Testing the Efficient Market Hypothesis in the Australian Share Market using a Differential Evolutionary Algorithm

Phil Simmons
School of Business, Economics & Public Policy
Faculty of the Professions
University of New England, Armidale, Australia, 2350

Abstract
The proposition that a relatively new technology such as a Differential Evolutionary Algorithm (DEA) can violate the weak form of the Efficient Markets Hypothesis is tested using daily data from the Australian share market from 2000 until 2008. An options trading strategy based on forecasts from a DEA is shown to perform better than a buy and hold strategy over parts of the sample space and, on average, over all of it. The paper concludes speculators may make supernormal profits from new methodologies however that such profits are unlikely to be sustained.

1. Introduction

The earliest published work on efficient markets was undertaken by Bachelier in 1900 who proposed that speculation should be a ‘fair game’ inasmuch as expected profits to the speculator should be zero. This proposition was developed by Fama (1969) who formalised such fair game theories generally as:

\[ E(\tilde{p}_{jt+1}|\varphi_t) = [1 + E(\tilde{r}_{jt+1}|\varphi_t)]p_{jt} \] (1)

where the tilde denotes random variables at \( t \), \( \tilde{p}_{jt+1} \) is price of \( j \) in \( t+1 \), \( \tilde{r}_{jt+1} \) is a one period per cent return similarly sub-scripted and \( \varphi_t \) is information at \( t \). It follows \( \varphi_t \) is fully utilised (or “fully reflected”) in determining next period price. Fama extended theory beyond the fair game model and, pertinent to this study, formalised the sub-martingale version of (1):

\[ E(\tilde{p}_{jt+1}|\varphi_t) \geq p_{jt} \text{ (or } E(\tilde{r}_{jt+1}|\varphi_t) \geq r_{jt}) \] (2)

This extension of theory has an important property, derived from the non-negativity property of changes in sub-martingales, that no trading rule based only on \( \varphi_t \) can produce greater profits than could be achieved with a simple “buy and hold” strategy in security \( j \). In addition, Fama introduced the concepts of weak form, semi-strong form and strong form tests of market efficiency. Weak form is where \( \varphi_t \) is historical prices, semi-strong form includes all publically available information and strong form includes publically and privately held information.

In this study, I argue weak form efficiency may be violated, at least temporarily, if a new methodology is developed suitable for forecasting prices and that the Differential Evolutionary Algorithm (DEA) method that emerged in the 1990s (Storn and Price, 1997) was such a development. Malkiel (2003) argued predictable patterns in prices can exist however seem to disappear when they are published in the finance literature and, presumably, become widely arbitraged. He used the January effect as an example which disappeared shortly after publication of Haugen and Lakonishok’s book The Incredible January Effect in 1988. In the same vein, new methodologies, artefacts of the research process, may reveal market inefficiency until speculators become aware and arbitrage them.

In this study, a DEA is used to examine the efficiency of the index for the top 200 stocks in the Australian share market (ASX200) using daily closing prices between 6th January 2000 and 1st of February 2008. Risk-adjusted returns from a daily trading strategy based on forecasts from a DEA are compared with those from a buy and hold strategy over the same period. Evidence is found that the Australian Stock market was weak form inefficient at times in the sample period and, on average, weak form inefficient for the whole period.

In the next section, the DEA and forecasting procedure is outlined and in Section 3 the performance of the DEA as a simple forecasting tool ignoring transaction costs is reported. In Section 4, the financial results from a trading strategy that uses the DEA with non-zero transaction costs are compared with returns from a buy and hold strategy in the ASX200. Discussion and conclusions are provided in Section 5.
2. Methodology

My interest in numerical algorithms for solving problems in economics was sparked by Szpiro (1997) who solved an investment problem with uncertainty using a genetic algorithm based on Holland (1975). Szpiro’s result was important for several reasons, both pertinent to this study. First, his algorithm was solved in a stochastic environment. The investing environment changed each period so the solution had to be robust enough to deal with a broad range of investment conditions. The algorithm appeared to go through a learning process improving its performance with the accumulation of experience. The second feature of Szpiro’s algorithm was it appeared to exhibit ‘caution’ in its behaviour. Ostensibly it was maximising the value of an asset over time however its behaviour turned out not to be profit maximising at all. The algorithm learned to avoid ‘over extending’ during volatile periods and appeared to be survivalist rather than maximising. Szpiro’s study led to Cacho & Simmons (1999), an application of those ideas to farm investment using a Traditional Genetic Algorithm from Goldberg (1989) then, later, Simmons & Cacho (2005) applied an evolutionary algorithm to a pollution tax problem requiring caution on the part of government. In this study, we use a similar methodology to examine market efficiency.

The DEA was used to obtain values for the coefficients, $\tilde{\alpha}_{i}(i=1,2)$ in the forecasting Eq.:

$$x_jo_{t+1}^{\tilde{F}} = \tilde{\alpha}_1 x_jo_t + \tilde{\alpha}_2 xsp_t$$

where $F$ denotes the forecast in the next period, the tildes indicate the variable character of the coefficients over time, $x_jo$ is daily closing value of the ASX200 and $xsp$ is daily closing value of the Standard & Poors 500 index from the New York market. New York closes four hours before Sydney opens and the time subscript, $t$, refers to closing bell in Sydney on say, Wednesday, and simultaneously, the closing bell on Tuesday in New York. The forecast is for $t+1$, Thursday in Sydney. Weekends are assumed not to exist and values for weekday public holidays are the average of adjacent trading days.

The DEA obtains coefficient values for each period by solving:

$$\min_\alpha J = MOTAD(x_jo_{t+1} - \tilde{\alpha}_1 x_jo_t - \tilde{\alpha}_2 xsp_t).$$

Data from $t=1$ were applied to Eq. 3 to obtain a forecast for $t=2$ and then the forecasts were used to obtain new values for coefficients $\alpha$, using ranking and crossover based on fitness measures from Eq. 4. New values of $\alpha$ were then applied to $t=2$ data to obtain the forecast for $t=3$ and so on until $t=1999$, providing 1999 forecasts in all. Twenty five candidate values for $\alpha$ were used in each ‘generation’ resulting in 25 forecasts, one for each candidate, in each of $t=1$ to $t=1990$. The 25 forecasts were averaged for each value of $t$ and the whole experiment repeated 500 times to remove noise by averaging the forecasts from each ‘run’.

Readers are referred to Storn and Price (1997) for a detailed description of a DEA that is similar to the one used in this study and an excellent discussion of how these algorithms perform in a range of tasks. A brief description of the DEA using some of the terminology from Storn and Price’s study is provided below.
Step 1 Agents $\alpha_i$ ($i = 1,2$) are initialised by creating two populations, $i$, ($i = 1,2$), both $x_{ij}$ ($j = 1,2,25$) where $x_{ij}$ are uniformly distributed random numbers representing the candidates for $\alpha$ such that the search space is $-1 \leq x_{ij} \leq 1$.

Step 2 Run loop:

Loop step 1 Three candidates $a$, $b$ and $c$ are randomly chosen from $x$ and combined as

$y_i = a_i + F(b_i - c_i)$ where, after experimentation, $F$ was set to 0.2.

Loop step 2 ‘Crossover’ occurs based on a user specified crossover probability, around 0.8, so 20% of $y$ are randomly selected to be replaced by members of $x$. Following experimentation with different values, the crossover probability was set at unity so none of the previous population found its way into $y$.

Loop step 3 A pair-wise comparison of the fitness of candidates in corresponding positions in $x$ and $y$ was undertaken and the fitter of each pair added to a new vector $z$.

Loop step 4 The $z$ candidates were ranked on the basis of fitness.

Loop step 5 Candidates in $z$ are used in Eq. 3 to obtain forecasts for later analysis.

The loop is repeated 1999 times with $z$ becoming $x$ in each successive iteration and the fitness function being updated with new data values and new $\alpha$ values until the end of the data series is reached in February 2008.

In fact, a fair bit of experimenting occurred with specifications and constraints. Crossover and scaling ($F$) were set at unity and 0.2 respectively after a number of alternative values were tested. The population size, or number of candidates, was also tested. Populations greater than 25 performed either the same or worse while smaller populations tended to give noisier results. Several ideas were explored to impose bounds on the candidates. First, candidates violating bounds were replaced with the values of the boundaries violated. Second, various penalties were applied to the objective function to discourage bounds violations. Finally, after neither of these specifications made any difference, the simple expedient of imposing the search space at initialisation and ignoring bound violations in later generations was adopted.

Another area of experimentation was in various data transformations where five differencing schemes were considered. The logic was that forecasting power might be located in particular parts of the frequency spectrum of the data and these parts might be enhanced through different types of differencing. We undertook the following differencing: no differences, conventional first differences, first percentile differences, first differences followed by additional percentage differencing and conventional second differences. The percentile differences performed best (and were thereafter used exclusively) conventional differencing was next best, no differencing next best and any form of double differencing was useless.

---

1 All computations were undertaken using Wolfram’s Mathematica Versions 6 and 7.
Finally, we started the algorithm at 6th January 2000 and did not record the forecasts until 2nd June 2000. This took on board possible ‘learning’ for the candidate population at the start of the run. The run was stopped at 1st February 2008.

3. Forecasting Results

The algorithm was ‘cautious’ in its predictions in the sense discussed earlier with the size of forecast changes only 25% the size of actual changes. An arbitrary proportional correction was made to all forecasts so that expected forecast changes equalled expected actual changes. This correction turned out to be unimportant as the results from the study eventually depended on only the algorithm’s ability to predict directions of, rather than magnitudes of, price change.

The first test was Theil’s Inequality Coefficient (Theil, 1966) where the coefficient is defined (Koutsiannis, 1981):

\[
U = \frac{\sqrt{\sum (P_i - A_i)^2}/n}{\sum A_i^2/n}
\]  

(5)

Where \( P \) and \( A \) refer to predicted and actual price changes respectively. \( U \) is bounded, \( 0 \leq U \leq \infty \) and, if \( U = 0 \), forecasts are perfect and if \( U > 1 \) the model predicts worse than a naive forecast. That is, worse than the rule \( price_t = price_{t-1} \). The result from the DEA was \( U = 7.52 \) which was not encouraging. However, Theil’s U statistic can be decomposed to reveal sources of variation in errors. \( U_m \) measures bias proportion, \( U_s \) measures variance proportion and \( U_c \) measures covariance proportion where \( U_m + U_s + U_c = 1 \). The results for these measures follow:

\[
U_m = \frac{(\overline{P} - \overline{A})^2}{\sum |P_i - A_i|^2/n} = 0.00
\]

(6.1)

\[
U_s = \frac{(\overline{S_P} - \overline{S_A})^2}{\sum |P_i - A_i|^2/n} = .265
\]

(6.2)

\[
U_c = \frac{2 \cdot 1 - \frac{r_{PA}}{\overline{S_P} \overline{S_A}}}{\sum |P_i - A_i|^2/n} = .735
\]

(6.3)

where \( S_P \) and \( S_A \) are the standard deviations of \( P \) and \( A \) respectively and \( r_{PA} \) is the correlation coefficient between predicted and actual changes. \( U_m \) is zero because bias was corrected as explained above. More than a quarter of variation in errors arises from variance errors, \( U_s \). The latter result is consistent with the DEA possibly forecasting the direction of change well and not inconsistent with the market having weak form inefficiency providing trading rules exploit predicted directions of price changes rather than their magnitudes.

The DEA made 982 predictions of price rises and 1017 predictions of price falls where, over the sample, there were in fact 1085 rises and 914 falls. Of these, 54% of the predicted downward movements were correct and 51% of upwards movements. In total, 52% of the predictions of direction were correct. It is pertinent whether the DEA predicted bigger price movements better since these would yield higher speculative returns. It turns out big price...
movements were predicted better than small ones. The first quartile (by absolute size) of actual price movements were predicted correctly 51% of the time while the fourth quartile were predicted correctly 58% of the time.

A strategy of using the predicted direction of price movements from the model to invest $1 in either short or long positions at close of business each day and holding the position for precisely 24 hours yielded an average return of 3.5 cents per day with zero transaction costs. The cumulative returns over the sample are reported in Figure 1 where it was assumed the dollar value of the index was invested for 24 hours each day at close of business allowing cumulative profits on the vertical axis to be interpreted either as dollars or as index points.

![Figure 1: Cumulative profit from 1st Feb 2000 to 1st Feb 2008](image)

The return is variable and rises relatively quickly over the first 300 observations (60 weeks), then performs poorly for a while then improves steadily as it enters the period of the global stock market boom until November 2007 when the Global Financial Crisis (GFC) started to impact on stock prices. Interestingly, the 9-11 Terrorist Attack, observation 441, and GFC (the last 30 observations) both reduced the forecasting power of the DEA dramatically for a number of periods. In Figure 2, the ASX200 index is superimposed on Figure 1 and these two special events are pinpointed.
4. Results for Efficient Markets Hypothesis

The Sharpe Ratio, also called the Reward to Variability Ratio (RVAR), is:

\[ RVAR_p = \frac{ar_p - ar_f}{\rho_p} \]  

where \( ar_p \) is the portfolio return, \( ar_f \) is the risk free rate and \( \rho_p \) is the standard deviation of the portfolio return (Sharpe, 1966). The annual return from buying and holding the ASX200 is calculated for the RVAR measure by multiplying the average daily return, 0.0348%, by 252 and turns out to be 8.77% p.a. ignoring dividends. The risk free rate was the 90 day Bank Accepted Bill rate from the Reserve Bank of Australia website which averaged 5.7% over the period. The annualised standard deviation, \( \rho_p \) was found by dividing the standard deviation of the nominal daily return by \( \sqrt{1/252} \). It was 13.9%. \( RVAR_p \) turned out to be 0.24. Incorporating dividends, assumed for simplicity to be risk free, increases this ratio to 0.44.

Predicted directions of price movements from the DEA were used in an option trading strategy with either a call or put being purchased at close of business each day and closed out after 24 hours. Based on industry advice, transaction costs of 0.5% per day were assumed to cover bid ask spreads, brokerage and capital costs. Volatility was measured for 10 working days prior to each trading day and interest rates were set to 5.7%. Option contracts were assumed to have 60 days to maturity and strike prices were set one per cent below spot prices at the time the position was taken. This information was used to solve Black and Scholes’ option valuation formula (Cox & Rubenstein, 1985) to obtain returns from the strategy. An annualised net return of 126% with a standard deviation of 527% was achieved. The RVAR value was 0.53, higher than the buy and hold strategy result of 0.44. This indicates the presence of weak form inefficiency over the sample period.

5. Discussion and Conclusions
The DEA produced supernormal profits over parts of the sample period and, on average, over all of it. The performance of the DEA wavered, being sometimes strong and sometimes weak. Performance appeared to improve during periods of rising prices suggesting the DEA might have acted as a momentum trader. However, this was not the case. The DEA took as many short positions as long positions during these boom periods. The other aspect of performance was the DEA not only did not perform well following the disasters of the 911 terrorist attack and setbacks from the GFC, it actually went into reverse and started losing money (Figure 2). A speculator would have done well using the DEA in the aftermath of these events by doing the exact opposite of what it suggested! DeBondt and Thaler (1995) argue investors are subject to periods of optimism and pessimism that cause systematic components in price series in the long run. These results indicate this may also be true of the very short run. During optimistic periods, speculators are likely to overvalue information leading to exploitable opportunities from both momentum and related downward corrections in the very short run. The DEA would rapidly detect such non-random patterns in prices and, given its variable coefficients, adapt to them. One might conjecture waves of pessimism might have exactly the opposite effect and result in information being undervalued and that the DEA, which has “learned” to operate during more optimistic times, might get things exactly wrong! However, it is not clear how the dynamics of such price movements in pessimistic times might work in this context. Further research is needed.

6. References


Cacho, O. and Simmons, P.R., A genetic algorithm approach to farm investment, Australian Journal of Agricultural and Resource Economics, 43:3, (1999), 305-323.


