Journal of Financial Economics 5 (1977) 309-327 © North-Holland Publishing Company

ESTIMATING BETAS FROM NONSYNCHRONOUS DATA

Myron SCHOLES and Joseph WILLIAMS*
University of Chicago, Chicago, IL 60637, U.S.A.

Received November 1976, revised version received October 1977

Nonsynchronous trading of securities introduces into the market model a potentially serious econometric problem of errors in variables. In this paper properties of the observed market model and associated ordinary least squares estimators are developed in detail. In addition, computationally convenient, consistent estimators for parameters of the market model are calculated and then applied to daily returns of securities listed in the NYSE and ASE.

1. Introduction

Central to contemporary theory in finance is the fundamental concept of systematic risk or beta for a security. Not surprisingly, much current empirical work in finance focuses on the associated problem of estimating this systematic risk or beta. Although to date almost all estimates have used monthly returns on common stocks, recently daily returns have become available.[1] With this new data more powerful empirical tests are now possible.

Unfortunately, the use of daily data introduces into the market model a potentially serious econometric problem. In particular, many securities listed on organized exchanges are traded only infrequently, with few securities so actively traded that prices are recorded almost continuously.[2] Because prices for most securities are reported only at distinct random intervals, completely accurate calculation of returns over any fixed sequence of periods is virtually impossible. In turn this introduces into the market model the econometric problem of errors in variables. With daily data this problem appears particularly severe.

*We thank Michael Jensen and the referee, G. William Schwert, for comments on a previous draft, Marvin Lipson for programming assistance, and the Center for Research in Security Prices, University of Chicago, for financial support.

[1]Daily data for total returns on approximately 4000 securities listed on the New York and American Stock Exchanges between July 2, 1962 and December 31, 1975 have been collected at the Center for Research in Security Prices, Graduate School of Business, University of Chicago.

[2]Fama (1965) and Fisher (1966) first recognized the nontrading of securities as a potentially serious empirical problem. Some of the results in this paper appear in a set of unpublished notes by Oldrich Vasichek, Graduate School of Management, University of Rochester. Related results appear in Cohen, Maier, Schwartz, and Whitcomb (1977).

In section 2 of this paper, properties of the market model with nonsynchronous data are developed in detail Assuming that true instantaneous returns on securities are normally distributed, it is shown that variances and covariances of reported returns differ from corresponding variances and covariances of true returns Under plausible restrictions on trading processes, measured variances for single securities overstate true variances, while measured contemporaneous covariances understate in absolute magnitude true covariances Also, reported returns on single securities appear serially correlated and leptokurtic relative to actual returns Additional properties for measured returns on portfolios of securities are reported

With errors in variables in the market model, ordinary least squares estimators of both alphas and betas for almost all securities are biased and inconsistent In this particular problem with errors from nonsynchronous trading of securities, there remains within reported returns sufficient structure to identify both the direction and magnitude of any asymptotic bias Accordingly, in section 3 it is shown that securities trading on average either very frequently or very infrequently have ordinary least squares estimators asymptotically biased upward for alphas and downward for betas By contrast, for those remaining securities with more average trading frequencies, least squares estimators of alphas and betas are asymptotically biased in the opposite directions

In section 3 computationally convenient, consistent estimators for coefficients in the market model are constructed The consistent estimator for beta is calculated as a combination of ordinary least squares estimators Specifically, the sum of betas estimated by regressing the return on the security against returns on the market from the previous, current, and subsequent periods is divided by one plus twice the estimated autocorrelation coefficient for the market index In turn these consistent estimators of alpha and beta are shown to be equivalent to instrumental variables estimators which use as an instrument the moving sum of measured rates of return on the market for the previous, current, and subsequent periods In addition, asymptotic standard errors for these estimators are calculated

In section 5 the consistent estimators are applied to daily returns from securities listed on the New York and American Stock Exchanges between January, 1963 and December, 1975 Estimates of alphas and betas are calculated and then compared to the corresponding ordinary least squares estimates for portfolios comprised of securities selected by trading volume Almost without exception, these estimates are consistent with the predictions from section 4 All basic results cited in the text are derived in the appendix

2. The problem

Implicit behind the basic capital asset pricing model in continuous time is the assumption that all risky securities have prices distributed as infinitely divisible, lognormal random variables [3] With this assumption the continuously compounded returns r_{nt} on risky securities $n = 1, \ldots, N$, as calculated over any intervals $[t-1, t]$, $t = 1, \ldots, T$, are joint normally distributed with the constant means μ_n, constant variances σ_n^2, and constant covariances σ_{nm}, $n \neq m$, $n, m = 1, \ldots, N$ In turn this implies that the corresponding rates of return $r_{Mt} \equiv \sum_{n=1}^{N} r_{nt} x_{nM}$ on the market index M are normally distributed with the constant mean μ_M, constant variance σ_M^2, and constant covariances σ_{nM}, $n = 1, \ldots, N$ Here x_{nM} represents the constant percentage weight of security n in the market index M Collectively, these assumptions imply the simple market model

$$r_{nt} = \alpha_n + \beta_n r_{Mt} + \varepsilon_{nt}, \tag{1}$$

where $\alpha_n = \mu_n - \beta_n \mu_M$ and $\beta_n = \sigma_{nM}/\sigma_M^2$ are the constant coefficients alpha and beta The residual ε_{nt}, orthogonal to r_{Mt}, is normally distributed with a zero mean plus constant variances and covariances

In practice, however, the market model (1) is not continuously observable Because most securities trade at discrete, stochastic intervals in time, with prices recorded only at points of actual trades, (1) cannot be observed at all times $0 \leqq t \leqq T$ This discontinuous trading of securities introduces into the market model the common econometric problem of errors in variables

Specifically, consider any sequence of distinct, uniformly spaced points in time $t = 1, \ldots, T$ with the corresponding intervals $[t-1, t]$ During any such interval there occurs either no trade, no observed price, and hence no calculated return, or, alternatively, at some random time $t - s_{nt}$, $0 \leqq s_{nt} \leqq 1$, a last trade and consequently a reported closing price Here s_{nt} represents the residual portion of trading period t during which no trades in security n occur If over any two consecutive intervals closing prices are reported, then a rate of return r_{nt}^s can be calculated for the corresponding period $[t-1-s_{nt-1}, t-s_{nt}]$ (See figure 1) Collecting all such measured rates of return, and ignoring periods over which no trades occur, generates the sampling sequence $\{r_{nt}^s\}$ differing in general from $\{r_{nt}\}$ [4] In turn this implies for the market index a measured sequence $\{r_{Mt}^s\}$ of returns, $r_{Mt}^s \equiv \sum_{n=1}^{N} r_{nt}^s x_{nM}$, also differing from $\{r_{Mt}\}$ [5] Errors in variables result when measured returns are used as proxies for true unobservable returns

[3]For a discussion of this pricing process see Merton (1973) More generally, the results of this paper require only that the pricing process be infinitely divisible with independent increments This permits, for example, compound Poisson processes

[4]All information about returns over days in which no trades occur is ignored This greatly simplifies the subsequent estimators

[5]Some variability over time in the market index is introduced by excluding from the index securities not trading during the particular period Presumably, these effects are minor Variability in the market index has no effect on the results of section 2 nor on the major results (16) and (17)

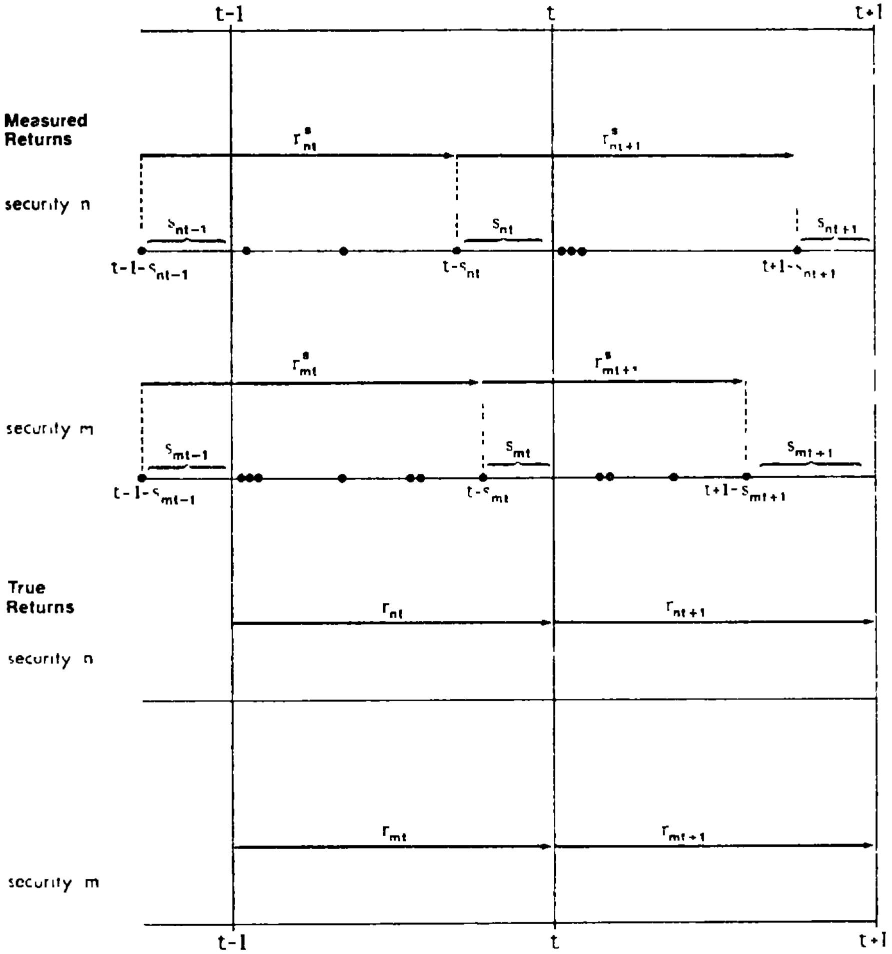

Fig 1 Measured returns versus true returns for securities n and m over periods t and $t+1$

With these presumably unavoidable errors in variables, measured returns are no longer generated by a simple Gaussian process Instead, measured returns arise from a subordinated process with parameters dependent upon realizations of the directing process determining actual times between trades In terms of the market model (1), this subordination implies

$$r_{nt}^s = \alpha_n^s + \beta_n^s r_{Mt}^s + \varepsilon_{nt}^s, \tag{2}$$

with the coefficients

$$\alpha_n^s = E[r_{nt}^s] - \beta_n^s E[r_{Mt}^s], \tag{3}$$

and

$$\beta_n^s \equiv \frac{\mathrm{cov}(r_{nt}^s, r_{Mt}^s)}{\mathrm{var}(r_{Mt}^s)} \tag{4}$$

Generally (3) and (4) differ from the corresponding coefficients α_n and β_n in (1) Again the residual ε_{nt}^s has a zero mean and zero covariance with the regressor r_{Mt}^s In (2) no restrictions are placed on the sequence of nontrading periods $\{S_t\}$, $S_t = (s_{1t}, \quad , s_{Nt})$

All differences between the observed market model (2) and the true model (1) reflect differences between measured returns and true returns In general, with errors of observation, measured returns deviate from normality with moments depending on properties of both true returns and nontrading periods Means, variances, and covariances for measured returns are derived in the appendix, part (i) In the plausible special case with all nontrading periods S_t independently and identically distributed over time, these moments simplify as follows

$$E[r_{nt}^s] = \mu_n, \tag{5}$$

$$\mathrm{var}(r_{nt}^s) = \{1+2\,\mathrm{var}(s_n)/v_n^2\}\sigma_n^2, \tag{6}$$

with the coefficient of variation $v_n \equiv \sigma_n/\mu_n$,

$$\mathrm{cov}(r_{nt}^s, r_{mt}^s) = \{1-E[\max\{s_n, s_m\}-\min\{s_n, s_m\}] \\ +2\,\mathrm{cov}(s_n, s_m)/\rho_{nm}v_n v_m\}\sigma_{nm}, \tag{7}$$

with the correlation coefficient $\rho_{nm} \equiv \sigma_{nm}/\sigma_n\sigma_m$,

$$\mathrm{cov}(r_{nt}^s, r_{nt-1}^s) = -\{\mathrm{var}(s_n)/v_n^2\}\sigma_n^2, \tag{8}$$

and

$$\mathrm{cov}(r_{nt}^s, r_{mt-1}^s) = \{E[\max\{s_n-s_m, 0\}] \\ +\mathrm{cov}(s_n, s_m)/\rho_{nm}v_n v_m\}\sigma_{nm} \tag{9}$$

In this special case all remaining covariances of various lags disappear

The properties of (5) through (9) are interesting With nontrading periods S_t distributed independently and identically over time, expectations of measured returns (5) for single securities always equal true mean returns By contrast, for single securities, measured variances (6) overstate true variances, while measured autocovariances (8) of lag one appear negative Also, measured contemporaneous covariances (7) differ from true covariances, while measured covariances (9) of lag one deviate from zero Finally, all remaining covariances for lags greater than one vanish in the absence of data from periods during which no trading occurs

With daily data additional simplifications in (6) through (9) are possible. Specifically, for daily returns on NYSE and ASE common stocks, the coefficient of variation v_n exhibits an average value roughly in the range of 30 to 40. This implies that for single securities measured variances (6) closely approximate true variances while measured autocovariances (8) closely approximate zero In addition, for single securities measured contemporaneous covariances (7) understate in absolute magnitude true covariances, while measured covariances (9) of lag one share with true contemporaneous covariances the same sign but a smaller absolute value Moreover, the magnitudes of these effects on covariances are greater for securities trading on average less frequently More precisely, in (7) the discrepancy of measured contemporaneous covariances from true covariances is greatest when one security trades on average very frequently while the remaining security trades very infrequently Similarly, in (9) the discrepancy is greatest when security n trades very frequently while security m trades very infrequently [6]

In turn these properties have implications for daily returns from large portfolios Suppose that in practice returns on individual securities are predominantly positively correlated Also recall that for large portfolios variances are primarily determined by the covariances of returns among component securities From (6) and (8) this implies that measured variances for daily returns on large portfolios typically understate true variances For portfolios more heavily weighted with securities trading on average less frequently – e g , an equally weighted portfolio – these effects are even more pronounced Clearly, both these properties contrast sharply with the corresponding results for single securities

For individual securities reported returns deviate from normality This deviation is measured in part by the kurtosis of reported returns Assuming as before that the nontrading periods S_t are independently and identically distributed over time, the kurtosis $\kappa(r^s_{nt})$ of measured returns can be written as[7]

$$\kappa(r^s_{nt}) = 3(1+2\,\mathrm{var}(s_n))+O(1/v_n^2). \tag{10}$$

In (10) the notation $O(1/v_n^2)$ identifies terms of order $1/v_n^2$ Given (10), for values of v_n plausible with daily data, the kurtosis exceeds 3, the kurtosis of a normal variate As a result, measured daily returns for single securities appear leptokurtic relative to actual, unobservable returns Again this effect is more pronounced for securities trading less frequently

[6]Detailed results for Poisson trading processes appear in a previous working draft of this paper The results are available from the authors upon request

[7]The derivation of (10) appears in the appendix, part (ii) This generalizes a previous result for zero drift processes in Clark (1973)

3. Ordinary least squares

With errors in variables in the observed market model, ordinary least squares applied directly to (2) generates estimators with unattractive properties Not surprisingly, the ordinary least squares estimators (OLSE) a_n and b_n of the coefficients α_n and β_n in (1) are biased and inconsistent This bias occurs because, as is typical in models with errors in variables, the regressor r^s_{Mt} in (2) covaries with the residual ε^s_{nt} In fact, in this model it is straightforward to show that

$$\operatorname{plim} a_n = \alpha^s_n \neq \alpha_n \tag{11}$$

and

$$\operatorname{plim} b_n = \beta^s_n \neq \beta_n \tag{12}$$

In at least one important special case, the direction of the asymptotic bias in (11) and (12) can be identified explicitly Suppose as before that the non-trading periods S_t are distributed independently and identically over time Also define the new regression coefficients

$$\beta^{s-}_n \equiv \frac{\operatorname{cov}(r^s_{nt}, r^s_{Mt-1})}{\operatorname{var}(r^s_{Mt-1})} \tag{13}$$

and

$$\beta^{s+}_n \equiv \frac{\operatorname{cov}(r^s_{nt}, r^s_{Mt+1})}{\operatorname{var}(r^s_{Mt+1})}, \tag{14}$$

plus the autocorrelation coefficient

$$\rho^s_M \equiv \frac{\operatorname{cov}(r^s_{Mt}, r^s_{Mt-1})}{\operatorname{std}(r^s_{Mt})\operatorname{std}(r^s_{Mt-1})}, \tag{15}$$

where std( ) represents the standard deviation In this case, as derived in the appendix, part (iii), the coefficients α^s_n and β^s_n in (3) and (4) satisfy

$$\alpha^s_n = \alpha_n + (\beta_n - \beta^s_n)\mu_M \tag{16}$$

and

$$\beta^s_n = \beta_n - (\beta^{s-}_n + \beta^{s+}_n - 2\beta_n\rho^s_M). \tag{17}$$

The relationship in (16) and (17) between measured coefficients and true coefficients can now be identified Examine the top two lines of fig 1, focusing on securities n and m with positive betas Suppose security m is traded on average about as frequently as the average security in the index, where securities in the index are ranked by average trading frequencies If, relative to security m, security n is traded on average only infrequently, then the overlap between periods of measurement for r^s_{nt} and r^s_{mt-1} is typically large This implies from (9) and (13) a relatively large lagged beta β^{s-}_n and hence from (16) and (17) the inequalities $\alpha_n < \alpha^s_n$ and $\beta_n > \beta^s_n$ Similarly, if, relative to security m, security n

is traded on average quite frequently, then it is possible, although not necessarily likely, that the overlap between r^s_{nt} and r^s_{mt+1} is on average large Again this implies $\alpha_n < \alpha^s_n$ and $\beta_n > \beta^s_n$ Overall, measured alphas and betas equal on average true alphas and betas

$$\sum_{n=1}^{N} \alpha^s_n x_{nM} = 0 = \sum_{n=1}^{N} \alpha_n x_{nM}$$

and

$$\sum_{n=1}^{N} \beta^s_n x_{nM} = 1 = \sum_{n=1}^{N} \beta_n x_{nM}$$ [8]

As a result, most remaining securities – that is, those trading neither very frequently nor very infrequently – exhibit the reverse inequalities $\alpha_n > \alpha^s_n$ and $\beta_n < \beta^s_n$

Together, (11), (12), (16), and (17) identify the asymptotic biases for a_n and b_n Securities trading very infrequently, plus possibly some trading very frequently, have estimators asymptotically biased upward for a_n and downward for b_n By contrast, most remaining securities have OLSE asymptotically biased in the opposite directions Overall, the estimators a_n and b_n equal on average across securities the true parameters a_n and β_n

Similar observations are possible for the autocorrelation coefficient ρ^s_m of the market index Specifically, from (15) and (A9) in the appendix, part (iii), it follows that

$$\rho^s_m = \tfrac{1}{2}\left[\frac{\operatorname{var}(r_{Mt})}{\operatorname{var}(r^s_{Mt})} - 1\right] \tag{18}$$

If, as argued in section 2, the measured variance $\operatorname{var}(r^s_{Mt})$ on the market index understates the true variance $\operatorname{var}(r_{Mt})$, then the measured autocorrelation coefficient ρ^s_m appears positive Because the sampling estimator $\hat{\rho}_M$ of the autocorrelation coefficient ρ^s_m is a consistent estimator of (18), this guarantees in large samples that $\hat{\rho}_M$ is also positive Again this estimator differs from the true autocorrelation coefficient for the market index, which has a value of zero.

4. Consistent estimators

From (11), (12), (16), and (17), computationally convenient consistent estimators of the coefficients α_n and β_n, $n = 1, \quad , N$, are immediate Let b^-_n, b_n, and b^+_n represent the OLSE associated with (13), (4), and (14), respectively Similarly, let $\hat{\rho}_M$ represent the sampling estimator associated with (15) With this new notation, the previous results imply the consistent estimators

[8] These results follow immediately from the definitions of α_n, β_n, α^s_n, and β^s_n

$$\hat{\alpha}_n \equiv \frac{1}{T-2}\sum_{t=2}^{T-1} r^s_{nt} - \hat{\beta}_n \frac{1}{T-2}\sum_{t=2}^{T-1} r^s_{Mt} \tag{19}$$

and

$$\hat{\beta}_n \equiv \frac{b_n^- + b_n + b_n^+}{1+2\hat{\rho}_M} \tag{20}$$

Simple consistent estimators are also possible for residual variances and covariances

Computationally, (19) and (20) have two important advantages First, $\hat{\alpha}_n$ and $\hat{\beta}_n$ are constructed from two sample means plus a sum and quotient of standard OLSE These OLSE are easily calculated from available data Second, the estimators do not depend on detailed assumptions about the probability distribution generating the sequence of nontrading times $\{S_t\}$ Instead, (19) and (20) require only that S_t is independently and identically distributed over time This latter property is especially important because $\{S_t\}$ is largely unobservable That is, not only are most detailed assumptions essentially unverifiable, but also any simple, analytically tractable distribution – e g , a homogeneous Poisson process for trading times – is unlikely to fit the limited data For example, because information accumulates between consecutive closings and openings of the NYSE and ASE, trades appear on average unevenly spaced over the trading day

In at least one important special case, the estimators (19) and (20) simplify still further With monthly data the problem of nonsynchronous trading of securities has for most common stocks little impact on recorded returns Because true returns on common stocks are uncorrelated over time, this implies for monthly data an estimator $\hat{\rho}_M$ of the autocorrelation coefficient for the market index close to zero However, if the asset in question – for example, stock exchange seats on the NYSE or ASE – is traded only infrequently, then for that asset errors of observation can be important Suppose n indexes the specific asset, while m represents some security trading on average about as frequently as the average security in the index, where again securities in the index are ranked by average trading frequencies In this case, from the top two lines of fig 1, the measurement period for r^s_{nt} overlaps the measurement periods for r^s_{mt} and r^s_{mt-1}, but not r^s_{mt+1} From (4), (11), (12), (13), and (14), this suggests that only b_n and b_n^- differ significantly from zero In this case (20) simplifies to $\hat{\beta}_n = b_n + b_n^-$ and (19) simplifies accordingly [9]

Not surprisingly, the estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$ are asymptotically equivalent to instrumental variables estimators For these new estimators additional notation is necessary Denote by the bold italic letters $\boldsymbol{r}$ deviations from sample means –

[9]This estimator appears in Schwert (1977)

for example,

$$r_{nt}^{s} \equiv r_{nt}^{s} - \frac{1}{T-2} \sum_{t=2}^{T-1} r_{nt}^{s}, \qquad t = 2, \quad \ldots, T-1.$$

Also identify by $r_{Mt3}^{s} \equiv r_{Mt-1}^{s} + r_{Mt}^{s} + r_{Mt+1}^{s}$ the sum of the rates of return on the market for the previous, current, and subsequent periods. Using this notation, the instrumental variables estimators become

$$a_n^* \equiv \frac{1}{T-2} \sum_{t=2}^{T-1} r_{nt}^{s} - b_n^* \frac{1}{T-2} \sum_{t=2}^{T-1} r_{Mt}^{s} \tag{21}$$

and

$$b_n^* \equiv \frac{\frac{1}{T-2} \sum_{t=2}^{T-1} r_{M3t}^{s} r_{nt}^{s}}{\frac{1}{T-2} \sum_{t=2}^{T-1} r_{M3t}^{s} r_{Mt}^{s}}, \tag{22}$$

for the instrument, the sequence of moving sums $\{r_{M3t}\}$. The equivalence between (19) and (21) plus (20) and (22) arises in the probability limit as $T \to \infty$.

Asymptotic standard errors for (19) and (20) or, alternatively, (21) and (22) are easily calculated. Assume as before that the nontrading periods S_t are independently and identically distributed over time. In this case, as can be verified from (5) through (9), the residuals ε_{nt}^{s} in (2) are stationary with a mean of zero, the variance $\omega_n^{s^2} \equiv \mathrm{var}(\varepsilon_{nt}^{s})$, the first-order autocorrelation $\rho_n^{s} \equiv \mathrm{cov}(\varepsilon_{nt}^{s}, \varepsilon_{nt-1}^{s})/\omega_n^{s^2}$, and all other autocorrelations zero. With the additional notation

$$\rho_{M3}^{s} \equiv \frac{\mathrm{cov}(r_{M3t+1}^{s}, r_{M3t}^{s})}{\mathrm{std}(r_{M3t+1}^{s})\,\mathrm{std}(r_{M3t}^{s})} \tag{23}$$

and

$$\beta_{M,M3}^{s} \equiv \frac{\mathrm{cov}(r_{Mt}^{s}, r_{M3t}^{s})}{\mathrm{var}(r_{M3t}^{s})}, \tag{24}$$

the asymptotic standard errors then become[10]

$$\left\{\frac{1}{T-2} \operatorname{plim} [\sqrt{(T-2)}(\hat{\alpha}_n - \alpha_n)]^2\right\}^{1/2}$$
$$= \left\{\frac{1}{T-2} \operatorname{plim} [\sqrt{(T-2)}(\hat{\beta}_n - \beta_n)]^2 \mu_M^2 + \frac{\omega_n^{s^2}}{T-2}(1 + 2\rho_n^{s})\right\}^{1/2} \tag{25}$$

[10] For the derivation see the appendix, part (iv).

and

$$\left\{\frac{1}{T-2}\operatorname{plim}\left[\sqrt{(T-2)}(\hat{\beta}_n-\beta_n)\right]^2\right\}^{1/2} = \left\{\frac{\omega_n^{s^2}}{T-2}\frac{1+2\rho_n^s\rho_{M3}^s}{\beta_{M\,M3}^{s^2}\operatorname{var}(r_{M3t}^s)}\right\}^{1/2}. \tag{26}$$

Consistent estimators of (25) and (26) are computed by replacing all population moments with the respective sample moments

5. Daily returns

In this section the above estimators are applied to daily returns from all stocks listed on the New York and American Stock Exchanges between January 1963 and December 1975 Both consistent estimates and ordinary least squares estimates are calculated for coefficients of five specially constructed portfolios comprised of securities selected by trading volume

The compositions of the five portfolios are as follows For each calendar year 1963 through 1975, each stock listed on the NYSE and ASE was ranked according to the total number of shares of that security traded during the year [11] Based on this ranking five portfolios were formed with portfolio 1 consisting of the 20 percent of securities with the lowest trading volume, portfolio 2 with the next 20 percent, etc Daily returns on each portfolio, including the value weighted market portfolio, were calculated as the logarithm of one plus the arithmetic average of returns on all securities within that portfolio [12] If in any given day a security was not traded, then no return for that security was included in any portfolio for both that day and the subsequent trading day This procedure excluded less than 2 percent of all available securities, where securities available for inclusion in the five portfolios ranged from a minimum number of 1487 in 1963 to a maximum of 2626 in 1973, with an average over 13 years of 2305 During the 13 years there were on average approximately 251 trading days per year

In these calculations trading volume was used as a proxy for the correct, but unobservable variable, the number of distinct trades in a security Because about 61 percent on the NYSE occur in round lots of 100 and 200 shares,[13] trading volume is likely to produce an accurate assignment of securities to the five portfolios Moreover, as a proxy this variable is clearly superior to the only other available index, the dollar volume of trades

[11] Incomplete data on trading volume precludes the inclusion during earlier years of all securities listed on both exchanges

[12] For daily data any differences between this definition of market returns and the definition of section 2 are small. Also, some nonstationarity is introduced in the market index by using value weights

[13] See the NYSE Market Data Systems Monthly Memoranda (July 1976)

All estimates for portfolios 1, 2, and 5 appear in tables 1 through 3 To save space the corresponding estimates for portfolios 2 and 4 are deleted In these tables the column headings reflect the notation of section 4 For example, in table 1 the column headings are, from left to right, the consistent estimates $\hat{\alpha}_1$ and $\hat{\beta}_1$ of alpha and beta for portfolio 1, the ordinary least squares estimates α_1 of alpha, the OLSE b_1^-, b_1, and b_1^+ of the lagged current, and lead betas, the sampling estimates $\hat{\rho}_M$ of the first-order autocorrelation coefficient for the market index, and the sampling estimates $\hat{\omega}_1$ and $\hat{\rho}_1$ of the residual standard deviation and associated first-order autocorrelation coefficient Throughout the tables asymptotic standard errors appear in parentheses below the corresponding estimates [14]

Examining table 1, it is clear that the portfolio of securities trading at the lowest levels of volume generates estimates $\hat{\beta}_1$ uniformly larger than the corresponding least squares estimates b_1 This discrepancy is reduced in table 2 and the inequality reversed in table 3 for portfolios of securities trading at progressively higher levels of volume As predicted by the previous theory, the result holds if the value-weighted market portfolio is heavily weighted with securities trading on average relatively frequently In this likely situation, portfolio 5 has an OLSE for beta asymptotically biased upward By contrast the predicted relationship between the consistent estimator and the OLSE for alphas cannot be verified directly from tables 1 through 3 The standard errors are too large

The relationship in tables 1 through 3 between consistent estimates of betas and ordinary least squares estimates is partly obscured by the apparent simultaneous relationship between true betas and trading volume In particular, in the tables larger consistent estimates of betas are associated with larger trading volumes For portfolios 1, 3, and 5, the consistent estimates of betas averaged across all 13 years are 0 674, 1 116, and 1 368, respectively Adjusting for these differences in consistent estimates by computing the ratios $(b_{nt}-\hat{\beta}_{nt})/\hat{\beta}_{nt}$, $n = 1, 3, 5$, for all years $t = 1963, \quad , 1975$, then gives a rough measure of the relationship between asymptotic bias and trading volume The results are displayed in figure 2 As indicated, average values for the ratios differ dramatically, increasing from $-0\ 218$ for portfolio 1 to $-0\ 082$ and $0\ 050$ for portfolios 3 and 5, respectively

The remaining estimates in tables 1 through 3 also appear consistent with the previous theory In switching from portfolio 1 to portfolio 3 to portfolio 5, the lagged betas b_n^- decrease, the lead betas b_n^+ increase, and the residual autocorrelation coefficients $\hat{\rho}_n$ decrease By contrast, for the residual standard errors $\hat{\omega}_n$ no clear trend is evident Finally, for the market index the autocorrelation coefficients $\hat{\rho}_M$ are generally significantly positive Throughout,

[14]The calculation of the standard errors for a_n, b_n^-, b_n, and b_n^+ is similar to the derivation of (25) and (26) The authors will provide details upon request

Table 1

Daily returns on low-volume portfolio regressed on value-weighted market returns

Year	$\hat{a}_1$	$\hat{\beta}_1$	a_1	b_1^-	b_1	b_1^+	ρ_m	$\hat{\omega}_1$	ρ_1
1963	−0 0000	0 544	0 0002	0 130	0 303	0 049	−0 058	0 003	0 012
	(0 0002)	(0 065)	(0 0002)	(0 035)	(0 031)	(0 036)	(0 063)		
1964	0 0002	0 561	0 0003	0 216	0 391	0 090	0 122	0 002	−0 019
	(0 0001)	(0 065)	(0 0001)	(0 045)	(0 040)	(0 047)	(0 063)		
1965	0 0005	0 647	0 0006	0 352	0 524	0 045	0 212	0 002	0 025
	(0 0002)	(0 052)	(0 0002)	(0 043)	(0 037)	(0 049)	(0 062)		
1966	−0 0002	0 581	−0 0003	0 391	0 426	0 102	0 291	0 003	0 055
	(0 0002)	(0 040)	(0 0002)	(0 032)	(0 031)	(0 040)	(0 061)		
1967	0 0011	0 651	0 0012	0 267	0 556	−0 015	0 120	0 003	0 161
	(0 0002)	(0 054)	(0 0002)	(0 045)	(0 035)	(0 047)	(0 063)		
1968	0 0009	0 775	0 0010	0 462	0 600	0 125	0 266	0 003	0 248
	(0 0003)	(0 061)	(0 0003)	(0 048)	(0 045)	(0 057)	(0 065)		
1969	−0 0010	0 872	−0 0010	0 620	0 749	0 183	0 390	0 004	0 216
	(0 0003)	(0 049)	(0 0003)	(0 044)	(0 041)	(0 062)	(0 060)		
1970	−0 0005	0 809	−0 0005	0 565	0 679	0 185	0 383	0 006	0 416
	(0 0005)	(0 056)	(0 0005)	(0 045)	(0 043)	(0 058)	(0 059)		
1971	−0 0000	0 993	0 0001	0 526	0 848	0 232	0 308	0 004	0 314
	(0 0003)	(0 066)	(0 0003)	(0 063)	(0 052)	(0 071)	(0 061)		
1972	−0 0003	0 661	−0 0002	0 364	0 596	0 121	0 317	0 003	0 536
	(0 0003)	(0 076)	(0 0003)	(0 057)	(0 054)	(0 059)	(0 061)		
1973	−0 0011	0 657	−0 0012	0 435	0 499	0 071	0 265	0 006	0 308
	(0 0004)	(0 056)	(0 0004)	(0 041)	(0 041)	(0 048)	(0 062)		
1974	−0 0002	0 431	−0 0003	0 307	0 346	0 024	0 284	0 006	0 493
	(0 0006)	(0 053)	(0 0005)	(0 032)	(0 033)	(0 036)	(0 061)		
1975	0 0017	0 577	0 0019	0 402	0 415	0 057	0 258	0 006	0 465
	(0 0005)	(0 067)	(0 0005)	(0 044)	(0 045)	(0 047)	(0 061)		

Table 2

Daily returns on intermediate-volume portfolio regressed on value-weighted market returns

Year	$\hat{a}_3$	$\hat{\beta}_3$	a_3	b_3^-	b_3	b_3^+	$\hat{\rho}_M$	$\hat{\omega}_3$	$\hat{\rho}_3$
1963	0 0001	0 905	0 0002	0 005	0 785	−0 039	−0 058	0 003	0 111
	(0 0002)	(0 065)	(0 0002)	(0 059)	(0 034)	(0 058)	(0 063)		
1964	0 0003	0 851	0 0003	0 233	0 754	0 071	0 122	0 002	0 001
	(0 0001)	(0 066)	(0 0001)	(0 060)	(0 040)	(0 063)	(0 063)		
1965	0 0007	1 202	0 0007	0 418	1 119	0 174	0 212	0 003	0 114
	(0 0002)	(0 059)	(0 0002)	(0 077)	(0 044)	(0 082)	(0 062)		
1966	0 0002	1 149	0 0001	0 565	1 005	0 248	0 291	0 004	−0 027
	(0 0002)	(0 045)	(0 0003)	(0 064)	(0 036)	(0 075)	(0 061)		
1967	0 0012	1 112	0 0012	0 212	1 123	0 045	0 120	0 003	0 222
	(0 0002)	(0 065)	(0 0002)	(0 079)	(0 046)	(0 081)	(0 063)		
1968	0 0008	1 274	0 0008	0 501	1 187	0 264	0 266	0 003	0 298
	(0 0003)	(0 065)	(0 0003)	(0 084)	(0 053)	(0 091)	(0 065)		
1969	−0 0005	1 330	−0 0006	0 690	1 257	0 421	0 390	0 003	0 079
	(0 0002)	(0 040)	(0 0002)	(0 077)	(0 035)	(0 089)	(0 060)		
1970	−0 0004	1 305	−0 0004	0 638	1 248	0 418	0 383	0 005	0 350
	(0 0004)	(0 047)	(0 0004)	(0 076)	(0 041)	(0 087)	(0 059)		
1971	0 0000	1 386	0 0001	0 516	1 296	0 428	0 308	0 003	0 105
	(0 0002)	(0 044)	(0 0002)	(0 084)	(0 035)	(0 088)	(0 061)		
1972	−0 0003	1 021	−0 0003	0 381	0 989	0 299	0 317	0 003	0 372
	(0 0002)	(0 059)	(0 0002)	(0 072)	(0 049)	(0 074)	(0 061)		
1973	−0 0007	1 142	−0 0009	0 564	0 983	0 199	0 265	0 006	0 142
	(0 0004)	(0 051)	(0 0004)	(0 064)	(0 042)	(0 075)	(0 062)		
1974	0 0001	0 830	−0 0001	0 462	0 724	0 116	0 284	0 007	0 175
	(0 0005)	(0 050)	(0 0005)	(0 047)	(0 036)	(0 058)	(0 061)		
1975	0 0010	0 996	0 0012	0 481	0 857	0 171	0 258	0 006	0 342
	(0 0005)	(0 061)	(0 0004)	(0 060)	(0 046)	(0 067)	(0 061)		

Table 3

Daily returns on high-volume portfolio regressed on value-weighted market returns

Year	$\hat{a}_5$	$\hat{\beta}_5$	a_5	b_5^-	b_5	b_5^+	$\hat{\rho}_M$	$\hat{\omega}_5$	$\hat{\rho}_5$
1963	−0 0000	1 336	−0 0002	−0 217	1 495	−0 097	−0 058	0 003	0 045
	(0 0002)	(0 067)	(0 0002)	(0 100)	(0 039)	(0 101)	(0 063)		
1964	0 0001	1 290	0 0001	0 049	1 355	0 199	0 122	0 002	0 217
	(0 0002)	(0 077)	(0 0002)	(0 096)	(0 055)	(0 095)	(0 063)		
1965	0 0010	1 501	0 0009	0 204	1 597	0 337	0 212	0 003	0 121
	(0 0002)	(0 066)	(0 0002)	(0 110)	(0 053)	(0 107)	(0 062)		
1966	0 0008	1 564	0 0008	0 297	1 725	0 452	0 291	0 005	0 107
	(0 0003)	(0 056)	(0 0003)	(0 115)	(0 051)	(0 109)	(0 061)		
1967	0 0011	1 369	0 0009	−0 122	1 602	0 219	0 120	0 004	−0 010
	(0 0003)	(0 075)	(0 0003)	(0 113)	(0 055)	(0 109)	(0 063)		
1968	0 0003	1 468	0 0002	0 281	1 520	0 449	0 266	0 004	0 240
	(0 0003)	(0 065)	(0 0003)	(0 110)	(0 055)	(0 104)	(0 065)		
1969	−0 0001	1 501	−0 0001	0 458	1 531	0 682	0 390	0 003	−0 004
	(0 0002)	(0 032)	(0 0002)	(0 102)	(0 027)	(0 090)	(0 060)		
1970	−0 0003	1 437	−0 0003	0 418	1 473	0 647	0 383	0 004	0 101
	(0 0003)	(0 033)	(0 0003)	(0 096)	(0 029)	(0 086)	(0 059)		
1971	−0 0000	1 441	−0 0000	0 349	1 445	0 535	0 308	0 002	0 125
	(0 0002)	(0 033)	(0 0002)	(0 095)	(0 027)	(0 087)	(0 061)		
1972	−0 0004	1 267	−0 0004	0 240	1 314	0 516	0 317	0 002	0 262
	(0 0002)	(0 048)	(0 0002)	(0 091)	(0 042)	(0 081)	(0 061)		
1973	−0 0001	1 314	−0 0001	0 316	1 318	0 375	0 265	0 003	0 308
	(0 0002)	(0 030)	(0 0002)	(0 085)	(0 026)	(0 084)	(0 062)		
1974	0 0002	1 120	0 0002	0 320	1 134	0 303	0 284	0 003	0 283
	(0 0003)	(0 025)	(0 0003)	(0 071)	(0 020)	(0 072)	(0 061)		
1975	0 0002	1 172	0 0001	0 290	1 174	0 312	0 258	0 003	0 161
	(0 0002)	(0 026)	(0 0002)	(0 075)	(0 021)	(0 075)	(0 061)		

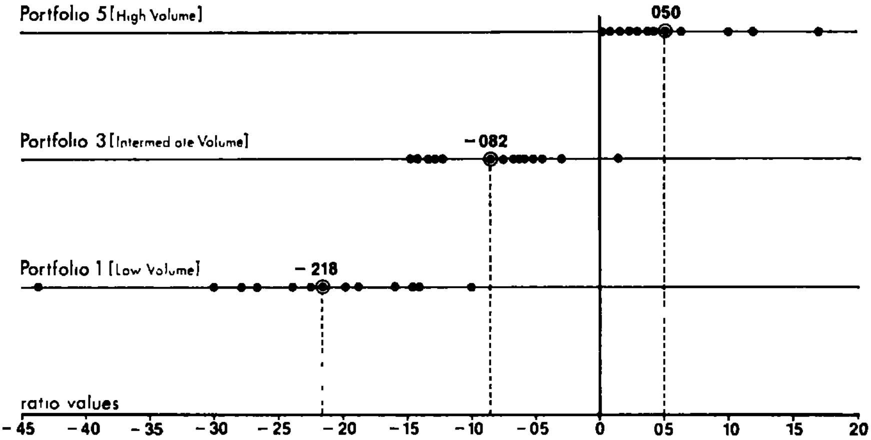

Fig 2 Measured betas vs true betas Values of the ratios $(b_{nt}-\beta_{nt})/\beta_{nt}$ for portfolios $n = 1, 3, 5$, and years $t = 1963, \ldots, 1975$

these results are consistent with errors of observation which are more significant for portfolios trading at lower levels of volume.

6. Summary

Given the recent availability of daily data, more powerful tests of the capital asset pricing model are now possible With daily data, however, there appears a potentially serious econometric problem. Because reported closing prices typically represent trades prior to the actual close of the trading day, measured returns often deviate from true returns. The resulting nonsynchronization of measured returns across different securities in turn introduces into the market model the econometric problem of errors in variables With daily data this problem is especially severe

As expected, with nonsynchronous trading of securities, ordinary least squares estimators of coefficients in the market model are both biased and inconsistent In this paper the directions and magnitudes of these asymptotic biases are specified in detail and then used to construct consistent estimators of alpha and beta These consistent estimators are subsequently applied to daily returns from specially constructed portfolios comprised of all stocks listed on the New York and American Stock Exchanges

Appendix

(i) Means, variances, and covariances for measured returns are

$$E[r^s_{nt}] = (1 - E[s_{nt} - s_{nt-1}])\mu_n, \tag{A1}$$

$$\begin{aligned}\operatorname{cov}(r^s_{nt}, r^s_{mt}) = {} & (1 - E[\max\{s_{nt}, s_{mt}\} - \min\{s_{nt-1}, s_{mt-1}\}])\sigma_{nm} \\ & + \operatorname{cov}(s_{nt} - s_{nt-1}, s_{mt} - s_{mt-1})\mu_n\mu_m,\end{aligned} \tag{A2}$$

$$\begin{aligned}\operatorname{cov}(r^s_{nt}, r^s_{mt-1}) = {} & E[\max\{s_{nt-1} - s_{mt-1}, 0\}]\sigma_{nm} \\ & + \operatorname{cov}(s_{nt} - s_{nt-1}, s_{mt-1} - s_{mt-2})\mu_n\mu_m,\end{aligned} \tag{A3}$$

and

$$\operatorname{cov}(r^s_{nt}, r^s_{mt-\tau}) = \operatorname{cov}(s_{nt} - s_{nt-1}, s_{mt-\tau} - s_{mt-\tau-1})\mu_n\mu_m, \tag{A4}$$

for $\tau = 2, \ldots, t-1$ and $n, m = 1, \ldots, N$.

To verify (A1) examine fig. 1. The dots on each line indicate times of actual trades, with the time of the last trade in each period labeled explicitly. Conditional on S_t and S_{t-1} the length of the trading period for r^s_{nt} is $1 - s_{nt} + s_{nt-1}$. With an infinitely divisible pricing process, this implies

$$\begin{aligned}E[r^s_{nt}] &= E[E[r^s_{nt} \mid S_t, S_{t-1}]] \\ &= E[(1 - s_{nt} + s_{nt-1})\mu_n],\end{aligned}$$

which equals (A1).

The variances and contemporaneous covariances (A2) are calculated as follows. Conditional on S_t and S_{t-1}, the length of the period of overlap between r^s_{nt} and r^s_{mt} is

$$\begin{aligned}&\min\{1 - s_{nt}, 1 - s_{mt}\} + \min\{s_{nt-1}, s_{mt-1}\} \\ &\quad = 1 - [\max\{s_{nt}, s_{mt}\} - \min\{s_{nt-1}, s_{mt-1}\}].\end{aligned}$$

With an infinitely divisible pricing process, this implies

$$\begin{aligned}\operatorname{cov}(r^s_{nt}, r^s_{mt}) &= E[\operatorname{cov}(r^s_{nt}, r^s_{mt} \mid S_t, S_{t-1})] \\ &\quad + \operatorname{cov}(E[r^s_{nt} \mid S_t, S_{t-1}], E[r^s_{mt} \mid S_t, S_{t-1}]) \\ &= E[(1 - [\max\{s_{nt}, s_{mt}\} - \min\{s_{nt-1}, s_{mt-1}\}])\sigma_{nm}] \\ &\quad + \operatorname{cov}((1 - s_{nt} + s_{nt-1})\mu_n, (1 - s_{mt} + s_{mt-1})\mu_m),\end{aligned}$$

which equals (A2). The derivations of (A3) and (A4) are similar.

(*ii*) The computation of the kurtosis $\kappa(r^s_{nt})$ is somewhat more lengthy. A sketch of the derivation is as follows. Again suppose $\{S_t\}$ is stationary. In this case it follows from (16) that

$$\operatorname{var}(r^s_{nt}) = \sigma^2_n + \operatorname{var}(s_{nt} - s_{nt-1})\mu^2_n. \tag{A5}$$

Similarly, because r_{nt} is symmetric, it can be verified that

$$\begin{aligned}\mu_4(r^s_{nt}) &= E[E[(r^s_{nt}-E[(r^s_{nt}|S_t, S_{t-1})])^4|S_t, S_{t-1}]]\\ &\quad +6E[\mathrm{var}(r^s_{nt}|S_t, S_{t-1})(E[r^s_{nt}|S_t, S_{t-1}]-E[r^s_{nt}])^2]\\ &\quad +E[(E[r^s_{nt}|S_t, S_{t-1}]-E[r^s_{nt}])^4]\\ &= (1+\mathrm{var}(s_{nt}-s_{nt-1}))\mu_4(r_{nt})\\ &\quad +6(\mathrm{var}(s_{nt}-s_{nt-1})-\mu_3(s_{nt}-s_{nt-1}))\mu_n^2\sigma_n^2\\ &\quad +\mu_4(s_{nt}-s_{nt-1})\mu_n^4, \end{aligned} \tag{A6}$$

where $\mu_3(\ )$ and $\mu_4(\cdot)$ represent third and fourth central moments Squaring (A5), dividing the result into (A6), rearranging terms, and recognizing $\mu_3(s_{nt}-s_{nt-1}) = 0$ yields (10)

(*iii*) To derive (16) and (17) recognize that (6) and (7) imply

$$\begin{aligned}\mathrm{cov}(r^s_{nt}, r^s_{mt}) &= (1-E[\max\{s_n, s_m\}-\min\{s_n, s_m\}])\sigma_{nm}\\ &\quad +2\mathrm{cov}(s_n, s_m)\mu_n\mu_m\\ &= \sigma_{nm}-\{E[\max\{s_n-s_m, 0\}]\sigma_{nm}+\mathrm{cov}(s_n, s_m)\mu_n\mu_m\}\\ &\quad -\{E[\max\{s_m-s_n, 0\}]\sigma_{nm}+\mathrm{cov}(s_n, s_m)\mu_n\mu_m\}\\ &= \mathrm{cov}(r_{nt}, r_{mt})-\mathrm{cov}(r^s_{nt}, r^s_{mt-1})-\mathrm{cov}(r^s_{nt-1}, r^s_{mt}) \end{aligned} \tag{A7}$$

Multiplying both sides of (A7) by x_{nM} and summing over $n = 1, \ \ , N$ yields

$$\begin{aligned}\mathrm{cov}(r^s_{nt}, r^s_{Mt}) &= \mathrm{cov}(r_{nt}, r_{Mt})-\mathrm{cov}(r^s_{nt}, r^s_{Mt-1})\\ &\quad -\mathrm{cov}(r^s_{nt-1}, r^s_{Mt})\end{aligned} \tag{A8}$$

Multiplying in turn both sides of (A8) by x_{nM} and again summing over $n = 1, \ \ , N$ gives

$$\mathrm{var}(r^s_{Mt}) = \mathrm{var}(r_{Mt})-2\mathrm{cov}(r^s_{Mt}, r^s_{Mt-1}) \tag{A9}$$

Dividing (A8) by (A9), rearranging terms, and exploiting the definitions (4), (13), (14), and (15) yields the desired result

(*iv*) The asymptotic standard errors (25) and (26) are calculated as follows From (19) it follows that

$$\frac{1}{T-2}\operatorname{plim}[\sqrt{(T-2)}(\hat{\alpha}_n-\alpha_n)]^2$$

$$=\frac{1}{T-2}\left\{\operatorname{plim}[\sqrt{(T-2)}(\hat{\beta}_n-\beta_n)]^2\left[\operatorname{plim}\frac{1}{T-2}\sum r^s_{Mt}\right]^2\right.$$

$$-2\operatorname{plim}(\hat{\beta}_n-\beta_n)\operatorname{plim}\left[\frac{1}{T-2}\sum_t\sum_\tau r^s_{Mt}\varepsilon^s_{n\tau}\right]$$

$$\left.+\operatorname{plim}\left[\frac{1}{T-2}\sum_t\sum_\tau \varepsilon^s_{nt}\varepsilon^s_{n\tau}\right]\right\}$$

$$=\left\{\frac{1}{T-2}\operatorname{plim}[\sqrt{(T-2)}(\hat{\beta}_n-\beta_n)]^2\right\}\mu^2_{Mt}+0+\frac{1}{T-2}\omega^{s^2}_n(1+2\rho^s_n),$$

which is equivalent to (25) Similarly, (20) and (22) imply

$$\frac{1}{T-2}\operatorname{plim}[\sqrt{(T-2)}(\hat{\beta}_n-\beta_n)]^2$$

$$=\frac{1}{T-2}\;\frac{\operatorname{plim}\dfrac{1}{T-2}\sum_t\sum_\tau r^s_{M3t}\varepsilon^s_{nt}\varepsilon^s_{n\tau}r^s_{M3\tau}}{\left[\operatorname{plim}\dfrac{1}{T-2}\sum_t r^s_{Mt}r^s_{M3t}\right]^2}$$

$$=\frac{1}{T-2}\;\frac{\omega^{s^2}_n\operatorname{var}(r^s_{M3t})+2\omega^{s^2}_n\rho^s_n\operatorname{cov}(r^s_{M3t},r^s_{M3t-1})}{[\operatorname{cov}(r^s_{Mt},r^s_{M3t})]^2},$$

which, with (23) and (24), is equivalent to (26)

References

Clark, P, 1973, A subordinated stochastic process model with finite speculative prices, Econometrica 41, 135–155

Cohen K, S Maier, R Schwarts and D Whitcomb, 1977, On the existence of serial correlation in an efficient market, Working Paper no 199, Graduate School of Business Administration, Duke University

Fama E, 1965 Tomorrow on the New York Stock Exchange, Journal of Business 38, 285–299

Fisher, L, 1966, Some new stock market indices, Journal of Business 39, 191–225

Merton R, 1973, An intertemporal capital asset pricing model, Econometrica 16, 868–887

NYSE Market Data Systems Monthly Memoranda, Statistical Reference Tables, July 1976, Office of Economic Research, Securities and Exchange Commission

Schwert, G, 1977, Stock exchange seats as capital assets, Journal of Financial Economics 4, 51–78