Chapter 4

4.0 Conclusions, Recommendations, and Other Matters.

4.1 Conclusions. In this chapter we review the important conclusions and status of the hypotheses. Recommendations for further research are made along with suggested improvements.

a. The simugram portfolios handsomely outperform the benchmark indexes. For both of the sub-portfolios Ω^0 = SP-100 and SP-500, the simugram portfolios outperform the applicable market indexes under all reasonable constraint sets by at least a factor of about 2. This multi-year performance is measured by average annual return percent, annualized return percent, or cumulative terminal value (TV) in dollars.

Table 4.1 Market benchmark outperformance by Simugram portfolios, 33- and 26-years study periods

	33 year Study Period, 1970-2002
	Simugr

As table 4.1 shows, in the 33-year study period (exclusively the SP-100 simugram results), all return multipliers are on the order of 3, meaning the simugram performance is 3 times that of the benchmark. The TV multipliers averaged 45; i.e., for every $1 million the market earned, the simugram returned $45 million. In the 26-year study period, when both the simugram SP-100 and SP-500 target portfolios were studied, the SP-100 continued to perform similarly to the 33-year study, and the SP-500 DAC-3 Sum 1 simugram returns averaged about twice the returns as the market benchmarks, with a TV multiplier of about 6.

b. The simugram portfolios exhibit better trading statistics than the benchmark market indexes, as measured by absolute drawdown, and number of losing trades. The absolute drawdown is the sum of all the negative returns. The SP-100 was better in this statistic than the market indexes in both study periods; the SP-500 had a higher absolute drawdown, but this is most likely due to the performance problems referred to throughout, caused by the combination of optimizer stalling and the loss of portfolio integrity inherent in the splitting process. These are summarized in table 4.2.

Table 4.2 Trading statistics for the simugram portfolios

	33 year Study Period, 1970-2002			
	Sim-100	Dow 30	SP500	Wilshire
# years total	33	33	33	33
# neg. years	8	10	9	11
Sum negative returns	-1.145	-1.119	-1.230	-1.277
Avg Negative return	-0.143	-0.112	-0.137	-0.116
Sum all returns	7.976	2.839	2.789	3.198

	26 year Study Period, 1970-2002				
	Sim-100	Sim-500	Dow 30	SP500	Wilshire
# years total	26	26	26	26	26
# neg. years	5	7	8	7	9
Sum negative returns	-0.655	-0.979	-0.678	-0.759	-0.752
Avg Negative return	-0.131	-0.140	-0.085	-0.108	-0.084
Sum all returns	7.003	5.073	2.463	2.487	2.799

The traditional trading statistics for percent profitable trades, average profitable trade, etc, are not tabulated since they have been the subject of this dissertation, and are summarized elsewhere. As with any good trading system, the positive returns overcome the unprofitable ones. From either a trader's or investor's perspective, this system has the traditional characteristics of a legitimate investing or trading program for two main reasons: (i) it takes time to make money, in this case, half a lifetime; satisfying a primary maxim of successful investing. And (ii), as a trading system it takes discipline and patience to follow. As presented, it is an annual trading system, which has unique challenges and benefits. The temptation would be to trade more frequently, but this must only be attempted under a discipline suggested in section 4.2.2(m).

c. The simugram portfolios have lower correlations than the benchmark indexes. It would be preferred in devising a portfolio that purports to outperform market indexes for it to have a lower simple correlation than the comparable indexes, which the simugram indeed does.

Table 4.3 Simugram portfolio correlation with market indexes

	Geomkt	Wil5000	Dow30	SP500
Simugram 100	0.874	0.874	0.820	0.882
Wilshire 5000	0.953	1	0.881	0.960
SP500	0.994	0.960	0.951	1

d. The Nelder-Mead (N-M) optimization procedure exhibits stalling in dimension greater than about 135-175, which causes problems implementing the simugram procedure for larger portfolios. On the other hand, we have demonstrated its effectiveness in handling portfolios the size of the most widely watched market indexes such as the Dow, the NDX, FTSE-100, OEX, etc.

e. The sub-portfolio experiments showed that the underperformance of the size K=500 portfolio is most likely due to optimizer stagnation. The workaround for this was the portfolio-splitting described in chapter 3, which introduced another drag on return, that due to loss of portfolio element synergy. The simugram process seems to work on the correlation structure of the entire portfolio. When this is fragmented, suboptimal weightings result, as seen in the epilogue DAC-4 and DAC-5 results (appendix F). The N-M algorithm is indeed having trouble in higher dimensions, resulting in not attaining more global maxima. This is evidenced by the results obtained with DAC-1 and DAC-2, and to a lesser extent with DAC-3. The reduction in dimension from 500 to 200-250 was not enough to improve the global maximum, as seen in tables 3.1 through 3.6. However, with K= 135-165 under DAC-3, statistically significant improvements were seen, even though some stalling appears to be happening also.

It is believed that the optimizer effect is greater than the lost synergy effect, at least to a point, as noted by the increase in returns with DAC-3. Using DAC-4 generates portfolios of size K=100-125, which we know can be solved to obtain over $400M TV in the case of the SP-100. Thus it is apparent that with portfolio-splitting into an equivalent number of stocks, the synergy effect takes over with a vengeance, so that by DAC-5 all the optimization benefits are subsumed by the loss of synergy, as seen by the lower overall returns than even the original DAC-1 stalled versions.

f. The distribution of terminal values, though left-skewed, may be considered Gaussian. In contrast, the distributions of within-year simugram returns are mostly non-normal. And, increasing the number of simulations M and hence the simugram length, changes the return distributions in a yet-unknown manner, resulting in reduced variance and also possibly the mean.

g. We cannot conclude that the 2-pass optimization approach is without merit. If used with less fragmented portfolio splitting, better results might be achieved.

We now turn to more detailed conclusions regarding the status of the hypotheses, the possible uniqueness of certain sub-portfolios, and the outlining of the recommended procedure arising in the process of this research.

4.1.1 Hypothesis Review.

To see whether we met our objectives, we must review the status of our hypotheses. We shall cover them in order. Many others could have been suggested, such as statements about the volatility of results, but these are the ones that were documented.

The first hypothesis was the major test of the efficacy of the simugram portfolio selection system. In it we had a distance function for which we use the scalar $r(\cdot) = TV(R)$.

Hypothesis 1.

$$H_0: \quad \left\| r(P^A) - r(P^0) \right\|_T = 0$$
$$H_A: \quad \left\| r(P^A) - r(P^0) \right\|_T > 0$$

In the nomenclature of chapter 1, Ω^0 =SP-100 or the SP-500. Using the result of our baseline SP-100 simugram in CS0, our test statistic is simply $\Delta TV = 476.8$ or 474.20, depending on which market baseline we use, either the Dow or Wilshire 5000. (The Dow is chosen for this example because it has the biggest tails.) Under the null hypothesis, $F_X^A = F_X^0 \doteq F_X^M$, and $F_r^A = F_r^0 \doteq F_r^M$. It is true that the distribution of our ΔTV statistic would be difficult to establish parametrically, since TV is that of the product of a number of translated, dependent, and not necessarily identically distributed, non-normal random variables. However, it is straightforward to simulate its distribution, either for the Dow or the Wilshire 5000. In the case of the SP-100 simugram P^A, it is very simple to read the test statistic's p-value directly from the simugrams in figure 4.1; it is one.

The SP-500 simugram TV test statistic in relation to the Wilshire is also easy to calculate from the graph, it is one also. However, when one uses the Dow's ΔTV distribution function, care must be used since our ΔTV statistic is as low as 13 in the case of DAC-5 Sum5, CS0. Simple percentile calculations from the Dow ΔTV empirical distribution show the lowest p-value for this comparison is 0.38, so in all cases, at least in reference to the Dow and Wilshire 5000, the null hypothesis must be rejected.

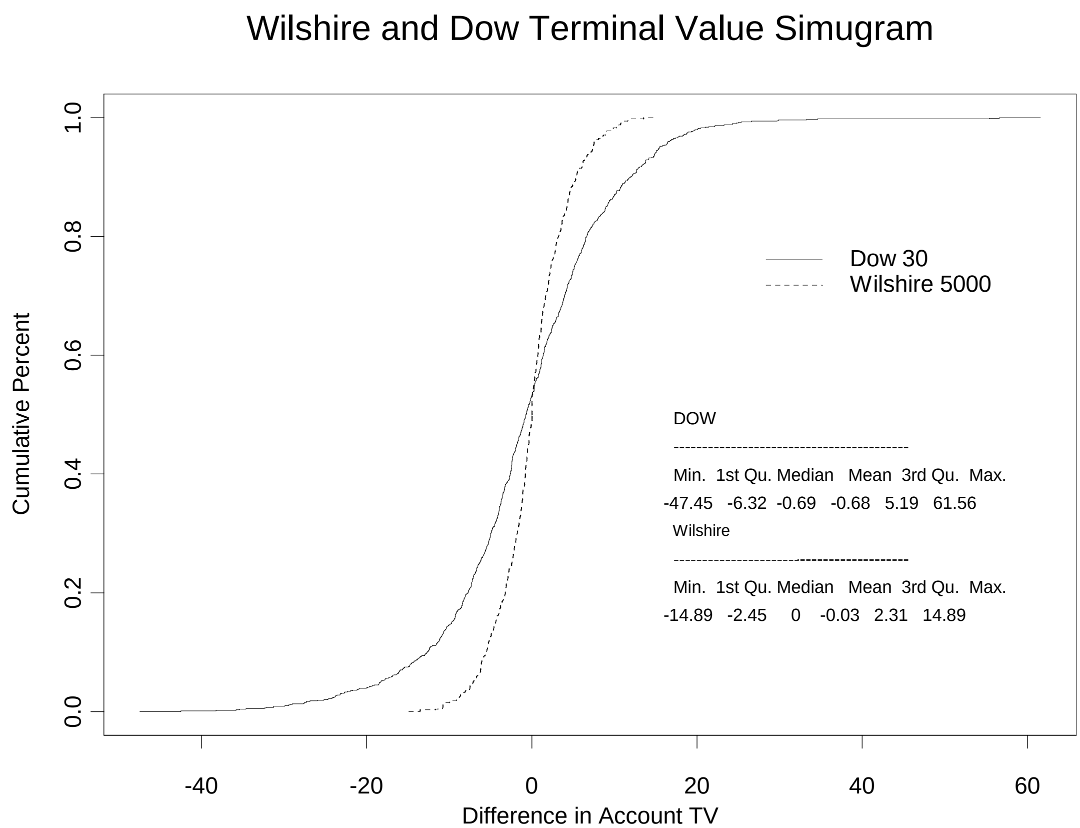

Figure 4.1 Simugram for null hypothesis distribution, M=1.000 Simulations of terminal value, N=33 years, for both Wilshire 5000 and the Dow 30.

The second hypothesis dealt with relative simugram performance and the maximum allocation $\psi_3 = w_i \le a$. On the basis of the alternate constraint set (ACS) curves, it tested that the simugram return would increase for certain increased allocation size. Recall hypothesis 2: Let $\psi_3 = w_i \le a$, be distinguished as $\psi_3^0 = w_i \le a_{cs0}$ and $\psi_3^1 = w_i \le a'$, $a_{cs0} < a' \le b$. Then hypothesis 2 was:

$$H_0: \quad \left\| r(P^A \mid \psi_3^1) - r(P^A \mid \psi_3^0) \right\|_T = 0$$
$$H_A: \quad \left\| r(P^A \mid \psi_3^1) - r(P^A \mid \psi_3^0) \right\|_T > 0$$

Based upon the results of the ACS curves, we have $b \in (.10,.30)$. We have sample data on $b = \{.20,.25,.30\}$, and smatterings of other values. Again we have empirical distribution functions, based on the realized TV for each of $n \cdot t$ samples, giving us a length *n* cdf. We could do much better by generating better representations of possible TV's simply by resampling within each years *n* simugram returns and calculating the TV. It is legal under the null hypothesis to pool all the non-CS0 samples with the CS0 samples. However, any honest statistician would immediately observe the strong bimodal result and admit there is something out of the ordinary going on, and would permit the use of separate populations.

So, for our purposes we just use the length 105 TV's, take samples of length 100 from those, do this 1,000 times, and generate a simugram of length 100,000 for the ΔTV 's. With observations of $a' = 0.20$, $r(P^A \mid \psi_3^1) = 5165$ and $r(P^A \mid \psi_3^0) = 487$, $\Delta TV = 4678$. This has a p-value of 1 since the $\max\left|\Delta TV\right| = 233$, also making us reject H_0 for all the non-CS0 summary results in table 2.15.

For the SP-500, we have a small problem. There are just not enough samples of CS0 TV's to permit a robust resampling approach. We note from table 3.10 that variance of the various SP-500 CS0 simugram returns is less than or equal to the SP-100 baseline. We might be able to parametrically simulate some ΔTV if we could argue for Gaussian; unfortunately, the distribution derived above miserably fails the Kolmogorov-Smirnoff goodness of fit test for normality, with a statistic value of more than 0.15. Since some of our SP-500 results are borderline anyway, we certainly wouldn't want to pool all of our samples. So, we used every SP-500 CS0 complete series of 26 years samples at hand, obtaining 24 samples in all. Clearly here one would have to resample on the individual year data, but for now this is sufficient. We could only bootstrap 20 at a time, but we obtained an ecdf for ΔTV, which ranged from ± 34.7, quartiles of ± 8.8, and a median of 0.06. Since we used all CS0 TV's, whether for DAC-2, -3, or whatever, and one could conceivably test those against a similar null, to be fair we use the corresponding CS0 TV for the comparison. So for example, $r(P^A \mid \psi_3^1)$ for DAC-3 CS25 = 113.4, and for CS0 is 54.8 ($r(P^A \mid \psi_3^0)$). Then ΔTV =58.6, which has an empirical p-value of 1 given the summary statistics on that series. The same result applies with ΔTV =38 for Dac-2 Sum 1. However, ΔTV for Dac-2 Sum2 is 3.4, with p-value of about .50 (based on an eyeball simugram since the lookup function is hindered by the paucity of data). We could go on but we see that in most cases we reject H_0, but in some we don't, which was called to our attention earlier we reviewed the data. This is one example of where work needs to be done, but the general idea, which is borne out more times than not by the hypothesis test, is that the ACS curves' indication of maximum allocation is significant.

Finally, let us address hypothesis 3. This was a prediction on the SP-500 portfolio-splitting approach, and was that the results for DAC-k $\Sigma w = 1$ would be exceed those for $\Sigma w = k$. This was stated as:

$$H_0: \quad \left\| r(P_{s1}{}^A) - r(P_{sk}{}^A) \right\|_T = 0$$
$$H_A: \quad \left\| r(P_{s1}{}^A) - r(P_{sk}{}^A) \right\|_T > 0$$

Unfortunately, there is not sufficient data to parametrically address this hypothesis. Complete 26-year comparative data is only available for DAC-2, with $TV_{DAC-2} = (TV_{s1}, TV_{sk}) = (44.4, 43.5)$, $\bar{r}_{DAC-2} = (\bar{r}_{s1}, \bar{r}_{sk}) = (18.7\%, 18.1\%)$ and

$\hat{r}_{DAC-2} = (\hat{r}_{s1}, \hat{r}_{sk}) = (17.9\%, 17.8\%)$, keeping in mind $\sigma_{TV} \sim 4$ and $\sigma_{\bar{r}} \sim 20\%$. None of these distance measures are significantly different. Sum k results for DAC-3 and DAC-4 are unavailable, and the DAC-5 Sum 5 numbers are so far below even the original stalled results that they do not add to the data needed. Most of the effort expended was toward developing a viable portfolio-splitting methodology, with the end goal of having Σ(w)=1. We can, however, attack the status of this hypothesis using stochastic ordering concepts given the algebraic relationship between the two portfolio returns. Recall from section 3.2 that if k portfolios P_k have returns r_k, and funds are allocated equally with $W_k = 1/k$, and $\Sigma w_k = 1$ inside each P_k, then the total return on capital C is $r_C = \overline{r_k}$. If, however, $\Sigma w_k = 1/k$, then the total return $r_C = k\overline{r}$, or Σr_k. Since $k > 0$, regardless of the distribution of r, if we define $Y = r_{s1} = kX = kr_{sk}$, then $F_Y \le F_X$, and strictly so unless $r_{sk} = 0$. This is expressible in the usual stochastic order as $r_{s1} \ge_{st} r_{sk}$, which should be easily extended to stronger multivariate stochastic orders such as stop-loss or supermodularity. Since the norms used are derived from $r_{\bullet}$, rejection of H_0 is uniform, unless the amount allocated to each by W_k are unequal, which could only occur if discretionary portfolio weighting under an approved trading strategy in *S* were implemented. In our hypothesis, we assumed the W_k were deterministic, so turning W_k into r.v.'s would require restatement of the hypothesis(es), and more complicated result metrics and analysis.

4.1.2 Performance-uniqueness for sub-portfolios.

There has been an unresolved question throughout the process when pondering the relative outperformance of the SPSS applied to the SP-100 vs. that when applied to random portfolios in the DAC methodology. Further research could show this to be a fluke, or that other "random" portfolios such as the Nasdaq 100 (NDX), or possibly the Dow-30, might show similar preferential returns. Indeed, it is well known that each of these indexes is a "sweetheart" index, under the laudable stated intent of maintaining a representative, or "state-of-the-art" composition. For instance, the most recent announced[1] changes to the Dow include deletion of AT&T, International Paper, and Eastman Kodak, and addition of American International Group (AIG), Pfizer (PFE), and Verizon Communications (VZ); Xerox Corporation was among those deleted during the last adjustment in 1999. Traders and anomaly practitioners are aware of the coincident price collapse of these securities and receive such announcements with a jaded eye.

We would expect that market-proxy indexes should perform about the same with respect to each other. When this is done against the OEX, we would expect a mean trace around zero cumulative outperformance. Using year-end return data, it appears that the SP-500 is below the OEX most of the time, and that the Wilshire 5000 (TMW) was generally above it until around 1994. In the latter case, since these are cumulative relative returns, this means that the TMW had to severely underperform in order to have lost that much cumulative percentage. Others could argue these charts do not have enough data to reject an hypothesis of straight line.

---

[1] Dow Jones & Company Press release dated 4/1/04.

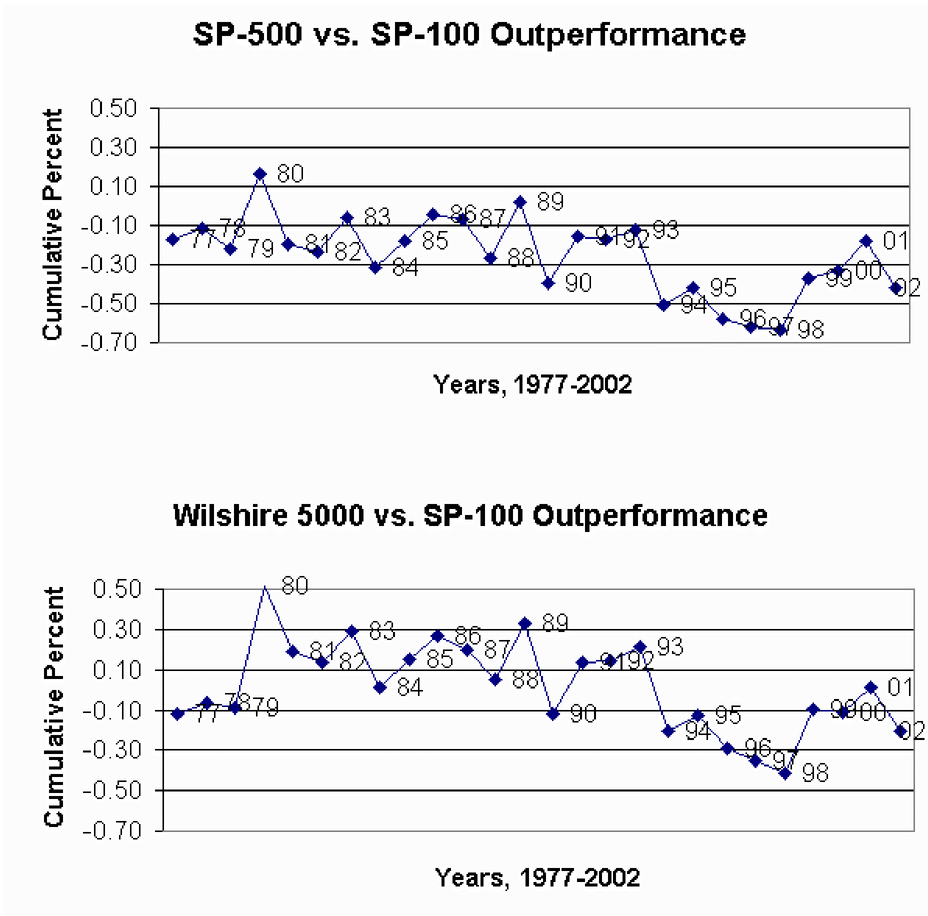

Figure 4.2 Annual cumulative percentage outperformance for the SP-500 (SPX) and Wilshire-5000 (TMW) against the SP-100 (OEX), 1977-2002.

Long-series daily data show a different picture. We use 5,048 daily returns to compute a similar comparison, and include the Dow 30 as well. While it might be argued that the Wilshire 5000 could be considered meandering around the zero line, it would be hard to make that case for either the SP-500 or the Dow.

Table 4.4 Summary statistics for daily returns of TMW, SPX, DOW, OEX and the Geoindex, 1983-2002.

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
r.geo	-0.230	-0.004	0.000	0.000	0.006	0.086
r.oex	-0.238	-0.005	0.000	0.000	0.006	0.085
r.spx	-0.229	-0.005	0.000	0.000	0.006	0.087
r.dow	-0.256	-0.005	0.001	0.000	0.006	0.097
r.wil	-0.197	-0.004	0.001	0.000	0.005	0.074

We did see in section 4.1.1 that the Dow has larger tails due to the slightly higher returns it experiences, which is discernable from table 4.1 only in the maximum values since the 10/90 percentiles are not reported. We also saw what a difference this made in terminal values in the calculation of our hypothesis p-values.

The simple correlation analysis of chapter 2 also does not reveal any discrepancies regarding uniqueness of distribution. Portions of table 2.3's 1970-2003 annual market correlations with respect to the Wilshire 5000, SPX and OEX are reproduced below:

```
Wilshire    1          SP-500      1
--------  -----        ------    -----
SP-100    0.970        OEX       0.980
SP-500    0.960        DOW       0.949
Geomkt    0.953
Dow 30    0.881        SP-100      1
                       ------    -----
                       DOW       0.946
```

With the exception of the Dow to the Wilshire, most of the proxy market indexes show a very high correlation with each other. However, the statistics of quantitative goodness of fit tests tell us otherwise. These sort of tests are described in any statistics book, although sources such as Kendall [51] and Thompson [58] have a more insightful treatment. Simple, 2-sample Kolmogorov-Smirnov tests were conducted to test the likelihood that these indexes belonged to a common market index. The results are summarized in table 4.5.

Table 4.5 Two-Sample Kolmogorov-Smirnov tests for common distribution for TMW, SPX, DOW, OEX, the Geoindex, 1983-2002.

Two-Sample Kolmogorov-Smirnov Test

Reference the Geoindex	Reference the Wilshire 5000	Reference the SP-500
data: r.geo and r.spx ks = 0.0087, p-value = 0.9889	Data: r.wil and r.spx ks = 0.0275, p-value = 0.0419	data: r.spx and r.geo ks = 0.0087, p-value = 0.9889
data: r.geo and r.dow ks = 0.0156, p-value = 0.5551	Data: r.wil and r.dow ks = 0.0351, p-value = 0.0038	data: r.spx and r.dow ks = 0.0135, p-value = 0.738
data: r.geo and r.wil ks = 0.023, p-value = 0.1346	Data: r.wil and r.oex ks = 0.0436, p-value = 0.0001	data: r.spx and r.oex ks = 0.0192, p-value = 0.3008
data: r.geo and r.oex ks = 0.0248, p-value = 0.0874		data: r.spx and r.wil ks = 0.0275, p-value = 0.0419

As noted before, all the constituents and the Geoindex appear to share the same distribution, again presumably by construction, although the OEX is borderline. The Wilshire 5000 tells another story. In spite of their high simple correlation, none of them share the same distribution as TMW. This is undoubtedly due either to sheer number of additional stocks comprising the TMW, or it could be due to the statistical havoc that arithmetic indexing causes in its construction (see section 1.4).

So although sub-portfolios such as the OEX, SPX, and DOW do not appear to made of the same cloth as the Wilshire 5000, they at least appear to among themselves, which leads us to believe that the sub-universe effect is not as important as the possible Nelder-Mead limitations in explaining the simugram’s SP-500 vs. SP-100 underperformance. However, the tantalizing question really remains unanswered.

g. Schedule execution. Unwinding or establishing the 2005 positions will take some time if economies are to be realized. Under the hedging program in section 4.2.1, January expirations are recommended, so there is some time to position the trades.

h. Complete prior-year closeout, reporting and any required regulatory filings.

4.2 Suggestions for Further Research

There are two classes of such suggestions, those required to fully understand this problem, and the those needed to further understanding in the field.

4.2.1 Required Research

a. Complete DAC-4 and DAC-5 Sum 1 tests. It is apparent that the direction of improvement is in this direction, with DAC-3 Sum 1 giving a TV of 55. If this is successful, it would indeed indicate the limitations of the N-M routine. And, the k=5 portfolio is still feasible for larger indexes, which is one of the ultimate aims of this research.

b. Resolve the equal-weighted outperformance issue. This phenomenon alone decries the EMH/CAPM market portfolio and must be better understood, or more universally traded, at least.

c. Implement enhanced strategies such as selling covered calls and hedging the portfolio. As an annual system, once the positions are established, it only makes sense to extract the time value of options during the rest of the year. This strategy is really deferred to the recommended research section, since it is not required for risk or money management. Hedging, however, is a required consideration. Since drawdowns are the only main problem when the market goes down in the forecast year, there should be a trade-off between preserved gain and the cost for insurance to hedge to a 15% loss (which might cost 5%). The following treatment indicates the general direction this leads; the cost estimates are rough, but based on real prices.

Hedging can be accomplished in a variety of ways, through vanilla or complex derivatives, traded on an exchange or over-the-counter (OTC). Options are the preferred instrument: "portfolio insurance," in vogue in the 1990's, was just a program which was supposed to sell futures if it looked like the market was tanking. The problem with this strategy is the occurrence of liquidity holes described in brutal detail in Taleb [51], when the trades simply cannot execute in a useful fashion. So if futures are the instrument of choice, all that can be done is to sell them before these holes appear, which is a mistake since that is equivalent to closing the position, when it was intended to be held for the year.

This leaves options are the vehicle of choice, but which type must be carefully considered. Appropriate varieties include vanilla index or single stock puts, and portfolio, or passport, options. For many reasons, single stock options do not make sense to hedge an index portfolio; and, index options are not appropriate for the sort of baskets of stocks that funds generate, due to correlation risk and the possible presence of non-optionable stocks in the portfolio.

Hedging either is or can be more expensive than most people realize. We briefly attempt to provide an estimate to hedge a $1 million SP-500 index or some other subset portfolio. The most expensive technique is to use vanilla puts; less expensive would be passport options, which are written on the entire portfolio. These are cheaper since it is unlikely that all the stocks would collapse at the same time. Less expensive still would be knock-in puts on the portfolio, for similar reasoning. These are described in any standard text such as Hull [29], or Banks [3]. For the purposes of this vignette, we will work with the vanillas, which would be on the high end of the scenario. Of course, single stock options were not available as exchange-traded instruments until 1973 (on the PHLX), and SP-500 index futures options were not traded until much later, but now the issue is moot.

Unfortunately, the cumulative costs of hedging over the 33 year study period are large and do not pay off unless something catastrophic happens. For single stocks, a 13-month, at-the-money put, in a period of low volatility for a moderately priced blue-chip stock, costs about $4-6; let us use the $5 figure. We will estimate a percent cost to hedge a known percent loss, resulting in a net cost, which has to be subtracted from year's portfolio return. This is done for both SP-500 index (SPX) options, and for a representative stock option (XYZ). Table 4.6 lists the option contract, the number of contracts needed, the gross loss based on the strike vs. the cost basis, the percent cost of the premium, and net loss hedged. This net loss then becomes the floor in any year's portfolio return; the cost is what is subtracted from any year's gain before breakeven is realized. Many characteristics of options are surprising; until one has the path-dependent results, one cannot tell *a priori* which option would be the best choice: one that locks in a profit, a breakeven, a small loss, or a moderate loss. These are presented in the table, and once this "new" optimization problem is completed, we are able to select the set of options that gives the highest terminal value after hedging costs. The optimal choices are indicated by shading in the table.

Table 4.6 Costs to hedge, 33 years, 13-month vanilla SPX index vs. single stock options

Gross		Round # contracts	9				Hedged Term. Value	
Loss	**1119**	**SP-500 Index**	Premium	$ Cost	% cost	Net Loss	SP500	Sim500 cs0
0.051	1175	SXGRD Je 05 1175 put	114.3	102,870	0.103	-0.052	2.94	169.8
0.028	1150	SXGRY Je 05 1150 put	101.4	91,260	0.091	-0.063	3.61	186.9
0.006	1125	SXGRC Je 05 1125 out	89.7	80,730	0.081	-0.075	3.29	198.8
-0.017	1100	SPLRC Je 05 1115 put	87	78,300	0.078	-0.095	2.84	174.0
-0.039	1075	SXGRO Je 05 1075 put	69.9	62,910	0.063	-0.102	3.82	227.5
-0.061	1050	SXGRJ Je 05 1050 put	61.7	55,530	0.056	-0.117	3.96	235.8
-0.084	1025	SXGRC Je -5 1025 put	54.3	48,870	0.049	-0.132	4.07	242.7
-0.106	1000	SXGRZ Je 05 1005 put	49.1	44,190	0.044	-0.150	4.00	237.4
						Initial TV	9.56	487.2

Gross		Round # contracts	250				Hedged Term. Value	
Loss	**40**	**XYZ Stock**	Premium	$ Cost	% cost	Net Loss	SP500	Sim500 cs0
0.000	40	XYZRH Je 05 40 put	5	125,000	0.125	-0.125	0.7	51.1
-0.063	37.5	XYZRU Je 05 37.5 put	4.01	100,250	0.100	-0.163	0.9	63.2
-0.125	35	XYZRG Je 05 35 put	3.04	76,000	0.076	-0.201	1.3	84.9
-0.188	32.5	XYZRZ Je 32.5 put	2.23	55,750	0.056	-0.243	2.0	126.3
-0.250	30	XYZRF Je 30 put	1.68	42,000	0.042	-0.292	2.7	159.3
						Initial TV	9.56	487.2

Three things are evident from this data. First, it costs a lot of money to hedge. For the SP-500 index, over half the terminal value is gone; the net annualized gain has been reduced to 4.3%, with $4.8M in terminal value going to the option sellers. We recall from table 1.4, with reluct, that the annualized returns from 35 years of T-bills is 6.6%, and are led to conclude that the hedged position on the EMH portfolio indeed returns, on average, the risk-free rate (4-5%). Only in the net simugram return do we see any inkling of capitalistic profit, and that was decimated by over 50% as well. Second, we observe that the index options are a much more cost-effective hedge than the vanilla stock options. Unfortunately, the latter would be required for the simugram portfolio, unless passports are implemented. Third is the surprising nature of the index option chosen: with the SPX 1125 put, one could actually hedge a small profit at a cost of 8%. However, the terminal value for the SPX 1025 put is almost 25% higher, but the option locks in an 8.4% loss for a total loss of 13%.   This leads us to conclude that the hedge program is only useful to protect against rare, unfortunate events.

These sobering costs are the result of running both the SP-500 and the simugram portfolios over the 33-year study period with their respective historical returns. Let us examine the impact of worse years than 1974 and 2002. We insert bad returns of –10%, -25%, -50%, -75% and –90%, into either 1973, 1983 or 1993 and see the effect of the hedged vs. the unhedged programs. This is done using both the SPX index and the XYZ options' cost/loss floor. Timing is indeed important. Persons investing $1M in the stock market in 1924 saw a maximum drawdown in 1930 of about 30%, and were back to their $1M within 3 years; those who invested in Summer 1929 saw maximum drawdowns of 90% within 18 months, and had still not broken even a third of a century later. It would be discomforting indeed to have our simugram outperform the market by 15% when the market was down 85%.

Table 4.7 shows the results for one bad year using the cost figures from SPX index options. Recall that these are the "cheap" ones. The terminal values for the SP-500 and simugram portfolios should be easy to differentiate.

Table 4.7 Effect of 1 bad year, SP-500 and Simugram 33-year terminal value, unhedged vs. hedged with SPX index options

	Bad Loss in 1973				Bad Loss in 1983				Bad Loss in 1993			
Loss	Unhedged		Hedged		Unhedged		Hedged		Unhedged		Hedged	
-0.9	1.2	52.3	4.1	238.6	0.8	35.8	3.1	160.4	0.9	32.6	3.5	145.7
-0.75	2.9	130.8	4.1	238.6	2.0	89.5	3.1	160.4	2.2	81.5	3.5	145.7
-0.66	3.9	177.9	4.1	238.6	2.8	121.7	3.1	160.4	3.0	110.9	3.5	145.7
-0.6	4.6	209.3	4.1	238.6	3.3	143.1	3.1	160.4	3.6	130.4	3.5	145.7
-0.5	5.8	261.6	4.1	238.6	4.1	178.9	3.1	160.4	4.5	163.0	3.5	145.7
-0.33	7.7	350.5	4.1	238.6	5.5	239.7	3.1	160.4	6.0	218.5	3.5	145.7
-0.25	8.7	392.4	4.1	238.6	6.1	268.4	3.1	160.4	6.7	244.5	3.5	145.7
-0.1	10.4	470.8	4.1	238.6	7.3	322.0	3.1	160.4	8.0	293.4	3.5	145.7
	9.6	487.2			9.6	487.2			9.6	487.2		

One of the novel things about hedged positions is their absolute stability, even when the investor thinks the position should be worth more than it is. Another is the pain of lost opportunity. In most of these scenarios, one is not glad one has hedged until the unimaginable event happens, such as the 66% negative return. It can be seen that the best time to have one bad year is in the middle of the time period, assuming the time period is long, and that if one has this bad year only 10 years ago, it would be best if the years after the bad year were high return years like in the 1990's.

The next table shows the impact of the more expensive single-stock hedging costs, which is even more breathtaking. The obvious feature is the $2M terminal value for the SP-500 index portfolio. These sort of numbers are why almost no one hedges; or, more accurately, why most investors feel it is more cost-effective to self-insure.

Table 4.8 Effect of 1 bad year, SP-500 and Simugram 33-year terminal value, unhedged vs. hedged with single stock options

	Bad Loss in 1973		Bad Loss in 1983		Bad Loss in 1993	
Loss	Unhedged	Hedged	Unhedged	Hedged	Unhedged	Hedged

-0.9	1.2	52.3	2.4	126.8	0.8	35.8	1.7	85.4	0.9	32.6	1.8	77.6
-0.75	2.9	130.8	2.4	126.8	2.0	89.5	1.7	85.4	2.2	81.5	1.8	77.6
-0.66	3.9	177.9	2.4	126.8	2.8	121.7	1.7	85.4	3.0	[illegible]	[illegible]	77.6
-0.6	4.6	209.3	2.4	126.8	3.3	[illegible]	[illegible]	85.4	3.6	130.4	1.8	77.6
-0.5	5.8	261.6	2.4	126.8	[redacted]							

	Bad Loss in 1973 and 74				Bad Loss in 1983 and 84				Bad Loss in 1993 and 94			
Loss	Unhedged		Hedged		Unhedged		Hedged		Unhedged		Hedged	
-0.9	0.2	7.6	4.1	238.6	0.1	3.6	2.8	148.6	0.1	3.1	3.2	125.7
-0.75	1.0	47.5	4.1	238.6	0.5	22.7	2.8	148.6	0.6	19.3	3.2	125.7
-0.66	1.9	87.8	4.1	238.6	0.9	42.0	2.8	148.6	1.0	35.8	3.2	125.7
-0.6	2.6	121.5	4.1	238.6	1.3	58.1	2.8	148.6	1.5	49.5	3.2	125.7
-0.5	4.1	189.9	4.1	238.6	2.0	90.8	2.8	148.6	2.3	77.3	3.2	125.7
-0.33	7.4	341.0	4.1	238.6	3.6	163.1	2.8	148.6	4.1	138.8	3.2	125.7
-0.25	9.3	427.3	4.1	238.6	4.5	204.3	2.8	148.6	5.1	174.0	3.2	125.7
-0.1	13.3	615.3	4.1	238.6	6.5	294.2	2.8	148.6	7.3	250.5	3.2	125.7
	9.6	487.2			9.6	487.2			9.6	487.2		

Unhedged vs. Hedged with Single Stock Options

	Bad Loss in 1973 and 74				Bad Loss in 1983 and 84				Bad Loss in 1993 and 94			
Loss	Unhedged		Hedged		Unhedged		Hedged		Unhedged		Hedged	
-0.9	0.2	7.6	2.4	126.8	0.1	3.6	1.2	64.1	0.1	3.1	1.4	54.3
-0.75	1.0	47.5	2.4	126.8	0.5	22.7	1.2	64.1	0.6	19.3	1.4	54.3
-0.66	1.9	87.8	2.4	126.8	0.9	42.0	1.2	64.1	1.0	35.8	1.4	54.3
-0.6	2.6	121.5	2.4	126.8	1.3	58.1	1.2	64.1	1.5	49.5	1.4	54.3
-0.5	4.1	189.9	2.4	126.8	2.0	90.8	1.2	64.1	2.3	77.3	1.4	54.3
-0.33	7.4	341.0	2.4	126.8	3.6	163.1	1.2	64.1	4.1	138.8	1.4	54.3
-0.25	9.3	427.3	2.4	126.8	4.5	204.3	1.2	64.1	5.1	174.0	1.4	54.3
-0.1	13.3	615.3	3.6	186.2	6.5	294.2	1.8	94.2	7.3	250.5	2.0	79.8
	9.6	487.2			9.6	487.2			9.6	487.2		

Again, the unhedged results are the same in both tables. There is a significant difference with the double-year hits. The unhedged returns drop by 40% when a 2-year 10% decline occurs in the last half of the study period. For the index option hedge, “breakeven” has been reduced to about a 40% decline in both years, with about a 60% decline for the single stock options. For 1973-74, the unhedged simugram TV is reduced by 20% relative to the index-option hedged version when the decline increases from 33% to 50%, and by over 40% in 1983-84 and 1993-94. The hedged approach does need gains, as we see in comparing the 1- vs. 2-year single stock hedged TV’s, which are 25-30% lower in the 2-year decline periods.

It is possible that the OTC passports or knock-in passports might realize an additional 5% or more discount from the index option cost, but we would not expect an overall 50% savings in any case.. Some substantial price improvement would have to be negotiated in an institutional deal. This would almost have to be the case to justify the sort of fully hedged program outlined above. The concepts of dynamic hedging, though of academic interest, are not applicable for this static portfolio, and are beset with innumerable difficulties, as best explained by practitioners with statistical background. But even the worst of tables in 4.9 show that the hedge program would pay off handsomely in the event of two consecutive 75% or higher declines; in that case, the additional $20-50M would be warmly received.

4.2.2 Suggested Research

These are not listed in order of importance, but are natural outgrowths of the results in this thesis.

a. Study the temporal nature of optimal r* (such as in table 2.8), to explain the reason for the apparent cyclic fluctuation. This can probably be accomplished with traditional time-series analysis in the Box-Jenkins [12], or the Maddla and Kim [35] generalized methodologies.

b. Conduct more complete time- and sampling-persistence studies and derive estimates for the VSL.

c. Investigate the positive impact of taking advantage of the "edge effect." Recall this takes advantage of the VSL slack to incorporate the larger returns when the sum of weights constraints are violated. If it can be shown that a constraint violation within the VSL slack results in a significantly higher return, even if only 0.5%, then it should be exploited to enhance returns.

d. See if there is an enhanced universe effect on other small target indexes like the Dow 30 or the Nasdaq 100 (NDX). Also, try "small-portfolio" optimization for enhanced universe effects on the Dow 30 vs. randomly selected sub-portfolios. This will require some adjustments to r* since it was discovered for small portfolios that the resulting return IS a function of r*, unlike the general character of the of the curves in figures 2.9-12. This is again due to the optimizer attempting to satisfy the tail return criterion with a limited number of good elements to choose from.

e. Develop a multivariate market outperformance distance measurement, perhaps along the lines of the practical ideas in Scott [47]. Figure 1.5 showed a graphical outperformance measure when compared to one market benchmark; a technique should be developed to summarize comparisons with multiple benchmarks. In part, the geomarket index was developed to accommodate this multivariate approach while still using the scalar comparison measures in chapter 1.

f. Improve the geomarket index with perhaps a better choice of weights. Also, argue for the inclusion in the geoindex an index of mutual funds, since they are now a non-reversible proportion of the total market capital (as opposed to what they were in 1991).

g. Conduct theoretical analysis of the DAC-k Sum k portfolios. For example, for DAC-2, if group A is the top market capitalization, and group B is the bottom, are they different from randomly chosen groups.

h. Study the result of including cumulative returns for all previously selected stocks. Currently the program just resamples from the prior year returns. If the collection of past returns were accumulated, then these prior returns would better define the empirical distribution of the stocks' returns. However, if there is a temporal short-memory effect in addition to the well known long-memory effect, then we would expect the simugram returns to not do as well, and the portfolio might revert to the market returns.

i. Continue studies in the distribution of simugram returns. We saw in chapter 2, in exploring the variance of simugram returns that about 60% of the years exhibited decidedly non-Gaussian distributions. We also saw in section 2.6.3 that increasing M changes the distribution in subtle ways, not necessarily negatively. But these should be understood better and explained.

j. Modify for system for baskets of commodities. The system needs to be redesigned to handle short (i.e., the Black) portfolios. If it can weed out bad stocks, it should be able to weed out good ones, too. This important capability would allow the same principles to be used on selecting any other portfolio of commodities which have the same trending features as stocks do with respect to time, such as the currencies, Benchmarks are easy to obtain, or could be constructed with the appropriate geoindex.

k. Conduct and compare parametric investigations, including Poissonian jumps, versus these non-parametric studies. Obtaining weights via parametric modeling, or perhaps using Bayesian updating, might better incorporate the expected outlier behavior for the individual portfolio elements, and hence result in higher returns.

l. Implement an enhanced strategy of covered call writing, with estimates of additional terminal value. This requires a time-varying position model, which has already been developed by the author under prior research. It will be found that not all the calls expire into free money; one reason the simugram returns are so high is that the stocks selected go on to perform very well in the next year. It would be nice if this exceptional performance were due to slow, steady increases over the year; however, and unfortunately for option-sellers, many times this is accomplished by several greater than 3-sigma jumps, followed by a settling down to the steady growth average. The model would take some losses to keep the stock, but it is believed this results in an additional 3-5% return, which would be especially important to help offset hedging costs implemented under subsection 4.2.1(c).

m. Time-scaling of simugram forecasts. What is the proper granularity for shorter-term forecasts? For monthly forecast, is 2-hour data sufficient? For 1-week forecast, is 15-minute data sufficient? There are many possibilities here, including monthly portfolio rebalancing.

n. Perform the notation and analysis necessary to show that the simugram risk profile is a coherent risk measure as in Artzner [1].

4.3 Recommendations for Improvements

a. Investigate use of better high-dimensional direct search optimization algorithms which can directly handle portfolio sizes of K=500.

b. Improve the code and data collection process. This would increase the size of the portfolios by preserving more stocks presently excluded because of missing data. This would include obtaining index membership data from further back for deeper backtesting.

c. Recommendations for improving portfolio splitting.

1. Try discretionary splitting. E.g., put all the OEX into one group, DOW- and SP-30 into another, etc.

2. Try market capitalization splitting and other approaches suggested by subsection 4.2.2(f).

3. Try a 2-pass optimization on the stalled SP-500 1-pass or DAC-2 Sum 2 optimization, i.e., use an allocation of 0.025 and reoptimize on the VSL. This would result in a new portfolio size of approximately 75 to 150, while retaining most of the target portfolio member synergy. Preliminary testing on this approach was undertaken just prior to publication, with a maximum allocation of 5%; with 2 samples, the resulting terminal value was 46, about $10M less than the DAC-3 results. This indicates that the synergy effect might be large, and that split-group allocations for 2 groups should be about 2.5%, giving the combined portfolio as indicated above.

4.4 Practical Issues

This system was developed with an eye to implementation. That is why the VSL was developed, and why emphasis was placed on the drawdown effect as impacting investors, clients or stakeholders. Even with that in mind, according to the numbers, a fund starting with $1M on this program would be on the order of $50M in less than 20 years. At this level, some of the execution capability we take for granted in the discussion become more important. These issues include:

a. Difficulty in unwinding positions with large size indicates the need for and advantage of block and negotiated OTC trades among other funds. For enhanced funds, there is the efficacy of selling calls to get rid of the stock and risks associated with not getting called away.

b. Scheduling/Timing. These include trading schedules, timing of LEAPS purchases and exercise, implementing the new weights, etc. Although it would be advisable to precess the sales of securities in a clever fashion so as to adjust for short-term vs. long-term capital gain (ST/LTCG) so that STCG occur only every 20 years or so, in reality the unwinding of large positions might have to occur over a number of days, complicating this tax strategy. Some of the schedule items are discussed in section 4.1.3.

c. Depending on the size of the business, the trading in a business organized around the SPSS will have staff dedicated to investment review, execution, risk management, and back-office function. For example, the risk management function might insist that the investment program is hedged whether the overall strategy requires it or not.

d. There are efficiencies when an entire symbol does not need to be eliminated, but only adjusted.

Finis