for either event studies or performance measurement" (Dybvig and Ross 1985b:415). They end their paper by restating essentially what Roll (1977:138) had already said, "[t]o put the matter bluntly, SML analysis tells us more about the index we are using than it tells us about the portfolio we are evaluating" (Dybvig and Ross 1985b:415).

It would appear that Roll (1977) is correct. The CAPM cannot be tested nor should it be used. Recently, Grinblatt and Titman (1989) introduced a model of portfolio evaluation that gives accurate rankings of performance. It appears that Grinblatt and Titman (1989) are using CAPM type analysis, but are not using the CAPM itself. This can be seen from their benchmark portfolio. Grinblatt and Titman (1989:411) argue that the most appropriate benchmark is a portfolio which contains those assets that could have been included into the portfolio under consideration. This implies that the portfolio manager should be evaluated relative to the asset universe that he invests in. This is a relative concept that does not measure the efficiency of the portfolio under evaluation vis-a-vis the market portfolio. The Grinblatt and Titman (1989) benchmark portfolio, while being an excellent idea, is a monument to the failure of more (objective) market portfolio's. It is final evidence that the CAPM's second testable hypothesis cannot be exploited in order to observe superiority in portfolio evaluation.

Some may wish to make the argument that Roll (1977) overemphasized his case. Yet in the final analysis, we cannot escape from the fact that to date the CAPM remains untested and unusable. This is not to say that the CAPM is unused, however the pitfalls of using the CAPM are known.

**4.1.4 CAN BETA EXPLAIN RETURNS AT ALL?**

Fama and French (1992:427) argue that there are two consequences of mean variance efficiency within the CAPM context:
- there is a linear relationship between expected return and $\beta$; and,
- $\beta$ should describe the cross section of average returns to the exclusion of other variables.

They test the latter implication with respect to size and book to market equity ratios.

Their results shed light on numerous issues in finance (both corporate and investment) and highlight a potential bias in previous studies. Fama and French (1992:432) report that they are able to find a positive relationship between average return and $\beta$ only when portfolios are formed on the basis of size. When size is controlled for in the test, there is no relationship between average return and $\beta$.

They report even stronger results negating the CAPM. When the size effect is isolated and removed from the test the relationship between return and $\beta$ has a slope equal to zero, even when there are no other variables in the regression (Fama and French 1992:438). There are two possible explanations of this result:

- The $\beta$'s are systematically related to other variables which obscure the relationship between measured $\beta$ and returns; and,
- noise in the data obscures the relationship (Fama and French 1992:438).

The evidence is so overwhelmingly against the CAPM, that Fama and French (1992:440) extend their test period from 1963-1990 to the fifty year period 1941-1990. They report for the entire period that there is no relationship between return and $\beta$, but that for the period 1941-1965 such a relationship does exist. However this latter "anomaly" disappears once the size effect is controlled for.

Fama and French (1992:444) test leverage and book to market equity ratios as

---

*In the previous chapter we saw that SML's were flat (when not negative). In those regressions, no ranking was made on the basis of size.*
alternate explanations for security returns. They are able to conclude that, "[i]n a nutshell, market β seems to have no role in explaining the average returns on ... stocks for 1963-1990 while size and book to market equity capture the cross sectional variation in average stock returns" (Fama and French 1992:445).

The arguments presented in this section seem to indicate that the CAPM is not internally valid. In the next section, we will examine whether the CAPM is externally valid.

4.2 THE EXTERNAL CONSISTENCY OF THE CAPM

In the early and mid-1970's the CAPM became to be generally accepted as being valid. It is towards the end of that decade that we see the first real cracks in the theory. It is interesting to note that the first anomalies were interpreted as being violations of the EMH. It seems that the EMH has been strictly interpreted and that this strict interpretation has led to it being discredited. As we saw in chapter two, the EMH should be broadly interpreted. It is with this thinking in mind, that market anomalies will be discussed, i.e. we assume the EMH to be correct.

4.2.1 ANOMALIES WITHIN THE CAPM

In his article on anomalies, Thaler (1987a:197) begins with the following quote from Thomas Kuhn: "[d]iscovery begins with the awareness of anomaly, i.e., with the recognition that nature has somehow violated the paradigm-induced expectations that govern normal science." According to the CAPM, investors are rewarded for bearing systematic risk, the measure of systematic risk is β and there is a linear relationship between β and return. Any systematic divergence from this situation is an anomaly.

Beginning with Basu (1977) a whole pantheon of anomalies have been
identified. These anomalies can be broadly categorised into:
- size effects;
- calendar effects.

In this section, the evidence in support of each of these two effects will be discussed. Cross linkages between the two effects will then be considered, i.e., are we really looking at only one anomaly that manifests itself in various guises?

4.2.1.1 SIZE EFFECTS

In addition to models which relax the assumptions of the CAPM, various *ad hoc* models have been investigated (Keim 1986a:20). One of the early investigations was that of Banz (1981). He investigated whether or not there was a relationship between market capitalization of the firm and return. His reason for doing so is, that while there were no conclusive direct tests of the CAPM (Banz 1981:3), "recent evidence suggests the existence of additional factors which are relevant for asset pricing".

Banz (1981:5) tested all stocks on the NYSE (subject to data availability) for the period 1926-1975. In order to "avoid" the Roll (1977) problem, he used three indices, the CRSP equally and value weighted stock indices and a composite stock-bond index (this he refers to as the "market index"). Stocks were assigned to portfolios firstly on the basis of their market value and then on the basis of their $\beta$. The model employed in the analysis was as follows:

$$ R_i = \gamma_{0i} + \gamma_{1i}\beta_i + \gamma_{2i}S_i + \epsilon_i $$

where $S_i = [(\phi_i - \phi_m)/\phi_m]$  
$\phi_i$ = market value of firm $i$  
$\phi_m$ = average total market value.

Banz (1981:8) states that the results of the study are essentially identical for all three indices, this he interprets as meaning that the choice of index does not affect the result and the Roll (1977) problem is circumvented. This result is similar to that of Stambaugh (1982).
He finds that the CAPM is probably misspecified: \(\gamma_0\) differs from the risk-free rate and there appears to be a negative relationship between firm size as measured by market capitalization and return (Banz 1981:8). The average excess return on holding a "small firm portfolio" would be 19.8%, however this would be offset by then having a poorly diversified portfolio (Banz 1981:16).

The longevity of the size effect is such that it is unlikely to be due to market inefficiency. He further states that he is unable to tell whether or not he has discovered a pure size effect or whether size is a proxy for some other variable.

One possible explanation is that size is proxy for P/E ratios (this is discussed below). The argument that Banz (1981:17) favours is that there is a positive relationship between size and information generation. Thus investors may not wish to hold the stock of small firms (we will also return to this argument below).

Elton and Gruber (1991:430) state that the market folklore has been to buy low price/earning (P/E) ratio stocks. Keim (1986a:25) traces this back to Graham and Dodd (1940:533) who argue that investors should buy stocks with "reasonable" P/E ratios. Basu (1977) was the first person to research this issue within a CAPM framework. Basu (1977) tests for the P/E effect as a test of the EMH, however he recognises that his tests are also a joint test of the CAPM. He writes (Basu 1977:670) that, "to the extent these models do not reflect the equilibrium risk-return relationships in capital markets, the related evaluative measures also do not appropriately measure the performance of the various P/E portfolios". Basu (1977) recognises that the CAPM may be misspecified.

---

46 Basu (1977:663) references others who have done work in this area, but argues that they have not adjusted for the following factors: risk, transaction costs, and taxes. He further argues that these studies suffer from various defects that make their findings questionable.
Basu (1977:663) writes that there are those who argue that prices are biased and that P/E ratios reflect this bias. Basu (1977) sets out to test whether or not such a bias exists on a risk adjusted basis. Basu (1977) formed five portfolios on the basis of P/E ratios. He then tests them on the following basis:

\[ R_{pt} - R_n = \delta_{pt} + \beta_{pt}(R_{mt} - R_n) \]

where \( \delta_{pt} \) is the differential return given the CAPM, i.e., Jensen's measure.

Basu (1977:666) finds that the two lowest P/E portfolios earned on average 13.5% and 16.3% over the test period (1957-1971) and the highest P/E portfolio earned 9.5%. These differentials were not related to the \( \beta \) differentials of the portfolios. Thus, low P/E portfolios outperform high P/E portfolios.

Basu (1977:670) also tests whether \( \delta_{pt} \) is systematically related to its \( \beta \). This is a similar test to that of Black, Jensen and Scholes (1972). This test confirms the previous study, the higher the \( \beta \) the lower the \( \delta \) and vice versa. Basu (1977:670) writes that when \( \beta \) is held constant, "\( \delta \) seems to depend on its P/E class". Basu (1977:671) argues that his results are consistent with one of the two following propositions:

- the CAPM is misspecified and relevant variables have been omitted; or,
- if the CAPM is correct, then P/E ratios carry information that has not been incorporated into price, thus the EMH is not valid.

Perhaps it should be mentioned that these two propositions are not mutually exclusive - they may both be correct. Basu (1977) seems to prefer the explanation that the EMH is not valid. Banz (1981:17) prefers the argument that the EMH is valid and the CAPM is misspecified.

Reinganum (1981a) investigates the size and P/E effects in order to determine whether or not they are related to each other. In trying to do so, he first tests whether or not the EMH is valid. He does so by first testing whether or not abnormal returns can be earned by forming portfolios on the basis of unexpected earnings - he finds that this is not possible. This result, Reinganum (1981a:24) argues is consistent with the CAPM and provides evidence of the validity of the EMH.
The next stage of the analysis is to determine whether or not there is a P/E effect and a size effect. He confirms the existence of a P/E effect in both quarterly and annual data (Reinganum 1981a:34 and 38) and concludes that given the evidence that the EMH is valid, that the CAPM is misspecified. The approach that he then takes is of interest here. He establishes a relationship between the size effect and the P/E effect.

Reinganum (1981a:41) categorises firms by both their size and their P/E and finds that firms with low P/E's (high E/P's in his terminology\textsuperscript{47}) tend to have low market capitalizations. A possible explanation of this is that one is a proxy for the other or that they are both proxies for some other factor. Reinganum (1981a:42) then controls for the P/E effect by creating twenty five portfolios on the basis of P/E ratios and size. He then calculates the risk adjusted returns for each portfolio and finds that the smallest firms in each P/E category outperform the large firms. This he interprets as being evidence of the size effect persisting irrespective of the existence of the P/E effect (Reinganum 1981a:43). Thus he concludes (Reinganum 1981a:44) that the size and P/E effects are proxies for the same "set of factors" that are missing from the specification of the CAPM. The size effect is a better proxy for this variable(s) than the P/E effect. Thus there is no double anomaly, the size and P/E effects are only one anomaly. Basu (1983:26) corroborates this point, with one major difference: he argues that the size effect is subsumed by the P/E effect.

Cook and Rozeff (1984) also undertook a study to determine whether or not the P/E effect or the size effect was dominant or whether they were two separate anomalies. They used nine methods for estimating abnormal returns, three portfolio formation rules and variety of statistical tests on the data (Cook and Rozeff 1984:450). Their data sample was taken from the CRSP tapes and covered the period 1964-1981.

\textsuperscript{47} Some of the literature refers to E/P ratios which are the inverse of the P/E ratios. We refer to P/E ratios throughout for the sake of continuity.
They report (Cook and Rozeff 1984:455) that there are three anomalies present within the data: a size effect, P/E effect and a January effect. They also report that there is an interrelationship between the three: size-January and P/E-January. The interrelationships between size-P/E and size-P/E-January are not significant. They further argue that there is strong evidence that the three effects are have an independent influence on returns. The extent of the P/E and size effects depends on the January effect - approximately half of these effects are reported in January (Cook and Rozeff 1984:464).

Cook and Rozeff (1984:464) conclude by stating that neither Reinganum’s (1981) or Basu’s (1983) views are consistent with their own findings. However, they do not reject the contention that both the size and P/E effects are symptoms of the same underlying effect.

4.2.1.2 CALENDAR EFFECTS

Calendar effects can be categorised into the following sub-effects:
- the January effect (turn of year effect);
- the monthly effect (turn of month effect);
- the weekend effect (Monday effect); and,
- the holiday effect.

All of these effects are related to the calendar, some seem to be related to calendar changes i.e. the change from one month to the next. There is no theoretical reason why this should be so. In this section we will review some of the literature pertaining to these effects. In the next section an attempt to relate these effects and the size effects to each other will be made.

---

48 The January effect will be discussed in greater detail below.

49 They argue (Cook and Rozeff 1984:449) that Reinganum’s (1981) result occurs because of a “fortuitous choice of methods” and Basu’s (1993) result is “sample specific”.
The first major investigation into seasonality or calendar effects was undertaken by Rozell and Kinney (1976). They refer to a previous literature, but argue that there is no consensus as to the existence of any seasonal effect (Rozell and Kinney 1976:379). In addition, in that previous literature there is no discussion of the possible impact of seasonality on financial models.

Rozell and Kinney (1976) conduct parametric and non-parametric tests on share price data drawn from the NYSE for the period 1904-1974. They report that a visual inspection of the data reveals that share returns are generated by a mean stationary stochastic process (Rozell and Kinney 1976:383). Due to high variability during the "Great Depression", they divide the sample into five periods: 1904-1928; 1928-1940; 1940-1974; 1904-1928 and 1941-1974; and the entire period.

They calculate summary statistics for each period and report that, "the most consistent feature of these statistics is the high January rate of return" (Rozell and Kinney 1976:387). The high January return is evident in all the statistics in all periods. They then conduct a Kruskal-Wallis test, in order to determine whether or not the monthly returns have identical population distributions. They are able to reject the null hypothesis at the five per cent level for the fourth period (1904-1928 and 1941-1974). This, Rozell and Kinney (1976:390) argue is "convincing evidence for seasonality". They further determine that January is the month that differs from the others. The parametric tests confirm the results of the non-parametric test.

Having established that a January effect does exist, Rozell and Kinney (1976:396-400) consider the implications this has for capital market theory. They argue (Rozell and Kinney 1976:396) that while their result invalidates the pure random walk theory, that it does not invalidate the EMH. The market may be efficient with respect to seasonal information. Of greater interest to us here, is the impact that the January effect has on the CAPM.
Using the data from the Fama and MacBeth (1973) study they replicate that study to determine whether or not the January effect is reflected in risk premia (Rozeff and Kinney 1976:398-400). They find that the risk-return trade-off in January (0.0450) is greater than in other months (0.0056). Thus, not only is there a greater return in January, but there is a greater risk premium in January. Rozeff and Kinney (1976:400-401) state that while it may be possible to rationalise a greater return in January, it is difficult to rationalise a seasonal effect in risk premia.

Tinic and West (1984) also investigate the January effect. While they agree with the Rozeff and Kinney (1976) tests, they argue that the original study was not an adequate reflection of the January effect (Tinic and West 1984:562). They also replicate the Fama and MacBeth (1973) tests and find similar results to that of Rozeff and Kinney (1976). They then run the following regression for the period 1935-1968 and for the sub-periods, 1935-1951 and 1951-1968.

\[ y_{jt} = \beta_1 + \Sigma \beta_i D_i + \epsilon_{jt} \]

where \( y_{jt} \) = an estimate of the intercept in a Fama-MacBeth (1973) type study when \( j = 0 \) and the risk premia when \( j = 1 \); 
\( D_i \) = dummy variables representing February to December; and, 
\( \beta_i \) = mean January intercept or risk premia.

This regression captures the difference in returns that non-January months provide as opposed to those of January. Tinic and West (1984:567) show that for each period under consideration, that the January \( y_{jt} \) is positive, the estimated value for the non-January month \( y_{jt} \) values are all negative. They also update the sample period to 1984, but find that the basic pattern prevails (Tinic and West 1984:569). The conclusion that can be drawn from this paper is that the CAPM is a January phenomena - the return to risk only occurs in January (Thaler 1987a:200). This study brings the entire theory of CAPM into serious question.

Ariel (1987) reports a similar effect. He states (Ariel 1987:161) that stocks only
earn positive returns around the beginning and first half of calendar months. He investigated the CRSP equally and value weighted indices for the period 1963-1981. For the purposes of the study he defined a trading month as the period from the last trading day of the previous month to the second last trading day of the month under consideration (Ariel 1987:162). Ariel (1987:164) reports that the cumulative return for the first half of the month significantly exceeds that of the second half of each month. He further states that this is not purely a January effect as even with January removed from the sample, the anomaly persists (Ariel 1987:171). As Ariel (1987:165) reports that the mean cumulative returns are negative for the second half of the month, it would seem that except for the first half of the month the Dow Jones Industrial Average (DJIA) falls (Thaler 1987b:173).

A similar (and related concept) is the weekend effect, discovered jointly by French (1980) and Gibbons and Hess (1981). French (1980) set out to determine whether or not the stock return generating process operated continuously or only during trading hours. In order to discover which process was at work, the calendar time hypothesis or the trading time hypothesis, he examined the share returns for the five trading days of the week. If the calendar time hypothesis is correct, the return for Monday should be three times as great than for other days (holding stocks over the weekend represents a three day investment). If, on the other hand, the trading time hypothesis is valid then the returns for Mondays should be no different than that of any other day (French 1980:56).

The results of studying the daily returns of the Standard and Poor composite portfolio for the period 1953-1977, were surprising. French (1980:56) reports that the data was inconsistent with both hypotheses. He divided the twenty five year period into five five-year sub-periods and found the same result for each period - the return for Monday was negative (on average) and lower than the average return for any other day (French 1980:57-59).
The obvious explanation to this phenomena would seem to be that the negative effects are related to a "closed market effect". French (1980:61-63) investigated this possibility by investigating the daily returns before and after holidays (where the market is closed) and was unable to find evidence supporting this hypothesis. He reports (French 1980:63) that where a holiday followed a weekend (i.e. the market was closed on Monday) that the return for Tuesday was negative - thus confirming some type of weekend effect. In the same section of his paper, hidden away in a footnote, he also reports during the second half of 1968, that the market was closed on Wednesdays and that the returns for Thursdays were negative during this period. This, French (1980:63) suggests is possible evidence of the weekend effect being related to the market being closed on a regular basis.

Gibbons and Hess (1981) report similar findings using a similar methodology over the period 1962-1978. The explanation that suggests itself to them is a "settlement effect". Gibbons and Hess (1981:588) argue that observed prices are not spot prices but are in fact forward prices, due to the fact that payment is only necessary several days after the transaction has occurred. The relationship between the spot and the forward price is such that the forward contract has a zero value. Thus the spot price needs to be adjusted by the riskless rate of interest. Before 1968 the settlement period was four (business) days, the investor who sold on Monday would be paid on Friday (four days), the investor who sold on Tuesday would be paid on the next Monday i.e. six days after the sale. This Gibbons and Hess (1981:590) argue could be responsible for the weekend effect. The fact that this could be the explanation is strengthened by the fact that the weekend effect is stronger in the pre-1968 period. Due to the fact that the weekend effect persists after 1968 (where the settlement period become five days) this cannot be the entire explanation of the phenomena.

The existence of a weekend effect raises an interesting question: "When exactly does the negative return arise?". Keim and Stambaugh (1984), Rogalski (1984)
and Smirlock and Starks (1986) have addressed this question.

The Keim and Stambaugh (1984) study investigates whether or not the weekend effect is a Friday to Monday effect or whether it is an end of trading week to beginning of trading week phenomena. They are able to investigate this question due to the fact that prior to 1952, the NYSE was open for trading on Saturdays. They investigate the weekend effect pre and post-1952 for the period 1928-1982.

In order to test the hypothesis that the return on Friday before a one day weekend is equal to the Friday return before a two day weekend Keim and Stambaugh (1984:824) estimate the following regression for the period 1941-1956:

\[ r_t = \theta_1 w_{1,t} + \theta_2 w_{2,t} + \theta_3 s_t + u_t \]

where \( r_t \) = Friday's return;
\( \theta_1 \) = dummy variable denoting a one day weekend;
\( \theta_2 \) = dummy variable denoting a two day weekend; and,
\( s_t \) = dummy variable denoting a return that occurs during June-September.

They find that the Friday return is larger if it is the last trading day of the week than if the Friday is the second last trading day \((\theta_2 > \theta_1)\). Keim and Stambaugh (1984:825) repeat this regression but substitute the Monday return for \( r_t \). They report that the extent of the negative return on Monday is independent of the length of the weekend, i.e. the weekend effect occurs between the close of the market on the last trading day and the close on Monday.

Rogalski (1984) goes one step further. He wishes to determine whether the weekend effect occurs during the day of Monday or whether it occurs during the period that the market is closed. Thus Rogalski (1984:1603) decomposes

---

50 This period is chosen because the market was closed on Saturdays for several months (usually June-September) of the year.
the weekend returns (Friday close to Monday close \([R_c]\)) into Friday close to Monday open returns \([R_o]\) and Monday open to Monday close returns \([R_d]\).

Rogalski (1984:1604) decomposes the returns for the DJIA index for the period 1974-1984 and the S&P composite index for the period 1978-1983. The results for each index are very similar. The \(R_c\) is negative, \(R_o\) is negative and the \(R_d\) is positive (Rogalski 1984:1605). Thus Rogalski (1984:1605) reports that the weekend effect arises from the lack of trading between the close on Friday and the open on Monday.

Smirlock and Starks (1986) replicate the Rogalski (1984) study with a much larger sample period (1963-1983) using the DJIA index. They report that the weekend effect has moved in time.

Smirlock and Starks (1986:199-200) divide their sample period into three sub-periods: 1963-1963; 1968-1974 and 1974-1983. They calculate the \(R_c\), \(R_o\), and \(R_d\), for each sub-period and over the whole period. Their results are as follows, (summarised for Monday only) (Smirlock and Stark 1986:201-202):

<table>
<thead>
<tr>
<th>Year</th>
<th>(R_c)</th>
<th>(R_o)</th>
<th>(R_d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963-1983</td>
<td>-.1108</td>
<td>-.0386</td>
<td>-.0739</td>
</tr>
<tr>
<td>1963-1968</td>
<td>-.1549</td>
<td>.0142</td>
<td>-.1692</td>
</tr>
<tr>
<td>1968-1974</td>
<td>-.2152</td>
<td>-.0264</td>
<td>-.1933</td>
</tr>
<tr>
<td>1974-1983</td>
<td>.0153</td>
<td>-.0714</td>
<td>.0551</td>
</tr>
</tbody>
</table>

If we compare the absolute values of these figures we will see that in the post-1974 period the non-trading weekend effect of Rogalski (1984) is valid.

\[51\] Rogalski (1984) does not use this notation. The notation has been taken from Smirlock and Starks (1986:205) which will be discussed below.