EXPECTED RETURN AND BETA ESTIMATION INCORPORATING BID/ASK DATA

by

Simon Wai Hung Ng

B.B.A.(HONS.), SIMON FRASER UNIVERSITY, 1986.

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF BUSINESS ADMINISTRATION

in the Faculty

of

Business Administration

© Simon Wai Hung Ng 1988

SIMON FRASER UNIVERSITY

August 1988

**APPROVAL**

Name:   Simon Wai Hung Ng

Degree: Master of Business Administration

Title of thesis:

    Expected Return and Beta Estimation Incorporating Bid/Ask Data

Examining Committee:

    Chairman: Prof. C. Ernest Love

    Supervisory committee:

Prof. Peter E. Kennedy
Senior Supervisor

Prof. John P. Herzog
Supervisor

Prof. Pao Lun Cheng
External Examiner

Date Approved: August 9,1988

## PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

Expected Return and Beta Estimation
Incorporating Bid/Ask Data.

Author:

(signature)

Simon Wai Hung Ng

(name)

15th Aug., 88.

(date)

## ABSTRACT

When shares are characterized by missing price data, researchers have chosen either to ignore information contained in the bid-ask prices or to assume that the unobservable true price is the mean of the bid and ask prices. The former technique produces an unbiased estimate, but the latter is only unbiased if the expected value of the unobservable stock price is in fact given by the mean of the bid and ask prices. In this paper, the difference between these two techniques are exploited to test the hypothesis that the mean of the bid and ask prices is equal to the expected value of the unobservable stock price, one of the most commonly used ad hoc assumptions in the financial literature. This test is supplemented by using a double-limit maximum likelihood estimation technique, since this technique also is only unbiased if the mean of the bid and ask prices is equal to the expected value of the unobservable stock price. The empirical results suggest that the hypothesis under test should be rejected.

## ACKNOWLEDGMENTS

# DEDICATION

To Mom, Dad and Grandmother.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

Whether or not Keynes was correct in his claims that policy makers are "distilling their frenzy" from economists, it appears that some economists have been distilling their policy implications from fables... . Thus to assume the state of the world to be as one sees fit is not even to compare the ideal with the actual but, rather, to compare the ideal with a fable.

The Fable of the Bees:

An Economic Investigation.

Steven N.S. Cheung (1973)

# I) Introduction

Financial researchers habitually overlook the implications of missing stock price data.[1] Except by employing some arbitrary rules or ad hoc assumptions to fill in the missing stock prices,[2] no serious attempt has been made to investigate or cope with this missing data problem. The existing treatments have been designed to facilitate the estimation problem rather than exploit the information carried by the fact that there are prices missing.

*Why is it so important for researchers to deal with missing stock prices?* Empirical investigations of stock price behavior and information effects of event studies often involve data sets characterized by infrequently-traded securities.[3] Price-adjustment delays caused by non-trading provide new insights into the existence of serial correlation in an efficient securities market. Empirical measurement of the systematic

---

[1]Stock price is considered as missing if there is no trade during a specified trading period. For example, a daily stock price is defined as missing if there is no trade for the entire day. Throughout the whole paper, phrases like thin-trading, infrequent trading or non-trading all refer to stocks with missing price data.

[2]A detailed review of the conventional ad hoc treatments of missing stock price data will be presented in section III of the paper.

[3]A number of studies have pointed out the pervasiveness of the infrequent trading phenomenon in different organized stock exchanges. For example, the American, Australian, British, Canadian, French, Japanese, O.T.C., and Scandinavian stock exchanges reported by Scholes and Williams (1977) and James and Edmister (1983), Officer (1975) and Ball, Brown and Finn (1977), Franks, Broyles and Hecht (1977) and Dimson (1979), Rorke and Love (1974a) and Fowler, Rorke and Jog (1980), Altman, Jacquillat and Lavasseur (1974), Lau, Quay and Ramsey (1974), Benston and Hagerman (1974) and Stoll (1978), and Berglund, Wahlroos and Örnmark (1983), respectively.

(or beta) risk for a security is subject to severe biases when shares are thinly-traded. Estimation of risk measures in the presence of non-trading provides a possible explanation of the alleged size-related anomalies ( or "small firm" effect), recently revealed by Banz (1981) and Reinganum (1981). Trading activity and the observed number of security returns available serves as an important proxy for measuring differential information which directly affects security returns and optimal portfolio choice. There are many financial puzzles related to the relevancy of missing stock price data! If the objective of any financial research is to understand the workings of the actual financial system, the phenomenon of missing stock price data should not escape our attention. We cannot change the real world to fit researchers' pure imagination by simply assuming that any imperfection is not important.

The purpose of this paper is to review the drawbacks of conventional treatments of missing stock price data, and present a maximum likelihood (ML) estimator which incorporates bid-ask spreads on the missing stock prices without imposing ad hoc assumptions on the true values of the missing stock prices. Calculation of the ML estimator provides an alternative means of testing the hypothesis that the mean of the bid-ask spread is equal to the expected value of the unobservable stock price,[4] one

---

[4]As with other tests of phenomena associated with the market model, the tests in this thesis, strictly speaking, must be viewed as joint tests of the particular hypothesis and the validity of the market model.

of the most commonly used ad hoc assumptions in the financial literature.[5]

The paper is organized as follows. In section II, a review of problems associated with the missing stock price data is provided. Section III discusses the conventional treatments for handling missing stock price data and their underlying weaknesses. Section IV presents the return generating process and the maximum likelihood (ML) estimator used in this paper to incorporate bid-ask spreads on the missing data, and constructs the associated testable hypothesis. Section V defines the data and methodology used in testing the stated hypothesis as well as comparing the maximum likelihood estimator with ad hoc variants of the ordinary least squares estimator. Section VI presents the empirical results. The findings suggest that the mean of bid-ask as a proxy of the expected value of the unobservable stock price is unjustified; and the different ways for handling missing stock prices have a sigificant influence on expected returns and beta estimates. Finally, section VII contains a summary and some concluding remarks.

---

[5]This hypothesis is conventionally taken for granted as the basic assumption by financial scholars in·building models, for example,Demsetz (1968), Blume and Stambaugh (1983) and Roll (1984). Also, the Center for Research in Security Price (CRSP) tape, which is extensively used by most of the empirical studies, substitutes the mean of the bid-spreads for the no trade prices and calculates the corresponding stock returns, leaving the user unable to distinguish a return calculated from the bid-ask spread (no trade) or one calculated from actual trades.

## II) REVIEW OF PROBLEMS ASSOCIATED WITH NON-TRADING

Several interrelated lines of research have suggested that missing stock price data or, in general, price-adjustment delays due to non-trading profoundly affect the observed return generating process. If correct, this has important implications for the observed return moments, risk assessment, pricing of financial assets and information processing. Therefore, there is a need to examine the impacts of non-trading on these related areas of research.

### Serial Correlation and Misspecification of Risk Measures[6]

An efficient market is one in which price fully reflects available information. Different levels of efficiency depend on the set of information that is said to be fully reflected in prices.[7] Samuelson (1965) develops the theory of "fair game" which shows that stock prices fluctuate randomly, and that the corresponding stock returns are independently distributed. Of course, the existence of serial correlation on individual or portfolio stock return series is inconsisent with the theory of "fair game". Fisher (1966) first identifies the non-trading problem and suggests that the delays of transaction (or closing) prices of individual securities

---

[6]Serial correlation in stock returns may be caused by factors other than non-trading, for example, bid-ask spreads, measurement error due to price recording and/or pricing averaging. See, for example, Niederhoffer and Osborne (1966) and Roll (1984), Officer (1975) and Praetz (1976), and Working (1960) and Daniels (1966) respectively. Therefore, non-trading is a suffcient but not a necessary cause of serial correlation.

[7]An in depth discussion of the theory of market efficiency and evidence for its existence is given by Fama (1970, 1976).

from the aggregate market introduces positive autocorrelation into the market index (or, in general, portfolios) which are constructed from such share price data.[8] Roll (1981) provides a simple intuitive illustration of the so called "Fisher effect" which states that "the [Fisher] effect is easy to see when an entire day passes without a trade; then that day's implicit return will be recorded on the day when its first subsequent trade takes place. The return is correlated, of course, with the returns of other firms which did register trades on the first days. The autocorrelation thereby induced in a portfolio of such securities is completely spurious and is simply the result of a defect in our record of prices" (p. 884). Building on Fisher's scenario, Scholes and Williams (1977) point out that since "prices for most securities are reported only at distinct, random intervals, completely accurate calculation of returns over any fixed sequence period is virtually impossible" (p. 309). Therefore, recorded closing prices typically represent trades prior to the end of a corresponding measuring interval, and this causes the measured returns to deviate from the true returns. An econometric problem of errors in variables is introduced into the market model which not only causes individual securities to display first-order *negative* serial correlation while diversified portfolios display *positive* serial correlation, but also

---

[8]Dimson (1979) also demonstrates that the existence of positive first order auotcorrelation in a market index is highly related to non-trading. By comparing an equally weighted (EW) index (which gives greater weight to infrequently traded shares) with the value weighted Financial Times-Activities (FTA) All Share Index (which gives less weight to the infrequently traded shares), he shows that the first order serial correlation coefficient for the FTA index (.07) is much lower than for the EW index (.25).

biases the systematic risk (or beta) estimates.[9] The most important

contribution of the Scholes and Williams study is their formal introduction

of how non-trading may cause misspecification of risk measures and their

refined method in coping with the non-trading problem. Dimson (1979, 1985);

Fowler, Rorke and Jog (1979a, 1979b);[10] Fowler and Rorke (1983); and Cohen

et al. (1980, 1983a, 1983b)[11] have demonstrated extensive development of

beta estimation under thin-trading by introducing multiple time lags to

their estimators. Ironically, none of the proposed estimators designed for

coping with the non-trading can actually estimate individual return series

with serious missing return data. For example, given a stock return series

with a specified sampling period of 100 days, but 50 days of this stock

have no closing price. How can a researcher input this return series into

any of the proposed estimators? This practical problem not only gives rise

to the conventional ad hoc treatments of missing stock price data which

will be discussed in section II, but also demonstrates the need for a new

estimator which can input the existing data without making any ad hoc

assumption on the true values of the missing data.

---

[9]Under the classical linear regression model, the ordinary least squares
estimator requires the regressor(s) to be uncorrelated with the random
error terms; however, with errors in variables the error terms are no
longer uncorrelated with the independent variable(s), which leads to biased
estimates.

[10]By applying the data on the Toronto Stock Exchange, which is subject to
seriously infrequent trading problem, Fowler et al. (1979b) test Scholes
and Williams' (SW) estimator, Dimson's AC estimator and their own estimator
which is an extension of the SW estimator, and show that in general, none
of the estimators is better than the ordinary least square estimator.

[11]Cohen et al. suggest that "Fisher effect" is only one of the factors
causing the price-adjustment delays, and in general, trading frictions like
information, decision, and transaction costs can also introduce biases in
beta estimates. They propose a three-pass regression procedure to correct
for the associated biases in beta estimates.

## Small Firm Effect and Differential Information Model

Banz (1981) employs a methodology similar to Fama and MacBeth (1973) and finds a statistically significant negative relationship between average returns to stocks and the market value of the stocks after adjusting for risk. Reinganum (1981) corroborates the existence of the "size effect" after controlling for other empirical anomalies exhibited in stock return data. In proposing a possible explanation of the "small firm" effect, Roll (1981) suggests that the abnormal returns earned by small firms might be attributable to the misspecification of risk measures. This is because the stocks of small firms are comparatively more thinly traded than the stocks of larger firms, causing the estimated beta for daily stock returns to be downward biased.[12][13] James and Edmister (1983) also hypothesize that firm size may merely be a proxy for trading activity; in their view, the higher risk adjusted returns observed for small firms may simply reflect higher transaction costs associated with inactively traded firms (i.e. liquidity

---

[12]See Scholes and Willams (1977), Dimson (1979) and Cohen et al. (1980, 1983a, 1983b).

[13]By applying Dimson's (1979) AC estimator, Roll runs a multiple regression of the equally-weighted index on the 5 leading and 21 lagged S&P 500 index. The empirical results show that first 9 of the 11 lagged coefficients turn out to be highly significant and even further lags are sometimes significant. Overall, he concludes that "Trading infrequency seems to be a powerful cause of bias in risk assessments with short-internal data. Rather horrendous bias is induced in daily data and the bias is still large and significant with returns measured over intervals as long as one month. The mis-assessment of risk has the potential to explain why small firms, low price/earnings ratio firms, and possibly high dividend yield firms display large excess returns (after adjustment for risk)" (p. 887).

7

premia).[14] By employing average daily trading volume and number of trades as measures of trading activity, their empirical results, consistent with Dimson (1979), show that trading activity and firm size are highly correlated; but no evidence is found consistent with the existence of a liquidity premium. Barry and Brown (1984) investigate another possible explanation of the small firm effect based on a model of differential information.[15] They argue that "securities for which there is relatively little information available may be preceived as riskier than are securities for which more information is available. Commensurate with that risk, participants in the market may rationally demand a premium to hold such securities. If, so, and if risk is measured empirically without regard to the amount of information available, then there may appear to 'abnormal' returns for low information securities. To the extent that low market value firms have less information available, it follows that there would appear to be 'abnormal' returns associated with small firms" (p. 284). In a later paper, Barry and Brown (1985) formally derive the equilibrium implications of the differential information and security market. They show that if the

---

[14]Demsetz (1968), Tinic (1972) and others have argued that transactions costs vary inversely with the trading activity (as measured by average number of shares traded). See Roll (1983), Schultz (1983), Stoll and Whaley (1983) and Amihud and Mendelson (1986) for detail investigation of transaction costs as possible explanation of small firm effect.

[15]Barry and Brown's differential information arguement as applied for the small firm effect is not without ancestors. Klein and Bawa (1977) present a model in which the set of information is limited to the historical data. In such differential information settings, low information securities which is characterized by fewer observations than other securities have relatively higher estimation risk (i.e., uncertainy about the true parameter of the return distribution), and this leads risk averse investors to diversify away from such securities. Banz (1981) in his concluding remarks, based on Klein and Bawa's model, argues that the small firm effect might be due to the relatively limited information available about the smaller firms.

number of observations available is employed as a proxy for the degree of relative estimation risk, limited information can have an effect on systematic risk, and therefore on required returns. In an equilibrium setting, portfolios of low information securities will appear to earn positive abnormal returns, and high information securities will appear to earn negative abnormal returns if their betas are measured without regard for differential information and if their average returns are consistent with a CAPM that reflects investor perception of differential information.[16]

To sum up, the overall related issues as summarized in Figure 1, support the important role of missing stock price data in affecting both the theoretical and empirical development of financial research. In the next section, we review the conventional treatments of handling missing stock price data in the existing financial literature.

---

[16]Besides using the number of security returns observations available, they also propose period of listing and divergence of analyst opinion to serve as additional proxies for the relative information. However, except the period of listing which has been tested by Ibbotson (1975) and Clarkson and Wharton (1987), no existing financial literature has employed the other two proposed proxies.

# Figure 1

## Problems Related to Missing Stock Price Data

```
┌─────────────────────┐
│   Fisher effect:    │
│  price-adjustment   │
│       delays        │
└─────────────────────┘
                        ╲

┌─────────────────────┐      ┌─────────────────────┐      ┌─────────────────────┐
│ Differential info.  │      │   Misspecification  │      │      Possible       │
│ : no. of observs.   │ ───→ │         of          │ ───→ │   explanation of    │
│      as proxy       │      │    risk measure     │      │  small firm effect  │
└─────────────────────┘      └─────────────────────┘      └─────────────────────┘

┌─────────────────────┐      ┌─────────────────────┐
│ Practical problem:  │      │   Misspecification  │
│    treatments of    │ ───→ │         of          │
│  missing returns    │      │  expected return    │
└─────────────────────┘      └─────────────────────┘
```
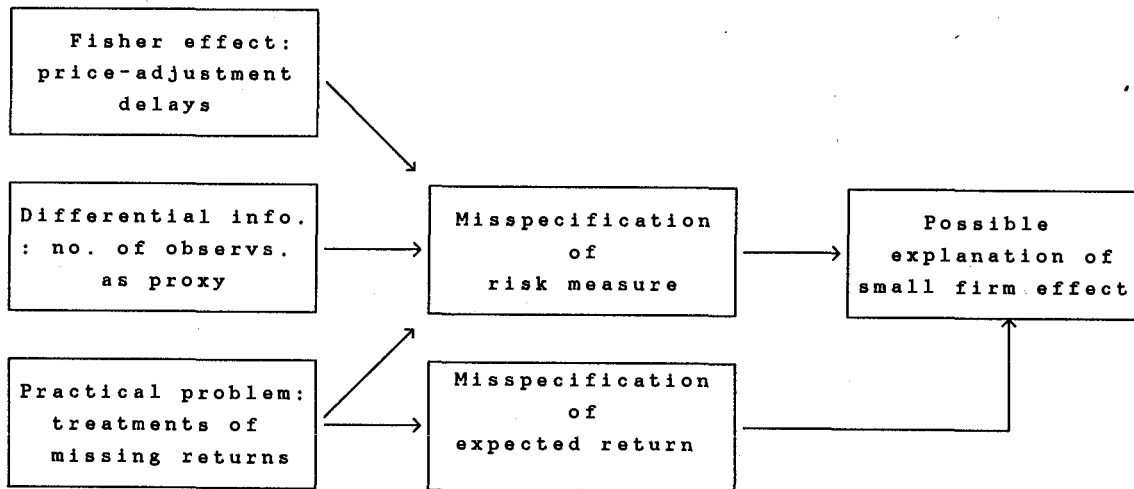
## III) CONVENTIONAL TREATMENTS OF MISSING STOCK PRICES

Let $P_{i,t}$ denote the transaction price (or closing price) of stock i for the trading period ending at time t. A single period return $(R_{i,t})$ of security i for period t is observed when there is a closing price on two successive trading dates, i.e., $R_{i,t} = (P_{i,t} - P_{i,t-1})/P_{i,t-1}$. If stock i has no trade for n-1 (n = {2,3,...}) consecutive periods, we will observe only $P_{i,t}$ and $P_{i,t+n}$ for the entire n+1 trading periods, and the associated n single period returns are undefined. There are three main approaches in the literature for handling missing stock prices.[17] [18]

The first main approach is simply to drop all returns associated with the missing stock prices and use only return data for which both successive trading prices, $P_{i,t-1}$ and $P_{i,t}$ are observed. Scholes and Williams (1977), Perry (1985) and Grube, Joy and Howe's (1987) approaches are some

---

[17]This list of common treatments in handling missing stock price data, of course, is not exhaustive. For pure statistical treatments, Maddala (1977, p.202-205) has suggested several other methods in handling missing observations for cross-section data; however, most of these cross-section data methods either assume all the missing values are equal or only provide estimates of the distribution moments, but not the estimate of the regression coefficient $\beta$. Since the main focus of this paper as well as the fundamental practice of the existing financial literature is concentrated on the use time-series return data to estimate the regression coefficient, $\beta$, none of these cross-section data methods is relevant or appliable. Also, given the stochastic nature of the stock returns, the assumption of equal value for all missing data is definitely undesirable.

[18]The first and second approaches are closely following the arguments suggested by Heinkel and Kraus (1985). In their paper, Heinkel and Kraus have constructed an iterative generating procedure to estimate and fill in the missing returns. Their fundamental argument "is whether the available data provide evidence about the information affecting that stock on that trading day could be used to provide a better estimate of return for that day than setting return to zero" (p.5).

examples. Scholes and Williams, in building their estimator to cope with the non-trading problem, explicitly state that

> "All information about returns over days in which no trades occur is ignored. This greatly simplifies the subsequent estimators" (p. 311 footnote 4).

Also, in specifying sample data for empirical investigation, they specified that

> "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" (p. 319).

Heinkel and Kraus (1985) point out that this approach "obviously suffers from the disadvantage of ignoring what is observed about stock's movement from [period] t to [period] t+n in the example above; at the extreme, a stock that was traded every second day would be ignored completely under this approach" (p.4). Dimson (1979) also argues that this approach leads to a loss of efficiency which can increase the size of the confidence interval about the beta estimates.[19] Given the fact that this approach ignores the multi-period price movements during the no trade periods, in general there is a loss of efficiency, although it provides consistent beta estimates.

---

[19]Based on efficiency as comparison criterion, Maddala (1977) p.201-207 points out that a least-squares estimator which discards all the missing observations is generally worse than any approaches based on some arbitrary rules to fill in the missing data. He suggests that it is not desirable to throw away the information associated with missing data.

The second approach is to assign the n-period return implied by $P_{i,t}$ and $P_{i,t+n}$ as single period returns for the period t+1 to t+n in some manner. One common practice is to allocate the entire n-period return to period t+n and correspondingly set all the single period returns for period t+1, t+2, ..., t+n-1 to zero.[20] This method implicitly assumes stock prices remain unchanged over the no trade periods. First, efficient market theory only suggests that security price reflects all available information, and new information causes the change in security price; however, no existing security market (or valuation) theory has implied that a change in asset values is necessarily accompanied by trading. Ignoring transaction costs, both Marshall (1974) and Rubinstein (1975) have demonstrated that under a state-preference economy setting, it is possible for new information to cause the change in security price without any trade being involved.[21] Second, if we take the operation of security markets into account, the bid-ask spreads on securities may widen in response to change in new information which prohibits any trade from taking place.[22] Building on an

---

[20]Fowler, Rorke and Joy (1980) and McInish and Wood (1986) are some of the examples in following this approach.

[21]Refer to Marshall (1974) for the case of pure exchange and Rubinstein (1975) for the case of non-speculative beliefs.

[22]Demestz (1968), Tinic and West (1974), Benston and Hagerman (1974) and Hamilton (1976) have shown that the bid-ask spread is inversely related to the volume of trading.

asymmetric information setting,[23] Glosten and Milgrom (1985) have shown that "if the insiders are too numerous and their information is too good relative to the elasticity of liquidity traders' supplies and demands, there will be no bid and ask prices at which trading can occur and the specialist can break-even. Then, the equilibrium bid price is set so low and the ask price so high as to preclude any trade" (p. 74). Copeland and Galai (1983) also suggest that "not all informed traders who arrive at the market place will consummate a trade. Non-traders are informed individuals who believe the post-trade price will fall between $K_A$ and $K_B$, the ask and bid prices, respectively" (p. 1461). Besides, Goldman and Sosin (1979) have demonstrated that given transaction costs, it is optimal for an investor or speculator to accummulate "new bits" of information for periodic review rather than continously to assess the import of each new bit as it arrives. Furthermore, even when assessments are made, Cohen et al. (1981) have shown that with transaction costs, an investor may find a "do-nothing" strategy to be optimal even if a trade would have been sought in a frictionless environment. With continuous monitoring of the market being prohibitively expensive, limit orders left with specialists can go "stale" without withdrawal. Therefore, it is possible for a security price to change in

---

[23]Asymmetric information is caused by traders' possessing special (or superior) information; they are commonly referred to as "insiders". Bagehot (1971) suggests that specialists always lose to pure inside information traders. The main idea ,is that a specialist faces an adverse selection problem, since the traders may trade at the specialist's bid or ask based on information which the specialist does not have. In order to recapture the loss suffered to insiders, the specialist widens up the bid-ask spread to incorporate the cost due to asymmetic information. This adversary information cost faced by specialists is modelled extensively by Grossman (1976), Jaffe and Winkler (1976), Copeland and Galai (1983), Klye (1985), Glosten and Milgrom (1985) and Glosten (1987). The existence of insiders has been empirically demonstrated by Lorie and Niederhoffer (1968), Scholes (1972) and Jaffe (1974).

response to new information, but the change in security price may not rise (or fall) enough to cause execution of the stale limit orders. In light of all this, it is clearly incorrect to conclude that security values do not change between trades, and that the associated returns are simply zero when there is no trade.

The third approach is to estimate the missing stock price data by interpolation. Instead of using the time series being interpolated, interpolation is often done by using a related series known for all relevant periods.[24] The most common practice of this general approach is to substitute the missing stock price with the average of the bid and ask prices.[25] Financial scholars often assume the average of the bid and ask prices as the underlying true value in building their models. For example, in modelling the transaction costs of stock exchange, Demsetz (1968) assumes that

"There exist two [market] equilibrium prices and not one - $A_i$ [the ask price] for immediate sales and $B_i$ [the bid price] for immediate purchase. $E_i$ [the conventional view of equilibrium price under no transaction cost] can be thought of as an arithmetic average of $A_i$ and $B_i$" (p. 37).

---

[24]Maddala (1977) points out that it is the dependence within the time series that allows the application of interpolation. However, given the stochastic and random nature of stock prices, it is unreasonable to use the remaining non-missing stock prices directly to interpolate the missing values. See Friedman (1962) for a detailed treatment and review of the existing literiture in dealing with interpolation of time series by related series.

[25]Blume and Stambaugh (1983), Theobald and Price (1984), Clarkson and Wharton (1987), Grube, Joy and Howe (1987) and the CRSP tape are some examples of following this approach.

Also, in modelling the effective bid-ask spread in an efficient market, Roll (1984) explicitly states that

> "We might think of "value" as being the center of the spread. When news arrives, both the bid and the ask prices move to different levels such that their average is the new equilibrium value. Thus, the bid-ask *average* fluctuates randomly in an efficient market" (p.1128).

In the same vein, in assessing the potential biases introduced by transaction prices on return calculation, Blume and Stambaugh (1983) provide a clear statement of their underlying common belief;

> "denote the true price as P and assume that the closing period is either a bid price, $P_B$, or an ask price, $P_A$, with equal probability. Since the expected closing price is assumed equal to the true price, P must be $(P_B + P_A)/2, ..."$ (p. 390).

Based on the above quotations, the conventional treatment of bid-ask average as the underlying true security price directly implies that:

1)    The causality of asset valuation flows from the transaction price to the underlying true asset value, i.e., the underlying true asset value can be determined by observing the transaction price, and

2)    transaction price is a random variable which only occurs in two states - the bid state and the ask state, with equal probability.

However, these two implications do not necessarily hold when shares are subjected to non-trading. When there is no trade, transaction

price is correspondingly undefined, and we only observe the bid and ask prices. Given there is no transaction price, the relevant question is whether we can unambigously estimate the true security value based on the bid and ask prices. First, a number of security market microstructure studies concerning the determination of bid-ask spread have pointed out that it is the market makers observing the market buy and sell orders, who try to estimate the underlying true security value and set the bid and ask prices accordingly.[26] It is clear that when there is no trade, the causality of security valuation flows from the formation of expected true security value to the determination of bid and ask prices, not from the opposite direction. At the same time, given the bid and ask prices per se, there is no unambigous way to specify the underlying true security values without making ad hoc assumptions on the return generating process, since these true values are simply unobservable. Second, when there is no trade, the true security price may equal any price straddled by the bid and ask prices. As Cohen et al. (1981) suggest "for the market, the spread is the difference between the lowest ask and the highest bid of all participants. In markets composed of many traders with heterogeneous beliefs and trading propensities, one might expect to have orders at virtually every permissible price in the neighborhood of equilibrium" (p. 288-9).[27] That is, instead of only occuring in two states - the bid price and the ask price, the true price can occur in an infinite number of states (or prices)

---

[26]See Bagehot (1971), Garbade and Silber (1979), Goldman and Beja (1979), Copeland and Galai (1983), Klye (1985) and Glosten and Milgrom (1985).

[27]This arguement is also well supported by Fisher(1966, p.198-9), Copeland and Galai (1983, p.1468), Gloston and Milgrom (1985, p.71 Propostion 1), Sharpe (1985, p.30) and French and Roll (1986, p.15).

between the quoted bid and ask prices. Also, the ad hoc assumption of equal probability in obtaining the bid and ask prices is unreasonable without careful empirical investigation. If the probability of obtaining the bid and ask prices is actually 50-50, the bid-ask averages should follow the "fair game" properties as proposed by Samuelson (1965), and correspondingly the return series contructed by these bid-ask averages should exhibit no autocorrelation. Hasbrouck and Ho (1987) provide direct empirical tests on the assumption that buy and sell orders arrive independently with equal probability and show that market buy and sell orders do not arrive independently but are in fact characterized by positive dependence. In addition, they also demonstrate that intra-day returns computed from the quote midpoints have exhibited strong serial correlation which is inconsistent with the "fair game" properties. They conclude that "in any event, the assumption of equiprobable buy and sell orders in the simple dealer model is quite suspect" (p.1041). Third, in assessing the true security value, several studies have pointed out that, in general, the relative position of the true security price within the bid-ask spread will change from time to time, instead of being fixed at the mean value. For example, in deriving the optimal pricing strategies of market makers Smidt (1971, 1979), Barnea (1974), Barnea and Logue (1975), Garman (1976), Stoll (1976, 1978), Ho and Stoll (1981) and Amihud and Mendelson (1980) have presented dynamic price-inventory adjustment policies which show how the specialists' quotes depend upon their inventory levels. They suggest that a specialist has a preferred inventory position and when his inventory level deviates from the optimal position, he will adjust the bid and ask price

18

relative to the "true" price to restore that optimal position.[28] As shown in Figure 2 (Ho and Stoll 1981, p.58 Fig. 3), instead of the absolute spread (s=a+b), it is the relative spread (d=a-b) that responds to the inventory changes. For a given expected true price (P), this implies that both the bid price ($P_B$) and ask price ($P_A$) fall in response to inventory increase and both rise in response to inventory decrease. It follows that unless the specialist's inventory level is always at the desired optimal level where no inventory adjustment is required, the bid-ask average is not a correct estimate of the true security value.

## Figure 2

### Dynamic Price-Inventory Adjustments
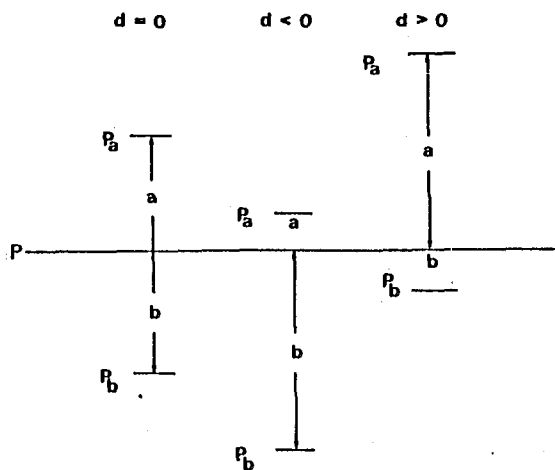


Fig. 3. Adjustment of bid price ($p_b$) and ask price ($p_a$) relative to 'true' price, $p$, with constant spread, $s = a + b$, and changing price adjustment, $d = a - b$.

___

[28]In this context, the "true" price (P) is the specialist's expected true value of the stock in the absence of transaction costs, at the time he set the bid-ask quotation. Denote a and b as the price of immediacy for a public buy order and a public sell order respectively. Instead of direct quotes a and b, the specialist quotes the bid ($P_B$ = P-b) and ask prices ($P_A$ = P+a).

Based on a totally different argument, Roll (1983) in assessing whether transaction costs can possibly explain the turn-of-the-year effect, points out that "most of tax sales near the year's end would be purchased by the specialist; the majority of transactions would occur near the low side of the bid/ask spread. After the new year, the trading would revert to the normal pattern of a roughly equal number of buyers and sellers and an average transaction price close to the center of the bid/ask spread" (p. 23).[29] Glosten (1987) also strikes a note of caution concerning the use of bid-ask spread data directly to correct for biases in estimators' of transaction return moments, and warns that "it would be inappropriate to take the average bid and ask spreads to adjust mean, variance, and serial covariance estimators without being assured that no part of the spread were due to presence (or believed presence) of traders with superior information" (p. 1294).[30] In addition, Working (1960), Cowles (1960), Daniels (1966), Rosenberg (1971), and Meyer and Corbeau (1975) have shown that the use of averages, particularly the first difference of the

---

[29]Keim (1983) provides new evidence on the seasonality of the "size effect" and points out that the average risk-adjusted return to a portfolio of small firm's stocks is much larger in January than the rest of the year. This new anomaly is often referred to as the turn-of-the-year effect or the January effect. A natural hypothesis to consider is that some investors sell securities at the end of the year to establish short-term capital losses for income tax purposes. If this 'selling pressure" depresses stock prices prior to the end of the year, then the stock prices during the first week of the subsequent year will rebound to a higer level. Also, see Brown and Keim (1983) and Reinganum (1983) for extensive arguements in related to the tax-loss selling hypothesis.

[30]Refer to footnote 23 of this paper for this asymmetric information argument.

midranges[31] will introduce spurious serial dependence into the original series. Therefore, the replacement of bid-ask averages for the no trade prices will introduce spurious autocorrelation to the original stock return series even if the original series is independently distributed. Based on all these arguments, it is clear that the use of bid-ask averages as no trade prices is not only inappropriate but undesirable.

In summary, the main criticism of the first approach (which skips all missing observations) is that although it provides unbiased estimates, there is an efficiency loss which increases the standard errors of the estimates. The main criticism of the second and third approaches (which use ad hoc assumptions to fill in the missing returns) is that although the incorporation of extraneous information based on the ad hoc assumptions will increase the efficiency of the estimation process, the residual errors associated with the missing observations are no longer random if this extraneous information is false. Thus, the direct application of the ordinary least squares estimation, given this false information, will provide inconsistent estimates of the parameter $\beta$. Clearly, there is a trade-off between efficiency and biasedness of the estimation process if the ad hoc assumptions in approximating the missing stock prices are false. Therefore, tests of the ad hoc assumptions for handling missing stock prices should provide important information for choosing the appropriate estimation procedure. The goal of this paper is to test the hypothesis that the mean of the bid-ask spreads is equal to the expected value of the

[31]Midrange is the average of the lowest and highest value of a variable in a given time interval. If we consider daily stock price data, the bid-ask average, which are the average of the respective lowest and highest price of a given trading day, can be treated as midrange.

unobservable stock price, one of the most commonly used ad hoc assumptions for handling missing stock prices and for building financial models. As shown in the next section, a double-limit maximum likelihood estimation technique is unbiased only if the mean of the bid and ask prices are equal to the expected value of the unobservable true price. Therefore, the testing of the biasedness of the maximum likelihood estimation technique provides an equivalent testing of the stated hypothesis.

## IV) TWO LIMIT MAXIMUM LIKELIHOOD ESTIMATOR

### True Return Generating Process : Market model

$$R^*_{it} = \alpha_i + \beta_i R^*_{mt} + \epsilon^*_{it} \tag{1}$$

The true returns on risky securities are assumed to be generated by the simple market model (1), where $R^*_{it}$ and $R^*_{mt}$ are the corresponding serially and cross-serially uncorrelated rates of return on the security i and the market index at time, t. $\beta_i = Cov(R^*_{it}, R^*_{mt})/Var(R^*_{mt})$ the stationary systematic risk of security i, and $\alpha_i$ is the constant intercept coefficient which is independent of time. The error term, $\epsilon^*_{it}$ is independently and normally distributed with a zero mean and constant variance, and is orthogonal to $R^*_{mt}$. If there is no missing price data, the estimates of $\alpha_i$ and $\beta_i$ are simply obtained by employing ordinary least squares estimator on the measured $R_{it}$ and $R_{mt}$. However, given are missing stock price data, some of the associated returns are not defined. As mentioned in the last section, the direct application of the ordinary least-square estimator, with ad hoc assumptions of the unobserved true returns, may produce biased estimates (if using the mean of the bid-ask) as well as inefficient estimates (if using only non-missing observations).

### Observed Return Generating Process with Bounds on Missing Stock Prices

Although the true returns associated with missing stock prices are unobservable, the corresponding upper and lower bounds can be determined by employing the bid-ask spreads on the missing stock prices. The

specification of the upper and lower bounds is to avoid ad hoc assumption on the true values of the unobserved returns. Consider the following observed return generating process:[32]

**Figure 3**

**Observed Return Generating Process**

$$P_t \longrightarrow P_{t+1} \quad \nearrow P^a_{t+2} \quad P^a_{t+3} \quad \searrow P_{t+4}$$

$$P^b_{t+2} \quad P^b_{t+3}$$

$|\longleftarrow$ Case 1 $\longrightarrow|\longleftarrow$ Case 2 $\longrightarrow|\longleftarrow$ Case 3 $\longrightarrow|\longleftarrow$ Case 4 $\longrightarrow|$
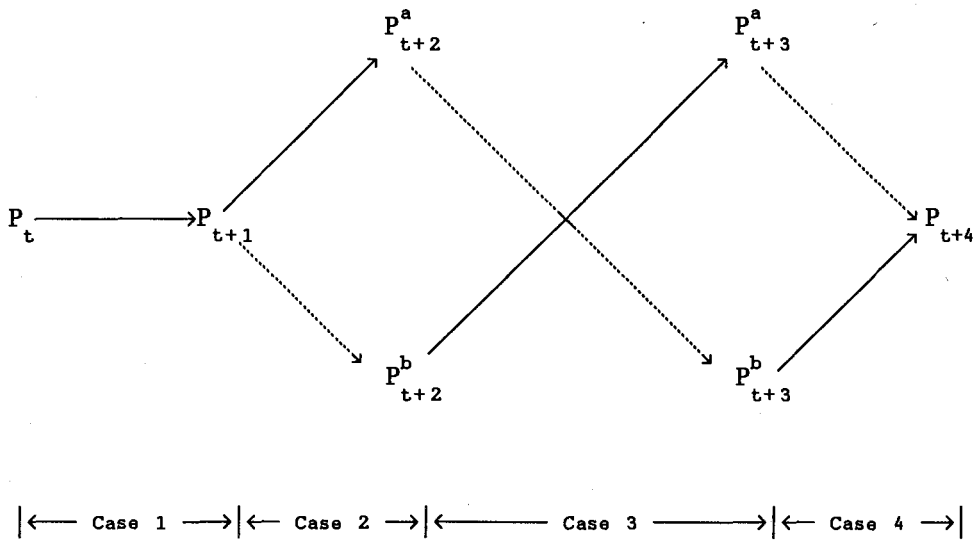
where $\nearrow$ and $\searrow$ are the respective maximum upward and minimum downward price movements,

$R^u_t$ and $R^\ell_t$ are upper and lower bounds of the observed returns associated with the maximum upward and minimum downward price movements at time t, respectively,

$P^a_t$ and $P^b_t$ are the respective ask and bid prices at time t, and

$P^*_t$ is the unobservable true price at time t.

---

[32]Throughout this paper, since we always focus on the estimation of beta for individual firms, the i subscript for individual firms will be supressed for notational convenience.

Assume the unobservable true price is straddled by the bid-ask spread, i.e., $P_t^b < P_t^* < P_t^a$, and the return on an individual security is defined as the logarithm of the successive price ratio, with all prices adjusted for dividends, stock dividends and stock splits, i.e., $R_t = \ln(P_t/P_{t-1})$. The upper bound $(R_t^u)$ and lower bound $(R_t^\ell)$ of the unobservable true return $(R_t^*)$ can be determined according to the specific forms of missing prices as follows:

Case 1: no missing both $P_t$ and $P_{t-1}$

$$R_t = \ln(P_t) - \ln(P_{t-1})$$

Case 2: only $P_t$ is missing, i.e., $R_t^\ell < R_t^* < R_t^u$

$$R_t^u = \ln(P_t^a) - \ln(P_{t-1})$$

$$R_t^\ell = \ln(P_t^b) - \ln(P_{t-1})$$

Case 3: both $P_t$ and $P_{t-1}$ are missing, i.e., $R_t^\ell < R_t^* < R_t^u$

$$R_t^u = \ln(P_t^a) - \ln(P_{t-1}^b)$$

$$R_t^\ell = \ln(P_t^b) - \ln(P_{t-1}^a)$$

Case 4: only $P_{t-1}$ is missing, i.e., $R_t^\ell < R_t^* < R_t^u$

$$R_t^u = \ln(P_t) - \ln(P_{t-1}^b)$$

$$R_t^\ell = \ln(P_t) - \ln(P_{t-1}^a)$$

## Two Limit Censored Regression Model

We assume that

$$R_t^* = \beta' X_t + \epsilon_t \qquad (2)$$

where $R_t^*$ is measured if there is no missing return, but if the return is missing, then $R_t^\ell < R_t^* < R_t^u$ and $R_t^*$ is measured by $A_t$, a value determined by an ad hoc assumption.

$$\beta = \begin{bmatrix} \alpha_0 \\ \beta_0 \end{bmatrix}, \qquad X_t = \begin{bmatrix} 1 \\ R_{mt} \end{bmatrix}, \qquad \epsilon_t \sim N(0, \sigma^2) \quad \text{i.i.d.}$$

Given the observed returns on individual firm $(R_t)$, market index $(R_{mt})$, upper bound $(R_t^u)$ and lower bound $(R_t^\ell)$ of the unobservable true returns generated by the bid-ask spreads on the missing stock price data, the simple market model can be replaced by a two-limit censored regression model which incorporates both missing and non-missing returns, to estimate $\alpha_0$, $\beta_0$ and $\sigma^2$. The intuitive rationale for the two-limit censored regression is very simple. When there is no missing price, the observed return is simply used as a measure of the true return generated by the market model. However, when there are missing prices, instead of using the stock returns based on market transactions which are unobserved, some arbitrary values determined by ad hoc assumptions are employed as a proxy for the true returns. Although the true returns are unobservable, their underlying values which are assumed to be generated by the market model are known to be bounded by $R_t^\ell$ and $R_t^u$. This implies that $R_t^\ell < \beta' X_t + \epsilon_t < R_t^u$ or $R_t^\ell - \beta' X_t < \epsilon_t < R_t^u - \beta' X_t$. Given the theoretical bounds in restricting the

values of the true return random errors, the estimation of this two-limit censored regression model can be done by using the maximum likelihood estimator on both unrestricted (i.e. non-missing return) and restricted (i.e. missing return) random error terms. Consider the regression model in equation (2), let $N_0$ be the number of missing observations and $N_1$ be the number of non-missing observations.

Define[33]

$$F(k) = F(\beta, \sigma^2) = \int_{-\infty}^{k} (2\pi\sigma^2)^{-1/2} e^{-(1/2\sigma^2)t^2} dt$$

$$f(t) = f(\beta, \sigma^2) = (2\pi\sigma^2)^{-1/2} e^{-(1/2\sigma^2)t^2}$$

where f(t) and F(k) are the normal density function and cumulative normal distribution with zero mean and variance $\sigma^2$ evaluated at t and k, respectively.

a) For the non-missing observations, we have:

$$\epsilon_t = R_t - \beta'X_t$$

$$\text{Prob } [\epsilon_t = R_t - \beta'X_t] = f(\epsilon_t) = (2\pi\sigma^2)^{-1/2} e^{-(1/2\sigma^2)\epsilon_t^2}$$

b) For the missing observations, we have:

$$\ell_t = R_t^{\ell} - \beta'X_t \text{ and } u_t = R_t^{u} - \beta'X_t$$

---

[33]The notation presented here is very similar to Maddala (1987) p. 151-152.

$$\text{Prob }[\ell_t < \epsilon_t < u_t] = \text{Prob }[R_t^\ell - \beta'X_t < \epsilon_t < R_t^u - \beta'X_t]$$

$$= \int_{R_t^\ell - \beta'X_t}^{R_t^u - \beta'X_t} f(\epsilon_t)\, dt$$

$$= \int_{-\infty}^{R_t^u - \beta'X_t} f(\epsilon)\, dt - \int_{-\infty}^{R_t^\ell - \beta'X_t} f(\epsilon)\, dt$$

$$= F(u_t) - F(\ell_t)$$

Hence, the likelihood function is:

$$L = L(\beta,\ \sigma^2 | R_t,\ X_t,\ R_t^u,\ R_t^\ell)$$

$$= \prod_1 (2\pi\sigma^2)^{-1/2} e^{-(1/2\sigma^2)(R_t - \beta'X_t)^2} \prod_0 [F(u_t) - F(\ell_t)] \tag{3}$$

where the first product ($\prod_1$) is over the $N_1$ observations for the non-missing $R_t$ and the second product ($\prod_0$) is over the $N_0$ observations for the missing $R_t$.

The corresponding log likelihood function is:

$$\ln L = -\frac{N_1}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_1 (R_t - \beta'X_t)^2 + \sum_0 \ln[F(u_t) - F(\ell_t)] \tag{4}$$

where the first summation ($\sum_1$) is over the $N_1$ observations for the non-missing $R_t$ and the second product ($\sum_0$) is over the $N_0$ observations for the missing $R_t$.

28

Maximize ln L with respect to $\beta$ and $\sigma^2$, we get the first-order conditions as:

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma^2}\sum_1 (R_t - \beta'X_t)X_t - \sum_0 X_t \left[ \frac{f(u_t) - f(\ell_t)}{F(u_t) - F(\ell_t)} \right] = 0 \qquad (5a)$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{N_1}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_1 (R_t - \beta'X_t)^2 - \frac{1}{2\sigma^2}\sum_0 \left[ \frac{u_t f(u_t) - \ell_t f(\ell_t)}{F(u_t) - F(\ell_t)} \right] = 0$$

Since $u_t = R_t^u - \beta'X_t$ and $\ell_t = R_t^\ell - \beta'X_t$

$$u_t f(u_t) - \ell_t f(\ell_t) = [R_t^u - \beta'X_t]f(u_t) - [R_t^\ell - \beta'X_t]f(\ell_t)$$

$$= [R_t^u f(u_t) - R_t^\ell f(\ell_t)] - \beta'X_t[f(u_t) - f(\ell_t)]$$

$$\therefore \frac{\partial \ln L}{\partial \sigma^2} = -N_1 + \frac{1}{\sigma^2}\sum_1 (R_t - \beta'X_t)^2 - \sum_0 \left[ \frac{R_t^u f(u_t) - R_t^\ell f(\ell_t)}{F(u_t) - F(\ell_t)} \right]$$

$$+ \sum_0 \beta'X_t \left[ \frac{f(u_t)' - f(\ell_t)}{F(u_t) - F(\ell_t)} \right] = 0 \qquad (5b)$$

Premultiplying (5a) by $\beta'$ and add the result to (5b), we get:

$$- N_1 + \frac{1}{\sigma^2} \sum_1 (R_t - \beta'X_t)\beta'X_t + \frac{1}{\sigma^2}\sum_1 (R_t - \beta'X_t)^2$$

$$- \sum_0 \left[ \frac{R_t^u f(u_t) - R_t^\ell f(\ell_t)}{F(u_t) - F(\ell_t)} \right] = 0 \qquad (6)$$

Rearrange (6) and solve for $\sigma^2$, to get:

$$\sigma^2 = \frac{\displaystyle\sum_1 (R_t - \beta'X_t)^2 + \sum_1 (R_t - \beta'X_t)\beta'X_t}{\displaystyle\sum_0 \left[ \frac{R_t^u f(u_t) - R_t^\ell f(\ell_t)}{F(u_t) - F(\ell_t)} \right] + N_1} \qquad (7)$$

Also, rearrange (5a) and solve for $\beta$, to get:

$$\beta^{MLE} = \frac{\displaystyle\sum_1 X_t R_t}{\displaystyle\sum_1 X_t'X_t} - \frac{\displaystyle\sum_0 X_t \sigma^2 \left[ \dfrac{f(u_t) - f(\ell_t)}{F(u_t) - F(\ell_t)} \right]}{\displaystyle\sum_1 X_t'X_t}$$

Since, $E(\epsilon_t | \ell_t < \epsilon_t < u_t) = \sigma^2 \left[ \dfrac{f(\ell_t) - f(u_t)}{F(u_t) - F(\ell_t)} \right]$

$$\therefore \quad \beta^{\text{MLE}} = \beta_1^{\text{OLS}} + \frac{\sum_0 X_t \, E(\epsilon_t \mid \ell_t < \epsilon_t < u_t)}{\sum_1 X_t' X_t} \tag{8}$$

where $\beta_1^{\text{OLS}}$ is the ordinary least-square estimator for $\beta$ obtained from using only $N_1$ non-missing observations on $R_t$, and $E(\epsilon_t \mid \ell_t < \epsilon_t < u_t)$ is the mean of a doubly-truncated normal distribution of the residual errors with upper bound $u_t$ and lower bound $\ell_t$, zero mean and variance $\sigma^2$.[34]

First, equation (7) shows that the variance of the residual errors estimated by the ML estimator, in general, will be smaller than the one estimated by using the OLS estimator with only all available observations. It is not difficult to verify this by taking the expected value of the estimated variance given in equation (7). The expected value of the numerator in equation (7) is simply the sum of squared residual errors (SSE) based on all available observations which is the same value given by the OLS estimator.[35] But, the denominator of equation (7), which includes an

_____

[34]Refer to Maddala (1987) Appendix on p. 365-367 for a detailed presentation of the results on truncated distributions.

[35]Based on the assumption of classical regression model, $E(\epsilon_t) = 0$ for all non-missing observations. Therefore, the expected value of the second term on the numerator of equation (7) is equal to zero. And, the expected value of the first term in the numerator is simply the SSE.

extra positive adjustment term[36] and so will reduce the overall value of the estimated variance of the residual errors.

Second, equation (8) shows that the maximum likelihood estimator for the parameter $\beta$ is determined by the ordinary least squares estimator obtained from the non-missing observations on stock returns plus an adjustment term which captures the expected value of the residual errors associated with the missing returns. This adjustment term is crucial. Taking the expected value of equation (8), we get the unbiased beta value ($\beta$) plus the expected value of this adjustment term. Clearly, if this expected value is not equal to zero, the beta estimates obtained from the ML estimator are biased. Indeed, the intuitive rationale for equation (8) is very simple. If an individual firm, on average, has responds positively to the market movements, the ML estimator which incorporates bid-ask data will underestimate the unbiased beta value. On the other hand, if an individual firm, on average, responds negatively to the market movements, the ML estimator which incorporates bid-ask data will overestimate the unbiased beta value. Here, positive response from the market movements means that if the market is up (i.e. $R_{mt} > 0$), we will expect the likelihood of obtaining an upward price movement to be higher than the likelihood of obtaining a downward price movement (i.e., $f(u_t) > f(\ell_t)$ and

---

[36]The adjustment term here refers to the first term on the denominator of equation (7). In general, this adjustment term is postive since $R_t^u$ is by definition larger than $R_t^{\ell}$ and correspondingly $R_t^u f(u_t)$ is larger than $R_t^{\ell} f(\ell_t)$ in most cases.

$E(\epsilon_t|\ell_t < \epsilon_t < u_t) < 0$). When the market is down (i.e. $R_{mt} < 0$), the likelihood of obtaining a downward price movement is higher than the likelihood of obtaining an upward price movement (i.e., $f(u_t) < f(\ell_t)$ and $E(\epsilon_t|\ell_t < \epsilon_t < u_t) > 0$). The negative response from the market movements is merely the opposite case of the positive one. However, the most interesting case is when an individual firm has a neutral response from the market movements. There the likelihood of obtaining both upward and downward price movements is, on average, the same (i.e. $f(\ell_t) = f(u_t)$ and $E(\epsilon_t|\ell_t < \epsilon_t < u_t) = 0$); in this case the beta estimates of the ML estimator and the OLS estimator based on all non-missing returns should not differ from each other. That is, both of the OLS and ML estimates are unbiased. Therefore, whether the beta estimates are different under these two estimators provides a mean for testing the hypothesis that the **mean of the bid-ask spreads is equal to the expected value of the unobservable stock price.**

For the expected return, in order to employ the sample mean of all return data as the proxy, we need to fill in missing returns by the estimates obtained from the two-limit censored regression model. The expected value of each missing return is given by the following equation:

$$E(R_t | R_t^{\ell} < R_t^{*} < R_t^{u}) = \beta'X_t + E(\epsilon_t | R_t^{\ell} - \beta'X_t < \epsilon_t < R_t^{u} - \beta'X_t)$$

$$= \beta'X_t + \sigma^2 \left[ \frac{f(\ell_t) - f(u_t)}{F(u_t) - F(\ell_t)} \right] \qquad (9)$$

And, the expected return for individual firm is simply the sample mean of all return data including both observed and estimated returns, that is:

$$\hat{E}(R) = \frac{\sum_1 R_t}{N_0 + N_1} + \frac{\sum_0 E(R_t | R_t^{\ell} < R_t^* < R_t^u)}{N_0 + N_1} \tag{10}$$

It is clear from equation (10), that the expected return on all observations (including both missing and non-missing) is affected by the specific values that we assign to the missing observations. Under the conventional treatments as described in section III, the second term on the left hand side of equation (10) is either ignored or replaced by some arbitrary values such as average returns based on mean of bid-ask spreads in substituting unobservable true prices. As with the same argument mentioned above, if the expected value of the second term in equation (10) is not equal to zero, the return estimates calculated by filling in ML estimated returns will be biased. Obviously, it is the incorporation of the randomness of the residual errors associated with the missing stock prices that allows us to visualize the biases introduced by the ML estimator. Additionally, it is the existence of these induced biases that allows us to test the maintained hypothesis.[37] Referring to both equation (8) and

_____

[37] Given the fact that the adjustment factors of the expected return and beta estimates shared the same structure, the testing of the hypothesis that the mean of the bid-ask equal to the unobserved true price not only can be done by comparing the beta estimates of the ML estimator with the OLS estimator with all non-missing observations, but also can be done by comparing the expected return calculated from the filled in ML estimated returns with the one that simply skipped all missing returns.

equation (10), the potential biases affecting the expected return and beta estimates share the same structure. However, in equation (8), given that there is a sum of squared term under the denominator of the adjustment factor and the expected values of the true random errors on the numerator are multiplied by the market returns which are measured in decimal places, the potential biases introduced by missing returns will influence the expected return more than the beta estimates. This is a important finding since most of the existing financial literature related to the alleged size-related anomalies or optimal portfolio choice is only concerned with correction of measurement errors in the covariance matrix (or the beta estimates). However, with but a few exceptions[38], not much work has been done in addressing the importance of perturbations in the mean vector (or the expected returns). In the next section we present an illustration of applying the ML estimator to a data set from the Center for Research in Security Price (CRSP) tape to demonstrate the biases introduced by missing returns and to test the hypothesis that the mean of the bid-ask price is equal to the expected value of the unobservable stock price.

---

[38]See Best and Grauer (1987, 1988).

# V) DATA AND METHODOLOGY

Monthly data on stock prices,[39] dividend payments and bid-ask spreads on missing prices for firms which are continuously listed and contain no more than 30 and no less than 6 missing stock prices, between the period of June 1957 to June 1962, are extracted from the monthly stock master file of the 1987 edition of the University of Chicago's Center for Research in Security Prices (CRSP) tape. Out of over three thousand firms listed on the New York Stock Exchange, 152 firms are admitted to the sample. The month June 1962 is chosen because the CRSP tape only contains bid-ask data up to this month. All months after June 1962 the bid-ask data are substituted by zeroes. June 1957 is picked as the starting month so as to provide enough price data for the construction of $60^{40}$ stock returns. The choice of missing prices ranging between 6 and 30 out of 61 price data points allows a reasonable 10% to 50% variability in the data set to illustrate the potential biases. The market return (or benchmark return) utilized in this study is the value-weighted index obtained from the CRSP monthly index tape over the same sampling period.

---

[39]The CRSP monthly stock master tape uses the last trading day price of each month to calculate the corresponding return. When a trade price is unavailable, the return is calculated using the average of the bid and ask prices.

[40]Beta estimates are empirically justified to be reasonably stationary within the period of 60 months. Therefore, the 60 months estimation period is widely used in the existing literature as well as in this paper.

To demonstrate the bias introduced by missing stock prices on the expected return and beta estimates, the above data set is applied to three different scenarios for handling missing data:

1)   the ML estimator incorporating bid-ask data

2)   the OLS estimator with the mean of bid-ask as a substitute for missing price, and

3)   the OLS estimator with all missing returns skipped.

Given the fact that the log likelihood function of the ML estimator is non-linear, there are no closed form solutions for the desired estimates. An iteration approach based on the Newton-Raphson method is employed to obtain the maximum likelihood estimates of $\beta$ and $\sigma^2$. The expected return and beta estimates generated by the OLS estimator with all missing observations skipped are compared to the estimates from the ML estimator as derived in section IV and the estimates from the OLS estimator based on the mean of the bid-ask spreads. The first comparison looks at the standard errors of the beta estimates. A second comparison looks at the estimated standard deviation of the residual error terms. A third comparison looks at the differences among these estimates by using the nonparametric sign test for median differences[41]. This statistic is given by the following steps:

---

[41]See Hollander and Wolfe (1973) for details of the procedure and Bradley (1968) for an in-depth explanation of the sign test.

Null hypothesis

$$H_0: \theta^{MLE} - \theta^{OLS}_{SKIP} = 0 \quad \text{and} \quad H_0: \theta^{OLS}_{B/A} - \theta^{OLS}_{SKIP} = 0$$

or, $\quad H_0: P(\theta^{MLE} - \theta^{OLS}_{SKIP}) > 0 = P(\theta^{MLE} - \theta^{OLS}_{SKIP}) < 0 = .5$

$$\text{and}$$

$$H_0: P(\theta^{OLS}_{B/A} - \theta^{OLS}_{SKIP}) > 0 = P(\theta^{OLS}_{B/A} - \theta^{OLS}_{SKIP}) < 0 = .5$$

Procedure

Define $\quad D_i = \theta^{MLE}_i - \theta^{OLS}_{SKIP_i} \quad OR \quad = \theta^{OLS}_{B/A_i} - \theta^{OLS}_{SKIP_i}$

$$\psi_i = \begin{cases} 0 & \text{if } D_i > 0, \\ 1 & \text{if } D_i < 0. \end{cases}$$

Set $\quad B = \sum_{i=1}^{n} \psi_i$

The statistic B is the number of negative D's.

Test statistic:

$$B^* = \frac{B - E_0(B)}{\left[ Var_0(B) \right]^{1/2}} = \frac{B - (n/2)}{(n/4)^{1/2}}$$

When $H_0$ is true, the statistic $B^*$ has an asymptotic $N(0,1)$ distribution.

The normal approximation to the binomial is

Reject $H_0$ if $B^* \geq Z_{(\alpha)}$,

Accept $H_0$ if $B^* < Z_{(\alpha)}$.

The rationale for this sign test is that given the parameters obtained from the ML estimator and ad hoc variants of the OLS estimator are only estimates of the underlying true parameters, these estimates are as likely to overestimate as to underestimate the unobservable true parameters.[42] This is equivalent to hypothesizing that for a randomly drawn[43] pair of observation on the two estimates, the difference-score $\theta^{MLE} - \theta^{OLS}_{SKIP}$ or $\theta^{OLS}_{B/A} - \theta^{OLS}_{SKIP}$ is as likely to exceed zero than be exceeded by it. This is equivalent to hypothesizing that the median of a continously distributed population of difference scores is zero.

The remaining part of this empirical application is to test the hypothesis that the mean of the bid-ask spread is equal to the expected value of the unobservable stock price. Three different methods are employed. First, a search approach based on the criteria of 1) minimizing sum of squared residual errors and 2) maximizing $R^2$, respectively, is performed to search for the optimal relative position $(\theta_i)$ of the true price within the bid-ask spreads for each individual firm i, given the

---

[42] Mathematically, we have

$$\theta^{MLE} = \theta^{TRUE} + \epsilon^{MLE} \qquad (1)$$

and, $$\theta^{OLS}_{SKIP} = \theta^{TRUE} + \epsilon^{OLS}_{SKIP} \qquad (2)$$

$(1) - (2) \quad \Rightarrow \quad \theta^{MLE} - \theta^{OLS}_{SKIP} = \epsilon^{MLE} - \epsilon^{OLS}_{SKIP}$

$\therefore$ Testing $\theta^{MLE} - \theta^{OLS}_{SKIP} = 0$ is equivalent as testing $\epsilon^{MLE} - \epsilon^{OLS}_{SKIP} = 0$.

[43] This test is weakened somewhat by the fact that the errors may be correlated across firms.

regression model $R_{it}^* = \alpha_i + \beta_i R_{mt}^* + \epsilon_{it}^*$. If the resulting frequency distribution of $\theta_i$ estimates is concentrated near 0.5, then using the mean of the bid-ask spreads as a proxy for the expected value of the unobservable stock price is not a bad approximation. If, however, this distribution is spread widely, we must conclude that the mean of the bid-ask spreads is an inappropriate proxy for the expected value of the unobservable stock price. Formally, the optimizing model can be stated as follows:

$$\text{Given} \quad R_{it}^* = \alpha_i + \beta_i R_{mt}^* + \epsilon_{it}^*$$

1) Minimize $\quad SSE_i$

2) Maximize $\quad R_i^2$

$$\text{subject to} \quad P_{it}^* = \theta_i P_{it}^a + (1 - \theta_i) P_{it}^b$$

$$\text{and} \quad 0 < \theta_i < 1$$

$$\text{or}^{44} \quad -.2 < \theta_i < 1.2.$$

The second method of testing the maintained hypothesis compares the ML estimates and the OLS estimates using the mean of bid-ask prices with the OLS estimates based on all non-missing observations. If the mean of the bid-ask spreads is a good approximation of the unobservable true price, the expected return and beta estimates based on the OLS estimators with all

---

[44]The second choice of bounds on $\theta_i$ is included to illustrate how wide the spread of $\theta_i$ estimates is and to demonstrate the relaxation of the assumption that the true price must be straddled by the associated bid and ask prices.

missing returns skipped should not be different from the ones provided by the ML estimator or from the ones provided by the OLS estimator using the mean of the bid-ask prices. By employing the sign test outlined at the beginning of this section, we can perform these two indirect tests on the maintained hypothesis.

The third method involves transforming all estimated returns from the ML estimator back to price space. By comparing whether the ML estimated prices are significantly different from the prices calculated by the mean of bid-ask spreads, we can perform the most direct test of the maintained hypothesis. In addition, since there are over 1700 missing prices for all firms as a whole, the power of the test should be high. If the sign test is able to reject the null hypothesis that the two vectors of price estimates are not different, and if the overall conclusion is consistent under three different methods, it is clear that the ad hoc assumption of treating the mean of bid-ask spreads as the expected value of the unobservable stock price is invalid. In the next section, we will present the empirical results.

# VI) EMPIRICAL RESULTS

The first issue addresses the biases introduced into the expected return and beta estimates under different treatments of the missing values given the presence of missing stock prices. Table 1 and Table 2 respectively summarize the expected return and beta estimates obtained from the ML estimator and two different treatments of the OLS estimator.[45] There are two important points. First, in terms of the mean, maximum and minimum values, it is clear that different ways of handling missing data have a significant influence on expected return and beta estimates. Table 1 compared to the OLS estimator with all missing returns skipped which is unbiased, both the ML estimator and the OLS estimator using the mean of the bid-ask prices, on average, underestimate the beta value. Although the average beta estimates under these three scenarios are quite close (mean $\beta^{MLE}$ = .9129, mean $\beta^{OLS}_{B/A}$ =.9299 and mean $\beta^{OLS}_{SKIP}$ = .9742), individual beta values can be way off. For example, the minimum beta estimates under these methods are very far from each other (min. $\beta^{MLE}$ = 0.2645, min. $\beta^{OLS}_{B/A}$ = -.0445 and min. $\beta^{OLS}_{SKIP}$ = -.80). Of course, if we group portfolios, which is a weighted average of all beta values, the misspecification of risk assessment (i.e. bias in beta estimates) in the presence of missing stock prices is greatly reduced. However, given some forms of financial research

---

[45]As mentioned in last section, the two forms of the OLS estimator are 1) substituting the mean of bid-ask spreads for the unobservable true price and 2) skipping all missing returns. In terms of notation, for all tables, the subscripts of B/A and SKIP are respectively denoting these two forms of treatments of the missing data; and the superscripts of ML and OLS are respectively denoting the maximum likelihood estimator and the ordinary least squares estimator.

like event studies[46] or capital budgeting which are based on many inputs from individual assets, clearly biases of this magnitude could seriously alter the conclusions drawn. Table 2 for the expected return, the findings are very similar to the above. On average, the OLS estimates with mean of bid-ask treatment and the ML estimates underestimate the unbiased estimates obtained from the OLS estimator with all missing returns skipped. However, if we consider the magnitude of the induced biases, the situation is clearly much worse. The mean expected returns under these three scenarios (mean $\mu^{MLE}$ = .2578%, mean $\mu^{OLS}_{B/A}$ = .160% and mean $\mu^{OLS}_{SKIP}$ = .3262%) are more than 50% away from each other. Undoubtedly, any financial research which uses expected return as an input, for example optimal portfolio choice based on the mean-variance framework, will find that this magnitude of bias significantly alters their optimal solutions or overall conclusions. In addition, the sign statistic ($B^*$), which is designed to test whether there is a statistically significant difference between the OLS estimates with all missing returns skipped and the other two estimators, suggests that expected return and beta estimates obtained from the OLS estimator with all missing returns skipped are significantly different from each other. For example, the hypotheses of no difference in expected return and beta estimates for the ML estimator and the OLS estimator with all missing return skipped are rejected at $\alpha$ = 5% level in both instances. The hypotheses of no difference in expected return and beta estimates for the OLS estimator with the mean of bid-ask spreads substituted for the

---

[46]In this study, we employ monthly data for the empirical illustration; however, for event studies which are mostly based on daily data, the missing data problem may be much more serious.

unobserved true price and the OlS estimator with all missing return skipped are also rejected at $\alpha = 5\%$ level in both instances.

A second important point, as mentioned in section IV, is that the biases introduced by the ad hoc treatments of the missing data should influence the expected return more than the beta estimates. This arguement is supported by the results of Table 1 and Table 2. For example, the mean absolute value of the difference in the ML and OLS estimates, which measures the magnitude of the average induced biases, shows that the expected returns are, on average, being influenced more than the beta estimates. Also, if we compare the mean, maximum and minimum values of the OLS estimates with all missing returns skipped with the ML estimates and the OLS estimates using the mean of bid-ask spreads respectively, it is clear that the differences in expected return are much larger than the differences in the beta estimates.

The second issue addresses whether the incorporation of the randomness associated with the missing returns will increase the efficiency of the estimation procedure. Table 3 and Table 4 respectively summarize the estimated standard errors of the beta estimates and the standard deviations of the residual errors obtained from the ML estimator and ad hoc variants of the OLS estimator. Using standard error of the beta estimates and standard deviation of the residual errors as measures of efficiency, the ML estimator and the OLS estimator using the mean of bid-ask spreads (which have incorporated extraneous information carried by the bid-ask data) outperform the OLS estimator with all missing returns skipped. Table 3 out

of the 152 firms in the sample, the ML estimator and the OLS estimator using the mean of the bid-ask spreads 141 and 149 times respectively, provide a lower estimated standard error of the beta estimates than the OLS estimator with all missing returns skipped. Over 92% of the time the two estimators which incorporate bid-ask data outperform the OLS estimator with all missing returns skipped based on this criterion. In addition, the sign statistic $(B^*)$ testing the hypothesis of no difference in the estimated standard errors of the beta estimates are rejected at $\alpha = 1\%$ level, in both instances. The results are almost identical when we consider the comparison using standard deviations of the residual errors. Table 4 out of the 152 firms in the sample, the ML estimator and the OLS estimator using the mean of the bid-ask spreads 116 and 132 times respectively, produces a lower standard deviation of the residual errors than the OLS estimator with all missing returns skipped. The sign statistics $(B^*)$ are also reject the null hypothesis of no difference in standard deviation of the residual errors at the $\alpha = 1\%$ level, in both instances. Also, if we consider the magnitude of the average absolute efficiency gain (i.e., $|S.E.^{MLE} - S.E._{SKIP}^{OLS}|$ or $|S.E._{B/A}^{OLS} - S.E._{SKIP}^{OLS}|$), the incorporation of missing data has greatly improved the estimation procedure.

Finally, the last issue addressed is to test the null hypothesis that the mean of the bid-ask spreads is equal to the unobservable true prices. First, let us consider the results of the search approach as described in the last section. Table 5a and Table 5b respectively show the frequency distributions of the optimal relative position $(\theta_i)$ of the unobservable true price within the bid-ask spreads for each individual firm i, under the

criteria of minimizing the sum of squared residual error (SSE) and maximizing $R^2$ subject to the constraints that $P_{it}^* = \theta_i P_t^a + (1 - \theta_i) P_{it}^b$ and $0 < \theta_i < 1$, given the regression model $R_{it}^* = \alpha_i + \beta_i R_{mt}^* + \epsilon_{it}^*$. Under both criteria, the frequency distributions of $\theta_i$ are bi-modal, with the highest frequencies at both ends of the restricted bounds (i.e. .01 and .99). The implication drawn from these frequency distributions is that the $\theta$ values are not concentrated close to the hypothetical value of .5, suggesting that the ad hoc assumption that the mean of the bid-ask spreads is equal to the expected value of the unobservable stock price is highly suspect. Table 6a and Table 6b present the results of the same search approach by dropping the assumption that the unobservable true price must be straddled by the associated bid and ask prices, by allowing the $\theta_i$ values move beyond the theoretical bounds of zero and one. For instance, we simply allow a 20% deviation from the theoretical bounds (i.e. -.2 and 1.2) The results are virtually identical to the previous ones. The frequency distributions of $\theta_i$ under both critieria are also bi-modal with the highest frequencies occurring at both ends of the restricted bounds (i.e. -.19 and 1.19). To sum up, the overall results of this search approach suggest that the ad hoc assumption of treating the mean of the bid-ask spreads as the unobservable true price is inappropriate.

The second method of testing the stated null hypothesis is done in the return space by using the structure of the ML estimator. If the beta estimates of the OLS estimator with all missing observations being dropped are statistically different from the ones obtained from the ML estimator, and the ones obtained from the OLS estimator with the mean of the bid-ask

spreads as proxy of the expected value of the unobservable stock price, we can reject the null hypothesis that the mean of bid-ask spreads is equal to the unobservable true price. From Table 2, the sign statistic ($B^*$) suggests that the null hypothesis is significantly rejected at $\alpha = 5\%$ level in both instances suggesting that the mean of the bid-ask spread as a proxy of the unobservable true is unjustified.

The third method of testing the stated hypothesis is done directly in the price space by comparing the ML estimated missing prices, which are converted from the ML estimated returns, with the prices substituted by the mean of the bid-ask spreads. Table 7 presents the summary of the estimated prices. The sign statistic ($B^*$) appied to all 1743 missing prices suggests that the stated null hypothesis is significantly rejected at $\alpha = 1\%$ level. In summary, all three methods of testing the stated hypothesis arrrive at the same conclusion that the mean of bid-ask spreads as a proxy of the expected value of the unobservable stock price is unjustified.

## VII) SUMMARY AND CONCLUdNG REMARKS

This study has tested the hypothesis that the mean of the bid-ask spreads is equal to the expected value of the unobservable stock price, one of the most commonly used ad hoc assumptions for handling missing stock prices and for building financial models. One test is done by comparing the biasedness of the OLS estimation techniques given the choice of either ignoring information contained in the bid-ask data or assuming that the expected value of the unobservable stock price is the mean of the bid-ask prices. The former technique produces an unbiased estimate of the expected returns and beta estimates, but the latter technique is only unbiased if the expected value of the unobservable stock price is in fact given by the mean of the bid and ask prices. This test is supplemented by using a double-limit maximum likelihood estimation technique which is also unbiased but only if the mean of the bid-ask prices is equal to the expected value of the unobservable stock price. By applying the maximum likelihood estimator and the ordinary least squares estimator under two different treatments for handling missing stock price to a data set which is characterized by a reasonable number of missing stock prices, the empirical results suggest that:

1) although the ML estimator incorporating the bid-ask data is biased, it reduces the standard errors of the beta estimates;

2) different treatments for handling missing prices in the estimation procedure have a profound influence on expected return and beta estimates of individual firms. These findings illustrate the important

48

point that the biases introduced by missing stock prices influence the expected return much more than the beta estimates. Given that most of the existing literature is only concerned with adjustment of the beta estimates, it is important for future research to pay more attention to the variability of the mean vector;

3) using the mean of bid-ask spreads as a proxy for the unobservable true price is rejected under three different methods: the search approach, calculation of the ML estimator in return space and direct comparison of estimated missing prices with the prices substituted by the mean of the bid-ask spreads.

The results of this study not only illustrate the importance of investigating the missing stock price problem, but also highlight an important issue related to the spirit of how empirical studies should be performed. As suggested by Steven Cheung (1978), using imaginary facts to support imaginary policies is the most common mistake committed by economists. A lot of economic research is built purely on hypothetical bases without testing whether or not the basic assumption is valid. For instance, the over-popularly used ad hoc assumption of treating the mean of the bid-ask spread as the expected value of the unobservable stock price is only one of the examples in demonstrating how researchers intend to change the real world to fit their pure imagination. This study has highlighted the need for careful empirical investigation in testing ad hoc assumptions. "The thesis is simply that any valuable application of economic theory must rest upon *careful empirical investigation to ensure that the facts are*

*true, that hypotheses are testable, and that tests are performed.*"[47]

Finally, although the double-limit maximum likelihood estimator developed in this paper provides biased estimates, it serves as a first step for parameterizing the bid and ask prices in the estimation procedure for future research. This future development not only will give rise to a more accurate estimation of the unobservable non-trading stock prices, but will also provide a means of understanding the interrelation between the function of the stock market and the determination of bid and ask prices. Suppose the stock market is rising (up market). The unobservable stock price should be closer to the ask price than the bid price. On the other hand, in the down market, the reverse should be true, while in the "flat" market the unobservable stock price should be equal to the mean of the bid and ask prices. This thesis provides a method for testing the above hypotheses by spliting the samples into up, down and "flat" markets for estimation. In addition, since the true returns on risky securities are assumed to be generated by the market model (a one factor model), the tests performed in this study can be extended via the arbitrage pricing model ( a multi-factor model) in future research. Also, it is valuable to apply the maximum likelihood estimator developed in this thesis to the small effect, the turn-of-the-year effect and the Monday effect, to illustrate how missing stock prices may affect the findings associated with these empirical anomalies exhibited in stock return data.

---

[47] Steven N. S. Cheung (1978) p. 67.

## Table 1

### Summary of the Beta Estimates

### from the ML estimator and ad hoc variants of OLS estimator

| Estimate | Mean value | Extreme Values' Maximum | Minimum | No. of negatives (B) | Statistic $B^*$ |
|---|---|---|---|---|---|
| $\beta^{MLE}$ | 0.9129 | 2.7925 | 0.2645 | | |
| $\beta^{OLS}_{B/A}$ | 0.9299 | 2.7918 | $-0.4452 \times 10^{-1}$ | | |
| $\beta^{OLS}_{SKIP}$ | 0.9742 | 3.3674 | $-0.8000$ | | |
| $\beta^{OLS}_{B/A} - \beta^{OLS}_{SKIP}$ | $-0.4437 \times 10^{-1}$ | 1.3990 | $-2.2377$ | 89 | 2.11* |
| $\left| \beta^{OLS}_{B/A} - \beta^{OLS}_{SKIP} \right|$ | 0.2029 | 2.2377 | $0.1079 \times 10^{-3}$ | | |
| $\beta^{MLE} - \beta^{OLS}_{SKIP}$ | $-0.6134 \times 10^{-1}$ | 1.3829 | $-2.2335$ | 89 | 2.11* |
| $\left| \beta^{MLE} - \beta^{OLS}_{SKIP} \right|$ | 0.1793 | 2.2335 | $0.1044 \times 10^{-3}$ | | |

Notes: $B^* = \dfrac{B - (152/2)}{(152/4)^{1/2}} \sim N(0,1)$ and $n = 152$ firms

* significant at 5% level

51

## Table 2

### Summary of the Monthly Expected Return (%)

### from the ML estimator and ad hoc variants of OLS estimator

| Estimate | Mean value (%) | Extreme Values (%) | | No. of negatives (B) | Statistic $B^*$ |
|---|---|---|---|---|---|
| | | Maximum | Minimum | | |
| $\mu^{MLE}$ | 0.2587 | 2.3263 | -3.0317 | | |
| $\mu^{OLS}_{B/A}$ | 0.1600 | 2.0947 | -3.0409 | | |
| $\mu^{OLS}_{SKIP}$ | 0.3262 | 3.9047 | -5.9410 | | |
| $\mu^{OLS}_{B/A} - \mu^{OLS}_{SKIP}$ | -0.1661 | 4.2717 | -3.6518 | 90 | 2.27* |
| $\left\| \mu^{OLS}_{B/A} - \mu^{OLS}_{SKIP} \right\|$ | 0.7722 | 4.2717 | $0.1964 \times 10^{-1}$ | | |
| $\mu^{MLE} - \mu^{OLS}_{SKIP}$ | $-0.6749 \times 10^{-1}$ | 4.3918 | $-0.2225 \times 10^{-1}$ | 89 | 2.11* |
| $\left\| \mu^{MLE} - \mu^{OLS}_{SKIP} \right\|$ | 0.6910 | 4.3918 | $0.1767 \times 10^{-1}$ | | |

Notes: $B^* = \dfrac{B - (152/2)}{(152/4)^{1/2}} \sim N(0,1)$ and n = 152 firms

*significant at 5% level

# Table 3

Summary of the Standard Error of the Beta Estimates

from the ML estimator and ad hoc variants of OLS estimator

| Estimate | Mean value | Extreme Values Maximum | Minimum | No. of negatives (B) | Statistic $B^*$ |
|---|---|---|---|---|---|
| $S.E.^{MLE}$ | 0.2909 | 0.7872 | 0.1156 | | |
| $S.E._{B/A}^{OLS}$ | 0.3101 | 0.8005 | 0.1175 | | |
| $S.E._{SKIP}^{OLS}$ | 0.4063 | 1.8976 | 0.1111 | | |
| $S.E._{B/A}^{OLS} - S.E._{SKIP}^{OLS}$ | $-0.9622 \times 10^{-1}$ | 0.3693 | $-1.1044$ | 141 | 10.54** |
| $\left\| S.E._{B/A}^{OLS} - S.E._{SKIP}^{OLS} \right\|$ | 0.1144 | 1.1044 | $0.4674 \times 10^{-2}$ | | |
| $S.E.^{MLE} - S.E._{SKIP}^{MLS}$ | $-0.1154$ | $0.2035 \times 10^{-1}$ | $-1.1131$ | 149 | 11.84** |
| $\left\| S.E.^{MLE} - S.E._{SKIP}^{OLS} \right\|$ | 0.1159 | 1.1131 | $0.2733 \times 10^{-2}$ | | |

Notes:  $B^* = \dfrac{B - (152/2)}{(152/4)^{1/2}} \sim N(0,1)$ and n = 152 firms

** significant at 1% level

53

Table 4

Summary of the Standard Deviation of the Residual Errors

from the ML estimator and ad hoc variants of OLS estimator

| Estimate | Mean value | Extreme Values | | No. of negatives (B) | Statistic $B^*$ |
|---|---|---|---|---|---|
| | | Maximum | Minimum | | |
| $\sigma^{MLE}$ | $0.8083 \times 10^{-1}$ | $0.2195$ | $0.3216 \times 10^{-1}$ | | |
| $\sigma^{OLS}_{B/A}$ | $0.8648 \times 10^{-1}$ | $0.2232$ | $0.3275 \times 10^{-1}$ | | |
| $\sigma^{OLS}_{SKIP}$ | $0.9111 \times 10^{-1}$ | $0.3529$ | $0.2930 \times 10^{-1}$ | | |
| $\sigma^{OLS}_{B/A} - \sigma^{OLS}_{SKIP}$ | $-0.4629 \times 10^{-2}$ | $0.1087$ | $-0.1317$ | $116$ | $6.49**$ |
| $\left\| \sigma^{OLS}_{B/A} - \sigma^{OLS}_{SKIP} \right\|$ | $0.1385 \times 10^{-1}$ | $0.1317$ | $0.8315 \times 10^{-4}$ | | |
| $\sigma^{MLE} - \sigma^{OLS}_{SKIP}$ | $-0.1028 \times 10^{-1}$ | $0.1466 \times 10^{-1}$ | $-0.1345$ | $132$ | $9.08**$ |
| $\left\| \sigma^{MLE} - \sigma^{OLS}_{SKIP} \right\|$ | $0.1145 \times 10^{-1}$ | $0.1345$ | $0.4160 \times 10^{-5}$ | | |

Notes:   $B^* = \dfrac{B - (152/2)}{(152/4)^{1/2}} \sim N(0,1)$ and n = 152 firms

** significant at 1% level

## TABLE 5A

Given $R^*_{it} = \alpha_i + \beta_i R^*_{mt} + \epsilon^*_{it}$

Frequency Distribution of $\theta_i$

Minimizing $SSE_i$

subject to $\quad P^*_{it} = \theta_i P^a_{it} + (1 - \theta_i) P^b_{it}$

and $\quad 0 < \theta_i < 1$

| $\theta_i$ value | Frequency | Percent | $\theta_i$ value | Frequency | Percent |
|---|---|---|---|---|---|
| 0.01 | 39 | 0.25658 | 0.54 | 1 | 0.00658 |
| 0.02 | 1 | 0.00658 | 0.55 | 1 | 0.00658 |
| 0.03 | 1 | 0.00658 | 0.56 | 1 | 0.00658 |
| 0.04 | 1 | 0.00658 | 0.57 | 1 | 0.00658 |
| 0.05 | 1 | 0.00658 | 0.60 | 1 | 0.00658 |
| 0.06 | 1 | 0.00658 | 0.63 | 1 | 0.00658 |
| 0.09 | 2 | 0.01316 | 0.64 | 2 | 0.01316 |
| 0.11 | 1 | 0.00658 | 0.66 | 2 | 0.01316 |
| 0.13 | 2 | 0.01316 | 0.67 | 1 | 0.00658 |
| 0.15 | 1 | 0.00658 | 0.68 | 2 | 0.01316 |
| 0.17 | 1 | 0.00658 | 0.70 | 2 | 0.01316 |
| 0.20 | 2 | 0.01316 | 0.71 | 4 | 0.02632 |
| 0.24 | 1 | 0.00658 | 0.74 | 2 | 0.01316 |
| 0.26 | 1 | 0.00658 | 0.78 | 2 | 0.01316 |
| 0.29 | 2 | 0.01316 | 0.82 | 1 | 0.00658 |
| 0.31 | 4 | 0.02632 | 0.83 | 1 | 0.00658 |
| 0.33 | 1 | 0.00658 | 0.84 | 1 | 0.00658 |
| 0.37 | 1 | 0.00658 | 0.85 | 2 | 0.01316 |
| 0.38 | 1 | 0.00658 | 0.86 | 1 | 0.00658 |
| 0.40 | 1 | 0.00658 | 0.87 | 1 | 0.00658 |
| 0.43 | 1 | 0.00658 | 0.88 | 1 | 0.00658 |
| 0.44 | 2 | 0.01316 | 0.90 | 1 | 0.00658 |
| 0.45 | 1 | 0.00658 | 0.91 | 2 | 0.01316 |
| 0.46 | 3 | 0.01974 | 0.93 | 1 | 0.00658 |
| 0.47 | 1 | 0.00658 | 0.95 | 1 | 0.00658 |
| 0.48 | 1 | 0.00658 | 0.96 | 1 | 0.00658 |
| 0.52 | 1 | 0.00658 | 0.99 | 39 | 0.25658 |
| 0.53 | 1 | 0.00658 | | | |

n = 152 firms   Median = 0.53   Mode = 0.99 or 0.01 with 39 observations

## TABLE 5B

Given $R_{it}^* = \alpha_i + \beta_i R_{mt}^* + \epsilon_{it}^*$

Frequency Distribution of $\theta_i$

Maximizing $R_i^2$

subject to $\quad P_{it}^* = \theta_i P_{it}^a + (1 - \theta_i) P_{it}^b$

and $\quad 0 < \theta_i < 1$

| $\theta_i$ value | Frequency | Percent |
|:---:|:---:|:---:|
| 0.01 | 71 | 0.46711 |
| 0.02 | 1 | 0.00658 |
| 0.03 | 1 | 0.00658 |
| 0.05 | 1 | 0.00658 |
| 0.08 | 1 | 0.00658 |
| 0.24 | 1 | 0.00658 |
| 0.25 | 1 | 0.00658 |
| 0.30 | 1 | 0.00658 |
| 0.35 | 1 | 0.00658 |
| 0.37 | 1 | 0.00658 |
| 0.46 | 1 | 0.00658 |
| 0.48 | 2 | 0.01316 |
| 0.49 | 2 | 0.01316 |
| 0.64 | 1 | 0.00658 |
| 0.65 | 1 | 0.00658 |
| 0.70 | 1 | 0.00658 |
| 0.71 | 1 | 0.00658 |
| 0.73 | 1 | 0.00658 |
| 0.77 | 2 | 0.01316 |
| 0.83 | 1 | 0.00658 |
| 0.88 | 2 | 0.01316 |
| 0.90 | 1 | 0.00658 |
| 0.92 | 1 | 0.00658 |
| 0.95 | 1 | 0.00658 |
| 0.96 | 1 | 0.00658 |
| 0.98 | 1 | 0.00658 |
| 0.99 | 52 | 0.34211 |

n = 152 firms   Median = 0.24   Mode = 0.01 with 71 observations

Given $R_{it}^* = \alpha_i + \beta_i R_{mt}^* + \epsilon_{it}^*$

Frequency Distribution of $\theta_i$

Minimizing $\underset{i}{SSE}$

subject to $\quad P_{it}^* = \theta_i P_{it}^a + (1 - \theta_i) P_{it}^b$

and $\quad -0.2 < \theta_i < 1.2$

| $\theta_i$ value | Frequency | Percent | $\theta_i$ value | Frequency | Percent |
|---|---|---|---|---|---|
| -0.19 | 28 | 0.18421 | 0.52 | 1 | 0.00658 |
| -0.16 | 1 | 0.00658 | 0.53 | 1 | 0.00658 |
| -0.14 | 1 | 0.00658 | 0.54 | 1 | 0.00658 |
| -0.13 | 2 | 0.01316 | 0.55 | 1 | 0.00658 |
| -0.11 | 1 | 0.00658 | 0.56 | 1 | 0.00658 |
| -0.09 | 1 | 0.00658 | 0.57 | 1 | 0.00658 |
| -0.08 | 1 | 0.00658 | 0.60 | 1 | 0.00658 |
| -0.06 | 1 | 0.00658 | 0.63 | 1 | 0.00658 |
| -0.03 | 1 | 0.00658 | 0.64 | 2 | 0.01316 |
| -0.02 | 1 | 0.00658 | 0.66 | 2 | 0.01316 |
| -0.01 | 1 | 0.00658 | 0.67 | 1 | 0.00658 |
| 0.02 | 1 | 0.00658 | 0.68 | 2 | 0.01316 |
| 0.03 | 1 | 0.00658 | 0.70 | 2 | 0.01316 |
| 0.04 | 1 | 0.00658 | 0.71 | 4 | 0.02632 |
| 0.05 | 1 | 0.00658 | 0.74 | 2 | 0.01316 |
| 0.06 | 1 | 0.00658 | 0.78 | 2 | 0.01316 |
| 0.09 | 2 | 0.01316 | 0.82 | 1 | 0.00658 |
| 0.11 | 1 | 0.00658 | 0.83 | 1 | 0.00658 |
| 0.13 | 2 | 0.01316 | 0.84 | 1 | 0.00658 |
| 0.15 | 1 | 0.00658 | 0.85 | 2 | 0.01316 |
| 0.17 | 1 | 0.00658 | 0.86 | 1 | 0.00658 |
| 0.20 | 2 | 0.01316 | 0.87 | 1 | 0.00658 |
| 0.24 | 1 | 0.00658 | 0.88 | 1 | 0.00658 |
| 0.26 | 1 | 0.00658 | 0.90 | 1 | 0.00658 |
| 0.29 | 2 | 0.01316 | 0.91 | 2 | 0.01316 |
| 0.31 | 4 | 0.02632 | 0.93 | 1 | 0.00658 |
| 0.33 | 1 | 0.00658 | 0.95 | 1 | 0.00658 |
| 0.37 | 1 | 0.00658 | 0.96 | 1 | 0.00658 |
| 0.38 | 1 | 0.00658 | 1.00 | 1 | 0.00658 |
| 0.40 | 1 | 0.00658 | 1.02 | 2 | 0.01316 |
| 0.43 | 1 | 0.00658 | 1.06 | 1 | 0.00658 |
| 0.44 | 2 | 0.01316 | 1.10 | 2 | 0.01316 |
| 0.45 | 1 | 0.00658 | 1.11 | 1 | 0.00658 |
| 0.46 | 3 | 0.01974 | 1.12 | 2 | 0.01316 |
| 0.47 | 1 | 0.00658 | 1.17 | 1 | 0.00658 |
| 0.48 | 1 | 0.00658 | 1.19 | 29 | 0.19079 |

n = 152 firms   Median = 0.53   Mode = 1.99   with 29 observations

Given $R^*_{it} = \alpha_i + \beta_i R^*_{mt} + \epsilon^*_{it}$

Frequency Distribution of $\theta_i$

Maximizing $R^2_i$

subject to $P^*_{it} = \theta_i P^a_{it} + (1 - \theta_i) P^b_{it}$

and $-0.2 < \theta_i < 1.2$

| $\theta_i$ value | Frequency | Percent |
|---|---|---|
| -0.19 | 66 | 0.43421 |
| -0.18 | 1 | 0.00658 |
| -0.17 | 1 | 0.00658 |
| -0.10 | 1 | 0.00658 |
| -0.05 | 1 | 0.00658 |
| -0.02 | 1 | 0.00658 |
| 0.02 | 1 | 0.00658 |
| 0.03 | 1 | 0.00658 |
| 0.05 | 1 | 0.00658 |
| 0.08 | 1 | 0.00658 |
| 0.24 | 1 | 0.00658 |
| 0.25 | 1 | 0.00658 |
| 0.30 | 1 | 0.00658 |
| 0.35 | 1 | 0.00658 |
| 0.37 | 1 | 0.00658 |
| 0.46 | 1 | 0.00658 |
| 0.48 | 2 | 0.01316 |
| 0.49 | 2 | 0.01316 |
| 0.64 | 1 | 0.00658 |
| 0.65 | 1 | 0.00658 |
| 0.70 | 1 | 0.00658 |
| 0.71 | 1 | 0.00658 |
| 0.73 | 1 | 0.00658 |
| 0.77 | 2 | 0.01316 |
| 0.83 | 1 | 0.00658 |
| 0.88 | 2 | 0.01316 |
| 0.90 | 1 | 0.00658 |
| 0.92 | 1 | 0.00658 |
| 0.95 | 1 | 0.00658 |
| 0.96 | 1 | 0.00658 |
| 0.98 | 1 | 0.00658 |
| 0.99 | 1 | 0.00658 |
| 1.06 | 1 | 0.00658 |
| 1.09 | 1 | 0.00658 |
| 1.10 | 1 | 0.00658 |
| 1.13 | 1 | 0.00658 |
| 1.18 | 2 | 0.01316 |
| 1.19 | 45 | 0.29605 |

n = 152 firms   Median = 0.24   Mode = -0.19 with 66 observations

Table 7

Summary of the Estimated Missing Prices

from the ML estimator and the mean of the bid-ask spreads

| Estimate | Mean value | Extreme Values | | No. of negatives (B) | Statistic $B^*$ |
|---|---|---|---|---|---|
| | | Maximum | Minimum | | |
| $P^{MLE}$ | 29.846 | 157.11 | 2.0574 | | |
| $P_{B/A}$ | 29.845 | 157.00 | 2.0625 | | |
| $P^{MLE}-P_{B/A}$ | $0.9948 \times 10^{-3}$ | 1.0393 | -0.6359 | 941 | 3.33** |
| $\left\| P^{MLE}-P^{OLS}_{B/A} \right\|$ | $0.2688 \times 10^{-1}$ | 1.0393 | 0.0000 | | |

Notes: $B^* = \dfrac{B - (1743/2)}{(1743/4)^{1/2}} \sim N(0,1)$ and n = 1743 missing stock prices

** significant at 1% level

# Bibliography

Altman, E.I., B. Jacquillat and M. Levasseur. 1974. "Comparative Analysis of Risk Measures: France and the United States." *Journal of Finance,* 29 (December): 1495-1511.

Amihud, Y. and H. Mendelson. 1980. "Dealer Market." *Journal of Financial Economics,* 8 (March): 31-53.

————— 1986. "Asset Pricing and the Bid-Ask Spreads." *Journal of Financial Economics,* 17 (December): 223-249.

Bagehot, W. 1971. "The Only Game in Town." *Financial Analysts Jouranl,* 22 (March-April): 12-14.

Ball, R., P. Brown and F. Finn. 1977. "Share Capitalization Changes, new information, and the Austrialian Equity Market." *Australian Journal of Management,* 2: 105-125.

Banz, R. W. 1981. "The Relationship between Return and Market Value of Common Stock." *Journal of Financial Economics,* 9 (March): 3-18.

Barnea, A. 1974. "Performance Evaluation of New York Stock Exchange Specialists." *Journal of Financial and Quantitative Analysis,* 9 (September): 511-535.

Barnea, A. and D. Logue. 1975. "The Effect of Risk on the Market Marker's Spread." *Financial Analysts Journal,* 31 (November-December): 45-49.

Barry, C.B. and S.J. Brown. 1984. "Differential Information and the Small Firm Effect." *Journal of Financial Economics,* 13 (July): 217-235.

————— 1985. "Differential Information and Security Market Equilibrium." *Journal of Financial and Quantitative Analysis,* 20 (December): 407-422.

Benston, G.J. and R.L. Hagerman. 1974. "Determinants of Bid-Asked Spreads in the Over-the-Counter Market.", *Journal of Financial Economics,* 1 (December): 353-364.

Berglund, T., B. Wahlroos and A. Örnmark. 1983. "The Weak-Form Efficiency of the Finnish and Scandinavian Stock Exchanges." *Scandinavian Journal of Economics,* 85: 521-530.

Best, M.J. and R.R. Grauer. 1987. "Should Investor Hold Well-Diversified Portfolios?" Working Paper, Department of Economics, Simon Fraser University.

————— 1988. "Direct Conditions for Mean-variance Efficient Portfolio to Have Positive Weights: Theory and Evidence." Working Paper, Department of Economics, Simon Fraser University.

Blume, M.E. and R.F. Stambaugh. 1983. "Biases in Computed Returns: An Applications to the Size Effect." *Journal of Financial Economics*, 12 (November): 387-404.

Bradley, J. 1968. *Distribution-Free Statistical Tests*. New Jersey: Prentice-Hall.

Brown, P., D.B. Keim, A.W. Kleidon and T.A. Marsh. 1983. "Stock Return Seasonalities and the Tax-Loss Selling Hypothesis: Analysis of the Arguements and Australian Evidence." *Journal of Financial Economics*, 12 (June): 105-128.

Cheung, S.N.S. 1973. "The Fable of the Bees." *Journal of Law and Economics*, 16 (April): 11-33.

———— 1978. *The Myths of Social Costs: a Ciritique of Welfare Economics and the Implications*. London: The Institue of Economic Affairs.

Clarkson P. and R. Thompson. 1987. "Measurement of Beta Risk within the Differential Information Model." Working Paper, Faculty of Business Administration, Simon Fraser University.

Cohen, K.J., S.F. Maier, R.A. Schwartz and D.K. Whitecomb. 1981. "Transaction Costs, Order Placement Strategy, and Esixtence of the Bid-Ask Spread." *Journal of Political Economy*, 89 (April): 287-305.

Cohen, K.J., G.A. Hawawini, S.F. Maier, R.A. Schwartz and D.K. Whitcomb. 1980. Implications of Microstructure Theory for Empirical Reserach on Stock Price Behavior." *Journal of Finance*, 35 (May): 249-257.

———— 1983a. "Estimating and Adjusting for the Internalling-Effect Bias in Beta." Management Science, 29 (January): 135-148.

———— 1983b. "Fricition in the Trading Process and the estimation of Systematic Risk." *Journal of Financial Economics*, 12 (August): 263-278.

Copeland T.E. and D. Galai. 1983. "Information Effects on the Bid-Ask Spread." *Journal of Finance*, 38 (December): 1457-1469.

Cowles, A. 1960. "A Review of Previous Conclusions Regrading Stock Price Behavior." *Econometrica*, 28 (October): 909-915.

Daniels, H.E. 1966. "Autocorrelation Between First Differences of Mid-Ranges." *Econometrica*, 34 (January): 215-219.

Demsetz, H. 1968. "The Cost fo Transacting." *The Quarterly Journal of Economics*, 82 (February): 33-53.

Dimson, E. 1979. "Risk Measurement When Shares are Subject to Infrequent Trading." *Journal of Financial Economics*, 7 (June): 197-226.

————— 1985. "Friction in the Trading Process and Risk Measurement." *Economics Letter*, 18: 251-254.

Fama, E.F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, 25 (May): 383-417.

————— 1976. *Foundations of Finance: Portfolio Decisions and Securities Prices*. New York: Basic Books.

Fama, E.F. and J.D. MacBeth. 1973. "Risk, Return and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (May/June): 607-636.

Fisher, L. 1966. "Some New Stock-Market Indexes." *Journal of Business*, 39 (January): 191-225.

Fowler, D.J. and H.C. Rorke. 1983. "Risk Measurement When Shares are Subject to Infrequently Trading: Comment." *Journal of Financial Economics*, 12 (August): 279-283.

Fowler, D.J., H.C. Rorke and V.M. Jog. 1979a. "Thin Trading, Errors in Variables and Market Model." Working Paper #77-47, Faculty of Management, McGill University.

————— 1979b. "Heteroscedasticity, $R^2$ and Thin Trading on the Toronto Stock Exchange." *Journal of Finance*, 34 (December): 1201-1210.

————— 1980. "Thin Trading and Beta Estimation Problems on the Toronto Stock Exchange." *Journal of Business Administration*, 12 (Fall): 77-90.

Franks, J.R. J.E. Broyles and M.J. Hecht. 1977. "An Industry Study of Profitability of Mergers in the United Kingdom." *Journal of Finance*, 32 (December): 1513-1525.

French, K.R. and R. Roll. 1986. "Stock Return Variances: the Arrival of Information and Reaction of Traders." *Journal of Financial Economics*, 17 (September): 5-26.

Friedman, M. 1962. "The Interpolation of Time Series By related Series." *Amercian Statistical Association Journal*, (December): 729-757.

Garbade, K.D. and W. Silber. 1979. "Structural Oraganization of Secondary Markets: Clearing Frequency, Dealer Activity and Liquidity Risk." *Journal of Finance*, 34 (June):577-593.

Garman, M.B.. 1976. "Market Microstructure." *Journal of Financial Economics*, 3 (June): 257-275.

Glosten, L.R. 1987. "Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices." *Journal of Finance*, 42 (December): 1293-1307.

Glosten, L.R. and P.R. Milgrom. 1985. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogenously Informed Traders." *Journal of Financial Economics*, 14 (March): 71-100.

Goldman, B.M. and A. Beja. 1979. "Market Prices Vs Equilibrium Price: Returns' Variance Serial Correlation and the Role of the Specialist." *Journal of Finance*, 34 (June): 595-607.

Goldman, M.B. and H.B. Sosin. 1979. "Information Dissemination Market Efficiency and the Frequency of Transactions." *Journal of Financial Economics*, 7 (March): 29-61.

Grossman, S. 1976. "On the Efficiency of Competitive Stock Markets When Traders Have Diverse Information." *Journal of Finance*, 31 (May): 573-585.

Grube, R.C., O.M. Joy and J.S. Howe. 1987. "Some Empirical Evidence on Stock Returns and Security Credit Regulation in the OTC Equity Market." *Journal of Banking and Finance*, 11 (March): 17-31.

Hamilton, J.L. 1976. "Competition, Scale Economics, and Transaction Cost in the Stock Market." *Journal of Financial Quantitative Analysis*, 11 (December): 779-802.

Hasbrouck, J. and Thomas S.Y. Ho. 1987. "Order Arrival, Quote Behavior, and the Return-Generating Process." *Journal of Finance*, 4 (September): 1035-1048.

Heinkel, R. and A. Kraus. 1985. "Measuring Impacts in Thinly Traded Stock." Working Paper, Graduate School of Business, University of British Columbia.

Ho, T.S.Y. and H.R. Stoll. 1981. "Optimnal Dealer Pricing Under Transactions and Return Uncertainy." *Journal of Financial Economics*, 9 (March): 47-74.

Hollander, M. and D.A. Wolfe. 1973. *Nonparametric Statistical Methods*, New York: John Wiley.

Ibbotson, F.G. 1975. "Price Performance of Common Stock New Issues." *Journal of Financial Economics*, 2 (September): 235-272.

Jaffe, J.E. 1974. "Special Information and Insider Trading." *Jouranl of Business*, 47 (July): 410-428.

Jaffe, J.E. and R. Winkler. 1976. "Optimal Speculation Against an Efficient Market." *Journal of Finance*, 31 (March): 49-61.

James, C. and R.O. Edmister. 1983. "The Relationship Between Common Stock Returns, Trading Activity and Market Value." *Journal of Finance*, 38 (September): 1075-1086.

Keim, D. 1983. "Size Related Anomalies and Stock Return Seasonality: Further Empirical Evidence." *Journal of Financial Economics*, 12 (June): 13-32.

Klein, R.W. and V.S. Bawa. 1977. "The Effect of Limited Information and Estimation Risk on Optimal Portfolio Diversification." *Journal of Financial Economics*, 5 (August): 89-111.

Klye, A.S. 1985. "Continuous Auctions and Insider Trading." *Econometrica*, 53 (November): 1315-1335.

Lau, S.C., S.R. Quay and C.M. Ramsey. 1974. "The Tokyo Stock Exchange and the Capital Asset Pricing Model." *Journal of Finance*, 29 (May): 507-514.

Lorie, J.H. and V. Niederhoffer. 1968. "Predictive and Statistical Properties of Insider Trading." *Journal of Law and Economics*, 11 (April): 35-51.

Maddala, G.S. 1977. *Econometrics*. New York: McGraw-Hill.

———— 1987. *Limited-Dependent and Qualitative Variables in Econometrics*. New York: Cambridge Unverisity Press.

Marshall, J.M.. 1974. "Private Incentives and Public Information." *American Economic Review*, 64 (June): 374-390.

McInish, T.H. and R.A. Wood. 1986. "Adjusting for Beta Bias: An Assessment of Alternate Techniques: A Note." *Journal of Finance*, 41 (March): 277-286.

Meyer, C.F. and A.B. Corbeau. 1975. "The Application of Spectral Analysis to Demonstrate the Stochastic Distortion in Delta Midrange of Price Series." *Journal of Financial Quantitative Analysis*, 10 (June): 221-230.

Niederhoffer, V. and M.F.M. Osborne. 1966. "Market Making and Reversal on the Stock Exchange." *Journal of Amercian Statistical Association*, 61 (December): 898-916.

Officer, R.R. 1975. "Seasonality in Australian Capital Markets." *Journal of Financial Economics*, 2 (March): 29-51.

Perry, P.R. 1985. "Portfolio Serial Correlation and Nonsynchronous Trading." *Journal of Financial and Quantitative Analysis*, 20 (December): 517-523.

Praetz, P.D. 1976. "Some Effects of Errors on the Independence and Distribution of Stock Price Returns." *Australian Journal of Management*, 1 : 79-83.

Reinganum, M.R. 1981. "Misspecification of Capital Asset Pricing: Emprirical Anomalies Based on Earnings' Yield and Market Values." *Journal of Financial Economics*, 9 (March): 14-46.

———— 1983. "The Anomalous Stock Market Behaviour of Small Firm in January: Empriical Tests for Tax-Loss Selling Effects." *Journal of Financial Economics*, 12 (June): 89-104.

Rorke, C.H. and C.E. Love. 1974a. "A Note Concerning the Availability of the Study: Some Indices and Rates of Return to Canadian Equity Investments for the Period January 1956 to December 1967." *Journal of Business Administration*, 6 (Fall): 64-66.

Roll, R. 1981. "A Possible Explanation of the Small Firm Effect." *Journal of Finance*, 36 (September): 879-888.

———— 1983. "Vas ist das? The Turn-of-the-Year Effecet and the Return Premium of Small Firms." *Journal of Portfolio Management*, 9 (Winter): 18-28.

———— 1984. "A Simple Implicit Mearsure of the Effective Bid-Ask Spread in an Efficiency Market." *Journal of Finance*, 39 (September): 1127-1139.

Rosenberg, B. 1971. "Statistical Analysis of Price Series Obscured by Averaging Measures." *Journal of Financial Quantitative Analysis*, 6 (September): 1083-1094.

Rubinstein, M. 1975. "Securities Market Efficiency in an Arrow-Debreu Economy." *American Economy review*, 65 (December): 812-824.

Samuelson, P.A. 1965. "Proof that Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review*, 6 (Spring): 41-49.

Scholes, M.S. 1972. "The Market for Securities: Substitution Versus Price Pressure and the Effects of Information on Share Prices." *Journal of Business*, 45 (April): 179-211.

Scholes, S.M. and J. Williams. 1977. "Estimating Betas From Nonsynchronous Data." *Journal of Financial Economics*, 5 (December): 309-327.

Schultz, P. 1983. "Transtaction Cost and the Small Firm Effect: Comment." *Journal of Financial Economics*, 12 (June): 81-88.

Sharpe, W.F. 1985. *Investments*, New Jersey: Prentice-Hall.

Smidt, S. 1971. "Which Road to an Efficient Stock Markets: Free Competition Versus Regulated Monopoly." *Financial Analyst Journal*, 27 (September-October): 18.

————— 1979. "Continuous Vs. Intermittent Trading on Auction Markets." *Journal of Financial and Quantitative Analysis*, 14 (November): 837-866.

Stoll, H.R. 1976. "Dealer Inventory Behavior: An Empirical Investigation of NASDAQ Stocks." *Journal of Financial and Quantitative Analysis*, 11 (September): 359-380.

————— 1978. "The Supply of Dealer Services in Securities Markets." *Journal of Finance*, 33 (September): 1133-1151.

Stoll, H.R. and R.E. Whaley. 1983. "Transaction Costs and the Small Firm Effect." *Journal of Financial Economics*, 12 (June): 57-80.

Theobald, M. and V. Price. 1984. "Seasonality Estimation in the Thin Markets." *Journal of Finance*, 39 (June): 377-392.

Tinic, S.M. 1972. "The Economics of Liquidity Services." *Quarterly Journal of Economics*, 86 (February): 79-93.

Tinic, S.M. and R.R. West. 1974. "Marketability of Common Stocks in Canada and the U.S.A.: A Comparison of Agent Versus Dealer Dominated Markets." *Journal of Finance*, 29 (June): 729-746.

Working, H. 1960. "Note on the Correlation of the First Differences of Averages in a Random Chain." *Econometrica*, 28 (October): 916-918.