Chapter 1

1.0 Foundations.

A new means to select weights on a target benchmark stock portfolio is presented, which results in better performance than that of the target market index. This is accomplished using the simugram as described in section 1.3 and in chapter 2. The simugram applied to financial engineering problems is a time-indexed risk profile showing the entire distribution of outcomes of a stochastic experiment. Using the simugram to maximize a portfolio return objective function subject to risk-tolerance constraints, we show that the resulting simugram portfolios exhibit at least twice the return as the equivalent target market index, and up to 50 times the terminal dollar value over two study periods of at least 25 years. Important conclusions on the distribution of portfolio returns and terminal values are made, as well as on the effectiveness of the Nelder-Mead optimization algorithm in high dimensions. The effectiveness of the simugram portfolio as an investing or trading program is also demonstrated.

1.1 Problem Statement

This dissertation tests a major and several subsequent hypotheses regarding a new type of portfolio selection process. For brevity, the problem setup and statement uses mathematical language. Those interested in the direct hypothesis statement can skip to page 4.

Let M^S be the market for equity securities where they might be exchanged for money, and let and M^P be the market for all privately held, restricted, preferred, or for any other issue not thought of as common stock, and let M^C be the market for all publicly traded common equity securities (common stocks). Put *M*= $M^C = M^S \setminus M^P$ and let a "selection universe" Ω^M be the collection of all publicly traded common stocks, and suppose $\Omega^M \subset M$.

Let $X_i \in \Omega^M = X_i(t), t \geq 0$, be a common stock so that X^M can be arrayed as the *nxN* matrix $\begin{bmatrix} X_1 & X_2 & \cdots & X_N \end{bmatrix}$, $N < \infty$. Let *P* be a portfolio of stocks X_i given weights w_i with $P = \begin{bmatrix} w_1 X_1 & w_2 X_2 & \cdots & w_k X_k \end{bmatrix}$, $w^T 1_k = \sum w_i = c$. Usually, $c = 1$. Then $P^M = \begin{bmatrix} w_1 X_1 & w_2 X_2 & \cdots & w_N X_N \end{bmatrix}$, $w^T 1_N = 1$, is Sharpe and Lintner's price-adjusted, "net asset value" market portfolio (Markowitz, [37]). Markowitz makes clear in his review that their definition is strictly in terms of a row vector of weights, with $w_i = \dfrac{X_i \cdot Vos_i}{X \cdot Vos}$, or the ratio of the value of X_i to the total market value, as determined by the price, $X_i(t)$, and shares outstanding $Vos(\tau)$, where $\tau >> t$.

We must next define some stochastic processes. We can model $X_i(t)$ as a stochastic process $\left(X(t); t \geq 0\right)$[1] with distribution function F_{X_i}, and $X^M(t)$ is jointly distributed F_X^M. Let $r(\cdot)$ be a measurable "return function" $r(X(t), X(t - \Delta t))$ which generates some sort of differential price change, with $r^M = r(X^M)$. r^M can be modeled as a

[1] The probabilistic context is that $X(t)$ is defined on the underlying probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathrm{P})$. This notation should not appear again, except for the distributions F_X, F_r induced on the target space of *X*, which will be used in the portfolio setting.

stochastic process with distribution function F_r^M, which experience has shown is not quite multivariate normal (MVN). Also, let r_f be a parameter called the risk-free rate, a realization in time of the US Treasury-bill interest rate process. $r_f(t) \in \Omega^F \subset M^F$, where F denotes Fixed Income.

We now create some subset portfolios. Let $\Omega^0 \subseteq \Omega^M$ and $P^0 = (\Omega^0, w^0)$ be subsets of the market portfolio such that the correlation $\rho(X^0)(t) = \rho(X^M)(t) - \varepsilon$, with ε small, say 0.01. This will be the market portfolio proxy since the market portfolio is unobservable in practice. Examples of Ω^0, in order of increasing ε, would be the Wilshire 5000 Index (denoted Ω^{5000}), the Standard & Poor's (S&P) 500 (Ω^{500}), the S&P-100 (Ω^{100}), the Dow 30 (Ω^{30}), etc. Define $P^A = (\Omega^A, w^A)$ as an alternative portfolio composed of the same stocks as Ω^0 but with different weights. By market capitalization weighting, for example, $P^0 \sim P^M$.

Finally, let there be some aggregate measure of comparative performance over time, $\left\| r(P^A) - r(P^0) \right\|_T$, where T denotes a time interval. Candidates for this norm are discussed at some greater length in section 1.6 below.

Given the preceding assumptions, we are proposing an alternative $P^A = (\Omega^0, w^A)$ based on a standard market index universe, which is intended to outperform the benchmark index returns. We are now in the position to state our main hypothesis.

Hypothesis 1

$$H_0: \quad \left\| r(P^A) - r(P^0) \right\|_T = 0$$
$$H_1: \quad \left\| r(P^A) - r(P^0) \right\|_T > 0$$

Simply put, there should exist a set of weights on the elements of a benchmark universe that creates a portfolio which outperforms the benchmark index.

The null hypothesis has been well established for almost half a century by what is collectively known as the efficient market hypothesis (EMH), discussed below. Under whatever strength version one wishes to use, under the null hypothesis, $F_X^A = F_X^0 \doteq F_X^M$ and $F_r^A = F_r^0 \doteq F_r^M$.

The problem is described as follows. Let $R(\cdot)$ be a return function, and $\Omega = \Omega^0$, for $X_i, w_i, i = 1,k$, and let a set of environmental parameters, linear and non-linear constraints be denoted $\psi' = (\psi_P, \psi)$, where ψ_P are parametric, or variables exogenous to the optimization but required in the overall stochastic simulation, and ψ be the set of optimization constraints. The specific objective function, constraints, and other parameters are discussed in chapter 2.

Problem Statement

$$\max R(w, X, \psi')$$
$$\text{s.t. } \psi$$

This problem seems to closely resemble the standard Markowitz mean-variance "standard analysis with upper bounds." It is not the Tobin-Sharpe-Lintner (T-S-L) problem described in [38] because there is no borrowing/lending consideration; that investment allocation decision is assumed to have been made. Nor is it the Black problem since we are constraining ourselves to no short positions. In the Markowitz [37] formulation, the goal is to select weights for Ω^0 subject to the following constraints ψ:

$$\psi_1 : \sum w = 1$$
$$\psi_2 : w_i \geq 0$$
$$\psi_3 : w_i \leq U$$

where U is an upper bound on maximum percent allocation. His optimization problem is:

$$\min V = w^T \Sigma w$$

s.t.

$$\phi w = b$$
$$\mu^T w = R = r^T w$$

where Σ is the covariance matrix of returns, ϕ is a function of the sum of the weights, and R is a target portfolio return. After this paradigm was formalized, beginning in 1952, it quickly allowed a relatively easily calculated covariance matrix to substitute for *risk*. In the process of time, optimality results followed. In his 1987 review (*op cit*) of the follow-on studies, Markowitz finds “the mean-variance approximations provide almost maximum expected utility except for utility functions … which have pathological risk aversion.”

With a way to select P^M via the exciting new field of linear programming, soon Sharpe’s and Lintner’s capital asset pricing models (CAPM) began to mature, and with them the establishment of the EMH. Even a review of the reviews of the CAPM/EMH would not be helpful in testing Hypothesis 1. Good texts include Campbell, Lo and Mackinlay [14], Lo and MacKinlay [34], Malkiel [36], or Williams and Findlay [65]; an important and more poignant treatment is in Thompson, Williams and Findlay (TWF) [61]. The CAPM assumes all investors share the same beliefs, have access to the same information, are all rational, all seek $\mu - \sigma$ efficient portfolios, and all share the T-S-L constraint sets, and enjoy the encouraging result that P^M is an efficient portfolio. Abnormal returns above $\mu_i - r_f \propto \beta_i$ are not sustainable, or even real. EMH puts all admissible portfolios on a capital market line (CML), with P^M occupying the limit point position.

The EMH is a watchdog theory, which enforces the position of P^M on the CML. In George Lucas' film *Star Wars* [27] , there was an unspeakably powerful and unassailable offensive platform called the "Death Star," much like the EMH. Results that begin in the anomalies literature, such as Fama and French's Arbitrage Pricing Theory[2], eventually become "reconciled" - the EMH just gets "bigger."

Many, though, have employed Lucas' "below-the-radar" penetration strategy in attempting to earn a living either from within or without the "Death Star." These are called arbitrageurs. Even in the Black-Scholes-Merton [6, 7, 8, 9, 39] and Cox-Ingersoll-Ross (CIR) [16, 42] models, in which it is assumed that opportunities for arbitrage do not exist, within the same papers they elsewhere argue that any mispriced security, futures or option prices are expected to return to "equilibrium" via arbitrage activity, as it is in put-call parity, and with any other arbitrage argument. This means there is an astute industry of arbitrage agents which perform their function admirably, so well in fact that they must

---

[2] Fama E.F., French K.R., "The Cross-section of Expected Stock Returns", *Journal of Finance*, Vol. 47, No. 2, 1992. This study originally threatened the CAPM, but was subsumed into EMH.

constantly revise their strategy after they have cleared away the instant opportunity, or else go out of business.

1.2 A Note on Returns.

When studying changes in economic or financial time series, terminology has developed over the years, some of which might be confusing or counter-intuitive. The purpose of this section is to summarize some of this standard terminology regarding “returns”, and show alternative representations. We will introduce investor, mathematical, and geometric definitions of return.

"Return" as used in this thesis and in financial mathematics denotes the change in a deterministic or stochastic variable over its prior value. This change is usually defined to be relative to the prior value. In contrast, the businessperson thinks in terms of an investment returning a periodic stream of cash payments over time, which calculates to a "return on investment." And, in an election context, the political scientist's returns represent an important sample of voter preference and a in most cases the outcome of a candidate's bid for office. Neither of these routinely considers returns as representing the changes in stochastic variables at an arbitrary temporal granularity.

The political scientist we shall not address now, but consider the businessman *A* who invests *C* in a capital project, selected by another optimization skill-set outside the scope of this dissertation. In simplest terms, *A* expects a future value (FV) of accumulated earnings such that he realizes a rate of return (ROR) of say 18% over a time horizon *H*.

Since he is rational, he is indifferent to the FV so long as his ROR is realized. Let him invest for time $H = n\Delta T = 1$. His return is still a return, even if $\Delta T = 1$. This can be 1 year, as Keynes mused, or it could be one day. *A* is actually indifferent to the horizon, for if the project cash flow annualizes to his ROR by time 3*T* instead of time 10*T*, his expectation is met. He may choose to stick with the investment; or, remorse over opportunity cost may have set in and he is off to hunt again.

Horizon-shortening occurs, as in this windfall scenario: suppose at 9:35 a.m. ET on 9/11/01, investor A received execution on his annual trade long the SP500. His hurdle rate is 10%. Investor B received the same, but bought one stock, Northrop Grumman (NOC). Investor *B* could have purchased a defense-sector mutual fund or ADR, but was simply adding to a single-stock portfolio. While we held our breath when the markets reopened on the next trading day,[3] NOC opened up +13 points (a 6σ event in the right direction, a mere 4 1/2 years after a previous 10.7σ return). Rational investor *B* sells all, pockets the 15%, and sails to Haifa. His one day return is a real return. Investor *A* broke even 2 years later, on 9/30/03.

The best-know early 20th century author connecting the price behavior of financial and economic time series and their changes with Gaussian random variables was Louis Bachelier (1900) [2]. But is was H. Working [70, 71] at Stanford's Food Research Institute, who first firmly introduced the apparent universality of the notion as a subject for study into the economic and finance literature beginning in 1934. The concept slowly spread in popularity until Samuelson took up the idea of Brownian motion in the stock

[3] See the Findlay *et al* (2003) paper by this title [25].

markets in 1955, followed by the rich history as outlined in various sources such as Cootner [15] and the host of others. For log price series, the difference is indeed the mathematical return *r*. The principle is in fact an application of Pearson's and Fisher's idea of "transformation to normality."

All the common return measures are transformations of the increase/decrease concept, defined as $R_t = \frac{X_t}{X_{t-1}}$. We refer to this as the raw return *R*. Other measures are percent (investor) return $r_{\%}$, natural log (mathematical) return *r*, and a geometrical degree return $r_d = D(h)$, where *h* is a scaling parameter usually related to volatility.

Investor, or percentage return, is $r_{\%} = \frac{X_t - X_{t-1}}{X_{t-1}}$, and $r_{\%} = R - 1$. Mathematical return is $r = \ln(R)$. We note that $r \leq r_{\%}$; however, for small *r*, $r_{\%} \doteq r$ by nature of the log expansion. Handy identities between these returns include $r_{\%} = R - 1 = e^r - 1$, and $r = \ln(r_{\%} + 1)$.

Degree returns are a geometric concept based on the triangle whose origin is at $t_0 = 0$, with base of length $H = n\Delta t$, and height *r*. Let *r* at $n\Delta t$ represent an annual mathematical return. Any return between $t = (0, \Delta t, 2\Delta t, ..., H)$ on the ray of angle θ from t_0 to *r*, is scaled to the same annual return. The return angle θ is calculated as the arctan of the dimensionless $\frac{r_d}{k\Delta t / H}$. What makes this return useful in parametric work is that for Brownian motion processes, a simple scaling of *H* to *h* provides a 1-1 map between θ expressed in degrees, and the *z*-score of a normal random variate. Degree returns *D(h)* in terms of z-scores are quite useful. The most common scaling is for $1^\circ = 1 \cdot \sigma$; other simple scalings provide for $\theta = 45^\circ$ at $z = 1 \cdot \sigma$, etc.

In selecting a return measure for use, one must consider tradeoffs between ease of calculation and layman understanding, and the importance of its linearity and symmetry properties. Almost any return measure is linear and symmetric for small changes, say less than 5%, owing to the expansion properties for the log, sine, and arcsine functions. However, even a modest change of $\pm 20\%$ corresponds to a log return of $(+.18, -.22)$, a "20% spread" At greater changes asymmetries increase, since *R* and $r_{\%}$ are constrained to $(0, \infty)$ and $(-1.0, \infty)$, respectively, but $r \in (-\infty, \infty)$. *D(h)* is nicely symmetric between $(-90^\circ, +90^\circ)$. These are summarized as follows:

Locally Symmetric	Globally Symmetric
Raw return, R	z-scores, z
Percent return, r%	Degree return, D(h)
Log return, r	
z-scores, z	
Degree return, D(h)	

In parametric work we always use mathematical returns, especially when evaluating Gaussian or more general stochastic processes. But this work is non-parametric, and its theme is that of market outperformance, and since market performance is ubiquitously quoted in percentages, we use $r_{\%}$ throughout. Additionally, percent returns are sometimes better at expressing the "emotional" component. For instance, suppose one's

Pages 11-23 omitted

1.5 Market Performance and Outperformance.

In this section we propose several means to measure market performance, and, by comparison, outperformance. The most commonly quoted measure is that of annualized return, which is an internal rate of return achieved by the future terminal value from a one-time initial investment due to variable period returns. It provides a single number, expressed in percent, which would, like a bond, give the terminal value realized over the time horizon. It is always less than or equal to the mean period return; that is because it is the geomean of the period returns. We notate it as $\hat{r} = \left(\frac{A_T}{A_0}\right)^{1/N} - 1$. One can of course easily account for periodic investments such as dividends, cash contributions, etc. Unfortunately, it is a fictitious number, never actually realized except by calculation, but is useful for summary comparisons. It is also prone to misuse by those attempting to annualize short period returns to a greater scale.

A more realistic and fundamental measure is the terminal value (TV) amount, $A_T = A_0 \prod (1 + r_t)$, where r_t are the investor returns realized in each period. These returns can be historic or forecast, and A_T will always be in numeraire, a sometimes objective unit. The disadvantage with using terminal value is its susceptibility to haggling regarding the time value of money, discount rate to be used in the present value reduction, inflation effects, and simply the long periods over which most of these analyses take place. Assume by whatever means one is able to achieve an annualized internal rate of return of 10% over *T*=30 years. If one were able to increase the return each year by only 2 basis points (0.2%), it would make a $1M difference at time *T*, about

5% of the $19.2M the base account would earn. But an extra million is much more than the average Nobel Prize-winning co-author realizes in the twilight of his life. In the case of the Standard & Poor's (S&P) 500 portfolio optimization, we are dealing with T=26 years, and an additional 1% improvement in the simugram returns equates to an extra $10M in TV, or a 25% TV performance improvement.

Simple differences and cumulative differences also have some value in quantifying the relative performance of two series. These, like index ratios, provide a quick indication of the outperformance or rate of outperformance. Investments which are essentially the same give zero or flat presentation. Unfortunately, these values do not seem to have a simple mapping to TV, or to educe an intuitive meaning. For continual out- or under-performance, these values can accumulate to rather large positive/negative values.

It is generally believed that stocks outperform the risk-free rate. Implicit in that belief is the long time period required to base it on. Table 1.4 gives 35 years of data on annual returns of the Standard & Poor's 500 (SP-500) index alongside the corresponding 1-year US Treasury bill secondary market rate, taken from US Federal Reserve online resources. The fixed-income rates are synchronized to maturity with the last trading day of each year. A trading strategy of buying the SP-500 (with some rebalancing each year) vs. buying 1-year T-bills each year is compared.

Table 1.4 Comparison of returns, SP-500 vs. 1 year Treasury Bills, 1965-2000. Assumes all principal reinvested.

Comparative Market Returns - SP500 vs. 1-Year T-Bills

Year	SP500	T-Bills	Year	SP500	T-Bills	Year	SP500	T-Bills
12/31/65	0.091	0.041	12/30/77	-0.115	0.057	12/29/89	0.273	0.079
12/30/66	-0.131	0.051	12/29/78	0.011	0.077	12/31/90	-0.066	0.074
12/29/67	0.201	0.047	12/31/79	0.123	0.097	12/31/91	0.263	0.055
12/31/68	0.077	0.055	12/31/80	0.258	0.109	12/31/92	0.045	0.037
12/31/69	-0.114	0.068	12/31/81	-0.097	0.132	12/31/93	0.071	0.033
12/31/70	0.001	0.065	12/31/82	0.148	0.111	12/30/94	-0.015	0.050
12/31/71	0.108	0.047	12/30/83	0.173	0.088	12/29/95	0.341	0.056
12/29/72	0.156	0.048	12/31/84	0.014	0.099	12/31/96	0.203	0.052
12/31/73	-0.174	0.070	12/31/85	0.263	0.078	12/31/97	0.310	0.053
12/31/74	-0.297	0.077	12/31/86	0.146	0.061	12/31/98	0.267	0.048
12/31/75	0.315	0.063	12/31/87	0.020	0.063	12/31/99	0.195	0.048
12/31/76	0.191	0.055	12/30/88	0.124	0.071	12/31/00	-0.101	0.058

Terminal $ Value – SP500:	15.6	**Geomean % (annualized) - SP500:**	7.9%
Terminal $ Value - T-Bill:	9.9	**Geomean % (annualized) - T-Bills:**	6.6%

The various outperformance measures can be easily listed:

```
Difference in Terminal Value ($M)          5.7
Sum of outperformance (%)                  0.904
Cumulative sum                             0.904
Avg of outperformance (standard dev.)      0.025   (0.162)
TV of outperformance ($M)                  1.525
geomean of outperformance (%)              0.012
```

The traditional overlay graph as well as the cumulative outperformance percent is plotted in figure 1.4:

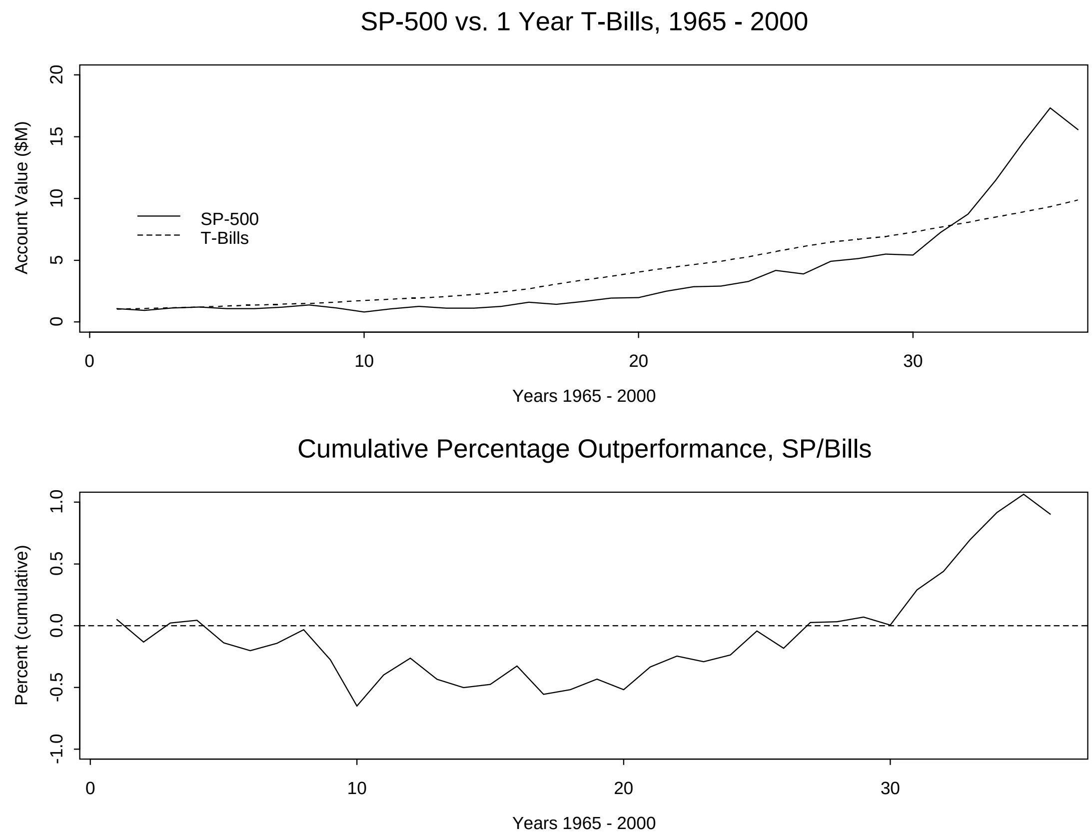

Figure 1.4 Comparison of returns, SP-500 vs. 1 year Treasury Bills (top panel); the bottom panel shows the cumulative percentage outperformance of the SP-500 vs. T-bills for the same period.

Over this time period, the outperformance of the SP-500 over that of T-bills is not that convincing. The stock return annualizes out with only a 1.3 percentage point edge; but we know that makes almost a $6M difference over 35 years. And, the 36% stock market decline in 2001-2002 will materially change the character of the plot.

When comparing a return stream to multiple indexes, other decisions come into play, such as what sort of norming needs to be employed in the comparison, or other distance measure candidates. Figure 1.5 illustrates the issue.

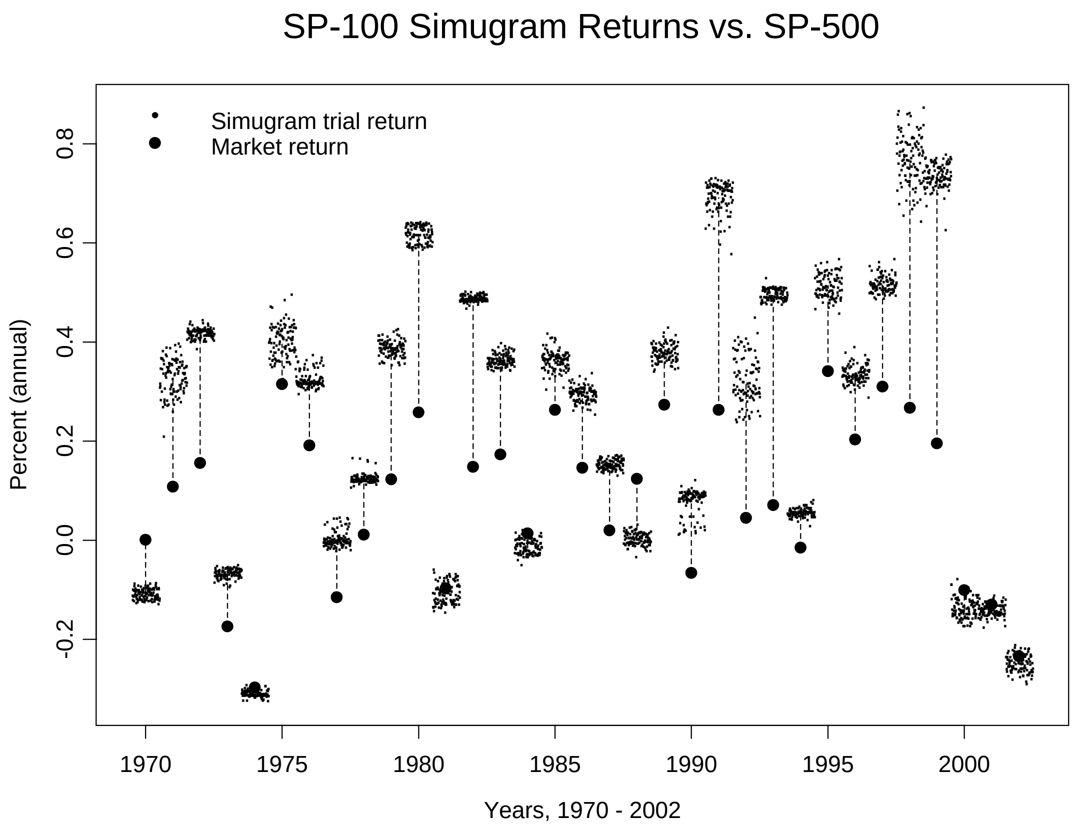

Figure 1.5 SP-100 simugram returns vs. SP-500 market returns; 105 simugram trials for each of years 1970-2002.

Simugram returns for the SP-100 portfolio from multiple trials are plotted relative to the actual market (SP-500) return for the year, with a “leader” drawn into the centroid of each year’s trials. If one were comparing these with several other market indexes, would one average the leader lengths, or calculate an appropriate normalized length or area? One could average in some way the difference in terminal values. Without a foray into index theory, in evaluating our results we will use either the Wilshire 5000, SP-500, or Geomarket Index as the appropriate benchmark, and the terminal value difference between the selected reference and our hard-earned portfolio returns.