Stock Return Anomalies and the Tests of the APT

Author(s): Mustafa N. Gultekin and N. Bulent Gultekin

Source: *The Journal of Finance*, Vol. 42, No. 5 (Dec., 1987), pp. 1213-1224
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328523
Accessed: 26/11/2009 08:13

THE JOURNAL OF FINANCE • VOL. XLII, NO. 5 • DECEMBER 1987

Stock Return Anomalies and the Tests of the APT

MUSTAFA N. GULTEKIN and N. BULENT GULTEKIN*

ABSTRACT

This paper shows that the empirical tests of the Arbitrage Pricing Theory (APT) model are very sensitive to the anomalies observed in January in the stock returns data. There is a strong seasonal pattern in the estimates of the risk premia from the APT model. The most important implication of the findings in this paper is that the APT model can explain the risk-return relation mostly for January. Once the January returns are excluded from the data, there is no significant relation between the expected stock returns and the risk measures predicted by the APT model.

IT IS NOW A very well-known phenomenon that stock returns exhibit a peculiar seasonal pattern. While such a seasonal pattern in stock returns is an interesting phenomenon in itself and has yet to be fully explained, there is a serious implication of the observed seasonality in stock returns on the tests of the asset-pricing models.[1]

Tinic and West [17], for example, report that the risk-return relation described by the Capital Asset Pricing Model (CAPM) also exhibits a similar seasonality in January, a result Rozeff and Kinney [16] report as well. More importantly, however, they show that only one month, January, shows a consistently positive and reliable relation between expected return and risk. Once the data for January are excluded from the analysis of the risk-return tradeoff, the estimates of risk premia are no longer reliably different from zero. Although Tinic and West do not offer an explanation for their findings, their results raise serious questions about the empirical tests of the CAPM and ultimately its validity as a viable model to explain the pricing of risky assets.

The purpose of this paper is to investigate the impact of stock return seasonality on the empirical tests of the Arbitrage Pricing Theory (APT). A number of researchers argue on both theoretical and empirical grounds that the APT is a good alternative to the CAPM. Although there is a growing body of empirical literature on the APT, the empirical tests of the APT thus far have produced inconclusive results and a good deal of controversy. It now seems appropriate to

* M. N. Gultekin is from The University of North Carolina at Chapel Hill; N. B. Gultekin is from The Central Bank of the Republic of Turkey and The Wharton School of the University of PennysIvania. The views expressed here are solely those of the authors and do not necessarily represent the views of the Central Bank. We thank Craig MacKinlay and Mark Grinblatt for their helpful comments. We also benefited greatly from the participants at the Conference on the Empirical Tests of the Arbitrage Pricing Theory held at the University of Southern California during November 21 to 22, 1985, at The University of North Carolina and Duke University Finance Seminar, and at the European Finance Association Meetings in Bern, Switzerland, August 28 to 31, 1985.

[1] See, for example, Wachtel [19], Officer [11], Banz [1], Reinganum [12], Keim [9], Roll [13], and Gultekin and Gultekin [7], among others.

investigate to what extent the stock return anomalies affect the previous tests of the APT.

Our paper is organized in the following order: Section I reviews the APT and produces empirical tests to be used as a benchmark for comparing the findings of this study; Section II investigates the impact of seasonality on the tests of the APT; Section III concludes the paper.

I. The APT Model and Review of Earlier Empirical Tests

This section provides a summary of the APT model and its empirical implications. Further details of the tests we use in this paper can be found in Roll and Ross (RR) [14], Dhrymes, Friend, and Gultekin (DFG) [3], Dhrymes, Friend, Gultekin, and Gultekin (DFGG) [4, 5], and Gultekin and Rogalski [8], among others.

The APT of Ross [15] assumes the following return-generating process:

$$r_{t.} = E_{t.} + f_{t.}B + u_{t.}, \tag{1}$$

where $r_{t.}$ and $E_{t.}$ are m-element row vectors containing the observed and the expected returns, respectively, for the m securities at time t. $f_{t.}$ is a k-element row vector of common but unobservable factors, assumed to have zero mean and unit variance. B is a $k \times m$ matrix of risk measures in terms of the sensitivity of securities to the common factors. $u_{t.}$ is the idiosyncratic error vector.[2]

Ross shows that, if the number of securities (m) is sufficiently large, there exists a $(k + 1)$-element row vector $c_{t.}$ such that

$$E_{t.} \approx c_{t.}B^*, \quad t = 1, 2, \cdots, T, \tag{2}$$

where $B^{*\prime} = [e{:}B']$ and e is an m-element column of ones. The result in (2) characterizes the no-arbitrage condition of the APT model, and, thus, (1) can be rewritten with any desired degree of approximation:

$$r_{t.} = c_{t.}B^* + f_{t.}B + u_{t.}, \quad t = 1, 2, \cdots, T. \tag{3}$$

Based on the specification in (3), the empirical tests of the APT model are based upon a two-stage factor-analytic approach. In the first stage, one determines the number of factors (k) and estimates the elements of B, say, by $\tilde{B}$.[3] In the second stage, using $\tilde{B}$ as "independent variables," one estimates a vector $c_{t.}$, the elements of which have the interpretation that c^*_{ti} is the "risk premium"

[2] Note that we assume that $\{u'_{t.}: t = 1, 2, \cdots\}$ is a sequence of identically independently distributed (i.i.d.) random vectors with $\mathrm{E}(u'_{t.}) = \mathrm{cov}(u'_{t.}, f'_{t.}) = 0$ and $\mathrm{cov}(u'_{t.}) = \Omega$, where Ω is a diagonal matrix. Regarding the common factors, we specify that $\mathrm{E}(f'_{t.}) = 0$ and $\mathrm{cov}(f'_{t.}) = I$. It is a consequence of these assertions that $\{(r_{t.} - E_{t.})': t = 1, 2, \cdots\}$ is a sequence of i.i.d. random vectors with $\mathrm{E}[(r_{t.} - E_{t.})'] = 0$ and $\mathrm{cov}[(r_{t.} - E_{t.})] = B'B + \Omega = \Psi$.

[3] In these tests, it is assumed that the sample size T (the number of trading days) is sufficiently large so that one estimates B by $\tilde{B}$ with negligible sampling error. This is clearly a strong assumption, and little is known about the sampling properties of factor loadings. DFG [3] and DFGG [4] document the effect of changing the sample size on the estimates of factor loadings. In this paper, however, our objective is not to provide conclusive tests of the APT per se but to demonstrate the sensitivity of previous results to the stock return anomalies.

attached to the ith factor, $i = 1, 2, \cdots, k$, while c_{t0} is the risk-free rate (or possibly the return on a zero-beta asset). The regression model used to estimate the vector $c_{t.}$ in the second stage for each t is the following generalized least-squares (GLS) model:

$$\tilde{c}'_{t.} = (\tilde{B}^* \tilde{\Psi}^{-1} \tilde{B}^{*'})^{-1} B^* \tilde{\Psi}^{-1} r'_{t.}, \quad t = 1, 2, \cdots, T. \tag{4}$$

Analogous to the CAPM tests, as in Fama and MacBeth [6], there are two critical testable hypotheses implied by the APT model. First, the intercept term in (4) is the risk-free (or zero-beta) rate. Second, there is a linear relation between the risk measures, B, and the expected returns.

In addition to the two hypotheses above, one could also test the APT model against a specific alternative by introducing a relevant explanatory variable in the model shown in (3). The respecified model in this context is

$$r_{t.} = c_{t.} B^* + d_{t.} P + f_{t.} B + u_{t.}, \tag{5}$$

where P is a vector of other "extraneous" variable(s). Equation (3) implies that these extraneous variables should not be related to expected rates of return. A commonly used variable for this purpose is either the total or residual variance because the central theme in asset-pricing theories asserts that variance is diversifiable and thus should not be priced.

The data used in this paper are daily stock returns from the Center for Research in Security Prices (CRSP) from July 3, 1962, to December 31, 1981. Securities with more than one hundred missing observations are deleted. This results in nine hundred stocks listed on the New York and the American Stock Exchanges, providing a sample size ranging from 4,793 to 4,893 daily observations of daily returns per security. These securities are ranked alphabetically and then formed into thirty and ten groups, each containing thirty and ninety securities, respectively. Using individual stocks, tests are replicated across each group.[4]

The significance tests of risk premia from the second stage are presented in Table I for the entire period 1962 through 1981.[5] Part A shows the chi-square test statistics for the vector of risk premia for groups of thirty stocks each.[6] Row numbers correspond to the thirty groups, each arranged alphabetically. Column (1) of Part A shows that, over the entire period, the risk-premium vector is significantly different from zero for only four groups (13.3 percent of groups) at the five percent level and for ten groups (33.3 percent of groups) at the ten percent level.

[4] DFG [3] and DFGG [4, 5] demonstrate that the empirical tests of the APT model depend on how the universe of securities is grouped and on the number of securities. They show that the number of factors increases with the group size. The grouping of securities in earlier research was introduced to get around the computational problems in factor analyzing a large group of securities. We demonstrate later that the anomalies in stock returns affect the empirical tests of the APT independently of the group size.

[5] The assumption of stationarity can be questioned for this length of time. However, varying the lengths of time periods does not produce substantial changes in the results and conclusions reported in this paper.

[6] Because factor loadings are identifiable only up to left multiplication by an orthogonal matrix, the significance tests for risk premia can be done only jointly for the vector of risk premia. (See DFG [3].) The relevant test statistic for the risk-premia vector $c^*_{t.}$ is

Table I

Two-Stage Tests of the Arbitrage Pricing Theory and Tests of Significance for Risk Premia against Specific Alternatives[a]

Independent Variables:	B[b]	B and Ω[c]		B[b]	B and Ω[c]		B[b]	B and Ω[c]	
Test Statistics:	χ^2	χ^2	$t(\Omega)$	χ^2	χ^2	$t(\Omega)$	χ^2	χ^2	$t(\Omega)$
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	A. Groups of 30 stocks, 7-factor model			B. Groups of 90 stocks, 17-factor model			C. Groups of 90 stocks, 7-factor model		
1	6.765	5.626	1.60	29.447**	18.410	2.55**	21.230**	8.857	3.13**
2	12.961*	7.366	0.90	27.061*	16.828	0.60	20.922**	15.723**	1.06
3	15.790*	5.953	2.15**	33.765**	18.313	2.22**	27.308**	14.179**	2.55**
4	9.854	5.403	1.03	21.121	13.258	1.77*	17.097**	7.965	1.68*
5	9.757	7.222	0.078	27.941**	16.051	1.92*	12.995*	7.081	2.84**
6	6.448	4.504	0.29	22.748	14.413	1.03	15.841**	8.932	1.50
7	13.848*	6.328	0.51	33.411**	8.590	1.80*	27.708**	5.752	2.64**
8	14.393**	7.749	1.33	24.258	13.751	2.33**	18.564**	8.953	2.65**
9	16.083**	7.457	0.94	23.611	13.589	3.51**	18.668**	10.583	3.70**
10	7.934	1.952	0.99	28.219**	14.690	1.47	21.312**	8.768	1.96**
11	12.717*	5.867	1.49						
12	7.325	4.393	0.71						
13	6.444	6.461	0.91						
14	8.979	8.259	1.88*						
15	5.945	3.283	0.62						
16	12.057*	5.943	1.20						
17	7.259	1.010	1.77*						
18	6.411	6.882	0.68						
19	10.904	2.563	2.05**						
20	19.027**	5.903	1.80*						
21	13.531*	3.270	1.16						
22	9.560	4.398	1.56						
23	6.479	2.330	0.85						
24	4.346	5.238	2.09**						
25	12.087*	11.551	2.66**						
26	6.421	5.483	1.94*						
27	9.433	3.260	1.23						
28	8.445	7.334	0.18						

29	11.420	6.159	1.25						
30	10.790	3.208	2.19**						
Number (percentage) of groups with significant statistics									
At 10%	10 (33)	0 (0)	9 (30)	6 (60)	0 (0)	7 (70)	10 (100)	2 (20)	8 (80)
At 5%	4 (13)	0 (0)	5 (17)	5 (50)	0 (0)	4 (40)	9 (90)	2 (20)	7 (70)

[a]Total standard deviation and residual standard deviations are used as alternatives to the APT model: summary statistics—July 3, 1962 through December 31, 1981.

[b] Chi-square value is computed by $T\bar{c}^* W^{-1}\bar{c}^{*\prime}$, where $\bar{c}^* = (1/T) \sum_{t=1}^{T} \tilde{c}_{t.}^*$ and $W = (1/T) \sum_{t=1}^{T} (\tilde{c}_{t.}^* - \bar{c}^*)' (\tilde{c}_{t.}^* - \bar{c}^*)$. $\tilde{c}_{t.}$s are the coefficients of the daily cross-sectional regressions using the GLS model in equation (4), $\tilde{c}_{t.}' = (\tilde{B}^*\tilde{\Psi}^{-1}\tilde{B}^{*\prime})^{-1}\tilde{B}^*\tilde{\Psi}^{-1}r_{t.}'$, $\tilde{c}_{t.} = (\tilde{c}_{t0}, \tilde{c}_{t.}^*)$; $B^{*\prime}$ is $[e{:}B']$, the augmented matrix of factor loadings (B) with unit vector (e). Chi-square tests the null hypothesis that the risk premia vector is null ($\tilde{c}_t^* = 0$). It is distributed as chi-square with degrees of freedom equal to the number of factors.

[c] Square root of residual (specific) variance (Ω) is included as an additional explanatory variable. Residual variances correspond to the diagonal elements of $\tilde{\Omega}$ in $\tilde{\Psi} = \tilde{B}'\tilde{B} + \tilde{\Omega}$. Chi-square tests the significance of the risk premia coefficients jointly and $t(\Omega)$ tests the coefficient of residual variance in equation (5). t-Ratio is given by $\sqrt{T}(\bar{d}/s)$, where $\bar{d} = \sum_{t=1}^{T} \tilde{d}_t$ and $s = \{(1/T) \sum_{t=1}^{T} (\tilde{d}_t - \bar{d})^2\}^{1/2}$. The minimum number of daily observations is 4511, and the maximum is 4730.

* Ten percent level of significance.

** Five percent level of significance.

Columns (2) and (3) present results to address the question of whether the use of the factor models exhausts the "explanation" of the factor-return process. We include the residual (specific) standard deviation as an alternative specification to the APT model. Once residual variance is included, the null hypothesis that the risk-premia vector is null cannot be rejected for any group at the five percent level for the entire period. Residual variance, however, is "priced" at least in five groups at the five percent level and in ten groups at the ten percent level for the same period.

Parts B and C of Table I include the summary statistics using groups of ninety stocks for seventeen- and seven-factor models. Results here are more favorable to the APT when factor loadings are used as explanatory variables alone (column (1) in Parts B and C). Risk premia are priced in five groups (fifty percent of groups) using a seventeen-factor model and in nine groups (ninety percent of groups) using a seven-factor model at the five percent significance level; corresponding numbers are six (sixty percent of groups) and ten (one hundred percent of groups) at the ten percent level.

However, once residual variance is included as an additional independent variable (column (2) in Parts B and C), risk premia are no longer priced at the five percent level, except for two groups for the seven-factor model in Part C. The residual variance itself, however, is priced more often (forty and seventy percent of the groups at the five percent significance level and seventy and eighty percent of the groups at the ten percent level for seventeen- and seven-factor models, respectively).

While the empirical results are at best mixed for the APT, they are not conclusive because of the problems associated with grouping and the restrictive assumptions used here and in most other empirical work.

II. Seasonality in Risk-Premia Estimates

Panel A in Table II presents the means of daily returns by month for the entire sample of nine hundred NYSE and AMEX stocks. The seasonal pattern in daily stock returns is evident: January returns are five to ten times larger than most other months. November is also a month with larger mean returns.[7] Groups of thirty or ninety stocks also exhibit a similar seasonal pattern. These groups are formed from an alphabetically ranked universe of NYSE and AMEX stocks; thus

---

$$T\bar{c}^* W^{-1} \bar{c}^{*\prime} \overset{A}{\sim} \chi_k^2,$$

where

$$\bar{c}^* = \left(\frac{1}{T}\right) \sum_{t=1}^{T} \tilde{c}_{t.}^* \quad \text{and} \quad W = \left(\frac{1}{T}\right) \sum_{t=1}^{T} (\tilde{c}_{t.}^* - \bar{c}^*)' (\tilde{c}_{t.}^* - \bar{c}^*).$$

The test statistic is chi-square with k degrees of freedom.

[7] We do not provide here formal statistical significance tests for the equality of means in Table II. The seasonality in stock returns in the U.S. and other countries is a well-documented phenomenon; Gultekin and Gultekin [7] provide such tests and report that January mean returns are statistically greater than those of most other months. Here, given the magnitudes of the January returns and the standard deviations (not reported here), formal statistical tests would confirm our observations.

Table II

Means for Month-to-Month Daily Stock Returns and Significance Tests for Means of Groups of Thirty and Ninety Stocks Each—July 4, 1962 through December 31, 1981

	January	February	March	April	May	June	July	August	September	October	November	December
Panel A. Means for month-to-month daily returns for 900 stocks[a]												
	0.2835	0.0408	0.0734	0.0746	−0.0033	0.0085	0.0641	0.0415	0.0297	−0.0072	0.1100	0.0707
Panel B. Groups of 30 stocks:												
Percentage of groups with individual security means equal to each other[b]												
10%	0	13	0	0	13	7	0	3	10	20	0	0
5%	0	3	0	0	7	3	0	0	3	13	0	0
Panel C. Groups of 90 stocks:												
Percentage of groups with individual security means equal to each other[c]												
10%	0	0	0	0	10	0	0	0	0	10	0	0
5%	0	0	0	0	0	0	0	0	0	0	0	0

[a] Means are the simple arithmetic average of daily returns over the entire period for nine hundred NYSE and AMEX stocks.

[b] Multivariate Hotelling T^2 test is used to test the hypothesis that means of individual securities are jointly equal in each group over the entire period. Percentages indicate the percentage of groups with equal means.

[c] Same test as in footnote b conducted for ten groups of ninety stocks each.

the possibility that the pattern observed in Table II is generated by a few outliers can safely be ruled out.

In addition, we tested for the equality of mean returns of individual securities in all groups of thirty and ninety stocks each for each month. These multivariate T^2 tests are summarized in Panels B and C of Table II. Mean returns for individual stocks are reliably different from each other for all groups of ninety stocks each for all months. For groups of thirty stocks each, we could not reject the null hypothesis in five months of the year for a very small number (three to thirteen percent) of the groups. These results rule out the possibilty that the findings we report in the remainder of the study are an artifact of the experimental design. Furthermore, since mean returns are not equal for securities, there is something to be explained by asset-pricing models.

Based on Table II, most of the reward for holding stocks is systematically realized in a few months of the year, primarily in January. It remains to be seen whether such a seasonal pattern affects the empirical tests of the APT model or whether the January seasonality is captured by the factor loadings of the APT.

Table III provides partial insight to this question. Here we re-estimate the GLS regression model in (4) for each calendar month separately and compute the corresponding test statistics from these month-to-month estimates of risk premia. Results for January are reported in column (1) of Table III. In column (2), we report corresponding results for the average of the other eleven months' results. The first row of Panel A reports the percentage of groups with significant risk premia using factor loadings as independent variables alone. Summary results in Table III correspond to the month-to-month version of the last two lines in Table I; for ease of presentation, they are shown in percentages.

A casual review of the first row of Panel A reveals that risk premia are *always* significant in January for each of the thirty groups but are rarely priced in other months. As in the case of the two-stage empirical tests of CAPM reported by Tinic and West [17, 18], the APT model also seems to hold only in January.[8]

The tests of the APT against a specific alternative using residual variance for each calendar month are summarized in the second row of Panel A for groups of thirty stocks. A similar seasonal pattern emerges here as well. Factor loadings are priced mostly in January; once the residual standard deviation is introduced, factor loadings are priced in fewer groups; finally, both risk premia and the residual standard deviation are priced rarely in other months.

Panels B and C of Table III contain similar results for groups of ninety stocks using seven- and seventeen-factor models, respectively. The first row of each panel shows that, in January, the risk-premia vector is priced for all of the ten groups, as in the case of groups of thirty stocks.[9] The tests of the APT against a

[8] So far we have demonstrated the seasonality in terms of the significance tests for risk premia. To test whether the estimates of risk premia themselves exhibit any seasonal pattern, we use standard-regression models with dummy variables for each month. Risk premia associated with individual factors also exhibit a seasonal pattern.

[9] Interestingly, February, June, and November here also have nonzero risk-premia vectors ranging from seventy to fifty percent for the seven-factor model. For the seventeen-factor model, the risk-premia vector is priced for half of the groups for February, indicating a seasonal pattern at the beginning of the year.

Table III

Seasonal Pattern in the Two-Stage Tests of the Arbitrage Pricing Theory: Test of Significance for Risk Premia against Residual Standard Deviation as an Alternate to the APT Model: Percentage of Groups with Significant Test Statistics at Five Percent Level—July 3, 1962 through December 31, 1981

Independent Variables	Test Statistics	January (1)	Other Months[a] (2)	Excluded Months: Jan and Dec (3)	Excluded Months: Feb–Nov (4)
Panel A.	Seven-factor model, 30 groups of 30 stocks each				
B Only[b]	chi-square for B	100[c]	9	3	90
B and Ω[d]	chi-square for B	60	6	10	17
	t-ratio for Ω	57	3	7	43
Panel B.	Seven-factor model, 10 groups of 90 stocks each				
B Only[b]	chi-square for B	100	27	10	100
B and Ω[d]	chi-square for B	90	21	10	90
	t-ratio for Ω	90	5	10	70
Panel C.	Seventeen-factor model, 10 groups of 90 stocks each				
B Only[b]	chi-square for B	100	15	10	90
B and Ω[d]	chi-square for B	100	12	0	80
	t-ratio for Ω	80	7	20	40

[a] Average of other eleven months' results.

[b] Factor loadings are the only independent variables. See Table I (footnote b) for the regression model. Chi-square tests the null hypothesis that the risk-premia vector is null.

[c] The percentages correspond to the percentage of groups for which the null hypothesis that the risk-premia vector is null is rejected; i.e., risk premia are priced.

[d] Residual standard deviations of securities are included as additional explanatory variables. Residual variance corresponds to the diagonal elements of Ω in $\Psi = B'B + \Omega$ in factor analysis. Chi-square tests the significance of the risk-premia coefficients and $t(\Omega)$ tests the significance of the residual-variance coefficient in regression model (5).

specific alternative also produce similar results. In the second row, risk premia are priced mostly in January. For other months, however, neither the risk-premia vector nor the residual standard deviation is often significant.

Findings in Table III have a disturbing implication for the empirical tests of the APT. Seasonality results are very robust with respect to the group size. While using group sizes of thirty versus ninety stocks per group produces dramatic changes in the test results of the APT per se (compare Part A with B or C in Table I), January anomalies affect the significance tests for all sizes of groups similarly.

We have by now established the existence of a seasonal pattern in the estimates of risk premia (see footnote 8) using the significance tests. Next we inquire whether risk premia are priced at all if we exclude the data for months, primarily January, that are responsible for the seasonality in stock returns.

Columns (3) and (4) in Table III provide an answer. Here we exclude January and December returns from the data and compute the significance tests using the returns from the remaining ten months. The decision to exclude December and January returns is based on Keim's [9] finding that the daily stock returns are the largest during the last week of December and the first two weeks of January.

Two interesting patterns emerge. First, in column (3), when January and December returns are excluded, the risk premia are priced in only one out of thirty groups at the five percent level (three percent); on the other hand, Table I (column (1)) shows the corresponding numbers using returns from all months: four groups (thirteen percent). Clearly the results are weaker without January and December returns. Second, as before, risk premia are priced in twenty-five groups at the five percent level (and, not shown in Table III, in twenty-seven groups at the ten percent level) using returns for January and December alone.[10]

Other results using residual variance in Table III provide further support for our observations. Without the January and December returns (in column (3)), residual variance is far less frequently priced in comparison to the corresponding results in Table I. Interestingly, factor loadings are priced about as frequently as standard deviation measures, in contrast to the results in Table I where we use returns from all months.

In order to investigate the robustness of findings, we conducted several additional experiments: (a) we repeated the cross-sectional regressions by excluding returns one month at a time and computed the corresponding test statistics using the returns from the remaining eleven months; (b) we repeated the tests in Table III for groups of sixty, ninety, and 120 stocks each; (c) we repeated results from Table III for twenty-five percent of the groups by re-estimating the factor loadings using the returns for each month separately.[11]

[10] February and November are two months in which significant test statistics occur more frequently; see footnote 9. We repeated the experiments in Table III by excluding the January and February and also January and November returns from the data. Results are similar to those reported in Table III. However, the test statistics that correspond to those in column (1) are significant more frequently and those that correspond to column (2) are significant less frequently.

[11] There is, however, a serious problem with experiments that exclude large amounts of observations. This procedure reduces the number of observations by one twelfth. DFG [3] demonstrate that the number of factors vary directly as the number of observations changes. Thus, one cannot make

The general pattern of the results gives the same impression, that more groups yield significant test statistics when January returns are included in the data.

In summary, when all results are considered together, they imply that most of the significant results are generated by the stock return anomalies around the turn of the year. Once the January returns are excluded from the data, the statistical significance of the results diminishes dramatically.[12]

III. Conclusions

Our results show that the two-stage empirical tests of the APT model are very sensitive to the anomalies in the stock return data. There is also a strong seasonal pattern in the estimates of the risk premia from the APT model that is similar to those reported by Tinic and West [17, 18] for the CAPM. The most important implication of the seasonality is that the APT model, like the CAPM, can explain the risk-return relation in January only. These results do not seem to depend on the group size that we use in the empirical tests.

Although we do not have a completely satisfactory explanation for the January seasonal, its effect on empirical research appears to be far more serious than previously realized. If a crucial test of an asset-pricing model is its ability to explain the turn-of-the-year effect, particularly as it relates to size, then we are left with a disturbing conclusion. The seasonal pattern in the stock return data is so strong in January that asset-pricing models based on covariance measures of risk are not likely to explain the turn-of-the-year effect or size-related anomalies. Furthermore, returns are so highly correlated with standard deviations in January that covariance risk measures are bound to fail tests where total or residual standard deviations are used as specific alternatives.

---

meaningful comparisons between results that use the entire set of observations and those that use only a small subset. We therefore do not report results where we use factor loadings estimated for each month separately. We obtain qualitatively similar significance-test results to those presented here; however, we do find fewer factors. See Cho and Taylor [2] for this approach.

[12] Miller and Scholes [10] point out that, if the returns are not normally distributed, the higher moments of a distribution will be correlated. This result might explain part of our findings. In order to overcome this problem, RR [14], DFG [3], and DFGG [4, 5] use noncontiguous days to estimate the parameters. Such experiments in our case did not produce results materially different from those already reported.

REFERENCES

1. R. W. Banz. "The Relationship between Return and Market Value of Common Stocks." *Journal of Financial Economics* 9 (March 1981), 3–18.
2. D. C. Cho and W. M. Taylor. "The Seasonal Stability of the Factor Structure of Stock Returns." *Journal of Finance* 42 (December 1987), 1195–1211.
3. P. J. Dhrymes, I. Friend, and N. B. Gultekin. "A Critical Reexamination of the Emprical Evidence on the Arbitrage Pricing Theory." *Journal of Finance* 39 (June 1984), 323–46.
4. P. J. Dhrymes, I. Friend, M. N. Gultekin, and N. B. Gultekin. "An Empirical Examination of the Implications of Arbitrage Pricing Theory." *Journal of Banking and Finance* 9 (December 1984), 73–99.
5. ———. "New Tests of the APT and Their Implications." *Journal of Finance* 40 (July 1985), 659–73.

6. E. F. Fama and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (May/June 1973), 607–36.
7. M. N. Gultekin and N. B. Gultekin. "Stock Market Seasonality: International Evidence." *Journal of Financial Economics* 12 (December 1983), 469–81.
8. N. B. Gultekin and R. J. Rogalski. "Government Bond Returns, Measurement of Interest Rate Risk, and Arbitrage Pricing Theory." *Journal of Finance* 40 (March 1985), 43–61.
9. D. B. Keim. "Size Related Anomalies and Stock Return Seasonality: Further Empirical Evidence." *Journal of Financial Economics* 12 (June 1983), 13–32.
10. M. H. Miller and M. Scholes. "Rates of Return in Relation to Risk: A Reexamination of Some Recent Findings." In Michael L. Jensen (ed.), *Studies in the Theory of Capital Markets*. New York: Preager, 1972, 47–48.
11. R. R. Officer. "Seasonality in the Australian Capital Markets: Market Efficiency and Empirical Issues." *Journal of Financial Economics* 2 (March 1975), 29–52.
12. M. R. Reinganum. "The Anomalous Stock Market Behaivor of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects." *Journal of Financial Economics* 12 (June 1983), 89–104.
13. R. Roll. "Vas ist das? The Turn of the Year Effect and the Return Premiums of Small Firms." *Journal of Portfolio Management* 9 (Winter 1983), 18–28.
14. ——— and S. A. Ross. "An Empirical Investigation of the Arbitrage Pricing Theory." *Journal of Finance* 35 (December 1980), 1073–1103.
15. S. A. Ross. "The Arbitrage Pricing Theory of Capital Asset Pricing Theory." *Journal of Economic Theory* 13 (December 1976), 341–60.
16. M. S. Rozeff and W. R. Kinney, Jr.. "Capital Market Seasonality: The Case of Common Stock Returns." *Journal of Financial Economics* 3 (October 1976), 379–402.
17. S. M. Tinic and R. R. West. "Risk and Return: January vs. the Rest of the Year." *Journal of Financial Economics* 13 (December 1984), 561–74.
18. ———. "Risk, Return, and Equilibrium: A Revisit." *Journal of Political Economy* (February 1986), 126–147.
19. S. W. Wachtel. "Certain Observation on Seasonal Movements in Stock Prices." *Journal of Business* 15 (April 1942), 184–93.