THE SENSITIVITY OF THE
EFFICIENT MARKET HYPOTHESIS
TO ALTERNATIVE SPECIFICATIONS
OF THE MARKET MODEL

by

Menachem Brenner
RESEARCH PROGRAM IN FINANCE AT THE
WALTER A. HAAS SCHOOL OF BUSINESS,
UNIVERSITY OF CALIFORNIA, BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH
Carl Shapiro, Director

The Institute of Business and Economic Research is an organized research unit at the University of California, Berkeley, whose mission is to promote research by faculty and graduate students in the fields of business and economics. The Institute carries out its mission by organizing programs and activities that enrich the research environment, administering extramural research awards, publishing working papers, and making direct grants for research.
THE SENSITIVITY OF THE EFFICIENT MARKET HYPOTHESIS TO ALTERNATIVE SPECIFICATIONS OF THE MARKET MODEL

by

Menachem Brenner*

*The Jerusalem Business School, Hebrew University and School of Business, University of California, Berkeley. I wish to thank Marshall Blume, Dave Downes, John McClain, Marshall Sarnat, Bernel Stone, Seymour Smidt and an anonymous referee for their helpful comments and suggestions. This research has been partially supported by a grant from the Lady Davis Foundation.

November 1978
THE SENSITIVITY OF THE EFFICIENT MARKET HYPOTHESIS TO
ALTERNATIVE SPECIFICATIONS OF THE MARKET MODEL

I. INTRODUCTION: EFFICIENT MARKETS AND EXPECTED RETURN MODELS

Tests of the Efficient Market Hypothesis (EMH) are always joint tests of market efficiency, an underlying equilibrium model and a related market model. The possibility that the EMH has not been rejected because the wrong market model was used was never adequately considered.

The main objective of this study is to empirically assess the importance of a correct specification of the market model in tests of the EMH. In other words, this study attempts to test the sensitivity of the conclusions of empirical studies on the efficiency of capital markets when various market models are used. This introduction proceeds with a formal presentation of the EMH and explains why a Capital Asset Pricing (CAP) model is needed in empirical testing.

The Efficient Market Hypothesis

Testing the efficiency of the market requires a formal model that defines more precisely the meaning of the word "efficient." In other words, to test whether all available information is reflected in market prices we require a model that describes how this information is reflected in prices. If we accept the basic assumption that market equilibrium can be stated in terms of expected return, we can formalize market efficiency by the so-called "expected-return model" described in Fama's [6] reply to LeRoy's [13] comment

\[
E^m(P_{t+1} | r^m_t) = [1 + E^m(R_{t+1} | r^m_t)]P_t
\]  

(1.1)
where \( P_{jt} \) is the price of security \( j \), set by the market, at time \( t \);
\( E^m(\widetilde{R}_{jt+1}|I^m_t) \) is the equilibrium expected return on security \( j \) implied by
the market assessed density function for \( P_{t+1} \) (the vector of prices of
securities at time \( t + 1 \)); \( I_t \) represents all the available information at
time \( t \); \( E^m(\widetilde{P}_{jt+1}|I^m_t) \) is the market assessed expected value of the price of
\( j \) at \( t + 1 \).

If \( E(\widetilde{P}_{jt+1}|I_t) \) and \( E(\widetilde{R}_{jt+1}|I_t) \) are, respectively, the true expected
price and return, then market efficiency implies that

\[
E(\widetilde{P}_{jt+1}|I_t) - E^m(\widetilde{P}_{jt+1}|I^m_t) = 0 \quad (I.2)
\]

and

\[
E(\widetilde{R}_{jt+1}|I_t) - E^m(\widetilde{R}_{jt+1}|I^m_t) = 0 \quad (I.3)
\]

This difference can be also stated as

\[
E(\widetilde{z}_{jt+1}|I_t, I^m_t) = E[(\widetilde{R}_{jt+1}|I_t) - E^m(\widetilde{R}_{jt+1}|I^m_t)] = 0 \quad (I.4)
\]

where \( \widetilde{z}_{jt+1} \) is the rate of return deviation. \( (I.4) \) implies that

\[
E(\widetilde{z}_{jt+1}|z_{jt}) = 0 \quad (I.5)
\]

That is, the serial covariance of residual returns is zero.

To test the EMH by \( (I.4) \) and/or \( (I.5) \) requires a model for \( E^m(\widetilde{R}_{jt+1}|I^m_t) \).

For example, the following model, where \( \alpha_{jt+1}, \beta_{jt+1} \) are parameters based
upon \( I^m_t \) and \( R_{mt+1} \) is the value of a market index,

\[
E^m(\widetilde{R}_{jt+1}|I^m_t) = \alpha_{jt+1}(I^m_t) + \beta_{jt+1}(I^m_t)E^m(\widetilde{R}_{mt+1}|I^m_t) \quad (I.6)
\]

does not necessarily correspond to any existing equilibrium theory but can
be used to test the EMH. In general, the more specific and complete the
equilibrium model, the stronger should be the tests that utilize this model.
At the same time, however, a more specific model requires more restrictive underlying assumptions. It may turn out, therefore, that a less theoretically developed but more realistic model provides stronger tests.

Since the theoretical foundations are not uniformly agreed upon and the empirical evidence does not lend overwhelming support to any specific model, which model one chooses to use is partly a matter of taste. The validity of the test, therefore, may depend upon the validity of the model. The question is: To what extent?

Most tests of EMH used a market model that is, implicitly or explicitly, associated with the Sharpe-Lintner (S-L) CAP model (Sharpe [20], Lintner [14]). More recent tests have used a market model linked to the more general derivation of the CAP model suggested by Black [3]. Although the behavior of the residuals seems to be consistent with efficient markets, three alternative interpretations can be provided: (1) The market is efficient with respect to the information set used. (2) The observed "efficiency" is due to an inadequate equilibrium model. (3) The market model is misspecified.

Empirical tests of the EMH have used different market models to study the efficiency of capital markets. Some of these tests, but not all, used market models that are explicitly linked to an existing CAP theory. Employing a certain market model and predictions from its associated CAP model might yield different results, if the market model is misspecified or if the true CAP model is different. If tests based on different market models result in different conclusions about market efficiency, then a correctly specified model is of vital importance. On the other hand, if the conclusions about market efficiency are insensitive to the different market models used, then we may have a robust theory (of efficient markets).
This study employs a sample of stock splits to test the sensitivity of empirical studies on the efficiency of capital markets to various market models. The remainder of this study is organized as follows. Section II presents the market models that are used in this study. Section III presents the hypotheses. Section IV provides the methodology, experimental design, and analysis of results using a stock splits sample. Section V gives a summary and conclusions.

II. MARKET MODELS USED IN EMPIRICAL TESTS OF THE EMH

Tests of the EMH generally proceed in two stages: first, we estimate the relevant parameters using a certain market model; second, we use the estimated parameters for prediction and use the prediction errors, also called "residuals," to test market efficiency. The different market models used in these tests and examined in our study are described next.

Most tests of the EMH (e.g., Scholes [18]) use Sharpe's one-factor market model to describe the process generating returns. The one-factor market model stated by Sharpe [19] and used in a multiperiod context by adding the subscript t is restated as

\[ \tilde{R}_{jt} = \alpha_j + \beta_j \tilde{R}_{mt} + \tilde{u}_{jt} \]  (II.1)

where \( \alpha_j \) and \( \beta_j \) are parameters pertinent to security \( j \) and \( \tilde{u}_{jt} \) is a disturbance term. This market model can be fully consistent with different equilibrium models (e.g., S-L, Black). Miller and Scholes [16] use the slightly modified model, that is explicitly linked to the S-L model,

\[ \tilde{R}_{jt} = R_{ft} (1 - \beta_j) + \beta_j \tilde{R}_{mt} + \tilde{e}_{jt} \]  (II.2)
in replicating Lintner's tests on the S-L CAP model. Kaplan and Roll [12] follow a similar practice to test the EMH via changes in accounting techniques. If, however, there are market factors, other than $R_{ft}$, in addition to $R_{mt}$ that influence the returns on all securities, then those should be used as in

$$\tilde{R}_{jt} = \tilde{\gamma}_{0t} (1 - \tilde{\beta}_j) + \tilde{\beta}_j \tilde{R}_{mt} + \tilde{\nu}_{jt} \quad \text{(II.3)}$$

where $\tilde{\gamma}_{0t}$ in (II.3) plays the role of $\tilde{R}_{zt}$ in Black's model. Ball [1], Jaffe [11], and Mandelker [15] were the first studies that used (II.3) to test the EMH.

In these models systematic risk ($\beta$) is the only measure of risk that affects returns of securities. There are, however, some researchers who hold that in addition to systematic risk, independent risk also affects returns on securities. We thus have

$$\tilde{R}_{jt} = \tilde{\gamma}'_{0t} + \tilde{\gamma}'_{1t} \tilde{\beta}_j + \tilde{\gamma}'_{2t} \tilde{s}_j + \tilde{\nu}_{jt} \quad \text{(II.4)}$$

where $s_j$ is a proxy for unsystematic risk.

Many performance evaluation studies that were merely adjusting for changes in the market have been criticized for not taking into account differential risk. They are simply using the residual ($\tilde{\nu}_{jt}$) from the following market model

$$\tilde{R}_{jt} = \tilde{R}_{mt} + \tilde{\nu}_{jt} \quad \text{(II.5)}$$

This model is a special case of (II.1) where $\alpha_j = 0$ and $\beta_j = 1$. It is
clear that for $\bar{\beta} = 1$, $E(\hat{\gamma}_{jt+1}) = E(\tilde{\epsilon}_{jt+1})$ if (II.2) is the true model. If $\bar{\beta}$ is not equal to 1 then, generally, we will have a downward bias when $\bar{\beta} < 1$ and an upward bias when $\bar{\beta} > 1$.

III. THE MAIN HYPOTHESIS: THE ROBUSTNESS OF THE TESTS ON THE EMH

In this section we present the main issue: how sensitive are tests of the EMH to alternative specifications of the market model? Since the correct market model is unknown we cannot separate truly inefficient market behavior from an observed indication of inefficiency due to biases. One may assume the market to be efficient and choose the model or models that coincide with it to be the correct one. Alternatively, one may assume a particular market model is true and test the efficiency of the market using this model. But it is possible that none of the existing market models are the correct ones for all companies for all periods.

Therefore,

the sensitivity of tests of the EMH is basically an empirical question. If we find that, using alternative specifications, we cannot reject the EMH, we may conclude that the results do not depend on the model used. If we find that using a certain market model we reject the EMH we may claim that this is due to a bias introduced by not using the correct model. The question of market efficiency will then remain unanswered.

The "Robustness" Hypothesis

For a long time, tests on the EMH used the one-factor market model, while more recent studies use the two-factor market model. Given the evidence in favor of the EMH on one hand, and some evidence against the S-L model on the
other hand, the question raised here is: are the tests and conclusions regarding the EMH dependent on the particular market model used? In other words, will the conclusions on the validity of the EMH be robust to different specifications of the market model?

The question can be formulated and tested in two ways. 1) The first test applies every model, separately, to the data and tests whether

$$E(\tilde{\eta}_m) = 0 \quad \text{for} \quad m = 1, \ldots, 5$$  \hspace{1cm} (III.1)

where \( \tilde{\eta}_m \) is the residual from model \( m \), identified above as II.1 to II.5, in the test period. 2) The second test is a pairwise comparison of the form

$$E(\tilde{\eta}_m - \tilde{\eta}_n) = 0 \quad \text{for} \quad m \neq n$$  \hspace{1cm} (III.2)

where \((m,n)\) represent any pair of models.

The interpretation of these two sets of tests goes as follows: the null hypothesis states that all models support the EMH and if validated we could say that the evidence for the EMH is not conditional on the underlying model. If, however, the conclusions vary from one model to the other, then questions are raised as to the validity of the existing evidence on the EMH, since that evidence is certainly conditional on what one believes is the true model.
The argument can also be reversed as follows: if one strongly believes that markets are efficient, then one can conclude that the true model is the one that supports the EMH.

The interpretation of the second test is related to the first but is less obvious. Even if all the models in the first test support the EMH, the test in (III.2) might result in the conclusion that the different models produce significantly different residuals. This, however, is even stronger evidence as to the robustness of the tests of the EMH. On the other hand, validating (III.2) just means that the models are not really different.

It might be argued that, since the various market models are not very different, we should observe very similar results using these models. To see, therefore, that the answer to the question presented in the beginning of this section is not straightforward, we can consider an example which, although it is not very appropriate in this context, is convenient for descriptive purposes.

Assume an information generating event occurs at time t, and assume for a moment that $\beta_j$ is given and that $R_{mt}$ and $R_{zt}$ are market factors that are not constructed from $R_{jt}$. Assume that this event is pertinent to high $\beta$ stocks so that all our points $(R_{jt}, \beta_j)$ are to the right of
point M like the stars in the figure. If we use \( R_{FM} \) as our market line then the average residual across all firms \( \hat{u}_t = 0 \), while using \( R_{2M} \) as the reference market line will result in \( \tilde{u}_t > 0 \). Even if the entire range of \( \beta_j \) is represented, as in the figure above, it is clear that not only will the variation of \( \hat{u}_t \) around \( R_{2M} \) be much larger than the variation of \( \hat{u}_t \) around \( R_{FM} \), but that \( \hat{u}_{tz} \) (\( \hat{u}_t \) around \( R_{2M} \)) will also be correlated with \( \beta_j \). This may lead to wrong conclusions regarding market efficiency.

IV. EMPIRICAL TESTS BASED ON A SAMPLE OF SPLITS

A good way to test the differences among the models is to apply them to an event that conveys new information and test the difference in the residuals that are obtained from the application of the different models to the same event.

The information generating event chosen here is stock splits. For the purpose of this study, out of all available events, stock splits seemed the best one to use. The main reason is the large number of large splits, the large number of companies involved, and the distribution of the event over time. Stock splits seem to suffer less from any possible distortions in some fundamental structure of the company involved. Finally, stock splits are the first event studied extensively, by Fama et al. [8], that offered the "residual" methodology as a better way to study the EMH. Using the "simple" market model (Sharpe's) they conclude:

Thus the results of the study lend considerable support to the conclusion that the stock market is "efficient" in the sense that stock prices adjust very rapidly to new information [p. 20].

Given the theoretical and empirical support for the multifactor market models the question is whether the above conclusion still holds if another market model is employed.\(^5\)
A. Stock Splits: Experimental Design

All splits (including stock dividends) that increased the number of shares outstanding by 20 percent or more and that occurred in the period January 1935–June 1970 made up the initial sample. The period covered here is January 1935–June 1968 since estimates of the market factors \( \tilde{\gamma}_0, \tilde{\gamma}_0', \tilde{\gamma}_2' \) obtained from Fama and MacBeth were available for that period only. Since at least 12 months of data were required on each side of the split, only splits in the period January 1936–June 1967 were chosen. The number of splits in that period is 1,717 and involves 889 companies.

We then applied a minimum data requirement; a minimum of 61 months in total, where at least 24 months were available for parameter estimation and the remaining 37 months of the period around the split were available to compute residuals. The following example shows when the minimum was applied.

```
residual period = 37  estimation period = 24

no data 1  no data
-12 0 +24 +48
```

The maximum number of observations used is 120 months around the split date, where 48 months are used for residuals and 72 for estimation.

We then looked for adjacent splits. For many companies several splits cluster in time such that an observed gradual rise in price after a split is nothing more than the pre-split behavior of the next split. Including the first split may show an increase in the average residual and appear to justify
the conclusion that the market is inefficient. This conclusion is incorrect, however, if people cannot predict the occurrence of the next split. If, given a split, the next one can be predicted then the pre-split behavior of the next split should look like a post-split behavior and, excluding the split might bias the results in favor of EMH. Splits occurring too close in time would also create a problem in the estimation of the parameters. It is desirable to have the parameters estimated very close to the split, excluding the period analyzed, but the occurrence of another split close to the previous one prevents us from using that period for estimation. Since there is no indication that, given a split, subsequent splits can be predicted and since the exclusion of very close splits will provide better estimates of the company parameters we decided to exclude some splits very close in time. Excluding a nonrandom set of splits may introduce some other biases that may affect the conclusion about the EMH. Since, however, all our models use essentially the same data there is no reason to believe that the sampling procedures used will affect the tests that study the difference among the models. Applying these criteria, the final sample consisted of 1,188 splits and involved 830 companies. Out of 830 companies, 63 percent had one split and 37 percent had two splits or more.

The month of split, month zero, is the month containing the day on which the split is effective (not the announcement day) since these dates are the only ones available for such a large sample. As will be seen later, the market levels off about two months before the effective date and thus the difference is not of much concern.
B. Models and Residuals: Empirical Estimation

1. Parameter estimation

To estimate company parameters \((a_j, \beta_j, s_j)\) associated with splits we use the regression

\[ R_{jt} = \hat{a}_j + \hat{\beta}_j R_{mt} + u_{jt}. \quad \text{(IV.1)} \]

If \(\text{cov}(\tilde{u}_{jt}, \tilde{R}_{mt}) = 0\) during the period studied, then \(\hat{\beta}_j\) will be an unbiased estimate of \(\beta_j\), regardless of the true underlying model [i.e., (II.1) or (II.2) or (II.3)]. From (IV.1) we also obtain the proxy of independent risk \(s(\hat{u}_j)\) which is the standard error of estimate in (IV.1).

If, however, the correct model is not model (II.1) but (II.2) or (II.3) and \(\text{cov}(\tilde{u}_{jt}, \tilde{R}_{mt}) \neq 0\) we should use a regression on (II.2) or (II.3) in estimating \(\hat{\beta}_j\). In so doing we found the actual values of \(\hat{\beta}_j\) estimated from models (II.1), (II.2) and (II.3) to be so close that the residuals based on \(\hat{\beta}_j\) from (II.1) could not be distinguished from those based on the other estimates of \(\beta_j\). Since the results are not sensitive to the different estimates of \(\beta_j\), for clarity and brevity we will proceed with the analysis of the results using the \(\hat{\beta}_j\) obtained from (IV.1), the regression on (II.1), in all models.

2. Residual estimation

The Residual Period (RP) generally contains 49 months around the month of split (including the month of split). The observations in the RP are independent of the observations in the Estimation Period (EP) (the two periods are totally non-overlapping). Using the parameters obtained from (IV.1) and models (II.1) - (II.5) the following residuals in the RP were computed.

For each of these five residuals a matrix of 49 x 1188 is obtained.

\[ \hat{u}_{jt} = R_{jt} - \hat{a}_j - \hat{\beta}_j R_{mt} \]
\[ \hat{e}_{jt} = R_{jt} - R_{ft} (1 - \hat{\beta}_j) - \hat{\beta}_j R_{mt} \]
\[ \hat{\nu}_{jt} = R_{jt} - \hat{\gamma}_t (1 - \hat{\beta}_j) - \hat{\beta}_j R_{mt} \]
\[ \hat{\nu}_{jt} = R_{jt} - \hat{\gamma}_t (1 - \hat{\beta}_j) - \hat{\beta}_j R_{mt} \]
\[ - \hat{\gamma}_{zt} [ s_j (\hat{u}_j) - \hat{\beta}_j s (\hat{u}_j) ] \]
\[ \hat{\nu}_{jt} = R_{jt} - R_{mt}. \]

In estimating \( \hat{\nu}_{jt} \), from model (II.4), we have rewritten (II.4) in terms of \( R_{mt} \). It should also be noticed that \( j \) refers to the number of splits and not to the number of companies. \( \hat{\beta}_j \) is pertinent to the split and not to the company. Thus, to some extent, we account for the possibility of a change in \( \hat{\beta}_j \) if it did occur.

C. The "Robustness" of Tests of the EMH

In the first set of tests we use separately the residuals, from each model, to test the EMH. The null hypothesis, stated in Section III, is tested by comparing the conclusions, with regard to the EMH, derived from the application of the above models.

1. The test statistic CAR

Most studies of the EMH use the Cumulative Average Residual (CAR) to test market efficiency. The CAR is computed in two steps; first we compute the Average Residual (AR) for month \( t \) across all splits. For example,

\[ AR_t(u) \equiv \hat{u}_t = \frac{1}{n} \sum_{j=1}^{n} \hat{\nu}_{jt}. \tag{IV.2} \]

where the letter in the brackets of \( AR_t( ) \) refers to the model used. Although the efficiency of the market can generally be deduced from the \( AR_t \)'s we usually test the EMH using the CAR that is computed as:

\[ CAR_T(u) \equiv \bar{AR}_T(u) = \sum_{t=1}^{T} \hat{u}_t. \tag{IV.3} \]

At every \( T \) we sum all the \( AR_t \)'s from 1 to \( T \).
The behavior of $\text{CAR}_T$ of all five models is presented in Figure 1. If we examine every model separately, we see that the steady increase in $\text{CAR}_T$, starting at month -24, stops at month 0 (month of split) and approximately the same $\text{CAR}_T$ is observed from there on. However, to compare the five models more thoroughly, we present in Table 1 the $\text{CAR}_T$'s for three selective months for models 1 to 5, which correspond to (II.1) to (II.5).

<table>
<thead>
<tr>
<th>Month</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0)</td>
<td>30.1</td>
<td>33.1</td>
<td>30.8</td>
<td>29.9</td>
<td>28.4</td>
</tr>
<tr>
<td>(+12)</td>
<td>30.0</td>
<td>34.4</td>
<td>32.0</td>
<td>30.6</td>
<td>28.4</td>
</tr>
<tr>
<td>(+24)</td>
<td>27.7</td>
<td>33.4</td>
<td>31.0</td>
<td>29.4</td>
<td>26.6</td>
</tr>
<tr>
<td>$\Delta_1 = (+12)-(0)$</td>
<td>-0.1</td>
<td>+1.3</td>
<td>+1.2</td>
<td>+0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>$\Delta_2 = (+24)-(+12)$</td>
<td>-2.3</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-1.2</td>
<td>-1.8</td>
</tr>
<tr>
<td>$\Delta_3 = (+24)-(0)$</td>
<td>-2.4</td>
<td>+0.3</td>
<td>+0.2</td>
<td>-0.5</td>
<td>-1.8</td>
</tr>
</tbody>
</table>

Using model 1 we observe a drop in $\text{CAR}_T(u)$ of about 2.3 percent in the last 12 months and this drop is significantly different from zero. Though statistically significant, it may not be economically significant. Provided the model is correct, a trader who does not pay any transaction costs will make excess returns of 2.3 percent. With transaction costs, however, the drop of 2.3 percent may be economically insignificant. Conclusions about the EMH will depend on how one interprets the 2.3 percent drop.

If we accept the interpretation that a drop of 2.3 percent over 12 months is not indicative of an inefficient market, then none of the models
used can reject the EMH (the next largest drop is 1.8 percent in model 5). Thus, in spite of the differences among the models, the evidence in favor of the EMH is insensitive to the market model employed. If, however, a 2.3 percent drop is considered large and we use statistical significance tests, then models 1 and 5 are different from models 2, 3 and 4. While \( \Delta_2 \) (the change in the CAR in the second year after split) is significantly different from zero in models 1 and 5, it is not significant in models 2, 3 and 4. Thus, if models 1 or 5 are assumed we should reject the EMH while we cannot reject the hypothesis using models 2, 3 or 4. If we assume the market is efficient, then models 1 and 5 should be rejected as possible generating processes. Though consistent with the EMH, models 2, 3 and 4 exhibit a positive \( \Delta_1 \) and a negative \( \Delta_2 \) which may be interpreted as results of frictions in the market (e.g., over-reaction to new information).

In spite of the significant drop of 1.8 percent in \( \text{CAR}_T(w) \) (model 5), its performance is quite surprising (it "beats" model 1 and is quite close to the other models). It may lead to the conclusion that, as far as tests of the EMH are concerned, there is no need to take into account differential risk as well as other suggested parameters.

2. The Mean Square Error (MSE) Criterion

If predictability is a test criterion that should be used to distinguish among models, then on the average, predicting \( R_{jt} \) conditional on \( R_{mt} \) (model 5) gives smaller values of the MSE than other models, except the zero-beta model. Using the Mean Square Prediction Error (MSE) as a test statistic for predictability we computed for each month for each model

\[
\hat{\mathbf{\eta}}_{jt}^2 = \frac{1}{n} \sum_{j=1}^{n} \hat{\mathbf{\eta}}_{jt}^2. \quad (IV.4)
\]
The MSE, for all models, increases as we approach month 0, decreases after month 0 and remains at some lower level thereafter. To compare the overall MSE we ranked all models at each month, from 1 to 5 (1 being the lowest rank). The average rank, across all months is given by

$$\bar{RK}(i) = 1/49 \sum_{t=1}^{49} RK_t(i) \quad \text{(IV.5)}$$

where $RK_t(i)$ is the rank of a residual from model $i$ in month $t$. The following statistics were obtained.

| TABLE 2
AVERAGE RANK OF MSE FOR FIVE MODELS |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RK(1)</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>3.78</td>
</tr>
</tbody>
</table>

The model that accounts for total risk (model 4) ranks last and Sharpe's market model (model 1) is almost as bad. Model 3 (the zero-beta model) ranks first\(^{18}\) while the model that just accounts for the market (model 5), and does not account for differential risks, ranks second. This average ranking represents also a very consistent month by month ranking (i.e., the standard deviation of the ranks around the mean rank is very small).

To summarize this section, we have tested the sentitivity hypothesis using the CAR's of each model, separately, to test the EMH. The results of this section give no clear-cut answer. The answer depends on how economic efficiency is defined.

The above tests and conclusion are based mainly on the CAR. The CAR is an average across securities as well as time and thus may be eliminating
existing differences. As previously mentioned, it is possible that the distribution of $\beta_j$ is not far enough from 1.0 for the differences to show up in the CAR. Therefore, we performed some additional tests that deal with the following questions. Is there any difference among the individual residuals month by month and security by security? Will an examination of individual securities after the split date provide more information on the EMH and on the differences between models than the CAR?

D. Tests on the Differences Between the Models

In this section we use some parametric tests on the differences between two sample distributions. We test whether some basic statistics that characterize distributions (e.g., mean and variance) are significantly different from each other. The tests assume that the residuals are drawn from a normal distribution.

The following three statistics are used to test the pairwise differences:

1. $t(D)$ is the $t$-value of the mean difference between two residuals.
   This is actually a test on the differences between means.

2. $\hat{\rho}(\eta_m, \eta_n)$ is the estimated correlation between any two sets of residuals.

3. $P(\eta_m, \eta_n)$ is the Pitman statistic\(^{19}\) that tests whether two dependent variances belong to the same population.

The three statistics are computed as
\begin{equation}
t(D) = \frac{\bar{\eta} \, \hat{\rho}(\eta_m, \eta_n)}{\sqrt{n}} \frac{\hat{\rho}(\eta_m, \eta_n)}{\hat{\sigma}(\eta_m) \hat{\sigma}(\eta_n)}
\end{equation}

Using all splits in each month we computed these three statistics for
the 24 months after split, for eight pairs of models (those pairs that are
most likely to be different). Table 3 presents the average absolute \( t(D) \),
the average absolute \( \hat{P} \) and the average \( \hat{\rho} \) over all 24 months.

The significant differences between the means of Sharpe's model (1) and
the "Rt" model (2) and the "\( \gamma_0 \)" (3) model, as indicated by the t-statistic,
and the significant difference between the variances, as indicated by the
Fitman statistic, are consistent with the previous finding. The insignificant
differences between models 1 and 5 are also consistent with the previous anal-
ysis.

E. Testing for Serial Independence\(^{20}\) in the Residuals

Since the standard CAR test of the EMH deals with averages and may
contain some evidence about market inefficiency at the individual security
level, we have used a serial correlation test on each security for the 24
months after the announcement. For each split for which we had data 24 months
after the announcement we computed the serial correlation of the residuals.
For each split we computed the two following statistics: \( \hat{r}_0(t, t-1), \)
\( \hat{r}_M(t, t-1) \). The first estimates serial correlation assuming a zero mean while the second estimates serial correlation using the sample mean. Using the statistics \((\hat{r}_0, \hat{r}_m)\) in Table 4 we cannot reject the hypothesis that the serial correlation in all models is the same. The fact that \( \hat{r}_M(t, t-1) \) is significantly different from zero does not concern us since the magnitude of the serial correlation is too small to warrant further investigation. In fact, the negative serial correlation of these magnitudes is in line with many studies on the serial independence of returns or residual returns.

V. SUMMARY AND CONCLUSIONS BASED ON STOCK SPLITS

The results obtained in this study demonstrate that when data on stock splits are used to test market efficiency and statistical significance is used as the test criterion, different market models lead to different conclusions. If "practical" differences are required we may, depending on what is considered "practical," conclude that the different models do not lead to different conclusions. The results also question the validity of the models used here and seem to suggest that all models may be grossly misspecified.

To see whether these results are general or pertinent to events similar to stock splits we discuss the advantages and disadvantages of using stock splits in our study. If stock splits convey new information (about prospects), then we are directly testing different conclusions about market efficiency. If stock splits do not convey new information, then we are only testing differences between models.\(^{21}\) Another advantage (or rather disadvantage) is the fact that stock splits are a fairly "clean" event. Stock splits (compared to other events) seem to suffer less from possible distortions
in different company parameters. On the one hand this is an advantage because it does not require us to study the possible effect of simultaneous changes on differences between market models. On the other hand, relatively small differences observed here may be a result of this characteristic of stock splits. In more complicated events, the differences may be large enough to make definite conclusions about market efficiency dependent upon the market model that is used. For example, in a study on tender offers Smiley [21] uses models 1, 2, and 3 in estimating transaction costs and finds that the different models give substantially different results.

We conclude this study with a general recommendation for future studies testing the EMH. In tests of the EMH, different market models should be used. The differences should be analyzed to see whether the results for the EMH are conditional upon the validity of the market model employed.
APPENDIX

A Significance Test

Though we prefer not to base our analysis on conventional normality tests, we provide some statistics that may give us some feeling about the possibility that the observed phenomenon could have been obtained by chance. Assuming the residuals for each model are distributed normally (under the null hypothesis that the model is correctly specified) we compute for every relative month the cross-sectional standard deviation divided by \( \sqrt{n} \) (\( n \) is the number of splits in any month before or after the split). This standard deviation is given by

\[
SD_t(\eta) = \sqrt{\frac{n}{n(n - 1)} \sum_{i=1}^{n} (\hat{\eta}_{jt} - \bar{\eta}_t)^2}
\]  

(A.1)

where \( \hat{\eta}_{jt} \) represent a residual from any model. This statistic can be used to test whether the average residual \( AR_t(\eta) = \bar{\eta}_t \) is significantly different from zero. To test \( CAR_T \) we compute the SD for the cumulative average residual by adding the monthly cross sectional variances over \( t \) and divide by \( n \). The cumulative SD is computed as

\[
CSD_T(\eta) = \sqrt{\frac{\sum_{t=1}^{T} \sum_{j=1}^{n} (\hat{\eta}_{jt} - \bar{\eta}_t)^2}{(n - 1)n}}
\]  

(A.2)

For example, to find the confidence interval for the drop of 1.8 percent (model 5) in the second year after split (see Table 1) we add the cross sectional variances of the last 12 months and divide by \( n \) (number of splits).
The square root gives us the \( \hat{CSD}_{T-37}(\eta) \). The confidence interval (using a 5 percent level) for this 1.8 percent drop is given by

\[
\Delta_2(\eta) \pm 2 \cdot \hat{CSD}_{49-37}(\eta) = -1.8 \pm 1.36 \tag{A.3}
\]

where \( \Delta_2 = \text{CAR}_{49-37}(\eta) = \text{CAR}_{49}(\eta) - \text{CAR}_{37}(\eta) \). Table 2 presents the \( \hat{CSD}_T \) for all five models for three selective \( \Delta \)'s (the same \( \Delta \)'s given in Table 1).

**TABLE A**

**STANDARD DEVIATIONS FOR SUMS OF AVERAGE RESIDUALS FOR FIVE MODELS**

<table>
<thead>
<tr>
<th>Month</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta_1 = (+12)-(0) )</td>
<td>0.0069</td>
<td>0.0072</td>
<td>0.0071</td>
<td>0.0073</td>
<td>0.0071</td>
</tr>
<tr>
<td>( \Delta_2 = (+24)-(12) )</td>
<td>0.0069</td>
<td>0.0069</td>
<td>0.0068</td>
<td>0.0070</td>
<td>0.0068</td>
</tr>
<tr>
<td>( \Delta_3 = (+24)-(0) )</td>
<td>0.0097</td>
<td>0.0099</td>
<td>0.0098</td>
<td>0.0101</td>
<td>0.0098</td>
</tr>
</tbody>
</table>
FOOTNOTES

1 It should be noted that tests of the EMH are also "weak" tests since the null hypothesis has always been that the market is efficient and no alternative of inefficiency has ever been specified.

2 For a more detailed presentation, see Fama [5].

3 The first studies (e.g., Benston [2]; Fama, Fisher, Jensen, and Roll (FFJR) [8]; Scholes [18]) just use a market model with no reference to any CAP model.

4 For example, Jaffe [10], Mandelker [15], Smiley [22].

5 It should be emphasized that we are not studying stock splits per se. Stock splits are only used to study the difference between models. Nevertheless, we tried to improve upon their experimental design in two points that seemed important.

6 Although we used only splits in the period January 1936 to June 1967, we needed the entire period January 1935 to June 1970 to check for successive splits.

7 The estimates of the market factors $\tilde{\gamma}_0t$, $\tilde{\gamma}_1t$, $\tilde{\gamma}_2t$ and others were kindly provided to me by Eugene Fama and Jim MacBeth.

8 When the splits were more than 48 months apart we divided the period between the splits in half. When the splits were less than 48 months apart we have dropped the second split from the sample and treated the first one as any other split.

9 The drop in 59 companies, from 889 in the initial sample to 830 in the final sample, is mainly due to the minimum data requirement.
In a sample of 236 splits for which we had both dates, we found an average difference between announcement and effective dates of 55 days. Fama et al. [8] reported a similar number contained in a study by Jaffe [11].

It is possible that inefficiencies do exist in the very short run and using monthly data provides only a weak test of market efficiency. It is interesting to note, however, that Housman et al. [9] could not reject the EMH using daily data.

The effects of model misspecification on parameter estimates are treated in Roll [17] and in Brenner [4].

The detailed results, for all models, using the same \( \hat{\beta}_j \) (IV.1) as well as other \( \hat{\beta}_j \) are available upon request from the author.

Out of 1,188, only 96 splits had less than 49 months in the residual period. The number of months for these splits ranged from 48 down to a minimum of 37.

This equation is analogous to the equation for the four-variable case suggested by Fama and MacBeth [7], p. 634.

The behavior of \( \text{CAR}_t(u) \) in Figure 1 is quite similar to Figure 2a in Fama et al. [8] despite some major differences in sample composition, time periods, and experimental design.

The significance tests are discussed and presented in the appendix. Due to some objections to significance tests based on normality, however, we should be cautious in drawing conclusions whenever we use these tests.
The performance of model 3 (the "γ₀" model) relative to the other models here may be attributed to the fact that the estimate \( \hat{γ}_{0t} \) is obtained from the same data (at least some of it).

See Snedecor and Cochran [23], p. 196.

Strictly speaking, the test is for lack of serial correlation and not independence.

Testing the differences may help us in choosing a model for tests of the EMH.

The problem of interpreting significance tests is one-sided; if the hypothesis is rejected under normality, it still may hold under a stable distribution. If, however, the hypothesis cannot be rejected under normality, then we cannot reject it assuming a stable distribution.

It is claimed that residuals from different securities at the same calendar month are dependent and we should, therefore, use portfolio residuals (see Jaffe [10]). This may be a problem in studies that cluster in time where for many securities each relative month is also the same calendar month. With stock splits, this problem is negligible.
REFERENCES


### Table 3
Three Test Statistics for Difference between Models Using 24 Months

|       | $|\tilde{t}(D)|$ | $\tilde{p}(\eta_m, \eta_n)$ | $|\tilde{F}(\eta_m, \eta_n)|$ |
|-------|-----------------|-----------------|-----------------|
| 1-5   | 0.844           | .948            | .027            |
| 1-2   | 4.285**         | .991            | .060*           |
| 1-3   | 2.125*          | .962            | .066*           |
| 1-4   | 1.117           | .933            | .064*           |
| 4-5   | 0.890           | .915            | .079*           |
| 5-2   | 1.480           | .956            | .003            |
| 5-3   | 1.400           | .949            | .029            |
| 2-3   | 0.833           | .970            | .034            |

* Significant at the 5% level.

** Significant at the 1% level.
<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{r}_m(t,t-1)$</td>
<td>-.112</td>
<td>-.112</td>
<td>-.123</td>
<td>-.121</td>
<td>-.110</td>
</tr>
<tr>
<td></td>
<td>(.191)</td>
<td>(.191)</td>
<td>(.194)</td>
<td>(.196)</td>
<td>(.194)</td>
</tr>
<tr>
<td>$\hat{r}_o(t,t-1)$</td>
<td>-.052</td>
<td>-.069</td>
<td>-.081</td>
<td>-.078</td>
<td>-.068</td>
</tr>
<tr>
<td></td>
<td>(.201)</td>
<td>(.199)</td>
<td>(.198)</td>
<td>(.201)</td>
<td>(.201)</td>
</tr>
</tbody>
</table>

*Numbers in parentheses are standard deviations.*