		10	5.06	109.31	0.75	4.34	(3.41)	5.38	-1.05	7.12	(2.88)	6.2	55.6	4.49	7.47

Notes: Columns 1 and 2 provide the first and last years of the in-sample period for each case. Column 3 provides the number of rules that were evaluated in the in-sample period. Column 4 is the annualized out-of-sample excess return, net of transactions cost, to the uniform portfolio rule while column 5 is the annualized mean difference between average market returns on days that the rules were in the market and the days that they were out of the market. The portfolio mean return over the riskless rate, net of transactions cost, divided by the standard deviation of the portfolio return is in column 6. Columns 7 and 8 show the annualized X* risk-adjusted return measure, net of transactions cost and its standard error. Columns 9 and 10 show the X_{eff} measure of risk-adjusted returns to the uniform rule and the buy-and-hold strategy, respectively. Columns 11 and 12 show Jensen's alpha and its standard error. To avoid evaluating trivial rules, Jensen's alpha was not calculated if the proportion of long positions was less than one percent or greater than 99 percent. Such cells are marked "NA." The number of trades per year and the proportion of time spent in the market are in columns 13 and 14. The annualized buy-and-hold returns are shown in columns 15 and 16. The final row displays the means of the columns.

Table 3: Out-of-sample uniform portfolio results from the benchmark case

		# Rules	Excess Return	R_b-R_s	Sharpe	W=.5	Weight ExAnte	Weight ExPost	X*	(s.e.)	X_{eff}	X_{eff} B&H	alpha	(s.e.)	Trades per year	% long	B&H in	B&H out
1929	1934	10	-2.49	-19.89	-0.02	-0.03	-0.18	0.00	-0.46	(0.42)	-0.49	0.27	-0.70	(0.29)	1.9	16.5	-8.53	6.34
1934	1939	10	-0.83	-11.26	0.27	0.02	0.03	0.03	0.27	(1.15)	0.95	1.02	0.63	(0.79)	2.8	64.5	0.85	7.35
1939	1944	4	-3.47	-1.15	-0.13	-0.11	-0.31	0.00	-1.60	(1.12)	-2.05	0.47	-2.48	(1.20)	5.5	24.1	3.98	7.10
1944	1949	10	0.25	13.77	0.26	0.04	0.07	0.07	0.67	(1.55)	1.02	0.41	1.11	(0.85)	4.9	82.9	8.01	7.52
1949	1954	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	15.63	6.47
1954	1959	9	0.19	26.63	0.07	0.01	0.02	0.02	0.22	(2.17)	-1.07	-1.36	0.32	(0.35)	1.6	95.4	12.06	6.70
1959	1964	10	-0.74	36.26	-0.08	-0.03	-0.07	0.00	-0.77	(2.05)	-2.76	-2.33	-1.11	(1.22)	19.1	82.4	7.31	6.28
1964	1969	10	-1.56	24.42	-0.06	-0.05	-0.13	0.00	-1.28	(1.98)	-2.19	-1.28	-1.77	(1.32)	21.5	70.4	2.96	7.52
1969	1974	10	-3.29	115.48	-0.18	-0.06	-0.36	0.00	-1.19	(0.76)	-1.07	0.66	-1.38	(0.72)	8.3	21.6	-2.00	9.50
1974	1979	10	-3.42	30.09	0.01	0.00	-0.22	0.00	-0.01	(0.12)	-0.05	1.28	NA	NA	0.3	0.2	4.67	9.97
mean		8.3	-1.71	23.82	0.02	-0.02	-0.13	0.01	-0.46	(1.26)	-0.86	-0.10	-0.67	(0.84)	7.3	50.9	4.49	7.47

Notes: The column headings are as in Table 2 with the following exceptions. Column 3 provides the number of rules (out of 10 trials) that had positive training period returns. Columns 7 through 9 display the improvement in the out-of-sample Sharpe ratio, over a buy-and-hold strategy, of a portfolio constructed from the uniform trading rule and the buy-and-hold strategy. Column 7 shows the results with equal portfolio weights, column 8 requires that the portfolio weights for the two strategies be chosen on in-sample information and column 9 shows the maximum improvement obtainable with perfect foresight.

Table 4: Out-of-sample uniform portfolio results generated using the Sharpe ratio as the fitness criterion

		# Rules	Excess Return	R_b-R_s	Sharpe	W=.5	Weight ExAnte	Weight ExPost	X*	(s.e.)	X_{eff}	X_{eff} B&H	alpha	(s.e.)	Trades per year	% long	B&H in	B&H out
1929	1934	10	-2.47	-35.38	-0.39	0.00	-0.56	0.00	-0.02	(0.04)	-0.03	0.27	NA	NA	0.1	0.0	-8.53	6.34
1934	1939	10	-2.85	165.37	0.11	-0.01	-0.12	0.00	-0.09	(0.30)	-0.07	1.02	-0.08	(0.34)	1.5	11.1	0.85	7.35
1939	1944	1	-2.63	3.68	-0.03	-0.05	-0.21	0.00	-0.38	(0.92)	-0.95	0.47	-0.90	(1.04)	2.4	8.5	3.98	7.10
1944	1949	1	0.24	9.90	0.24	0.03	0.05	0.05	0.68	(1.68)	1.01	0.41	1.43	(1.42)	3.8	82.4	8.01	7.52
1949	1954	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	15.63	6.47
1954	1959	7	-0.26	5.82	0.07	0.01	0.01	0.02	0.18	(1.20)	-0.60	-1.36	0.35	(1.46)	5.4	42.4	12.06	6.70
1959	1964	5	0.09	11.91	-0.01	0.00	0.00	0.00	0.03	(1.45)	-1.31	-2.33	0.02	(1.34)	10.6	57.9	7.31	6.28
1964	1969	10	-1.32	22.20	-0.06	-0.04	-0.12	0.00	-0.76	(1.33)	-1.32	-1.28	-1.07	(0.89)	9.2	38.6	2.96	7.52
1969	1974	10	-2.69	210.18	NA	NA	NA	NA	-0.01	(0.13)	-0.06	0.66	NA	NA	1.0	0.4	-2.00	9.50
1974	1979	10	-3.42	NA	NA	NA	NA	NA	0.00	(0.08)	0.00	1.28	NA	NA	0.0	0.0	4.67	9.97
mean		6.4	-1.70	49.21	-0.01	-0.01	-0.17	0.01	-0.04	(0.79)	-0.37	-0.10	-0.04	(1.08)	3.8	26.8	4.49	7.47

Notes: See the notes to Table 3.

Table 5: Out-of-sample uniform portfolio results generated using the X* measure as the fitness criterion.

		# Rules	Excess Return	R_b-R_s	Sharpe	W=.5	Weight ExAnte	Weight ExPost	X*	(s.e.)	X_{eff}	X_{eff} B&H	alpha	(s.e.)	Trades per year	% long	B&H in	B&H out
1929	1934	10	-1.37	7.00	0.20	0.02	0.04	0.04	0.26	(0.69)	0.27	0.27	0.67	(0.78)	3.1	33.1	-8.53	6.34
1934	1939	10	-0.77	6.62	0.28	0.04	0.05	0.05	0.70	(1.10)	1.11	1.02	1.29	(1.17)	3.2	52.7	0.85	7.35
1939	1944	3	-3.64	-0.72	-0.17	-0.12	-0.35	0.00	-1.65	(1.02)	-2.10	0.47	-2.67	(1.09)	5.2	18.8	3.98	7.10
1944	1949	8	0.40	11.96	0.30	0.05	0.11	0.11	1.16	(1.44)	1.41	0.41	2.07	(1.22)	4.7	69.2	8.01	7.52
1949	1954	1	-0.49	7.40	0.04	-0.02	-0.03	0.00	-0.38	(1.99)	-1.42	-1.08	-0.53	(1.09)	4.7	88.1	15.63	6.47
1954	1959	4	-0.26	7.32	0.07	0.01	0.02	0.02	0.21	(1.13)	-0.49	-1.36	0.39	(1.26)	5.8	38.4	12.06	6.70
1959	1964	4	0.87	15.45	0.07	0.04	0.08	0.08	0.82	(1.80)	-0.77	-2.33	1.16	(1.56)	10.5	64.9	7.31	6.28
1964	1969	10	-1.40	20.23	-0.06	-0.04	-0.12	0.00	-0.96	(1.60)	-1.68	-1.28	-1.34	(1.08)	14.9	52.8	2.96	7.52
1969	1974	10	-2.35	4.40	0.10	-0.01	-0.08	0.00	-0.25	(0.75)	-0.11	0.66	-0.16	(0.61)	2.1	22.0	-2.00	9.50
1974	1979	10	-2.23	6.61	0.23	0.01	0.01	0.01	-0.03	(1.33)	0.46	1.28	0.12	(0.69)	2.5	35.8	4.67	9.97
mean		7	-1.13	8.63	0.11	0.00	-0.03	0.03	-0.01	(1.28)	-0.33	-0.19	0.10	(1.05)	5.7	47.6	4.49	7.47

Notes: See the notes to Table 3.

Table 6: Out-of-sample uniform portfolio results generated using the X_{eff} measure as the fitness criterion.

		# Rules	Excess Return	R_b-R_s	Sharpe	W=.5	Weight ExAnte	Weight ExPost	X*	(s.e.)	X_{eff}	X_{eff} B&H	alpha	(s.e.)	Trades per year	% long	B&H in	B&H out
1929	1934	10	-2.05	-20.08	0.13	0.00	-0.03	0.00	-0.10	(0.37)	-0.03	0.27	-0.09	(0.13)	0.2	20.1	-8.53	6.34
1934	1939	10	-0.87	6.23	0.29	0.04	0.05	0.05	0.54	(1.05)	1.03	1.02	1.15	(1.09)	4.3	54.6	0.85	7.35
1939	1944	7	-2.86	1.36	-0.06	-0.07	-0.24	0.00	-0.72	(0.96)	-1.24	0.47	-1.34	(1.08)	3.1	13.1	3.98	7.10
1944	1949	9	-0.01	10.99	0.24	0.03	0.05	0.05	0.48	(1.51)	0.85	0.41	1.02	(1.06)	5.4	80.3	8.01	7.52
1949	1954	1	0.00	NA	0.07	0.00	0.00	0.00	0.00	(2.10)	-1.08	-1.08	NA	NA	0.0	100.0	15.63	6.47
1954	1959	9	0.43	30.53	0.09	0.02	0.04	0.04	0.47	(2.14)	-0.80	-1.36	0.68	(0.37)	1.8	94.1	12.06	6.70
1959	1964	9	-0.49	42.63	-0.05	-0.02	-0.05	0.00	-0.51	(2.07)	-2.50	-2.33	-0.75	(1.41)	23.5	83.3	7.31	6.28
1964	1969	10	-2.43	17.14	-0.17	-0.09	-0.24	0.00	-2.08	(1.74)	-2.85	-1.28	-2.88	(1.33)	20.8	61.6	2.96	7.52
1969	1974	10	-4.26	97.01	-0.43	-0.12	-0.61	0.00	-1.98	(0.82)	-2.05	0.66	-2.69	(0.84)	10.2	15.0	-2.00	9.50
1974	1979	10	-3.42	NA	NA	NA	NA	NA	0.00	(0.08)	0.00	1.28	NA	NA	0.0	0.0	4.67	9.97
mean		9	-1.60	23.23	0.01	-0.02	-0.15	0.01	-0.39	(1.29)	-0.87	-0.19	-0.61	(0.91)	6.9	52.2	4.49	7.47

Notes: See the notes to Table 3.

Table 7: Results from ARIMA rules

In-sample Period	Search Criterion	AR Order	MA Order	Daily Dummy	Excess Return	R_b-R_s	Sharpe	X*	Trades per year	% long
1929-35	AIC	5	5	2	-34.19	15.36	-3.04	-32.62	145.56	0.575
	SC	1	0	1	-3.69	NA	NA	0.00	0.00	0.000
	Excess Return	1	2	1	-3.89	46.81	-0.07	-0.22	1.44	0.003
1934-40	AIC	2	2	1	-34.40	5.21	-3.00	-32.23	134.14	0.515
	SC	2	2	1	-34.40	5.21	-3.00	-32.23	134.14	0.515
	Excess Return	1	0	1	-0.17	-2187.39	0.32	-0.17	0.04	0.999
1939-45	AIC	5	5	3	-34.05	-10.33	-2.90	-32.27	118.79	0.549
	SC	1	0	1	-0.58	-610.46	0.25	-0.58	0.36	0.999
	Excess Return	1	0	1	-0.58	-610.46	0.25	-0.58	0.36	0.999
1944-50	AIC	5	4	2	-25.34	13.85	-2.03	-23.78	108.27	0.617
	SC	2	0	1	-24.35	-0.26	-1.78	-23.14	92.43	0.703
	Excess Return	1	0	1	-0.20	-2187.25	0.30	-0.20	0.04	0.999
1949-55	AIC	3	5	2	-30.43	15.35	-2.62	-29.54	131.82	0.670
	SC	2	0	2	-22.37	21.52	-1.68	-21.74	102.44	0.767
	Excess Return	1	0	1	-0.24	-1095.88	0.19	-0.24	0.10	0.999
1954-60	AIC	4	1	3	-28.12	18.57	-2.25	-27.29	125.33	0.685
	SC	2	0	2	-24.67	21.02	-1.87	-24.00	112.11	0.745
	Excess Return	1	0	1	-0.97	-107.73	0.12	-0.96	2.06	0.996
1959-65	AIC	2	3	2	-25.61	25.65	-2.28	-24.91	123.85	0.626
	SC	1	1	2	-23.71	22.48	-1.86	-23.27	109.16	0.768
	Excess Return	1	0	1	-7.68	-11.64	-0.46	-7.55	27.50	0.935
1964-70	AIC	5	5	3	-33.53	13.82	-3.04	-32.16	142.47	0.540
	SC	3	2	2	-30.84	10.80	-2.71	-29.56	128.99	0.571
	Excess Return	2	2	1	-16.05	-11.26	-1.38	-14.61	47.38	0.515
1969-75	AIC	5	5	2	-24.25	18.42	-1.96	-21.75	105.42	0.482
	SC	1	1	2	-26.49	11.81	-1.84	-25.36	109.95	0.765
	Excess Return	2	2	1	-11.67	5.59	-1.26	-7.14	29.71	0.060
1974-80	AIC	4	3	2	-14.23	-8.02	-0.75	-11.50	38.00	0.500
	SC	2	1	1	-58.92	0.18	-5.45	-56.17	224.89	0.495
	Excess Return	4	4	1	-14.06	-8.28	-0.74	-11.34	37.07	0.502

Notes: Column 2 shows the in-sample model selection criterion. Columns 3 and 4 show the chosen orders of the autoregressive and moving average components. Column 5 summarizes the deterministic component of the model: 1 indicates a simple constant, 2 indicates a weekend dummy on returns while 3 indicates that a full set of day-of-the-week dummies was used. For the other columns, see the notes to Table 2.

Table 8: Cumby-Modest tests of market timing

		Benchmark			Sharpe			X*			X_{eff}		
		Beta	(s.e.)	p-value	Beta	(s.e.)	p-value	Beta	(s.e.)	p-value	Beta	(s.e.)	p-value
1929	1934	0.41	19.07	0.49	NA	NA	NA	16.11	7.53	0.02	-44.33	62.40	0.76
1934	1939	13.81	6.74	0.02	21.06	15.28	0.08	11.07	4.88	0.01	11.37	4.79	0.01
1939	1944	-1.67	4.99	0.63	2.77	6.77	0.34	-3.48	5.84	0.72	0.64	6.40	0.46
1944	1949	24.35	6.97	0.00	11.23	5.10	0.01	18.02	5.40	0.00	18.66	6.22	0.00
1949	1954	NA	NA	NA	NA	NA	NA	7.53	6.48	0.12	NA	NA	NA
1954	1959	47.91	20.06	0.01	7.72	5.19	0.07	11.10	5.98	0.03	51.71	17.36	0.00
1959	1964	63.07	10.27	0.00	25.17	7.95	0.00	21.58	6.49	0.00	58.23	8.52	0.00
1964	1969	82.27	13.16	0.00	60.75	18.49	0.00	95.61	17.28	0.00	46.52	11.36	0.00
1969	1974	69.28	28.83	0.01	NA	NA	NA	15.12	24.30	0.27	17.68	18.32	0.17
1974	1979	NA	NA	NA	NA	NA	NA	26.08	26.16	0.16	NA	NA	NA
		7		6	6		3	9		6	7		5

Notes: The four panels show the results of Cumby-Modest tests (see equation (7)) on the four cases in which the fitness criteria were excess returns, the Sharpe ratio, the X* statistic and the X_{eff} statistic. The columns of each subpanel show the coefficient, its standard error and its p-value. The final row displays the number of positive betas and the number of p-values less than 0.05. To avoid evaluating trivial rules, the test was not performed if the proportion of long positions was less than one percent or greater than 99 percent. Such cells are marked “NA.”