In Y. Abu-Mostafa, J. Moody, P. Refenes, and A. Weigend, editors, *Neural Networks in Financial Engineering*, pages 303–316. World Scientific, London, 1996.

STOCK SELECTION USING RECON$^{\mathrm{TM/SM}}$

George H. John
Computer Science Department, Stanford University,
Stanford, CA 94305, USA
gjohn@cs.stanford.edu
http://robotics.stanford.edu/~gjohn/

Peter Miller
Lockheed AI Center, 3251 Hanover Street,
Palo Alto, CA 94304, USA
pmiller@aic.lockheed.com
http://hitchhiker.space.lockheed.com/~recon/

Randy Kerber
1422 Flora Avenue, San Jose, CA 95130, USA
kerber@best.com

Abstract

We approach the problem of stock selection from the perspective of *knowledge discovery in databases*: given a database of several years of quarterly information on over a thousand companies, discover patterns in the data that will allow one to predict which stocks are likely to have exceptional returns in the future. The database includes measures of trends in the stocks' prices as well as fundamental data on the companies. For this task we employed the Recon system, which is able to induce a set of classification rules or a neural network to model the data it is given. To evaluate Recon's performance in the stock selection task, we paper-traded a portfolio of the fifty stocks ranked highest by Recon. When trading costs were taken into account, Recon's portfolio had a total return of 238% over a four-year period, significantly outperforming the benchmark, which returned 93.5% over the same period. The performance is not attributable to growth/value or size effects alone. We conclude that Recon is a valuable tool for stock selection.

1 Introduction

In the stock selection problem, one is given a large database of historical information on many stocks. The problem is to select from this universe of stocks some portfolio which contains stocks likely to exhibit exceptional return

Recon$^{\mathrm{TM/SM}}$ is a trademark and service mark of Lockheed Missiles & Space Company, Inc., subsidiary of Lockheed Corporation, a Lockheed Martin Company.

over a future period of time. For believers in the efficient market hypothesis, any pattern in historical data which is useful in the above prediction task is an *anomaly.* Jacobs & Levy (1987) suggest that "An optimal investment approach aims to capture directly the excess returns produced by anomalies in a disciplined, coordinated manner." Such a coordinated search for patterns (anomalies) is precisely what many algorithms developed in the statistics, machine learning, and knowledge discovery in databases communities do. In this paper we discuss our use of the Recon rule induction system (Simoudis, Livezy & Kerber 1994, Kerber, Livezey & Simoudis 1995) to find patterns expressed as understandable if-then rules, and the use of these rules to form portfolios of high-return stocks.

The general class of problems that Recon addresses, of which stock selection is a member, is the supervised learning problem: given a database of some arbitrary number of records (stocks, in this case), and some distinguished fields which we would like to be able to predict given the remaining fields, discover some useful patterns in the database and express these patterns in some language. The language must be sufficiently expressive to allow a computer to fill in missing values in the distinguished field in future databases of the same type. The language should also be understandable so that an expert or analyst can verify that the discovered patterns make sense, and so that important new knowledge discovered by the system can be put to use by the analyst. A wide range of problems fit into this framework; stock selection, bond rating, credit scoring, and targeted marketing are just a few examples.

We obtained a database of quarterly information on 1,160 to 1,480 companies during the period 31 Dec 1987 to 30 Jun 1993. For each company and quarter, we have nearly one hundred different fields of information, such as market capitalization, price-earnings ratio and trend information. The data also includes each stock's return on investment over the following three month time period. Since we are interested in predicting whether or not a stock will exhibit exceptional return, we defined a target concept for Recon to learn, the *exceptional* concept: those stocks with returns in the top 20% in a given quarter were marked as exceptional and the rest were marked as unexceptional. Recon's job was to analyze a historical database and produce rules which would classify present stocks as exceptional or unexceptional future performers.

Since we have defined our problem as stock selection instead of portfolio management, this leaves us with the problem of how to evaluate the system. We chose to use a simple method to form a portfolio from the stocks selected by Recon, and then evaluate the portfolio using standard techniques for portfolio evaluation.

We present an overview of the Recon system in Section 2, showing an example of the program's output and describing how it captures salient statistics from a database. In Section 3, we describe our experiments in stock selection using Recon's predictions of future exceptional performance. We perform traditional analyses of the Recon portfolio and also show how the style of the portfolio varied over time in Section 4. Section 5 discusses related work, and Section 6 presents our conclusions drawn from this experience with Recon and stock selection.

2 The Recon System

Recon (Simoudis et al. 1994, Kerber et al. 1995) is an integrated system for the exploration and analysis of large databases. The exploratory tools allow a user to graphically define concepts and new features, and to visualize data. The analytical tools include AutoClass (Cheeseman, Kelly, Self, Stutz, Taylor & Freeman 1988), rule induction, and neural networks. In this paper we discuss only the rule induction component of Recon, but in future work we will apply neural networks as well.

The rule induction algorithm in Recon is a descendant of the systems described in Kerber (1991) and Kerber (1992). The algorithm searches through the space of rules to find those that are likely to be useful in classifying items in a future database. Some examples of rules that were found by Recon are shown in Figure 1. For example, the first rule reads "If `current assets` is less than 3.75 and `normalized cash` is between 0.0088 and 0.0272 then predict `unexceptional` (low) return, with strength 1.526." (These rules were learned from ValueLine data. We are not permitted to show rules learned from the database used in this paper, which was provided to us by an equity management company.)

Each rule has an associated *strength*. Since rules may overlap, the class (`exceptional` or `unexceptional`) that Recon finally predicts for a stock is a combination of the predictions made by each rule. A rule's prediction is weighted by the amount of evidence supporting the rule, allowing Recon to not only make a prediction but also to assign some measure of certainty or strength to the prediction.

When used to learn a set of rules from a database, Recon's rule induction algorithm begins by first discretizing all numeric features. For example, the price-to-book ratio might be segmented into three intervals which could be interpreted as an indicator that a stock is an aggressive-growth, growth, or value stock. All features are discretized using the ChiMerge algorithm described in Kerber (1992). ChiMerge uses the training data, including the class label, to build intervals such that the distribution of class values in neighboring intervals

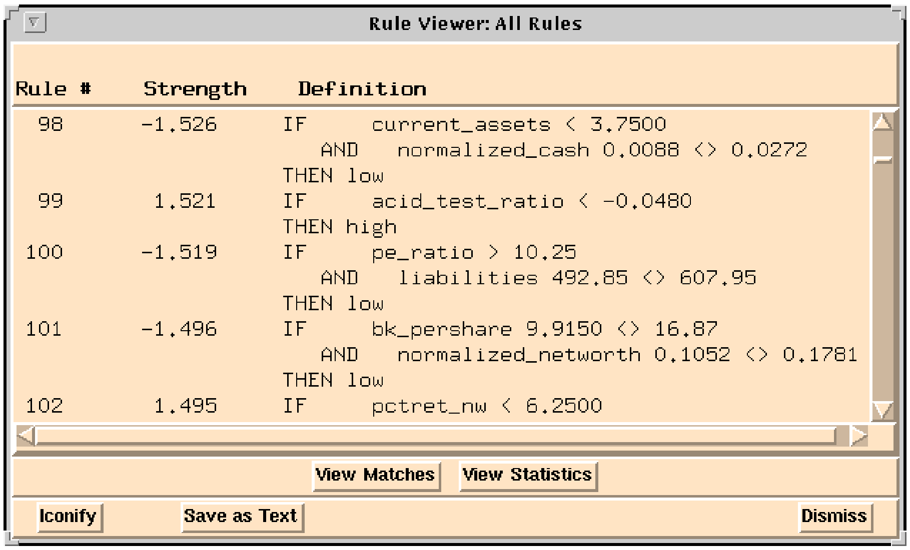

Figure 1: Example of rules learned by the Recon rule induction system.

is statistically dissimilar. The rule induction algorithm then begins exploring the space of all possible rules. It begins with all single-antecedent rules and then adds conditions using a combination of several heuristics to concentrate on interesting rules, ultimately yielding a set of rules with high predictive accuracy. Each rule is characterized by a set of statistics that describe the (training-sample estimates of the) conditional distribution of the class given whether or not the antecedent of the rule is true; thus, the search through the space of rules is essentially just a search through the space of the multivariate statistics of the data that are pertinent to the classification problem.

The Recon system allows an analyst to explore the discovered rules by viewing the training cases that led to a particular rule's creation, or viewing statistics about a rule. One can also determine which features are important by seeing which ones tend to participate in most of the rules. Similarly, features which are never tested did not provide information useful to the classification problem. Unrepresentative training data can yield rules that disagree with an analyst's knowledge of a problem; in such cases the analyst can edit the set of rules, perhaps removing the offending rules. (For example, a court ruling might cause pharmaceutical companies to perform well during one quarter, causing Recon to include rules favoring these companies. The analyst may

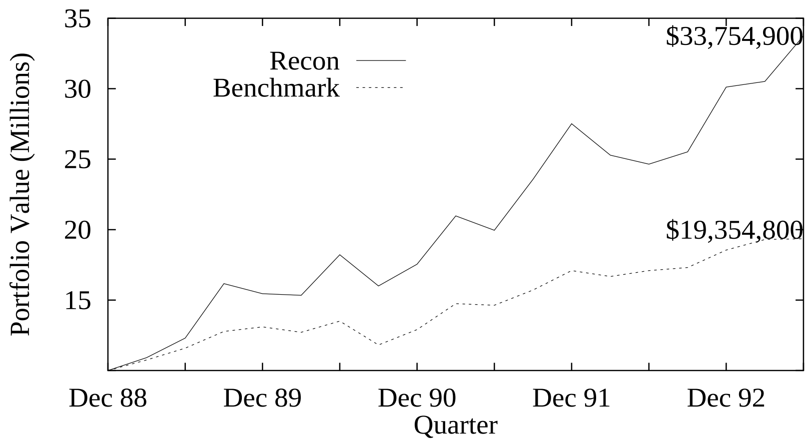

Figure 2: Out-of-sample returns (net of costs) for the period 31 Dec 1988 to 30 Jun 1993.

delete these rules if she believes the exceptionally favorable climate will not continue in the future.) Once the analyst is comfortable with the set of rules, they may be applied to an out-of-sample testing dataset (either future data, or currently-available data that was held out of the training sample). For each test case, Recon shows the relative strength of its membership in each class, and can show the set of rules participating in the prediction.

Recon performs a search through a database to find statistically salient rules, and presents them to an analyst in a comprehensible format. If the resulting rules can result in superior performance in stock selection (or credit scoring or other problems), a strong case exists for the use of a tool like Recon because of the understandability of the patterns it discovers.

3 Stock Selection Experiments

To evaluate the stock selection models built by Recon, we formed portfolios using its predictions and analyzed the resulting performance. Below we describe how Recon was used to determine a desired asset mix for a portfolio. To evaluate Recon's portfolio we use standard methods for portfolio evaluation, comparing its performance to the appropriate benchmark (Figure 2).

Since we have only quarterly data, all trading occurs at the beginning of each quarter and the resulting asset mix is held for the duration of the quarter. At the beginning of each quarter, we use Recon to build a set of rules that accurately predict exceptional return for stocks during the preceding four quarters. We then use the learned model to rank all present stocks.

For example, the portfolio held from 31 Mar 1993 through 30 Jun 1993 was selected in the following way: Recon was trained on the four preceding quarters beginning 31 Mar 1992, 30 Jun 1992, 30 Sep 1992, and 31 Dec 1992. During training, Recon found many rules that were good predictors of exceptional future return within the training sample. These rules were then applied to the stock data available on 31 Mar 1993 (which did not include the actual future returns, of course!), and the stocks were ranked according to Recon's certainty that they would exhibit exceptional returns during the coming quarter. The top fifty such stocks were selected and a market-capitalization-weighted portfolio of the stocks was purchased (because we did not have access to a portfolio optimizer). Table 1 shows the portfolio selected by Recon on 31 March 1993.

At the end of each quarter, a new desired portfolio is computed, and trades are executed to achieve the new desired asset mix. We assumed 1% transaction costs (round-trip). On average, 80% of the portfolio was traded each quarter. Compared with the appropriate benchmark, Recon's portfolio generates a cumulative return of 238%, while the benchmark returns 93.5%.

The key question in determining a benchmark is "If I didn't use this manager [Recon, in this case], what would I do?" (Sharpe 1995). We obtained the data through an analysis contract with an equity management company, and were not given a description of the screening process used in creating the database. Thus, our best answer to this question is that we would have bought a market-cap-weighted portfolio of all the stocks in the universe from which Recon was selecting. This is the "Benchmark" portfolio in Figure 2. We assumed 1% round-trip trading costs for the benchmark portfolio re-weighting.

Figure 3 shows "strictly out-of-sample" returns. We actually ran twelve versions of Recon with varying definitions for "exceptional return" and varying parameters to the learning algorithm, and produced out of sample returns on the fourteen quarters from Dec 1988 through Mar 1992 for each version. There were not significant differences among the returns of the different versions. When discussing empirical performance tests using historical data, one has to be very careful and forthcoming about the process that led to the results. Black (1995) gives a concise description of the sin of *data mining*, mentioning many of its incarnations. In statistics and finance the term *data mining* is meant to connote poorly done experiments, leading to an error often called "training on the test data" in neural networks.[a] Thus we used a *strict* out-of-sample set, the

[a] Unfortunately, the term *data mining* is overloaded: in the knowledge discovery in databases community, it refers to the virtue of processing large databases and discovering useful patterns and knowledge. To avoid confusion, we only use the term in its pejorative sense in this paper, but in our other papers we use the term only in its virtuous sense.

Table 1: Portfolio bought by Recon on 31 March 1993, showing name of security, return over the following 3 months, and percent of portfolio. Return of portfolio was 11.58%.

Name	Return	% of Portfolio
Federated Department Stores	11.05	11.41
Synoptics Communications	19.17	7.45
American Power Conversion	17.83	5.92
Safeway	7.96	5.87
Marvel Entertainment Group	51.00	5.08
Storage Technology Corp.	57.29	4.33
Midlantic Corp.	-3.43	4.24
Stone Container Corp.	-33.94	4.08
BioGen Inc.	9.75	3.86
Arrow Electronics	11.81	3.86
ARMCO Inc.	0.00	3.16
Borland International	4.02	2.41
Public Service Co. of New Mexico	13.54	2.11
Gensia Pharmaceuticals	45.31	1.89
Pacificare Health Systems	12.50	1.81
Advanta Corp.	24.66	1.76
Tseng Labs	-28.32	1.75
Data General Corp.	-18.09	1.68
Cragin Financial	-10.14	1.61
Showboat Inc.	-12.39	1.39
Bally Manufacturing Corp.	37.50	1.34
Symantec	3.88	1.29
Centocor Inc.	22.22	1.17
Air & Water Technologies	27.27	1.15
Ladd Furniture Inc.	-23.15	1.14
Synergen	9.41	1.11
Sizzler International	-12.40	1.08
Maxtor Corp.	-19.35	0.94
Novellus Systems	58.46	0.93
Western Co. of North America	27.00	0.93
Integrated Device Technologies	40.32	0.92
Dime Savings Bank of NY	-24.00	0.91
Smithfield Foods Inc.	29.81	0.89
Lifetime Corp.	46.24	0.86
Western Digital Corp.	-26.67	0.83
VLSI Technology Inc.	53.57	0.81
All 14 Remaining Stocks	-2.99	8.02

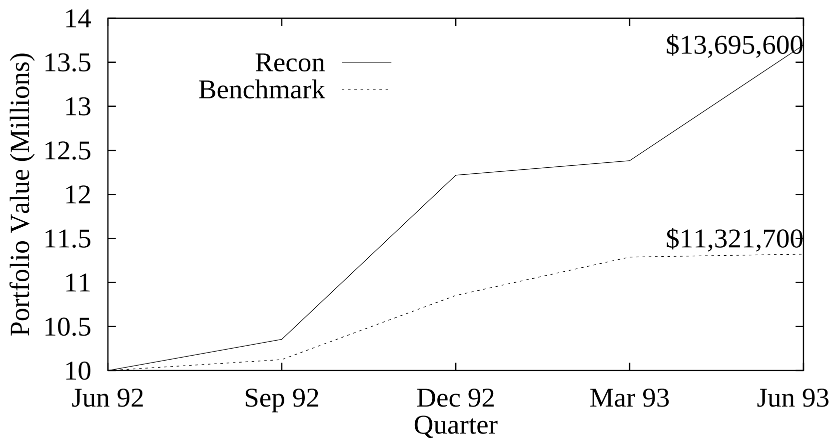

Figure 3: Strictly out-of-sample returns (net of costs) for the period Jun 1992 to Jun 1993.

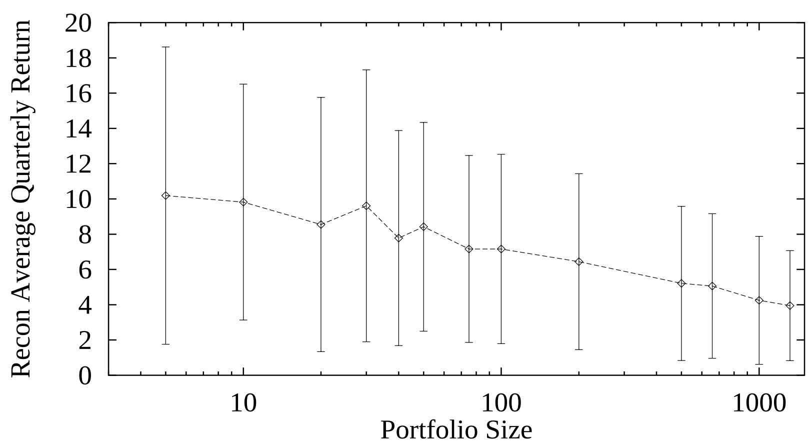

Figure 4: Mean quarterly returns for the varying-sized portfolios of stocks ranked by Recon over the period Dec 1988 to Mar 1992, showing a 68% (one standard deviation) confidence interval about each mean.

four quarters beginning Jun 1992 through Mar 1993. This data was only used *once*, after we had picked one of the twelve versions. In Figure 3 we begin with the same amount invested in Recon and the benchmark. Encouragingly, Recon outperformed the benchmark significantly during the strict out-of-sample set, returning over 32.2% on invested capital, versus 11.8% for the benchmark.

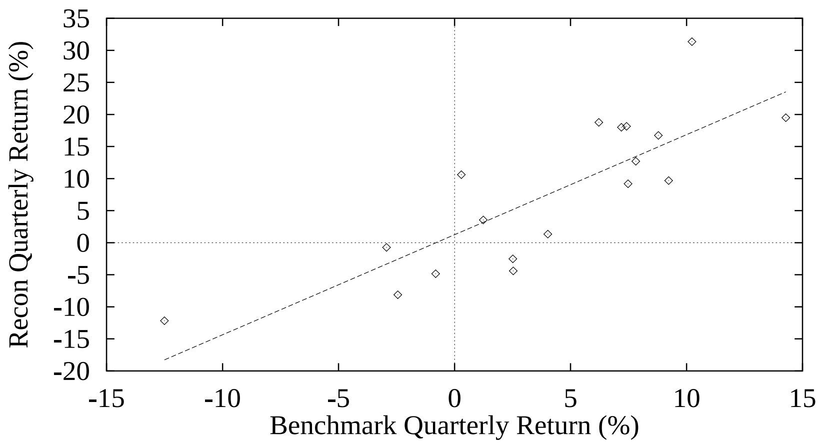

Figure 5: A linear regression of Recon's quarterly return versus the Benchmark.

Recall from above that Recon only produces a ranking on stocks, and it is up to a portfolio optimizer to do the actual selection and weighting. We decided *a priori* to use a simple portfolio that bought the top n stocks, weighted by market cap. We used Figure 4 to choose $n = 50$. The figure shows clearly that the portfolio of 50 stocks represented a good tradeoff of expected return and risk. (Of course, risk tolerance is ultimately a personal choice.)

Over the out-of-sample period, the Recon portfolio had a Beta of 1.56 and Alpha of 1.24 (Sharpe, Alexander & Bailey 1995). This means that the portfolio varied more widely than the benchmark, but also generated higher return. Figure 5 shows the scatter plot of Recon returns versus the benchmark and shows the fitted line with slope Beta and intercept Alpha. Looking only at those points where the return on the benchmark was negative, we can see that a fitted line would have slope even less than one, which means that in a down market, Recon had a small Beta, a highly desirable quality. The mean quarterly excess return over the benchmark for Recon was 3.47%. Recon's standard deviation was 11.8 versus 6.2 for the benchmark, so its portfolio was more volatile—were the excess returns worth the increased volatility? The Sharpe ratio (Sharpe 1994) for the Recon portfolio was 5.48, versus 1.85 for the benchmark (assuming a risk-free rate of 5.89% per annum, the average Treasury bill rate during the period). The results of paired t-tests show that Recon's quarterly return was higher than the benchmark at a 96% confidence level, and Recon's quarterly return was more than 1% higher than the benchmark at a 90% confidence level.

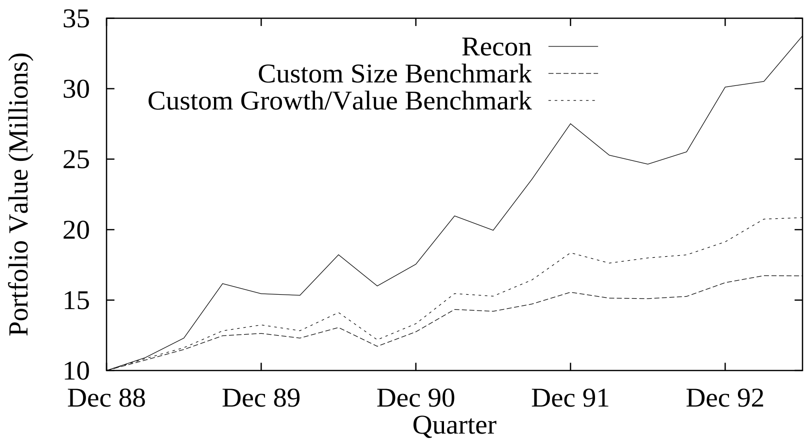

Figure 6: Return of the Recon portfolio versus a custom growth/value benchmark and market-cap benchmark. (No transaction costs were assessed on the benchmarks.)

4 Performance Attribution

There are several known effects which might result in superior performance of a portfolio, such as the size effect and growth/value effect (Sharpe et al. 1995, Elton & Gruber 1991). If Recon's performance is attributable solely to one of these factors, then it has only discovered previously-known patterns. We will show that this is not the case. Figure 6 shows the Recon portfolio's return and the returns of two custom benchmarks constructed to have the same growth/value and size distributions as Recon's portfolio. Recon still outperforms both benchmarks.

Fama & French (1992) and Capual, Rowley & Sharpe (1993) find that over various periods in history, value stocks, or stocks with low price-to-book ratios, tend to exhibit higher returns than growth stocks. We analyzed Recon's portfolio to determine to what extent its superior performance is due to its growth/value weighting. The stocks in Recon's portfolio were classified as growth or value stocks using the procedure described by Capual et al. (1993). At the beginning of each quarter, we sorted the universe of stocks by price-book ratio, and picked a threshold such that the total market value of all securities above and below the threshold was approximately the same. Stocks with a low price-book ratio were classified as value and the rest were classified as growth stocks. We then determined the percentage of Recon's portfolio invested in growth and value stocks. Figure 7 is an area chart showing the percentage of

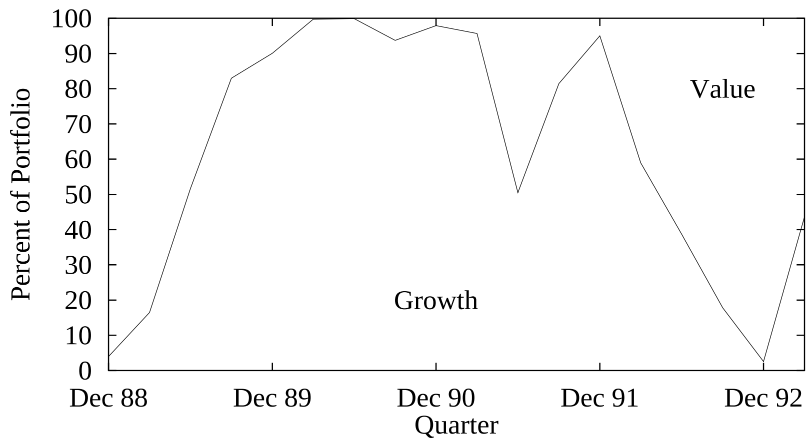

Figure 7: Area chart showing the percentage of the Recon portfolio invested in growth and value stocks.

Recon's portfolio invested in growth and value. Figure 6 shows the return from a benchmark with the same weight as Recon's portfolio in growth and value stocks each quarter. Notably, the growth/value custom benchmark still generates roughly the same return as the original benchmark, so Recon outperforms this customized benchmark by a wide margin. Thus, Recon's performance cannot be solely attributed to a value effect.

Banz (1981) shows that small-cap stocks (those stocks whose market capitalization is in the bottom quintile of all stocks) exhibit statistically significantly higher return than large stocks. One might hypothesize that Recon's performance is attributable to its selection of small-cap stocks. Figure 8 is an area chart showing the capitalization-weighting of Recon's portfolio changing over time. The results for each quarter were produced by sorting the universe of stocks by market capitalization and creating four bins such that the same number of stocks fell into each bin. For example, in 31 Dec 1988, stocks with market-cap < \$201M were classified as *micro*, those with < \$561M as *small*, those with < \$1744M as *medium* and the rest as *large*. We formed a customized benchmark by weighting each of the market-cap quartiles in the proportions as they appeared in the Recon portfolio. As Figure 6 shows, the market-cap custom portfolio gives even lower return than the original benchmark. Thus, Recon's performance cannot be solely attributed to size effects.

We have shown that Recon's performance is not solely attributable to growth/value effects or size effects, which supports the hypothesis that its

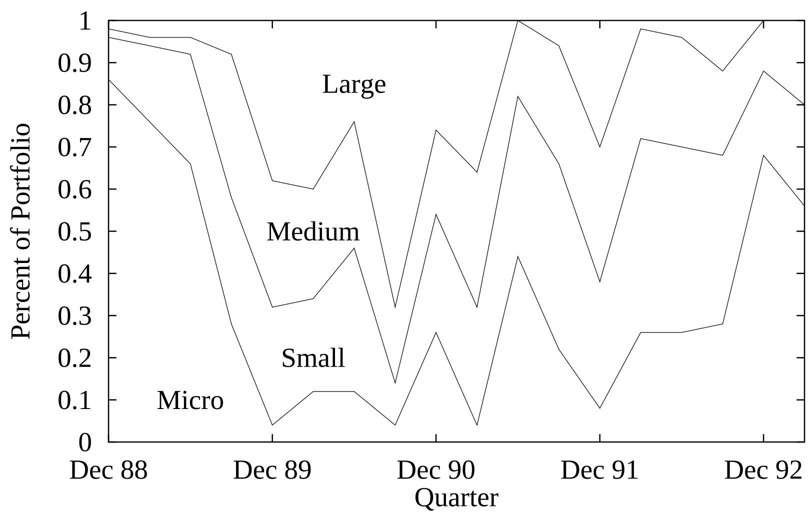

Figure 8: Area chart showing the percentage of the Recon portfolio invested in micro, small, medium, and large-capitalization stocks.

performance is due to stock selection within the styles that it chooses.

5 Related and Future Work

Much of the related work applying neural nets and other nonlinear modeling techniques to the capital markets has focused on a single asset (usually an index or derivative thereof), studying methods for using relatively high frequency timeseries data to predict future prices and generate trading signals (Trippi & DeSieno 1992, Hutchinson, Lo & Poggio 1994). Levin (1995) presents a study very similar to ours, using neural networks instead of rule induction to form a portfolio. In Levin's work, a portfolio optimizer was used to limit exposure to undesired factors and to contain turnover. Levin also used zero-investment strategy portfolios (Jacobs & Levy 1993), which can be designed to have effectively zero exposure to certain kinds of risk by holding long and short positions. This would be an intriguing addition to our method.

6 Conclusion

Any modeling tool used for stock selection should produce a comprehensible model. With millions of real dollars at stake, the ability to understand and interact with the model is a necessity. We believe that Recon's ability to discover patterns and present them as easily understood rules, explain why it included certain rules, interact with an analyst to modify the rules, and explain its predictions, makes it an excellent tool for financial modeling and stock selection.

Recon's rule-induction tool can be used to create portfolios of stocks with exceptional return. Its method of ranking stocks can be used to create an array of varying-sized portfolios, allowing investors to choose among them based on their risk tolerance. The portfolio we selected significantly outperformed the benchmark over four and a half years of trading.

Acknowledgments

The authors would like to thank Ramesh Subramonian and Ken Jung for their work in implementing the rule induction system, and Evangelos Simoudis for his role in obtaining the data and suggesting ideas on the application of Recon to the problem. Many of the portfolio analysis ideas came about while the first author was enrolled in a Portfolio Management course taught by William F. Sharpe. Thanks to Sherri Fisher and an anonymous referee for healthy skepticism that led to the growth/value and size-effect analysis. The paper also benefited from discussions with Archan Basu, Vida Tigrani and Pat Walker.

References

Banz, R. W. (1981), "The relationship between return and market value of common stock", *Journal of Financial Economics* **9**, 3–18.

Black, F. (1995), "Estimating expected return", *Financial Analysts Journal* **51**(1), 168–171.

Capual, C., Rowley, I. & Sharpe, W. F. (1993), "International value and growth stock returns", *Financial Analysts Journal* **49**(1), 27–36.

Cheeseman, P., Kelly, J., Self, M., Stutz, J., Taylor, W. & Freeman, D. (1988), Autoclass: A Bayesian classification system, *in* "Machine Learning: Proceedings of the Fifth International Workshop", Morgan Kaufmann, pp. 54–64.

Elton, E. J. & Gruber, M. J. (1991), *Modern Portfolio Theory and Investment Analysis*, fourth edition, John Wiley & Sons, New York.

Fama, E. F. & French, K. R. (1992), "The cross-section of expected stock returns", *Journal of Finance* **47**(2), 427–465.

Hutchinson, J. M., Lo, A. W. & Poggio, T. (1994), "A nonparametric approach to pricing and hedging derivative securities via learning net-

works", *Journal of Finance* **49**(3), 851–889.
Jacobs, B. I. & Levy, K. N. (1987), "Investment management: Opportunities in anomalies?", *Pension World* **23**(2), 46–47.
Jacobs, B. I. & Levy, K. N. (1993), "Long/short equity investing", *Journal of Portfolio Management* **20**(1), 52–63.
Kerber, R. (1991), Learning classification rules from examples, *in* "Proceedings of the AAAI-91 Workshop on Knowledge Discovery in Databases", AAAI.
Kerber, R. (1992), Chi merge: Discretization of numeric attributes, *in* "AAAI-92: Proceedings of the Tenth National Conference on Artificial Intelligence", AAAI Press/MIT Press, pp. 123–128.
Kerber, R., Livezey, B. & Simoudis, E. (1995), A hybrid system for data mining, *in* S. Goonatilake & S. Khebbal, eds, "Intelligent Hybrid Systems", John Wiley and Sons, pp. 121–142.
Levin, A. U. (1995), Stock selection via nonlinear multi-factor models, *in* "Advances in Neural Information Processing Systems", Morgan Kaufmann Publishers, San Francisco.
Sharpe, W. F. (1994), "The Sharpe ratio", *Journal of Portfolio Management* **21**(1), 49–58.
Sharpe, W. F. (1995), "Course notes: Portfolio management", Course Given in the Graduate School of Business, Stanford University.
Sharpe, W. F., Alexander, G. J. & Bailey, J. V. (1995), *Investments*, fifth edition, Prentice Hall, New Jersey.
Simoudis, E., Livezey, B. & Kerber, R. (1994), Integrating inductive and deductive reasoning for database mining, *in* "Proceedings of the AAAI Knowledge Discovery in Databases Workshop", Seattle, WA.
Trippi, R. R. & DeSieno, D. (1992), "Trading equity index futures with a neural network", *Journal of Portfolio Management* **19**(1), 27–33.