UNIVERSITY OF ROCHESTER
William E. Simon Graduate School of Business Administration

FIN 411 **Professor G. William Schwert**
Investments **Spring 1997**

TEACHING NOTE -- NOT FOR QUOTATION OR CIRCULATION

Statistical Behavior of Stock Prices

I. The Random Walk Model

The simplest form of the random walk model assumes that the natural logarithm of "with dividend" price, $ln\ P_t$, changes by random amounts through time:

$$ln\ P_t = ln\ P_{t-1} + \mu + \varepsilon_t, \quad (1)$$

where

P_t is the sum of the price plus any dividend payment made by the corporation to its stockholders in period t,

$\mu = E[ln\ (P_t/P_{t-1})]$ is the expected ***continuously compounded return***,[1] and

ε_t is the random change in the stock price from period t-1 to period t.

If ε_t is serially uncorrelated, then the ***changes*** in (logs of) stock prices are ***random***, and (logs of) price levels follow a ***random walk***. In other words, if the stock returns are random, then the price levels follow a random walk.

The simple rate of return to an asset is

$$R_t = (P_t - P_{t-1} + D_t) / P_{t-1} \quad (2)$$

which is close to the continuously compounded return r_t for small values of the returns.[2]

---

[1] Define $P_t = P_{t-1}\ exp(r_t)$, so that the beginning of period investment P_{t-1} grows exponentially at the rate r_t. Then $r_t = ln(P_t/P_{t-1})$ is the continuously compounded rate of return to the asset. The natural logarithm of the exponential function equals unity, $ln\ exp = 1$.

To understand how random walks behave, imagine leaving a party at the Distillery (or any other similar establishment). Draw a picture of the path you would follow if, at each step, you stagger either left or right randomly as you walk down the sidewalk. Flip a coin denoting heads as "right" and tails as "left". If you carry out this coin-flipping experiment, you should notice an apparent pattern in the your movement down the path -- once you stagger either left or right, there is no tendency for you to return back to the middle of the sidewalk. Given the picture of your path after 100 steps, an “expert” observer could probably conjure up a variety of theories to explain the pattern of your walking. Nevertheless, you and I both know that your “random walk” was purely the result of the ***random*** outcomes of your coin-flipping experiment.

To a large extent, the behavior of stock returns is similar to the walk of the inebriated student -- many observers could conjure up theories to explain the path of prices that they just observed, but none of these theories could explain the ***future*** course of stock prices because each future price change is uncorrelated with past changes. Equation (1) can be rewritten in terms of returns,

$$r_t = \mu + \varepsilon_t, \tag{3}$$

where the “unexpected” return in period t is the error term ε_t. As with the errors from regression models (e.g., APS 402), the errors should be random. Autocorrelation tests on returns are equivalent to testing whether the errors are indeed random.[3] Most evidence finds that common stock returns have very small autocorrelations for daily or monthly data for many lags k. Thus, these tests are consistent with the *random walk model* for stock prices.

---

[2] It is straight-forward to show that $r_t = ln\,(1 + R_t)$, and it is a well-known mathematical fact (that can be checked on most hand calculators) that: $ln\,(1+x) \approx x$ for $x < |.15|$. Thus, for returns less than 15 percent, these two measures are close to each other. The continuously compounded return is always less than the simple return. (Check this out.)

[3] Remember that an **autocorrelation** coefficient is just the correlation of the return in period t, r_t, with the return in period t-k, r_{t-k}.

II. Some Properties of the Random Walk Model

The mean and variance of returns are proportional to the length of the measurement interval k. So

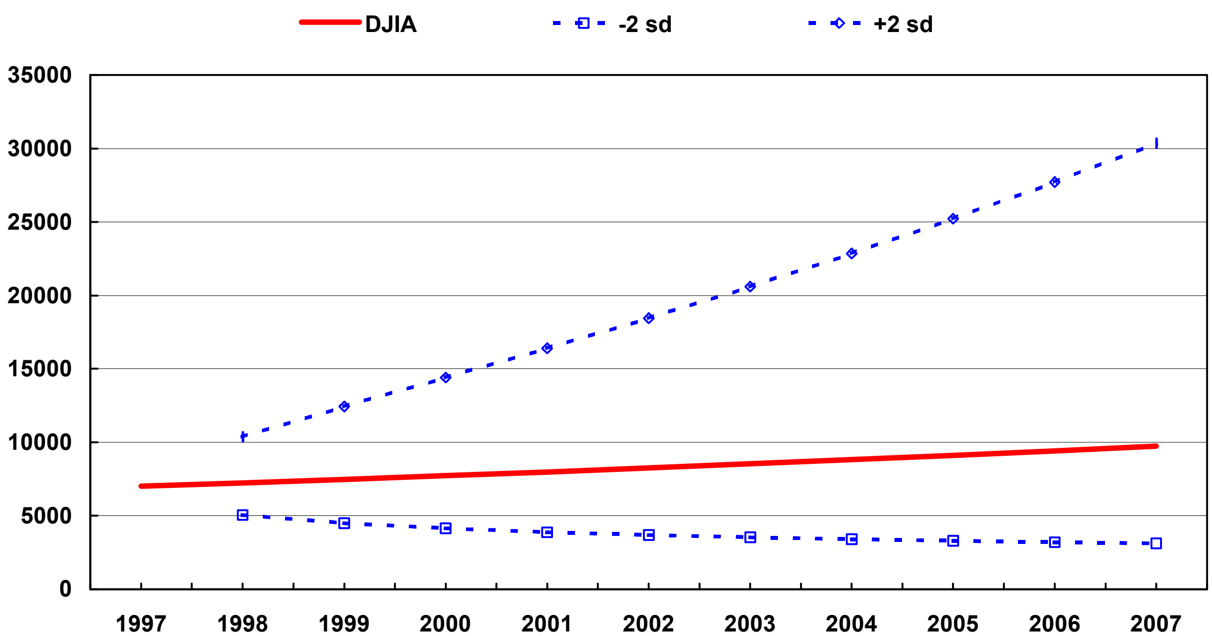
$$E\,[\mathit{ln}\,(P_t/P_{t\text{-}k})] \;=\; k\,\mu \tag{4a}$$

$$\mathrm{Var}\,[\mathit{ln}\,(P_t/P_{t\text{-}k})] \;=\; k\,\sigma^2. \tag{4b}$$

Thus, starting at some "early" point in time t = 0, both the mean and the variance of the value of the stock P_t grow larger as you look out further into the future. As an exercise, I calculate a 95% confidence interval around predictions of the Dow Jones Industrial Average (DJIA) if the current value is 7000, the annual expected continuously compounded capital gain return is 3.3 percent, and the standard deviation of the annual return is 18.0 percent (the historical sample statistics for 1834-1995.) These forecasts are constructed for 1, 2, and 10 years into the future in Figure 1.

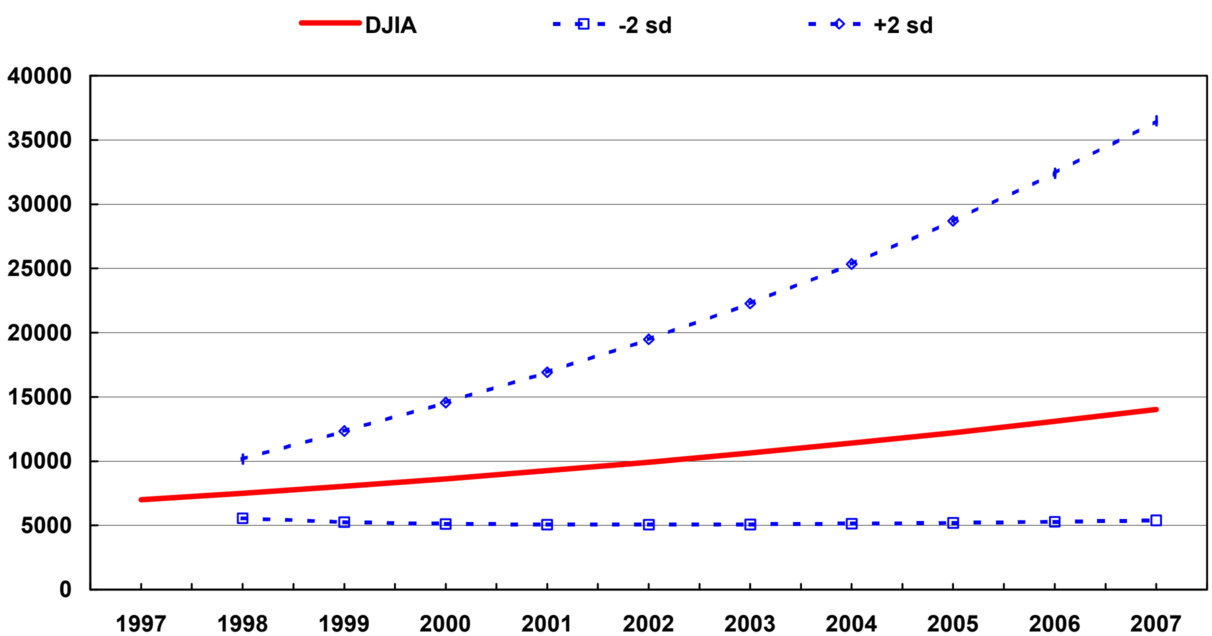
I also construct forecasts assuming the annual expected continuously compounded capital gain return is 7.0 percent, and the standard deviation of the annual return is 15.1 percent in Figure 2 (the historical sample statistics for 1946-95). You can see that the point estimate grows smoothly in both Figures 1 and 2, and that the 95 percent confidence interval gets much wider as you look further into the future. In Figure 2 the upper limit rises above *36,000* by the year 2007! The point estimate is above 14*,000* in 2007 in Figure 2. The continuously compounded returns are assumed to be Normally distributed (see below), so confidence intervals for *ln* P_t would be symmetric. However, the exponential function is nonlinear, so the confidence intervals for P_t are larger above the mean than below (this reflects the fact that prices cannot go below 0.)

Fig. 1 -- Prediction of Dow Jones Industrial Average
(Random Walk Model, 1997 = 7000)



Mean = 3.3%, Std Dev = 18.0% for Returns, 1834-1995

Fig. 2 -- Prediction of Dow Jones Industrial Average
(Random Walk Model, 1997 = 7000)



Mean = 7.0%, Std Dev = 15.1% for Returns, 1946-95

III. The Distribution of Stock Returns

A secondary issue concerns the distribution of the returns (or errors). For a number of reasons that will become clearer when we discuss portfolio theory, it is often convenient to assume that stock returns have a "Normal" distribution. This particular mathematical model has many nice properties (e.g., sums of Normals are Normal; the Normal distribution is completely characterized by its first two moments, the mean μ and standard deviation σ or variance σ^2; and lack of correlation is equivalent to independence for Normals). It turns out that monthly stock returns are reasonably well approximated by a Normal distribution, although there is a tendency for there to be too many extreme values ("outliers").[4] Distributions of daily returns are more "fat-tailed", but still look relatively bell-shaped and symmetric.

Figure 3 shows the plot of monthly stock returns from 1834-1996. Figure 4 shows the histograms for stock returns for the 1834-1996, 1834-1925, 1926-45 and 1946-96 sample periods. The time series plot of returns looks random around a small positive mean, with a substantial standard deviation. Note that the spread in the data seems much greater during the 1929-40 period of the "Great Depression." The histograms in Figure 4 look similar in the different samples: they are all symmetric around the mean and there are probably too many extreme observations (fat-tails).

It is useful to know that there are at least two reasons that histograms can look fat-tailed relative to a Normal. First, of course, the distribution generating the data could really be fat-tailed (e.g., the Student-t distribution, the Stable-Paretian distribution, etc.) Second, and I believe more important, if one samples from a Normal distribution many times, but the variance of the distribution changes during the sampling, the resulting histogram will seem fat-tailed relative to a Normal with constant mean and variance. In other words, if there is *heteroskedasticity* (non-constant variance of returns), the sample distribution will seem fat-tailed.

---

[4] This is often referred to as a fat-tailed, peaked, or leptokurtic distribution relative to the familiar bell-shaped Normal curve.

Fig. 3 -- Monthly Stock Returns, 1834 - 1996
Smith-Cole/Macaulay/Cowles/CRSP Data

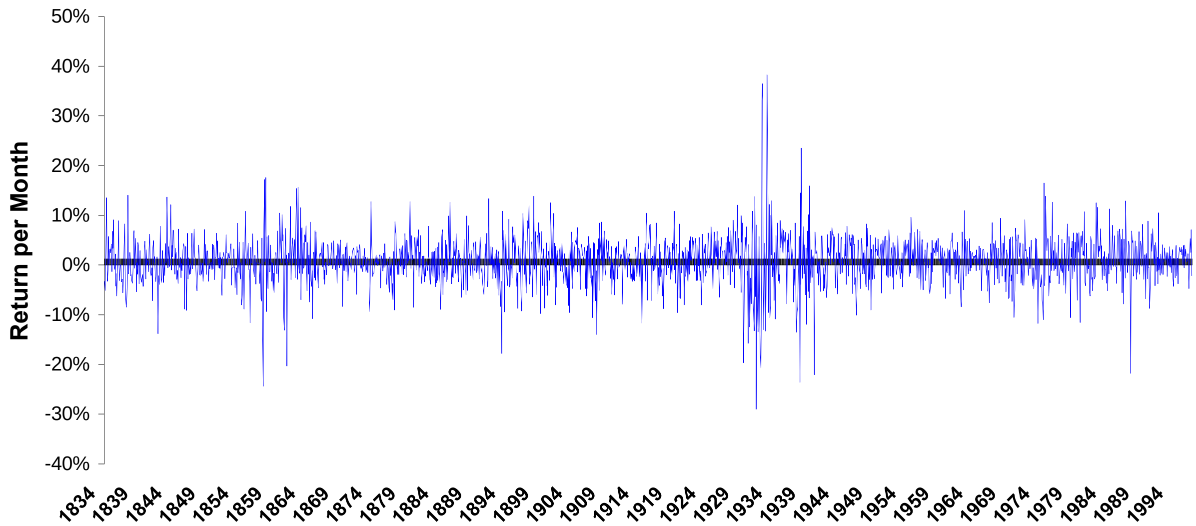


Fig. 4 -- Histograms of Monthly Stock Returns

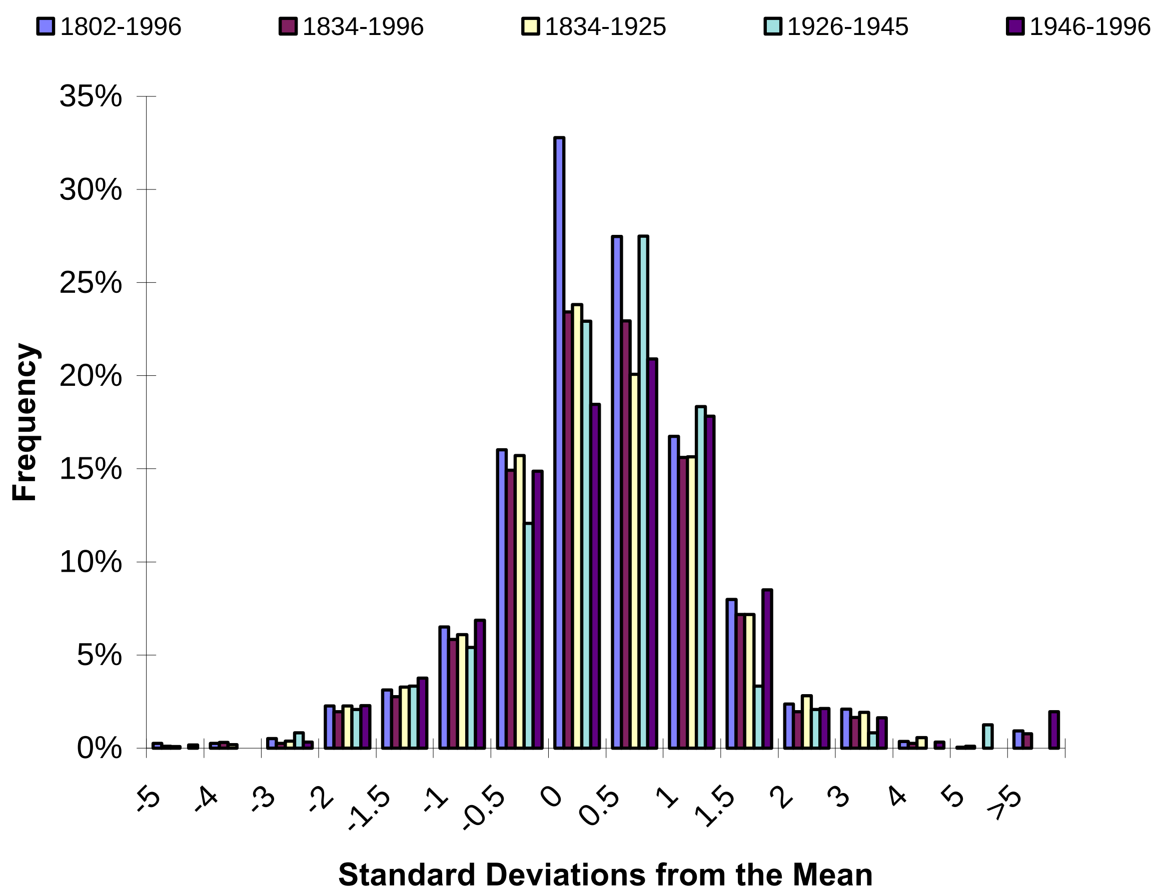


Table 1 contains summary statistics for monthly rates of return to a large portfolio of common stocks, to long-term corporate bonds, and to short-term commercial paper or Treasury bills for 1834-1995. Over the entire sample period, the mean return to stocks is about 0.79 percent per month (about 10.0 percent per year), with a monthly standard deviation of 4.89 percent (about 18.8 percent per year).[5]

$$\sum_{m=1}^{12} (1+R_m) - 1 = R_a$$

The annual variance is just 12 times the monthly variance if stock returns are not autocorrelated, so the annual standard deviation is 12 = 3.46 times the monthly standard deviation. Papers by Fama and French and Poterba and Summers show that this approximation is not always accurate. Mean returns to long-term bonds and to short-term bills are .41 and .37 percent, with standard deviations of 1.30 and .24 percent, respectively. Note that the standard deviation of bill yields is not a measure of risk (since these instruments have virtually no default risk), it merely measures changes in expected returns (interest rates) over time. Thus, stocks have higher mean returns and much higher standard deviations than long-term corporate bonds or bills. The Studentized Range (SR) statistics for stocks and bonds indicate fat-tailed distributions or heteroskedasticity. SR = {[Max – Min] / Standard Deviation}, which is the range of the data expressed in units of standard deviations – i.e., how big are the largest outliers? SR statistics around 5 or 6 are reasonable if the underlying data come from a normal distribution with a constant mean and variance. Larger SR values indicate either fat-tailed distributions or heteroskedasticity.

[5] The annual return R_a is the product of 1 plus the monthly returns R_m, minus 1. See section IV below for more details.

Table 1. Distribution of Stock and Bond Returns and Short-term Interest Rates, 1834-1995

	Stock Returns	***Long-term Bond Returns***	***Short-term Yields***
		1834 - 1995	
Avg	0.79%	0.41%	0.37%
Std	4.89%	1.30%	0.24%
Max	38.28%	8.89%	1.67%
Median	0.79%	8.89%	0.35%
Min	-29.00%	-7.35%	-0.02%
SR	13.75	12.47	7.10
T-stat	7.17	12.77	68.08
		1834 - 1925	
Avg	0.67%	0.46%	0.42%
Std	4.36%	1.04%	0.20%
Max	17.59%	6.92%	1.67%
Median	0.52%	0.42%	0.38%
Min	-24.37%	-7.35%	0.13%
SR	9.62	13.75	7.54
T-stat	5.08	12.87	68.78
		1926 - 1945	
Avg	0.85%	0.29%	0.09%
Std	8.06%	1.32%	0.13%
Max	38.28%	5.06%	0.43%
Median	1.31%	0.42%	0.02%
Min	-29.00%	-7.83%	-0.02%
SR	8.35	9.75	3.56
T-stat	1.63	3.44	10.93
		1946 - 1995	
Avg	1.01%	0.48%	0.38%
Std	4.08%	1.72%	0.26%
Max	16.53%	9.84%	1.35%
Median	1.21%	0.31%	1.35%
Min	-21.81%	-7.20%	0.03%
SR	9.40	9.90	5.18
T-stat	6.07	6.90	34.83

IV. Returns Over More than One Period

For the purpose of analyzing returns over more the one time period it is often convenient to use continuously compounded returns, $r_t = ln(1+R_t)$, since these returns add up over time:

$$ln\,(P_{t+k}/P_t) = \sum_{i=1}^{k} ln\,(P_{t+i} / P_{t+i-1}) = \sum_{i=1}^{k} r_{t+i}.$$

If one were to use simple returns, it is the product of $(1+R_{t+i})$ that is used to measure the change in value:

$$P_{t+k} = P_t \prod_{i=1}^{k} (1+R_{t+i}),$$

where the product operator Π is analogous to the summation operator Σ. Thus, the k-period return (P_{t+k}/P_t)-1, is

$$(1+R_{t+1})(1+R_{t+2})...(1+R_{t+k}) - 1 = R_{t+1} + R_{t+2} + . . . + R_{t+k},$$

ignoring the cross-product terms. These cross-product terms can be important, however, as shown in Table 2 by the example of two period returns to three different assets, A, B and C. Even though the average returns to the three assets are the same, asset A has the highest value at the end of period 2, followed by asset B, then asset C. The assets with the highest standard deviation of return have the lowest terminal value for a given level of average simple returns. You should convince yourself that the average continuously compounded returns are *not* equal for these three assets.[6]

[6] The expected simple return $E(R) = \exp\{E(r) + ½\, \sigma^2(r)\} - 1$, where $E(r)$ and $\sigma^2(r)$ are the mean and variance of the continuously compounded returns, respectively, if the continuously compounded returns are Normally distributed.

Table 2

	Returns to Different Assets		
Period	**Asset A**	**Asset B**	**Asset C**
1	0.10	0.20	0.30
2	0.10	0.0	-0.10
Average 2 Period Return	0.10	0.10	0.10
Value of $1 Investment in Period 0 at the end of Period 2	$ 1.21	$ 1.20	$ 1.17

Table 3 below shows the sample statistics for the simple and continuously compounded monthly returns to the CRSP value-weighted portfolio for 1926-95 and two subperiods. The simple and continuously compounded returns are almost perfectly correlated, with the continuously compounded returns being smaller than the simple returns (the means are lower, as are the maximums and minimums.) The regression intercept is negative, and the slope is essentially 1. This illustrates the differences between continuously compounded and simple returns.

Table 3

Distribution of CRSP Value & Equal-weighted Stock Returns, 1926-95

	Cont Comp CRSP VW Returns	Simple CRSP VW Returns	Simple CRSP EW Returns
	1926 - 95		
Avg	0.80%	0.96%	1.31%
Std	5.52%	5.53%	7.61%
Max	32.41%	38.28%	65.51%
Min	-34.25%	-29.00%	-31.23%
SR	12.08	12.16	12.71
t-test	4.22	5.03	4.98
α	-0.15%		0.10%
β	0.994		1.257
$t(\beta=1)$	-3.87		5.69
R^2	0.993		0.835
	1926 - 45		
Avg	0.53%	0.85%	1.64%
Std	7.99%	8.04%	11.66%
Max	32.41%	38.28%	65.51%
Min	-34.25%	-29.00%	-31.23%
SR	8.34	8.37	8.29
t-test	1.02	1.63	2.18
α	-0.31%		0.49%
β	0.989		1.365
$t(\beta=1)$	-0.47		5.71
R^2	0.990		0.886
	1946 - 95		
Avg	0.91%	1.00%	1.18%
Std	4.13%	4.12%	5.16%
Max	15.33%	16.56%	29.92%
Min	-25.48%	-22.49%	-27.10%
SR	9.89	9.48	11.05
t-test	5.43	5.97	5.58
α	-0.09%		0.08%
β	1.001		1.095
$t(\beta=1)$	0.15		2.57
R^2	0.998		0.762

V. Returns for More than One Security (i.e., a Portfolio)

On the other hand, simple returns are easiest to use when measuring the returns to many different securities at the same point in time -- in other words the returns to a ***portfolio***. Define a portfolio return as the weighted average of the returns to the securities in the portfolio:

$$R_{pt} = \sum_{i=1}^{N} w_{it} R_{it}, \qquad \text{where} \quad \sum_{i=1}^{N} w_{it} = 1.$$

The portfolio weights w_{it} represent the proportion of wealth invested in asset i at the end of period t-1 (the beginning of period t). If an investor put equal dollar amounts in each of N securities, this would be an *equal-weighted* portfolio ($w_{it} = w_i = 1 / N$). If one invested in proportion to the outstanding market value (i.e., price times shares outstanding) of each of N securities, this would be an *value-weighted* portfolio (w_{it} = value of asset i / total value of all N assets.) The return to an equal-weighted portfolio is the average return to the assets in the portfolio in period t so it is easy to compute, but it is hard to maintain an equal-weighted portfolio through time because you must rebalance every period as the value of the holdings change. On the other hand, the return to a value-weighted portfolio is difficult to compute because the value-weights change every period, but it is easy to maintain a value-weighted portfolio since it requires *no rebalancing* (except to account for new issues or retirements of securities.) The S&P 500 and the CRSP value-weighted portfolios are example of value-weighted portfolios.

By construction, a value-weighted portfolio places larger weight on large firms, and therefore smaller weight on small firms relative to an equal-weighted portfolio. A comparison of the returns to the CRSP value- and equal-weighted portfolios of all NYSE stocks from 1926-95 shows this effect. For 1926-95, the mean monthly returns are .96% and 1.31% for the value and equal-weighted portfolios; the monthly standard deviations of returns are 5.53% and 7.61%. These differences are greater for 1926-45, when the mean monthly returns are .85% and 1.64% for the value and equal-weighted portfolios, and the monthly standard deviations are 8.04% and 11.66%, respectively. Thus, the risks and returns to small stocks are probably higher than for large stocks.

Table 3 also shows the market model regression using the CRSP value-weighted portfolio as the regressor,

$$R_{et} = \alpha + \beta R_{vt} + \varepsilon_t, \qquad t = 1,\ldots,T \tag{5}$$

where R_{et} is the simple return to the equal-weighted portfolio and R_{vt} is the simple return to the value-weighted portfolio in period t. The slope coefficient β ("beta") is a measure of the relative nondiversifiable risk of the asset which is the dependent variable in the regression (R_{et}) as part of the portfolio which is the independent variable (R_{vt}). The weighted average beta equals 1.0 by construction for the set of assets which make up the regressor portfolio (R_{vt}). The weighted average intercept in the market model regression (5) must equal 0.0 by construction.

If the Capital Asset Pricing Model (CAPM) is true, then the intercept $\alpha = (1 - \beta) R_f$, where R_f is the risk-free rate of return, so betas greater than 1 would typically be associated with α's less than 0. To see this relation (which we will discuss in more detail later in the course), take expected values of (5) and solve for the intercept, $\alpha = E(R_e) - \beta E(R_v)$. Next, use the CAPM to determine what the expected return to the equal-weighted portfolio should be: $E(R_e) = R_f + \beta [E(R_v) - R_f]$. Substituting this relation into the equation for (5) yields:

$$\begin{aligned} \alpha &= R_f + \beta [E(R_v) - R_f] - E(R_v) \\ &= R_f [1 - \beta]. \end{aligned} \tag{6}$$

For the 1926-95 period and the subsamples shown in Table 3, the beta of the equal-weighted portfolio is significantly above 1 (t-statistics of 5.69, 5.71 and 2.57). On the other hand, the intercept estimates are all positive, suggesting that the returns to the equal-weighted portfolio are higher than would be predicted by the CAPM. We will see more about this evidence later in the course -- it is sometimes called the “small firm effect”.

Table 4 shows estimates of the market model in risk-adjusted form for the equal-weighted portfolio:

$$(R_{et} - R_{ft}) = \alpha + \beta (R_{vt} - R_{ft}) + \varepsilon_t, \qquad t = 1,\ldots,T. \qquad (7)$$

Because the risk-free rate is subtracted from both the dependent and independent variables in the regression, the intercept should equal zero if the CAPM is true.

Table 4

Test of the CAPM Using Equal-weighted Stock Returns, 1926-95

Sample Period	α	**t(α=0)**	β	**t(β=1)**
1926-95	0.18%	1.86	1.258	5.72
1926-45	0.52%	2.23	1.366	5.73
1946-95	0.11%	1.12	1.094	2.58

Thus, you can see in Table 4 that the average monthly return to the equal-weighted index is between 0.1% and 0.5% per month higher than predicted by the CAPM. The effect is largest (and the t-statistic is more than 2) from 1926-45.