Final version appeared in
Tsang, E.P.K., Markose, S. & Er, H., Chance discovery in stock index option and future arbitrage, New Mathematics and Natural Computation, World Scientific, Vo.1, No.3, 2005, 435-447

New Mathematics and Natural Computation
© World Scientific Publishing Company

CHANCE DISCOVERY IN STOCK INDEX OPTION AND FUTURES ARBITRAGE

EDWARD TSANG

Department of Computer Science, University of Essex, Wivenhoe Park
Colchester, C04 3SQ, United Kingdom
edward@essex.ac.uk

SHERI MARKOSE

Department of Economics, University of Essex, Wivenhoe Park
Colchester, C04 3SQ, United Kingdom
scher@essex.ac.uk

HAKAN ER

Department of Business Administration, Akdeniz University, Dumlupinar Bulvari, Kampus
Antalya, 07058, Turkey
erhakan@akdeniz.edu.tr

Received Day Month Year
Revised Day Month Year

The prices of the option and futures of a stock both reflect the market's expectation of futures changes of the stock's price. Their prices normally align with each other within a limited window. When they do not, arbitrage opportunities arise: an investor who spots the misalignment will be able to buy (sell) options on one hand, and sell (buy) futures on the other and make risk-free profits. Historical data suggest that option and futures prices on the LIFFE Market do not align occasionally. Arbitrage chances are rare. Besides, they last for seconds only before the market adjusts itself. The challenge is to not only to discover such chances, but to discover them ahead of other arbitragers. In the past, we have introduced EDDIE as a genetic programming tool for forecasting. This paper describes EDDIE-ARB, a specialization of EDDIE, for forecasting arbitrage opportunities. As a tool, EDDIE-ARB enables economists and computer scientists to work together to identify relevant independent variables. Trained on historical data, EDDIE-ARB was capable of discovering rules with high precision. Tested on out-of-sample data, EDDIE-ARB out-performed a nave ex ante rule, which reacts only when misalignments were detected. This establishes EDDIE-ARB as a promising tool arbitrage chances discovery. It also demonstrates how EDDIE brings domain experts and computer scientists together.

Keywords: Keyword1; keyword2; keyword3.

1. Introduction

Chance discovery refers to machine discovery of situations that arise very rarely, but potentially have significant implications, either favourably or unfavourably. [34 35 1 2]

The scarcity of chances makes chance discovery particularly difficult in machine learning, as only a very small percentage of the data is of significant interest.

In the literature of machine learning, positive results are often presented. Hidden behind such results was a methodology and possibly failures. Failures were not reported partly because they may be considered uninteresting scientifically. In fact, the route to success is sometimes as important as the success itself. This paper reports a positive experience in chance discovery. This experience is probably applicable to other applications of similar nature.

The FTSE 100 spot index has a stock index futures and a European style index option traded on it. A *futures* contract requires the investor to buy or sell at certain prices at a future time. A *put option* contract give the holder the right to sell a certain quantity of an underlying security to the writer of the option, at a specified price (strike price) on or up to a specified date (expiration date). A *call option* contract gives the holder the right to buy a certain quantity of a security from the writer of the option, at a specified price (the strike price) on or up to a specified date (the expiration date).

The price of futures and options both reflect the investors' expectation of the price changes in the spot market, though they bear different risks and returns. More importantly, both options and futures should converge to the theoretical price, within some margin, which reflects transaction costs and interest rates. Therefore, if one of them is over-priced and the other is under-priced beyond certain margins, one can buy (or sell) options on one hand, and sell (buy) futures on the other, and obtain a guaranteed risk-free profit. Such operations are called *arbitrage*.

EDDIE (which stands for *Evolutionary Dynamic Data Investment Evaluator*) is a decision support tool for forecasting. [41 42 43 44] EDDIE is an interactive tool, based on genetic programming (GP) [20 22 23], a branch of evolutionary computation [19 18 31], for exploring the space of decision trees. The effectiveness of EDDIE in forecasting has been reported in [41 42 43]. Tsang et al described the development of EDDIE as a decision support system [44], thus establishing the practicality of EDDIE. As a tool, it helps to channel expert knowledge into the system for forecasting. This paper reports our experience in channelling economic knowledge into arbitrage forecasting.

2. The Put-call-futures Arbitrage Strategy

Arbitrage profit opportunities can be calculated theoretically based on futures and options prices, transaction costs and interest rates (e.g. See Ref. 32 10 45). The equilibrium pricing relationship between the index option and the stock index is based on the Stoll Put-Call (P-C) parity relationship. [38] Here, typically the option price is derived from the underlying stock index. However, it is well known from studies on the efficiency of index options (see, e.g. Ref. 11 15 16), that the high costs of hedging index option positions with a portfolio based on the constituent shares of the spot index may lead to market inefficiency. Therefore, options and futures are often used in arbitrage. [24] [13] [14] [6]

The Tucker [45] result that combinations of put and call options with the same exercise price can replicate payoffs of futures positions with the same expiration and on the same cash asset underpins the put-call-futures (P-C-F, for short) parity relationship. Arbitrage strategies based on the P-C-F parity involves time synchronized triplets of prices on a put, a call and a futures contract. For example, a risk free *short hedge arbitrage* portfolio can be constructed by the following transactions:

(i) Shorting a futures contract,
(ii) Buying a call option,
(iii) Shorting a put option, and
(iv) Borrowing the present discounted value of the futures price and lending the same for the exercise price.

The idea is to short futures (operation i), but protect it by a synthetic long futures position (operations b, c and d together). The upper bound of a futures bid price, denoted by F_{bt}, is given by

$$F_{bt}e^{-r_a(T-t)} \leq C_{at} - P_{bt} + Xe^{-r_b(T-t)\tau} + TC \tag{2.1}$$

Here, T is the expiration date and t is today, i.e. $T - t$ is the remaining time to maturity; C_{at} is the option's call premium at the ask, P_{bt} is the option's put premium at the bid. X is the exercise price for the option. TC is the transaction cost. ra is the interest rate on the borrowing to finance the futures. r_b as it has to lent. If the condition in (2.1) is violated then the arbitrageur by definition will make a risk free profit equal to:

$$F_{bt}e^{-r_a(T-t)} - [C_{at} - P_{bt} + Xe^{-r_b(T-t)} + TC] > 0 \tag{2.2}$$

by shorting the futures and by creating the synthetic long futures given on the R.H.S of (2.1).

Let the spot price at time T be S_T. Should S_T rise above X, the arbitrageur will exercise the call option (to buy at X) and honor the futures contract (to sell at S_T) and make a profit. If condition (2.2) does not hold, such profit will be sufficient cover all the other costs and earn the arbitrageur a profit. Should ST fall below X, the arbitrageur will have to honor the put option and buy at X. This is offset by the income in selling the futures at F_{bt}. Again, if condition (2.2) does not hold, the surplus will be sufficient to generate a net profit. In fact, the spot price plays no part in defining the net profit.

Likewise we have a risk free long hedge arbitrage portfolio by combining a long futures contract and a short synthetic futures position obtained by shorting a call at the bid, being long in the put at the ask and by lending the present discounted value of the futures price and borrowing the same for the exercise price. The lower bound of the ask futures price, denoted by Fat, is given by

$$F_{at}e^{-r_b(T-t)} \geq C_{bt} - P_{at} + Xe^{-r_a(T-t)} - TC \tag{2.3}$$

If the present value of the ask futures price is less than the R.H.S of (2.3), then the arbitrageur buys the futures and sells the portfolio on the R.H.S of (2.3). This nets a risk free profit equal to

$$[C_{bt} - P_{at} + Xe^{-r_a(T-t)} - TC] - F_{at}e^{-r_b(T-t)} \geq 0. \quad (2.4)$$

3. LIFFE Intraday Data for FTSE-100 Index Options and Futures

We studied the availability of arbitrage opportunities in the FTSE-100 Index options and futures market. If such opportunities are available, one may attempt to discover them ahead of others. Our research started with intraday historical tick data on the FTSE-100 European style index option traded on the LIFFE from March 1, 1991 to June 18, 1998. With inductive learning in mind, intraday tick data for a further period from June 30, 1998 to March 30, 1999 was earmarked for out-of-sample test to see whether the introduction of the LIFFEConnect electronic trading had changed the fundamental conditions. The tick data records contain both two way bid/ask quote data and the transactions/trade price data. Each record is time stamped with the following information: (i) option identity, i.e. whether it is a call or put; (ii) the maturity date; (iii) the exercise price; (iv) the price of contract (bid and ask for quoted prices and transaction/trade price if indicated as a trade record); (v) the time synchronized underlying FTSE-100 stock index price.

The risk free interest rate used for borrowing funds in the hedge portfolio is the London Interbank Offer Rate (LIBOR) maturing on the day closest to the expiration date of the option. The tick data on FTSE-100 is obtained from Data Stream. The bid or lending interest rate is conventionally taken to be 1/8 of the LIBOR rate. The mid interest rate is the mean of the above two rates.

We processed the tick data to identify put-call-futures parity, which indicate arbitrage opportunities. For trade data, we matched all calls and puts with the same exercise price and maturity and traded within the same minute. This pair is matched further with a futures contract traded within a minute of the time stamp of the call-put pair. In the LIFFE tick trade data from January 1991 to June 1998, 15,670 P-C-F arbitrage trade price triplets were identified out of the millions of tick data. Out of these triplets, before transaction costs, 8,073 profitable short arbitrage triplets and 7,410 profitable long arbitrage triplets were identified. Only 2% of the triplets had no profitable arbitrage opportunities. 1,641 of short arbitrage were followed up. In reality, one may assume transaction cost for a broker/market maker to be less than 0.1% of the value of the futures contract, which works out to be about £60 per P-C-F arbitrage. (Note, however, that as many of the options remain unexercised at maturity, the full inclusion of these round trip commission costs as percentage of the index option premium is an overestimation of costs.) Based on this assumption, 2,345 (or 29%) of the 8,037 triplets were profitable.

An ex post analysis of efficiency violations for short arbitrage positions was conducted using the 8,073 P-C-F price triplets for the sample period (1991-1998) at costs appropriate to the market maker/broker. Some 42% of the sample yields profitable short P-C-F arbitrage opportunities. The profits reported are those that accrue if the arbitrageur could have obtained as quotes the trade prices recorded at these points in time. In a real time setting, the arbitrageur may not be able to obtain the quoted prices (due to delay). Therefore, an arbitrageur may not be able to exploit all the profitable short arbitrage opportunities, especially if it reacts passively.

It is worth noting that as trading volume in the index options increased, we have a three fold increase of P-C-F arbitrage opportunities by 1994. By the end of 1998, there is over a tenfold increase since the inception of index options trading in LIFFE. On average the numbers of profitable P-C-F arbitrage opportunities are far outnumbered by loss making P-C-F opportunities in all years. However, in the years after 1995, the average total profitability of P-C-F positions become positive with the loss making arbitrages generating smaller losses than the gainful ones. In other words, the returns to P-C-F arbitrage are significant if the arbitrageur can successfully 'pick' the cherry.

Overall, our study suggests that on average for all periods to maturity the profits from the arbitrage are substantial and statistically significant [30]. The spot month nets on average £285 and the 20-50 day period to maturity yields over £316.

4. Defining our forecasting task

Having established that arbitrage opportunities exist in the FTSE-100 index market, our next task was to define our task: what should we aim to forecast. Building on EDDIE's track record, we defined our task to be forecasting whether the current moment is an arbitrage opportunity, which is a Boolean question. By looking at historical data, we can determine in hindsight whether each tick is an opportunity or not an opportunity. Such hindsight information is used for training EDDIE.

With our task defined, we defined the criteria for evaluating the quality of a set of forecast. Table 1 defines the notations that we use to summarize forecasting results. In Table 1, P_- is the number of times that a positive case is predicted. In our example, a positive case means an arbitrage opportunity. P_+ is the number of times that a negative case is predicted. A negative predication is a true negative (TN) if in reality the case is negative. It is a *false negative* (FN) if the case turns out to be positive. Similarly, a positive prediction is a *true positive* (TP) if in reality it is positive; otherwise it is a false positive (FP). The total number of positive cases in reality is R_+, and the total number of negative cases is R_-. Naturally, R_+ plus R_- is equal to P_- plus P_+, the total number of cases, N.

A perfect prediction would have TN plus TP equals N, i.e. all the cases are correctly predicted. Failing that, the problem is a multi-objective optimization problem. In EDDIE, three measures are used to evaluate a set of predictions: The *rate of correctness* (RC) measures the total number of correct predictions. The *rate of missing chances* (RMC) measures the percentage of actual positives classified as negative. The *rate of failure* (RF) measures the percentage of false positives over

Table 1. Contingency table for summarizing results in two-class classification predictions

	Predictions		
Reality	# of True Negatives (**TN**)	# of False Positives (**FP**)	**Actual Negative**($\mathbf{R}_-$) R_-=TN+FP
	# of False Negatives (**FN**)	# of True Positives (**TP**)	**Actual Positive**($\mathbf{R}_+$) R_+ = FN + TP
	Predicated Negative ($\mathbf{P}_-$) P_- = TN + FN	Predicated Positive ($\mathbf{P}_+$) P_+ = FP + TP	Total Number of Cases (**N**) $N = P_- + P_+ = R_- + R_+$
Performance Measures:		Rate of Correctness: RC = (TP + TN) ÷ N Rate of Missing Chances: RMC = FN ÷ R+ Rate of Failure: RF = FP ÷ P+	

the number of positive predictions. RF is important because false negatives potentially lead to loss in investments. Note that precision is used in the literature to refer to 1 - RF, *recall* used to refer to 1 - RMC.

5. Identifying independent variables and refinements

After defining the task, we need to decide what independent variables are relevant to predicting opportunities. Following naturally from equations (2.1) to (2.4), the following independent variables are obvious candidates:

- The exercise price of theoptions, which is also called the *strike price*
- The price of the security in the spot market, which is also called the *underlying price*
- The call premium
- The put premium
- The futures price
- The number of days to maturity
- The profit or loss after transaction cost, based on the model described in Section 2.

Unlike our forecasting work presented in [25] [26] [42], we input to EDDIE independent variables that are used in fundamental analysis. For convenience, we refer to this version of EDDIE which specializes in finding arbitrage opportunities EDDIE-ARB.

Initial runs found no arbitrage patterns. EDDIE-ARB failed to recommend any arbitrage opportunities. (Neither did it recommend any false opportunities.) We went through over a dozen of iterations to refine the independent variables. These refinements provide EDDIE-ARB with more meaningful variables, which are likely to be discovered repeatedly. Following are some of the major refinements:

(i) Call minus put option premiums is used to replace call and put premiums. This is because the spread between call and put plays a much more prominent role in determining arbitrage opportunities than the absolute values of call and put themselves.

(ii) Moneyness, the ratio of the spot price over the strike price (Spot ÷ Strike) is introduced. This variable is introduced as the in, at and out of the moneyness of call and put options has an impact on the arbitrage profits.

(iii) Basis, which is defined by the Futures price minus Spot price, is introduced. This variable helps to capture mis-pricing in the futures leg of the arbitrage. This is later divided by the Futures price to obtain a relative value. Sometimes a small change in this ratio could be significant (given that intraday data is used). However, when the numbers get too small, they are subject to the computer's precision problem. Therefore, these ratios are multiplied by 10,000. In other words, Basis = (Futures - Spot) × 10,000 ÷ Futures.

(iv) Profit or loss after transaction cost is replaced by the profit or loss divided by the futures price. This ratio is more meaningful than the absolute value of the profit or loss because the significance of the latter is relative to the futures price, but futures prices may fluctuate. This variable is then multiplied by 1,000,000 to improve precision in the calculation significantly.

Refinement in EDDIE-ARB was only possible when economists and computer scientists work closely together. The economists understand what are relevant to arbitrage opportunities and the computer scientists understand complexity of the search space. Some of the above refinements may look obvious or unimportant to non-experts in machine learning, who might have unrealistically high expectations what machine learning can do. The refinements helped EDDIE-ARB to focus on building blocks which are more likely to be relevant to the forecasting task.

6. EDDIE-ARB for arbitrage chance discovery

One typical characteristic of chance discovery applications is that chances are scarce. This applies to arbitrage chance discovery. In a situation where, say, 3% of the cases are actually positive, a predictor could secure a high rate of correctness by making no positive predictions, as shown in Table 2. For the predictions shown in Table 2, the rate of correctness of the prediction is 97%, which is very high, but the rate of missing chances is also high.

Table 2. Contingency table showing chances are scarce and no positive predictions are made

	Predictions		
Reality	97	0	97
	3	0	3
	100	0	3
Rate of Correctness: RC = (0 + 97) ÷ 100 = 97%			
Rate of Missing Chances: RMC = 3 ÷ 3 = 100%			
Rate of Failure: RF = 0 ÷ 0 (undefined)			

Table 3 shows the results of applying EDDIE-ARB with the constrained fitness function to LIFFE FTSE-100 index 1991 to 1998 data. Parameters were set to instruct EDDIE-ARB to generate decision trees with low rate of failure (RF). This is what was achieved. Although the RC dropped from 94.18% to 27.55% in the test data, RF was 80.38% in the training data, and 93.01% in the test data. Note that both the substantial fall of RC and modest increase of RF from the training to the test period were atypical to EDDIE. As explained in Section 3, the market saw huge increase of volume from 1994 to 1998. This suggests some underlying changes in the market behaviour which may have explained for such atypical observations here.

Table 3. Application of EDDIE-ARB to LIFFE FTSE-100 index data

Training 4178 Rows 29 January 1991 to 30 December 1996				**Testing** 3895 Rows 1 January 1997 to 18 June 1998			
	Negative Predictions	Positive Predictions			Negative Predictions	Positive Predictions	
Real Negative	3894	168	4062 (97.2%)	Real Negative	867	2743	3610 (92.7%)
Real Positive	75	41	116 (2.8%)	Real Positive	79	206	285 (7.3%)
	3969 (95%)	209 (5%)	4178		946 (24.3%)	2949 (75.7%)	3895
RC 94.18%, **RMC** 64.66%, **RF** 80.38%				**RC** 27.55%, **RMC** 27.72%, **RF** 93.01%			

The scarcity of chances and high rate of failure in previous runs led us to explore further actions to help EDDIE-ARB in its search. To reduce EDDIE-ARB's incentive to produce no positive predictions, we used the 1,641 triplets that were followed up. About 25% of these cases resulted in profits after transaction cost. This means that, with the reduced set of data, if EDDIE-ARB does not generate positive predictions, its RC can only be 75%.

The 1,641 rows of data were divided into training (1,000 rows, or 61%) and test (641 rows, or 39%) sets. The former contains 1,000 rows and the latter 641 rows. Data was picked *randomly* rather than chronologically to give us more samples for training and testing.

EDDIE uses a constrained fitness function. [25 43] First, EDDIE allows the user to give a weight to each of the three criteria. Given that negative positives potentially lead to losses, whereas missing chances do not cost the arbitrager (apart from the opportunity cost), we made sure that a non-trivial (negative) weight is given to rate of failure. We do not want to miss chances either, given then scarcity. However, false negatives do not lead to losses. Besides, missing chances can be compensated by generating more than one decision trees, with each of them picking up some opportunities. Therefore, a smaller weight is given to RMC.

Secondly, EDDIE allows the user to specify a preferred boundary for the percentage of positive predictions. The lower the percentage of recommendations, the lower the rate of failure (i.e. the higher the *precision*) tend to be. In the tests on the 1641 data set, we asked EDDIE to generate predictions with 5-10%, 10-15%, 15-20% and 20-25% of positive recommendations. This reflects the percentage of real opportunities in the data.

Table 4 shows the result of applying EDDIE-ARB with different recommendation percentages. By adjusting the constraint-based function, EDDIE-ARB could discover rules with 0% to 10% failure (i.e. 80% to 100% precision). The low rate of failure is achieved at the cost of missing opportunities. This result is consistent with what we found in the past. [43]

Table 4. Application of EDDIE-ARB with different windows of recommendation percentages

	RF		**RMC**		**RC**		Number of Recommend-ations	AVERAGE PROFIT LOSS
EDDIE-ARB	TRAIN	TEST	TRAIN	TEST	TRAIN	TEST	TEST	TEST
5-10%	0.00%	0.00%	59.80%	58.00%	85.60%	85.50%	67	957
10-15%	2.10%	1.00%	48.80%	47.40%	88.00%	88.00%	84	787
15-20%	22.90%	23.70%	38.90%	38.20%	86.30%	85.60%	129	491
20-25%	39.70%	37.50%	32.60%	32.60%	81.30%	81.50%	173	465

7. Comparing EDDIE-ARB with a naive arbitrage approach

Being able to achieve high RC and low RF do not guarantee high profit. To assess the commercial potential of EDDIE-ARB, it was tested against a naive arbitrage approach. The standard ex ante analysis of arbitrage [6] is based on the following scenario.

Scenario of a naive arbitrage approach

The naive premise is that the arbitrageur waits for a contemporaneous profit signal in the category of either short or long arbitrage (given by equations 2.1 to 2.4) above and then continues with arbitrage trades in the same direction in a given time interval. It is assumed that after a time execution delay of one minute, the arbitrageur proceeds to execute his arbitrage position within a time interval of nine minutes. He, thus, faces execution price risk, should the trio of the P-C-F prices diverge.

As explained in Section 3, not every contemporaneous P-C-F profit signal will lead to a profit, due to execution delay. The question is whether EDDIE-ARB can out-perform this naive arbitrage approach. The answer to this question heavily hinges on whether EDDIE-ARB can reliably anticipate arbitrage opportunities in advance.

The objective we set EDDIE-ARB is as follows: at every point in time with the occurrence of a matched P-C-F price triplet, predict whether profitable arbitrage is possible in the next 10 minute interval in a given direction (long or short) after a one minute of execution delay. Profit and loss of the recommended arbitrage positions are calculated by the criteria given in Section 2. It is important to point out that the naive rule only attempts to exploit profitable arbitrage in the next ten minutes. EDDIE is trained to anticipate arbitrage opportunities from any point in time. With electronic trading, any observed P-C-F arbitrage opportunity can be executed immediately. As quotes do not 'live' longer than 1-3 minutes even with electronic trading, the methodology given here for generating decision trees that anticipate profitable arbitrage with a ten minute lead is advantageous.

Performance of the best rule generated by EDDIE-ARB, which we call "GDT(4)" was compared with the naive rule above. Results are summarized in Table 5. The two methods were compared on training, test and out-of-sample periods. It is worth noticing that there are significantly more opportunities in the out-of-sample period, after June 1998.

GDT(4) compares on par with the naive rule on total profits. However, it has a higher average profit than the naive rule. (In fact, the average profits of all the rules generated by EDDIE-ARB were above the naive rule (£338), see Ref. 30 for details.) The reason that this is not translated to the total profit is because GDT(4) did not make as many positive recommendations as the naive rule. This is because we instructed EDDIE-ARB to generate rules with low rate of failure. With the constrained fitness function in EDDIE-ARB, we can generate many rules with precision as high as GDT(4). If each rule picks out arbitrage opportunities reliably and these opportunities do not form proper subsets of one another, then the set of rules together will have a good chance of out-performing the naive rule.

Table 5. Performance of EDDIE-ARB vs Naive Rule

Periods		Total Trade Signals	Total Profit / Loss	Average Profit / Loss
Training	GDT(4)	246	151,984	617.82
	Naive Rule	290	79,682	274.76
Test	GDT(4)	154	112,684	731.71
	Naive Rule	196	66,268	338.1
Out-of-sample	GDT(4)	249	259,795	1,043.35
	Naive Rule	273	260,135	952.88

8. Concluding Summary

Overall, this is a valuable experience in chance discovery. From millions of records, valuable information was retrieved for training EDDIE-ARB. Through close collaboration between economists and computer scientists, variables were refined and data were better pre-processed to help EDDIE-ARB. The latest version of EDDIE-ARB managed to generate predictions reliably. Such predictions helped us to beat a naive arbitrager on out-of-sample data. The pick up rate of EDDIE-ARB could potentially be improved by generating more rules (as long as the rules do not pick up the same patterns). Therefore, EDDIE-ARB is a promising tool for arbitrage chances discovery.

One may question who contributes to the positive results, EDDIE-ARB or the human researchers. The answer is both. Without human researchers, EDDIE-ARB will not be able to find useful patterns. That is why over a dozen iterations were needed to perfect EDDIE-ARB. On the other hand, without EDDIE-ARB, it is impossible for human users to find rules of such high precision (rules with 90% to 100% were found). Rules with high precision may form building blocks for practical, reliable arbitrage monitoring systems.

Acknowledgements

This project is partly funded by the University of Essex's Research Promotion Fund. This paper extends the paper published in Congress of Evolutionary Computation 2001. [29]

References

1. A. Abe and Y. Ohsawa, (eds.) (2003). *Special Issue on Chance Discovery. New Generation Computing*, **21(1)**, Berlin: Springer and Tokyo: Ohmsha, 2003.
2. Abe, A. and Ohsawa, Y. (eds.) (2003). Special Issue on Chance Discovery. New Generation Computing, 21(2), Berlin: Springer and Tokyo: Ohmsha.
3. F. Allen and R. Karjalainen (1995). Using genetic algorithms to find technical trading rules. Working paper at Rodney L. White Center for Financial Research.
4. S.S. Alexander, (1964). Price movement in speculative markets: trend or random walks, No. 2, in Cootner, P. (ed.), The random character of stock market prices, MIT Press, Cambridge, MA, 338-372.
5. Backus, J.W., (1959). The syntax and semantics of the proposed international algebraic language of Zurich, ACM-GAMM conference, ICIP, Paris, June.
6. Bae, K.H., Chan, K., Cheung, Y.L. (1998). The profitability of index futures arbitrage: Evidence from bid-ask quotes, Journal of Futures Markets, 18, 743-763.
7. Bauer, R. J. Jr., (1994). Genetic Algorithms and Investment Strategies. New York, John Wiley & Sons, Inc.
8. W. Brock, J. Lakonishok and B. LeBaron, Simple technical trading rules and the stochastic properties of stock returns, *Journal of Finance* **47**(1992) 1731-1764.
9. S-H Chen and C-H Yeh, Toward a computable approach to the efficient market hypothesis: An application of genetic programming, it Journal of Economic Dynamics and Control **21** (1997) 1043-1063.
10. B. Cornell and K. French,(1988) Taxes and the Pricing of Stock Index Futures *Journal of Finance* **38** 675-694.
11. J. Evnine and A. Rudd, Index Options - Early Evidence, *Journal of Finance* 1985.
12. E.F. Fama and M.E. Blume, Filter rules and stock-market trading, *Journal of Business* **39(1)**(1966) 226-241.
13. J. Fung, K.W. Chan and C. Kam, On the Arbitrage Free Relationship Between Index Futures and Index Options: A Note. *Journal of Futures Markets*, 14(1994) 957-962.
14. J. Fung, K.W. Chan, L. Cheng, T.W. and C. Kam, The Intraday pricing Efficiency of Hong Kong Hang Seng Index Option and Futures Markets, *Journal of Futures Markets* **17**(1997) 797-815.
15. Gemmill, G. (1993). Options Pricing, Maidenhead, UK, Mc Graw-Hill.
16. Goonatilake, S. & Treleaven, P. (ed.), (1995). Intelligent systems for finance and business, Wiley, New York.

17. O.P.Gwilm and M. Buckle, Volatility Forecasting in the Framework of the Option Expiry Cycle, *The European Journal of Finance* **5**(1999) 73-94.
18. Goldberg, D.E., (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley.
19. Holland, J.H., (1975), Adaptation in natural and artificial system, University of Michigan Press.
20. Kasabov, N. K., Foundations of neural networks, fuzzy systems and knowledge engineering, MIT Press, 1996.
21. J.R. Koza, *Genetic Programming: on the Programming of Computers by Means of Natural Selection* (MIT Press, 1992).
22. J.R. Koza, *Genetic Programming II: Automatic Discovery of Reusable Programs* (MIT Press, 1994).
23. Koza, J., Goldberg, D., Fogel, D. & Riolo, R. (ed.), (1996). Proceedings of the First Annual Conference on Genetic programming, MIT Press.
24. J.H. Lee and N. Nayar(1993), A transactions data analysis of arbitrage between index options and index futures *The Journal of Futures Markets* **13**(1993), 889-902.
25. J. Li and E.P.K. Tsang, Improving technical analysis predictions: an application of genetic programming, in *Proc. of The 12th International Florida AI Research Society Conference*, (Orlando, Florida, 1999) pp.108-112.
26. Li, J. & Tsang, E.P.K, (1999b). Investment decision making using FGP: a case study, Proceedings of Congress on Evolutionary Computation (CEC'99), Washington DC, USA, July 6-9 1999.
27. S. Mahfoud and G. Mani, Financial forecasting using genetic algorithms, *Applied Artificial Intelligence*, **10**(1996) 543-565.
28. S. Mahfoud and G. Mani, Financial Forecasting Using Genetic Algorithms, *Journal of Applied Artificial Intelligence* **10,6**(1997) 543-565.
29. S. Markose, E.P.K. Tsang, H. Er and A. Salhi,(2001). Evolutionary arbitrage for FTSE index options and futures, in *Proceedings, Congress on Evolutionary Computation, CEC* 2001, 275-282
30. Markose, S., Tsang, E. & Er, H., (2002). EDDIE for stock index options and futures arbitrage, in Chen, S-H. (ed), Genetic algorithms and genetic programming in computational finance, Kluwer Academic Press, Chapter 14, 281-308
31. Mitchell, M., (1996). An Introduction to Genetic Algorithms. MIT Press.
32. Modest, D. & Sunderesan, M., (1983). The relationship between spot and futures prices in stock index futures markets: some preliminary evidence, Journal of Futures Markets, 3, 15-41.
33. Neely, C., Weller, P. & Ditmar, R., (1997). Is technical analysis in the foreign exchange market profitable? A genetic programming approach, Journal of Financial and Quantitative Analysis, 32, 405-26.
34. Ohsawa, Y. & McBurney, P., (2003). Chance Discovery. Berlin: Springer Publishers.
35. Ohsawa, Y., McBurney, P. & Parson, S., (2002). Chance Discovery: The Discovery and Management of Chance Events. Papers form the AAAI Fall Symposium. Technical Report FS-02-01, ISBN 1-57735-172-x, Cambridge, MA: The AAAI Press
36. Oussaidene, M., Chopard, B., Pictet, O. & Tomassini, M., (1997). Practical aspects and experiences - Parallel genetic programming and its application to trading model induction, Journal of Parallel Computing Vol. 23, No. 8, 1183-1198.
37. Saad, E., Prokhorov, D. & Wunsch, D., (1998). Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks, IEEE Transactions on Neural Networks, vol. 9. 1456-1470.
38. Sornette, D. & Zhou, W.-X., (2002). The US 2000-2002 Market Descent: How Much

Longer and Deeper? Quantitative Finance 2 (6), 468-481
39. Stoll, H.R., (1969). The relationship between put and call option prices. Journal of Finance. 25: 801-824
40. Sweeney, R.J., (1988), Some new filter rule tests: Methods and results, Journal of Financial and Quantitative Analysis, 23, 285-300.
41. Tsang, E.P.K , Li, J., Markose, S., Hakan, E., Salhi, A. & Iori,G., (2000). EDDIE in Financial Decision Making, Journal of Management Economics, http://www.econ.uba.ar/www/servicos/publicaciones/journal3/index.htm.
42. Tsang, E.P.K., Li, J. & Butler, J.M., (1998). EDDIE beats the bookies, International Journal of Software, Practice & Experience, Wiley, Vol.28 (10), 1033-1043.
43. Tsang, E.P.K. & Li, J., (2002). EDDIE for financial forecasting, in Chen, S-H. (ed.), Genetic Algorithms and Programming in Computational Finance, Kluwer Academic Publishers, 161-174.
44. Tsang, E.P.K., Yung, P. & Li, J., (2004). EDDIE-Automation, a decision support tool for financial forecasting, Journal of Decision Support Systems, Special Issue on Data Mining for Financial Decision Making, Vol.37, No.4, 559-565.
45. Tucker, A.L., (1991). Financial Futures, Options and Swaps, West Publishing Company, St. Paul, MN.
46. Yadav, P.K and Pope, P., (1990). Stock Index Futures arbitrage: International Evidence, Journal of Futures Markets, 10, 573-603.
47. Zellner, A., (2004). Statistics, Econometrics and Forecasting, Cambridge University Press.