UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

LIQUIDITY PREMIUM AND INVESTMENT HORIZON

A research report on the influence of liquidity on the return and holding period of securities on the Johannesburg Stock Exchange

by

Barend Christiaan Vorster

Submitted in partial fulfilment of the requirements for the degree

MASTER OF BUSINESS ADMINISTRATION

in the Faculty of Economic and Management Sciences

University of Pretoria

Pretoria

Study leader

Prof M Ward

April 2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DECLARATION

"I, Barend Christiaan Vorster, herewith declare that the language of this research report has been edited by R Burger."

……………………………………..

Student's signature

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ACKNOWLEDGEMENTS

Professor Mike Ward for his insightful suggestions and guidance.

I-Net Bridge for providing the data at no cost.

LIST OF ABBREVIATIONS

CAPM:	Capital asset pricing model
JSE:	Johannesburg Stock Exchange
MRP:	Market risk premium
R^2:	Coefficient of determination
SML:	Security market line

LIST OF EQUATION SYMBOLS[1]

α	Intercept of the regression equation
β	Beta
Cap	Market capitalisation. This is calculated as the number of shares outstanding times the share price and is used as a proxy for company size
CL	Cost incurred due to illiquidity
cr	Cumulative return
D	Dummy variable. Used to separate different periods from each other in the regression equation
Δ	Change in a parameter
ε	Random variation in the dependant variable not captured by the regression equation. If the regression equation includes all the possible independent variables with explanatory power then ε represents non-systematic risk which is also known as firm specific risk
Hp	Holding period. This is calculated as average shares in issue over a period (mostly one month) divided by the trade volume over the same period
LP	Liquidity premium
MRP	Market risk premium
n	Number of periods
r	Return. Also used to denote expected return depending on the context in which it is used
rf	Risk free rate. The money market rate usually serves as a proxy
RV	Relative volume. The monthly trading volume divided by the number of shares in issue serves as a proxy for trading frequency
S	Relative spread. It is calculated as the average difference between bid and ask prices divided by the average share price over the same period
SII	Shares in issue
Var	Variance in share return
Volume	Volume of shares traded in a specific period
x	Weighting of an individual security in a portfolio

[1] The symbols of all published equations have been changed in order to standardise variables according to the symbols presented in the list of equation symbols.

List of subscripts used in equations

ave	Denotes an average
I	Denotes an index
i	Denotes a single security
n	Number of periods
m	Denotes the market. This is equivalent to the all share index (ALSI) on the JSE
p	Denotes portfolio
T	Denotes a starting point or period

List of superscripts used in equations

e	Used to denote return in excess of what is predicted by CAPM

EXECUTIVE SUMMARY

Liquidity is a measure of the ease with which an asset can be converted into cash. In a perfectly liquid market, conversion is instantaneous and does not incur costs. Amihud and Mendelson (1986:224) proposed that illiquidity increases the expected return on an investment (liquidity premium) and simultaneously lengthens the holding period. These two effects are known respectively as the “spread-return relationship” and the “clientele effect” and have theoretical as well as practical implications. From a theoretical perspective it may help to explain the gap between the capital asset pricing model (which assumes that markets are perfectly liquid) and the associated empirical evidence; which thus far has been rather poor. From a practical perspective, liquidity will influence stakeholders’ decisions and market competitiveness (Amihud & Mendelson, 1991:61-64). The relevant stakeholders are governments, stock exchange regulators, corporations, investors and financial intermediaries. Emerging economies such as the South African economy typically have less liquid markets than the developed world. While this may be attractive for investors looking for higher returns, Amihud and Mendelson (1991:61) are of the opinion that liquid markets are more generally favoured by investors. Constantinides (1986:842-858), also proposes a model for liquidity, but found the liquidity premium to be of lesser importance than that proposed by Amihud and Mendelson (1986:223-231) but also supports the suggestion that investors will favour liquid markets.

Although it is by no means a perfect proxy, a security’s bid-ask spread has been found to be an attractive and effective measure of liquidity. It has been found to correlate with beta as well as market capitalisation and several other variables commonly used in capital markets research. Because of this correlation the effect of the bid-ask spread cannot be studied in isolation when regression techniques are employed (Ramanathan, 1998:166). This is particularly problematic because empirical evidence for beta, which is arguably the most important independent variable in financial cross sectional relationships, is weak. Beta has to be estimated and so it is not clear if real markets do not support CAPM theory or if beta cannot be estimated with the required accuracy. All of the common independent variables used in empirical capital markets research are correlated to beta, and for this reason it cannot be established if these variables have a real effect or if they are simply serving as a proxy for the difference between the real and the estimated beta. Various strategies have been proposed to increase the accuracy of beta estimation and these are discussed in detail in this research. Successes with these strategies have been mixed. A second problem encountered in the empirical research base relating to the CAPM is that in the theory the cross-sectional relationship is between expected market return (which cannot be observed

due to the vast number of real investments beyond those listed on exchanges) and beta, whereas empirical research makes use of actual return on a market proxy and beta. In order for the actual return to approach the expected return, empirical studies have to be conducted over extended periods. Accurate data for such periods are generally lacking and severe macro-economic changes such as wars, may also affect rational economic behaviour. It has to be kept in mind that the entire CAPM theory flows from the simple assumption that investors aim to achieve the highest return per unit of risk, and so a rejection of beta is a rejection of rational investor behaviour. Liquidity however, addresses one of the assumptions of CAPM, namely that markets are perfectly liquid; which obviously is not met in real markets and so CAPM models expanded for liquidity should be a reasonably fundamental starting point for all empirical capital markets research.

The current empirical evidence for the spread-return relationship is inconclusive. While some researchers have found a significant relationship, others have questioned the ability of the methodology to differentiate a true relationship from the 'proxy for errors in the estimated beta' problem. Deductions (as explained in section 4.3) that have been made from the research of Marshall and Young (2003:176-186) in particular, provide strong evidence that at least some of the relationship is due to the 'errors in estimated beta' problem. Little empirical work has been done on the clientele effect. Atkins and Dyl (1997:318-321) found a significant relationship between holding period and bid-ask spread, although their approach was somewhat unorthodox in the sense that portfolio formation was not done and the effect of beta was not tested.

This study tests empirically both the spread-return relationship and the clientele effect on the Johannesburg Stock Exchange over the period stretching from January 2002 to June 2007. The methodology of Fama and Macbeth (1973:614-617) as well as the aggregated beta of Dimson (1979:203-204) were mainly used, with some modifications as suggested by other researchers. With regard to the spread-return relationship, the findings of this study do not support theoretical expectations. This may be due to the short time period that was used as well as the difficulty in estimating beta. To the contrary, very significant evidence for the clientele effect was found, with little to no influence from market capitalisation and beta, which is as expected.

Further investigation into the spread-return relationship is required. If a liquidity premium is not present, foreign investors will favour liquid developed markets above the JSE. This implies that efforts of exchange regulators and the government to decrease illiquidity will lead to foreign portfolio investment inflow into the South African economy.

CHAPTER 1: THE RELEVANCE AND IMPLICATIONS OF LIQUIDITY

1.1 Introduction

This chapter is a brief introduction to liquidity. It aims to define and outline the concept and give reasons for its academic as well as practical importance. The focus then shifts to the aims and importance of the current study in relation to previous research. It concludes with a brief outline of the rest of the dissertation. The reader is required to have an understanding of the theoretical concepts of the capital asset pricing model (CAPM).

1.2 Definition of liquidity

Liquidity is a measure of both the cost and ease with which an asset can be sold (Bodie, Kane & Marcus, 2005:297). In a perfect market the conversion between assets and cash and vice versa is instantaneous and at no cost. In the real world transactions proceed at varying degrees of ease and incur costs. This challenges one of the assumptions of CAPM which states that investors do not incur transaction costs.

1.3 Determinants of liquidity

Liquidity costs can be divided into explicit and implicit costs (Aitken & Comerton-Forde, 2003:46). Explicit costs refer to costs that are not incurred in the market making process of an exchange. This includes brokerage commissions, exchange fees and taxes. Implicit costs are incurred as a result of imperfections in the market making process of an exchange and represent the difference between the actual execution price and the intrinsic price. These are incurred because of imperfect information flow; imperfect supply and demand economies; imperfect competition between and within markets; delay and search costs; irrational investor beliefs; and exchange control and regulation. It is customary not to include explicit costs in liquidity theory. There are however, a number of reasons why this may lead to incorrect conclusions:

I. Brokerage fees are not the same for all investors. This could be especially important when comparing liquidity across different exchanges.
II. Performance appraisal of institutional investors is based on before tax performance, as opposed to individual investors who are more likely to consider tax implications in switching decisions.

III. Not all investment vehicles are subject to the same tax implications which could influence investment horizon decisions.

IV. Securities on which losses were incurred are not subjected to capital gains tax which makes them more liquid.

Thus, as in the case of the CAPM, where theoretical models assume conditions that are not always met in empirical research, so too will there be some degree of disagreement between assumed market conditions in theoretical models on liquidity, and actual market conditions experienced in empirical research.

1.4 Implications of illiquidity

An improved understanding of liquidity has theoretical and practical implications. From a theoretical perspective, liquidity theory may assist in narrowing the gap between theoretical expectations and actual outcomes. From a practical perspective an improved understanding of liquidity may assist stakeholders in reaching optimal decisions. The relevant stakeholders and their interest in liquidity are as follows:

1.4.1 Government

Liquidity influences market competitiveness. This can be seen from the fact that investors tend to favour liquid markets more than illiquid markets (Amihud & Mendelson, 1991:61). This is true for local, but especially for international investors. Since government can influence explicit and implicit liquidity costs via regulation, it has influence over investment flows into a stock exchange (Amihud & Mendelson, 1991:63, 65). Furthermore, since liquidity risk is not usually viewed separately from other sources of risk, it has the potential to increase the cost of capital for corporations. A government that is conscious of the effect of taxes and regulation on liquidity can increase portfolio inflows on the financial side and capital investment on the real side of the economy which will enhance the overall capital formation process.

1.4.2 Corporations

Liquidity influences the investor's expected return and likewise the corporation's cost of capital. It would therefore make sense for directors to take measures that would ensure liquidity. This could be achieved through a commitment to transparency and information flow to the public domain which could also be a natural deflection against insider trading. Closely held corporations are more likely to suffer from illiquidity. Thus while directors of such firms may be more shielded from the actions of shareholders, they are likewise exposed to a higher cost of doing business. This poses a greater risk for those who choose have their capital tied up in such closely held firms. Liquidity gains is also one of the incentives for privately held corporations to go public, despite the high costs involved in the process (Amihud & Mendelson, 1991:62-63).

1.4.3 Investors

Investors with longer investment horizons should invest in less liquid securities. The rationale behind this is explained in the next chapter. Fund managers can assist in this process by not only stating the appropriate index against which the fund is to be benchmarked, but also specifying the investment horizon of the fund. Financial advisors should consider both a client's risk aversion and expected holding period when tailoring investment strategies (Amihud & Mendelson, 1991:62).

1.4.4 Stock exchange regulators

For the same reason as mentioned in section 1.4.1, stock exchange regulators should strive to lower liquidity premiums by means of trading regulation, information dissemination and technology that supports market integration (Amihud & Mendelson, 1991:63-64).

1.4.5 Financial intermediaries and financial engineering

An increase in liquidity may well be one of the primary roles of financial intermediaries (Amihud & Mendelson, 1991:62-63). There are a number of situations in which this is evident. For example,

- Underwriters increase liquidity by disseminating company information and by bearing the underwriting risk (guaranteeing that all shares will be taken up by the market, even if it means that the underwriter has to buy the shares).
- Intermediaries of asset securitisation increase liquidity and thus lower the cost of borrowing.
- Derivatives increase trading in the underlying asset which leads to higher liquidity.
- Financially engineered instruments can also at times decrease liquidity and financial intermediaries need to be cognisant of this.

1.5 Research hypothesis and aim

The aim of this study is to investigate the validity of the clientele effect and the spread-return relationship (see next chapter) as proposed by Amihud and Mendelson (1986:228) through empirical research on the Johannesburg Stock Exchange (JSE). Knowledge of its validity is important for the reasons outlined in section 1.4.

1.6 Assumptions and limitations

1.6.1 Study design limitations

The bid-ask spread may not be the best measure of liquidity (Aitken & Comerton-Forde, 2003:58). It may also not be the only determinant of the holding period (Atkins & Dyl, 1997:318). This study makes the assumption that the bid-ask spread is an adequate measure of liquidity. Chapter 4 explains why this may or may not be the case. As for the holding period, this study tests only the bid-ask spread, beta and the market capitalisation as possible determinants. This study also suffers from the same limitations as other studies on liquidity: namely that it considers only implicit liquidity costs.

1.6.2 Data limitations

The data, which was provided by I-Net Bridge, does not include de-listings and share dividends. Not including de-listings has the effect of overestimating the return of small company portfolios because these are more likely to de-list due to insolvency. This is known as survivorship bias. To the contrary, the exclusion of dividends primarily affects large company portfolios as they are more likely to pay dividends. A high negative correlation has been found empirically between size and beta (Chan & Chen, 1988:317; Amihud & Mendelson, 1989:482). Thus small corporations will tend to have higher betas whereas large corporations will tend to have low betas. The combined effect of over- and underestimating the returns of small and large corporations respectively are an overestimation of the slope of the security market line (SML) in the second pass regression (see section 3.2.1). This overestimation could however be offset by the downwardly biased estimation of the beta of corporations (predominantly small companies) which do not trade synchronously with the market index (see section 3.2.2.2).

1.7 Unique contribution

This is, to the best of the author's knowledge, the first study of liquidity on the JSE. In addition, the author has also deviated from the methodology employed in previous studies by making use of portfolios rather than individual securities. Doing this greatly reduces the firm specific risk which lessens errors in the estimation of the regression parameters and improves the cross-sectional relations. Although this strategy has been used in empirical studies of the spread-return relationship, it has not been applied to the clientele effect (see section 2.2 for an explanation of the clientele effect).

The methodology of Fama and Macbeth (1973:614-617) was used in combination with the aggregated coefficient beta as suggested by Dimson (1979:203-204). Previous studies which showed that the Fama and Macbeth method does not support the Amihud and Mendelson model (Chen & Kan, 1995:6-8) did not make use of the aggregated beta coefficient estimation method.

The empirical model does not make the assumption that the liquidity factor is linear, instead it assumes that a logarithmic relation is a reasonable substitute for the concave piecewise linear relation, as described by Amihud and Mendelson (1986:230).

1.8 Plan of the study

In addition to this introductory chapter, this dissertation is composed of six chapters. Chapter 2 focuses on theoretical models of liquidity. The author also proposes a simplified model in addition to the technically complex models of Amihud and Mendelson (1986:225-231) and Constantinides (1986:844-858). This is followed by a general overview of the methodology used in empirical capital markets research in chapter 3, which is not limited to research on liquidity. In chapter 4 an overview is given of the most important empirical studies on liquidity. The proposed methodology is presented in chapter 5, and the results are discussed in chapter 6. The study concludes with a summary of the most important findings in chapter 7. Recommendations are made and opportunities for further research are outlined.

CHAPTER 2: THEORETICAL MODELS OF LIQUIDITY PREMIUMS

2.1 Introduction

Numerous studies have found poor empirical evidence for the CAPM, mainly due to an empirical beta which is flatter than the theoretical beta. Various explanations have been proposed to explain this phenomenon of which some relate to the methodology of the research (Fama & MacBeth, 1973:613-633) and some to limitations in the testability of CAPM in the real world (Roll, 1977:126-176). The existence of transaction costs is however a real shortcoming of the CAPM model, and a modified model that includes liquidity costs can be expected to narrow the gap between theoretical models and empirical research. The omission of an important variable like liquidity may cause biased estimations of the regression coefficients and constant which can lead to incorrect conclusions from significance tests (Ramanathan, 1998: 166). This chapter is a review of the two most widely accepted theories on liquidity and the expected outcome of empirical research if the theory holds.

2.2 The Amihud and Mendelson model

Amihud and Mendelson (1986:225-231) proposed a model to predict the effect of the bid-ask spread on asset pricing. The results of the model indicated that the bid-ask spread has implications for asset pricing, asset returns and holding periods and was summarised in the following two propositions:

"Proposition 1 (clientele effect): Assets with higher spreads are allocated in equilibrium to portfolios with (the same or) longer expected holding periods."

"Proposition 2 (spread-return relationship): In equilibrium the observed market (gross) return is an increasing and concave piecewise-linear function of the (relative) spread."

2.3 The Constantinides model

The Constantinides (1986:844-856) model arrives at virtually the same conclusions as the Amihud and Mendelson (1986:225-231) model; namely that increasing transaction costs lead to longer holding periods as well as an additional liquidity premium, although the premium is minor. In addition to this, this research paper also shows that liquidity models are limited by certain assumptions. More specifically liquidity premiums of single period models[2] are independent from portfolio variance whereas optimal trading models[3] are strongly positively correlated to portfolio variance. This again emphasises that models which are all to some extent based on homogeneous assumptions, can at most only approximately reflect real heterogeneous market conditions.

2.4 A simplification of liquidity premium theory

Bodie, Kane and Marcus (2005:297) provide a simplified model of liquidity. The following explanation of liquidity is along the same lines as their model, but attempts to give a better account of the holding period.

Consider two portfolios, A and B, with exactly the same systematic risk that is perfectly positively correlated and of which all non-systematic risk has been diversified. The supposition is made that no transaction costs or tax costs are incurred by them during trading. In an efficient market one can expect the two portfolios to be priced exactly the same and that an investor will be indifferent between the two portfolios. Now suppose that there is no bid-ask spread for portfolio A (a perfectly liquid portfolio) but that there is one for portfolio B (less than perfectly liquid portfolio). The returns of an investor in portfolio B will be penalised with the bid-ask spread whereas the returns of an investor in portfolio A will not incur any penalties. If investor behaviour is rational the increase in demand for portfolio A and decrease in demand for portfolio B should create a widening in price difference between the two portfolios until the market returns to an equilibrium state where the two portfolios are perceived to be of equal value. *Ceteris paribus*, portfolio B which is now traded at a discount can be expected to earn a higher rate of return than portfolio A. Stated otherwise, portfolio B can be expected to earn an illiquidity premium in addition to the return predicted by CAPM to compensate for its illiquidity in an efficient market.

Formally the cumulative return on a less than perfectly liquid portfolio is the market risk premium (MRP) multiplied by beta as predicted by CAPM as well as the liquidity premium minus the liquidation cost which can be expressed as:

---

[2] These models assume that investors enter the market randomly and plan to hold a portfolio of securities for a specific period.

[3] These models assume that investors plan to be infinitely invested in the market and aim to maximize utility by portfolio adjustments.

$$\left(1+r_p\right)^n - 1 = \left(1+\beta_p * MRP + rf + LP_p\right)^n - 1 - CL_p \text{ ....2.1}$$

Where:

r_p is the expected return on portfolio p,

β_p is the portfolio beta,

MRP is the market risk premium,

r_f is the risk free rate,

LP_p is the liquidity premium of the portfolio and

CL_p is the liquidity cost of the portfolio[4].

If r^e is defined as the excess return above what is predicted by CAPM, then equation 2.1 simplifies to:

$$(1+r_p^e)^n - 1 = (1+LP_p)^n - 1 - CL_p \text{ ....2.2}$$

and so the liquidity premium of portfolio p can be expressed as:

$$LP_p = \sqrt[n]{\left(1+r_p^e\right)^n + CL_p} - 1 \text{ ....2.3}$$

Suppose that all investors plan to hold a portfolio for one period. This is consistent with another one of the assumptions of CAPM, which states that all investors have the same investment horizon. In order to make the two portfolios, A and B, equally attractive, portfolio B will have to earn a liquidity premium equal to its liquidation cost. To see why this is so, consider the following: In a perfect market portfolio A can be expected to earn a return in addition to what is predicted by CAPM of 0%. In order to be equally attractive portfolio B also has to earn the same excess. By substituting $r^e{}_p = 0$ and $n = 1$ equation 2.3 simplifies to:

$$LP_p = CL_p \text{ ....2.4}$$

[4] Liquidity cost is usually incurred when a security is bought (the difference between the ask price and intrinsic price) as well as when the security is sold (the difference between the bid price and the intrinsic price). For the purpose of this discussion it is assumed that all the costs are incurred when the security is sold.

Extend the example above to include portfolio C which is even less liquid than B, but in all other aspects the same as A and B in a market where investors plan to hold a portfolio for either one or two periods. Based on the above reasoning investors who plan to hold a portfolio for one period will be indifferent between portfolio A and B.

Investors who plan to hold a portfolio for two periods will always favour portfolio B above portfolio A because an excess return is earned while the liquidity cost is now depreciated over two periods. The above argument can now be extended to portfolio C where the compounded expected excess return minus the liquidity cost over two periods has to equal that of portfolio B in order to ensure indifference between B and C. The argument can be generalised in the following way: If p portfolios are adjusted for risk and arranged in order of increasing bid-ask spreads from p_0 to p_p where p_0 represents a perfectly liquid portfolio and where the cumulative excess return on portfolio p in period n has to equal the cumulative excess return on p-1 in the same period, the expected liquidity premiums can be predicted in a sequential fashion in the following way:

$$(1+LP_p)^n - CL_p = (1+LP_{p-1})^n - CL_{p-1} \qquad \text{2.5}$$

which after rearrangement can be written as:

$$LP_p = \sqrt[n]{(1+LP_{p-1})^n + \Delta CL} - 1 \qquad \text{2.6}$$

ΔCL is the difference in liquidation cost between portfolio p and portfolio p-1.

In equation 2.5 the cumulative excess return on portfolio p (the left hand side of the equation) should equal the cumulative excess return on portfolio p-1 (the right hand side of the equation) which is the portfolio immediately prior to portfolio p in the range p_0 to p_p. As long as the liquidity premium of p_0 approximates a perfectly liquid portfolio with liquidity premium of zero the liquidity premiums of all other portfolios can be predicted in sequential fashion by sequentially applying equation 2.6.

The impact of this equation is illustrated in the following example: Consider 7 portfolios A to G, which have been adjusted for risk and have a range of liquidation costs from 0 to 30% of the bid price in a market where investors hold portfolios for 1 to 6 periods. The liquidity premium for portfolio B is calculated from portfolio A (the perfectly liquid portfolio). This premium is then used to calculate the compounded excess return in period 2 which is in turn used to calculate the liquidity premium of portfolio C according to equation 2.6. The process is repeated until the liquidity premiums of all portfolios have been calculated. The results are presented in Table 2.1.

Table 2.1 – Calculated liquidity premiums of portfolios A to G

	Portfolio						
	A	B	C	D	E	F	G
Liquidation cost	0%	5%	10%	15%	20%	25%	30%
Equilibrium period		1	2	3	4	5	6
Liquidity premium	0	5.00%	7.35%	8.78%	9.74%	10.42%	10.92%

The calculated liquidity premiums and liquidation costs are then used to calculate the cumulative return net of liquidation costs of all portfolios over 0 to 8 years. This is presented in Table 2.2 which also indicates the periods in which sequential portfolios are in equilibrium.

Table 2.2 – Expected cumulative excess returns over different periods

	Portfolio						
	A	B	C	D	E	F	G
0	0.00%	-5.00%	-10.00%	-15.00%	-20.00%	-25.00%	-30.00%
1	0.00%	0.00%	-2.65%	-6.22%	-10.26%	-14.58%	-19.08%
2	0.00%	5.25%	5.25%	3.33%	0.43%	-3.07%	-6.96%
3	0.00%	10.76%	13.73%	13.73%	12.16%	9.63%	6.48%
4	0.00%	16.55%	22.83%	25.03%	25.03%	23.66%	21.39%
5	0.00%	22.63%	32.59%	37.33%	39.16%	39.16%	37.92%
6	0.00%	29.01%	43.08%	50.70%	54.66%	56.26%	56.26%
7	0.00%	35.71%	54.34%	65.26%	71.67%	75.15%	76.61%
8	0.00%	42.75%	66.43%	81.09%	90.34%	96.01%	99.18%

The expected excess return per period net of liquidation cost for periods 1 to 8 is presented in Table 2.3.

Table 2.3 – Expected excess return over different periods

	Portfolio						
Period	A	B	C	D	E	F	G
1	0.00%	0.00%	-2.65%	-6.22%	-10.26%	-14.58%	-19.08%
2	0.00%	2.59%	2.59%	1.65%	0.21%	-1.55%	-3.54%
3	0.00%	3.47%	4.38%	4.38%	3.90%	3.11%	2.11%
4	0.00%	3.90%	5.27%	5.74%	5.74%	5.45%	4.96%
5	0.00%	4.16%	5.80%	6.55%	6.83%	6.83%	6.64%
6	0.00%	4.34%	6.15%	7.07%	7.54%	7.72%	7.72%
7	0.00%	4.46%	6.40%	7.44%	8.03%	8.34%	8.46%
8	0.00%	4.55%	6.57%	7.70%	8.38%	8.78%	8.99%

The expected excess return per period of portfolios A, B, C and G are graphically displayed in Figure 2.1 below which illustrates the relationship between expected excess return and the holding period. Note that progressively illiquid portfolios outperform the liquid ones over longer time periods.

Finally in Figure 2.2 the liquidity premium is plotted in relation to liquidation cost.

Figure 2.1 – Expected excess return as a function of the holding period for portfolios A, B, C and G

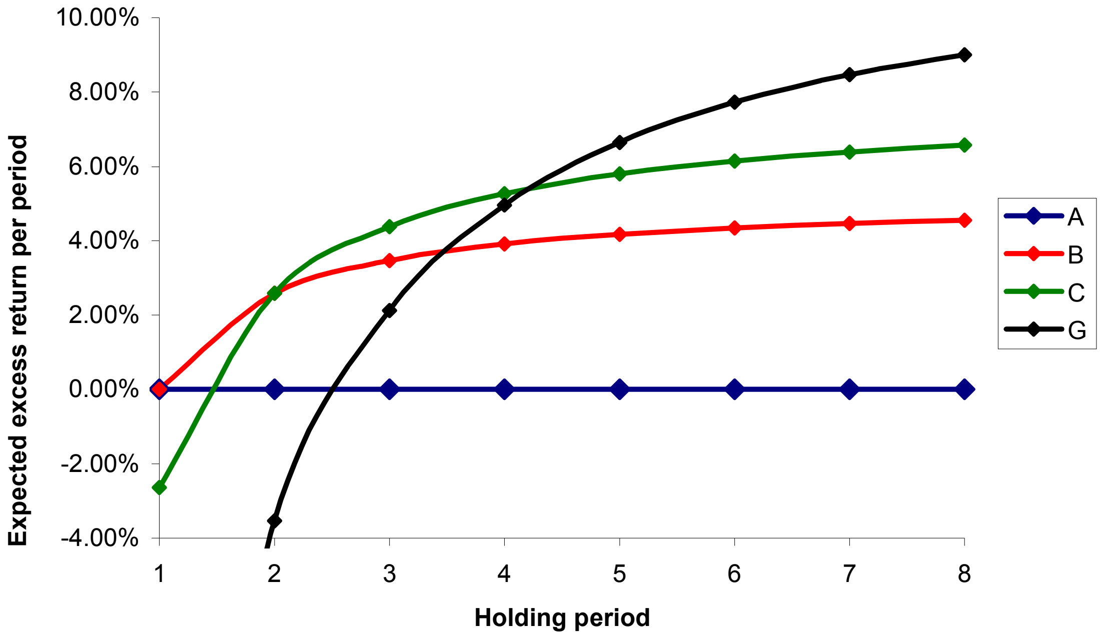
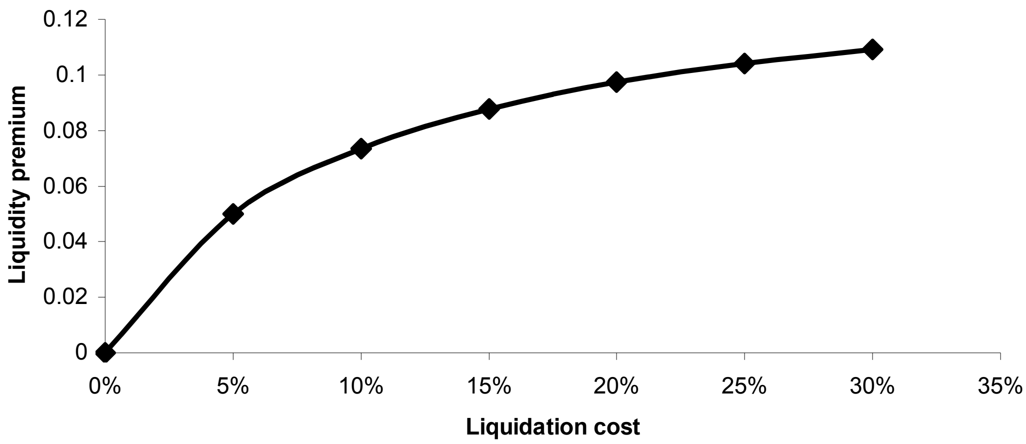


Figure 2.2 – Liquidity premium in relation to liquidation cost



2.5 Conclusion

The effect of liquidity can be summarised as follows:

I. In an efficient market assets that are less than perfectly liquid earn a liquidity premium to compensate for the asset's liquidation costs.
II. The less liquid the asset (the higher the liquidation cost) the higher the liquidity premium.
III. Investors with longer investment horizons have a preference for less liquid portfolios. This is equivalent to the clientele effect of Amihud and Mendelson (1986:228).
IV. The increase in the liquidity premium in relation to liquidation cost is not linear, but flattens out with increasing liquidation cost. This is equivalent to the spread-return relationship of Amihud and Mendelson (1986:228). This implies that more liquid assets are more affected by changes in liquidity than less liquid assets.

CHAPTER 3: METHODOLOGY USED IN EMPIRICAL TESTS OF MARKET EFFICIENCY

3.1 Introduction

The methodology used in tests of market efficiency mainly derives from the attempts of researchers to establish whether or not the CAPM is empirically supported. The CAPM (Sharp, 1964:425-442) is based on the assumption that investors seek to optimise the Sharp ratio, that is, that they seek to achieve the highest return per unit of risk (mean-variance efficiency) which is achieved through the Markowitz portfolio selection model (Markowitz, 1952:77-91). Under certain restrictive conditions (*viz:* investors are price takers, have an identical holding period, are limited to publicly traded financial assets, do not incur taxes or transaction costs and have the same economic outlook) it can be argued that all investors will select the same portfolio, which will be the market portfolio, and further that the expected return of a selected security (security i) is given by

$$r_i = \beta_i \left[r_m - rf \right] - rf \quad \text{...} 3.1$$

(Bodie, Kane & Marcus, 2005: 282).

From this the following empirically testable relationships can be formulated:

I. The risk-return relationship is linear

II. All systematic risk is captured by β

III. There is a positive expected return-risk ratio

IV. The empirically established rf (the intercept of the second pass regression – see section 3.2.1) is equal to the market risk free rate, for which a short term interest rate is used as a proxy.

V. The empirically established MRP (the slope of the second pass regression – see section 3.2.1) is equal to the weighted market portfolio minus the risk free rate proxy. The all share index of an exchange usually serves as a proxy for the weighted market portfolio.

Empirical tests of CAPM started in the sixties (Lintner, 1965:610-612; Miller & Scholes, 1972:53-71). The findings of these early studies were inconsistent with CAPM. Of the five testable hypotheses listed above, only hypotheses I and III had convincing empirical evidence. Roll (in Bodie, Kane & Marcus, 1995:424) states that all of the hypotheses commonly tested in empirical studies are directly derived from the mean-variance efficiency assumption. Rejection of this assumption, in other words, claiming that investors do not aim to achieve the highest return per unit of risk, would be a rejection of rational human behaviour and cannot be accepted.

Various explanations and propositions have been made to explain the paradox of the inconsistency between the model and the empirical findings which can broadly be categorised as propositions related to statistical methodology, and propositions that aim to improve or expand on the CAPM model. This chapter starts off with a general framework of the statistical methodology used in empirical research. It is followed by various improvements on the statistical methodology that have been suggested. It then moves on to the various expanded and also alternative models to CAPM, and concludes with some remarks on the different methodologies.

3.2 Statistical methodology of empirical tests

3.2.1 The basic methodological process of empirical research

The methodology that underlies most studies is a two stage linear regression procedure. On a selected exchange a number of periods (usually months) starting at say, month T and ending after n months are selected for the study. After exclusion of securities that are deemed inappropriate for the study, the betas of the remaining securities are estimated by regressing the serial security returns in excess of the risk free rate against the serial market proxy returns in excess of the risk free rate over a period extending from T_{-k} to T_{-1}, where k is usually 60 periods, or the 5 years prior to the starting time if periods are in months. Thus the first pass regression equation for security i is:

$$r_i - rf = \alpha + \beta_i (r_m - rf) + \varepsilon_i \qquad 3.2$$

If CAPM holds α should be zero. To see why this is so β is set to zero in which case equation 3.2 becomes:

$$r_i = \alpha + rf + \varepsilon_i \qquad 3.3$$

According to CAPM the return on r_i should now be equal to the risk free rate and thus α has to equal zero. By following this procedure the estimated betas of all securities included in the model are calculated. These betas now become the independent variables for the second pass regression against which serial security returns for the period T to T_{+n} are regressed. This second pass regression which now includes all securities, is the empirical equivalent of the SML in CAPM and is as follows:

$$r_i = \gamma_0 + \gamma_1 \beta_i + \varepsilon_i \qquad (3.4)$$

Now if CAPM holds, γ_o should equal the risk free rate and γ_1 should equal r_m-r_f which is the MRP. This is essentially the methodology that was followed by Miller and Scholes (1972:52-54).

3.2.2 Improving the estimates of the regression parameters

Roll and Ross (1994:101) argue that the expected return-beta relation is exact, so no other independent variable should have explanatory value. Thus one explanation for the poor empirical results could be that the model parameters are not estimated with the required accuracy. Their focus is mainly on the proxy used for the market portfolio. If the proxy is not a good estimation of the market portfolio, then the exact relation is lost and any other variable that is correlated with the inefficient market proxy will have explanatory power. This argument is essentially the starting point of the next section, but is mentioned here because what is equally important is that even if the market portfolio is estimated with the required accuracy, the exact relation will still not hold if the estimates of beta are inaccurate. Thus before the explanatory power of variables other than beta are assessed, it is essential that the estimates of beta be as accurate as possible.

3.2.2.1 Portfolios vs. individual securities

Here the use of portfolios as opposed to individual securities for the estimation of beta will be discussed.

Black, Jensen and Scholes (1972:85-91) and Fama and Macbeth (1973:614-618) used portfolios instead of individual securities in the second pass regression. The reasoning behind this is that the use of portfolios decreases non-systematic risk through diversification and thus allows for better estimations of beta. The beta of any portfolio can be calculated as:

$$\beta_p = \sum_{i=1}^{N} x_i \beta_i \quad \text{...3.5}$$

where x_i is the weight of the i[th] security in the portfolio (Fama & Macbeth,1973:614).

If each of the estimated betas of securities 1 to N is independently distributed around the true beta then the portfolio's estimated beta will be a better estimate than the individual security estimates because the independent estimation errors will tend to offset each other. The drawback is that the number of independent variables available for the second regression is reduced from the total number of securities available to the total number of portfolios that are formed. In order to compensate for this, portfolios are selected to cover a wide range of betas. It is well known that the correctness of the estimated parameters of the regression equation are more dependant on the range and spread of the independent variables than the number of measurements. The betas of the T_{-k} to T_{-1} period cannot be used for the portfolio formation process as betas with a positive sampling error will be allocated to higher beta portfolios and those with negative sampling error, to lower beta portfolios.

There are two possible remedies for this problem. The first approach, as used by Fama and Macbeth (1973:615-617), is to use a period prior to T_{-k} for the portfolio formation process. The second is to use a variable that is closely correlated to beta. Chan and Chen (1988:314-315) used market capitalisation for the portfolio formation process. They found that beta is highly correlated to size and that when beta is accurately estimated, size has no further explanatory power.

There is a further benefit to be derived from using portfolios as opposed to individual securities. There is evidence that the distribution of security returns conforms better to symmetric stable distributions than Gaussian distributions (Mandelbrot, Fama & Roll in Blume, 1970:166-173). This observation has important implications for forecasting and tests of hypotheses. Symmetric stable distributions are described by a mean, a dispersion parameter and a characteristic exponent which can take values between 0 and 2 (exclusive of 0 and inclusive of 2). The Gaussian distribution has a characteristic exponent of 2. The further the characteristic component is from 2 the greater the area that that falls under the tails of the distribution's density function. The characteristic exponent of expected security distributions is 1.7, but when portfolios are used a characteristic exponent between 1.7 and 2 does not improve the performance of tests of hypothesis (Blume, 1970:166-173). Thus a normal Gaussian distribution can be assumed when working with portfolios.

3.2.2.2 The estimation of beta using the aggregated coefficient method

A potential source of bias in the estimation of the betas of illiquid securities which are not traded frequently, is the discrepancy in the time of security return recording between these securities and the index (Dimson, 1979:197-224). A common scenario is where a security was traded somewhere between two recording intervals, whereas the index return is traded at the end, or quite close to the end of the interval. Since returns are calculated from the prices obtained during trading, the illiquid security return is lagging the index return. This effect is known as non-synchronous trading. The reverse situation, where a security is traded more frequently than the index is also possible, although much less common, and with lesser effect. The lagging of illiquid security returns behind that of the index has the following effects:

I. The covariance between illiquid securities and the market is underestimated.

II. The underestimation leads to a beta estimate that is downwardly biased.

III. Underestimating the betas of illiquid securities leads to overestimation of liquid security betas as the average beta has to be 1.

IV. Leading and lagging serial correlation is introduced into the data with the number of lagging and leading correlations dependant on the frequency of trade.

This type of bias is well characterised by the intervalling effect. This effect is the tendency of the coefficient of determination (R^2) to increase as the interval between consecutive measurements is increased. The average beta of the most frequently traded securities on the London Stock Exchange fell from 1.15 to 0.99 (R^2 improved from 36% to 48%) whereas the average beta of the most infrequently traded securities rose from 0.5 to 0.72 (R^2 improved from 6% to 20%) when the interval was lengthened from 1 month to 6 months (Dimson, 1979:214-215). Dimson (1979:197-224) suggested that the aggregated lagging and leading coefficients should be used for the estimation of beta. The aggregated beta coefficient is calculated as:

$$r_{i_n} - rf = \alpha + \sum_{k=-p}^{p} \beta_i \left(r_{m_{n+k}} - rf_{n+k} \right) + \varepsilon_i \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.6$$

where r_{i_n} is the return on security i in period n and beta is calculated as the sum of the coefficients of the current (k = 0) as well as $-p$ to p lagging and leading index returns regressed against r_{i_n} He showed that the periods $k = -4$ to $k = +1$ showed significant cross correlation, with the greatest correlation attributable to $k = -1$ and $k = 0$. The reproducibility of this finding is obviously related to market liquidity, but it is likely that the lagging coefficients will also be more important than the leading coefficients in other markets. Dimson (1979:216) also tested the aggregated coefficient methodology and found that the average beta coefficient of the most liquid securities changed from 1.16 to 0.93 (R^2 changed from 34.6% to 36.6%) and that of the most illiquid stocks changed from 0.47 to 0.91 (R^2 changed from 5.3% to 8.4%).

3.2.2.3 The use of feasible generalised least squares regression

Ignoring heteroscedasticity[5] in linear regression leads to invalid tests of hypotheses due to an overestimation of the variance in the least square regression procedure (Ramanathan, 2002:346). One remedy to this problem is to use feasible generalised least square regression. In this procedure each variable is divided by an estimation of the standard deviation of the error term, after which ordinary least square regression is applied. This leads to residuals that are homoscedastistic[6] and valid tests of hypotheses.

3.2.2.4 Pooling of the cross sectional and times series data and seemingly unrelated regression

This regression procedure recognises and takes into account the fact that cross sectional correlation exists between residuals at a given point in time (it is likely that macro-economic effects influence all cross sectional errors in a similar

[5] Heteroscedasticity implies that one of the assumptions of ordinary least square regression is violated, namely that the variances of the error terms are equal.

[6] This implies that the variances of the error terms are equal.

way). This is known as contemporaneous correlation (Ramanathan, 2002: 479). The main drawback of this methodology is that the MRP is assumed to be constant over time (Chen and Kan, 1995:1-11). This has the effect that any factor that is correlated to the true MRP (assuming that it is not constant over time) will now appear to have explanatory power.

3.2.2.5 The Cross-sectionally correlated and timewise autoregressive model

This regression model is in the same spirit as seemingly unrelated regression, except that it also takes into account longitudinal correlation between successive data points. This is known as serial or autocorrelation (Ramanathan, 2002:380). The drawback of this model is that it estimates only one beta for the entire model and thus requires portfolios to be randomly selected (same beta portfolios) as well as the beta of each portfolio to be constant over time (Marshall and Young, 2003:178).

3.3 Expanded CAPM models and alternatives

Arguably the most explored topic regarding portfolio selection during the last four decades has been the inclusion of variables other than beta in the estimation of expected returns. If an independent variable with a non-zero regression coefficient is excluded from a regression model, the estimation of the constant term will be biased as well as all other regression coefficients that are correlated to the omitted variable (Ramanathan, 2002:166). It should be remembered that the expected return-beta relation is exact given the restrictions. Thus, if the rational investor aims to optimise the sharp ratio; beta is measured accurately; the expected market portfolio is known and all restrictions of the model are met, then according to CAPM no other variable should have explanatory power. Nevertheless, a number of studies have proposed multifactor models to better explain security returns.

The proponents of multifactor models, which may or may not include beta, can be broadly divided into two groups. The first states that even if beta is an efficient explanatory variable, it is still insufficient, because it groups all macro-economic risk exposures into one risk measure. Since securities are not typically exposed to all risk factors in equal measure, multifactor models could enable better prediction of future returns. This is of course based on the assumption that macro-economic conditions can be predicted with the required accuracy and that a proxy that is highly correlated to the factor in question is readily available. But even if macro-economic conditions cannot be predicted, an understanding of degree of exposure to different factors may still be valuable in hedging strategies.

The second group, which in terms of the literature is the predominant group and mainly of interest in this study, is more concerned with the proving or disproving of the expected return-beta relationship. Since the aim is to disprove the theory, there is no prerequisite that the independent variables in the regression equation be proxies for macro-economic exposure, and so any variable with explanatory power can be used. If the expected return-beta relationship does not hold, then one is forced to conclude that either the mean-variance hypothesis does not hold uniformly, or that one of the prerequisites of CAPM does not hold in practice. Since investors do incur transaction costs and since there is a market for liquid, as well as illiquid securities, one is safe to conclude that CAPM should at least be expanded to include liquidity effects. The following section discusses a number of multifactor models which should be considered.

3.3.1 Fama and French (1992) model

Fama and French (1992:427-464) investigated the combined effect of several variables that had previously showed correlation with security returns. The study covered the period from 1963 to 1990 and was also expanded to include the period 1941 – 1990. The results, which were rather disturbing and sparked some frenzy in the financial media, were as follows:

I. When beta was the only independent variable in the second pass regression the monthly market risk premium was only 0.15% which was not statistically different from zero.

II. When market capitalisation/ firm size was included in the regression, beta did have statistically significant explanatory power, but the coefficient/slope was negative!

III. When market capitalisation and book to market equity were included in the regression, the other tested independent variables (leverage and earnings to price ratio) did not have explanatory power.

Table 3.1 – Disentangled effects of market capitalisation and beta

	All	Low-β	β-2	β-3	β-4	β-5	β-6	β-7	β-8	β-9	β-10
Average monthly returns (in percent)											
All	1.25	1.34	1.29	1.36	1.31	1.33	1.28	1.24	1.21	1.25	1.14
Small-Cap	1.52	1.71	1.57	1.79	1.61	1.50	1.50	1.37	1.63	1.50	1.42
Cap-2	1.29	1.25	1.42	1.36	1.39	1.65	1.61	1.37	1.31	1.34	1.11
Cap-3	1.24	1.12	1.31	1.17	1.70	1.29	1.10	1.31	1.36	1.26	0.76
Cap-4	1.25	1.27	1.13	1.54	1.06	1.34	1.06	1.41	1.17	1.35	0.98
Cap-5	1.29	1.34	1.42	1.39	1.48	1.42	1.18	1.13	1.27	1.18	1.08
Cap-6	1.17	1.08	1.53	1.27	1.15	1.20	1.21	1.18	1.04	1.07	1.02
Cap-7	1.07	0.95	1.21	1.26	1.09	1.18	1.11	1.24	0.62	1.32	0.76
Cap-8	1.10	1.09	1.05	1.37	1.20	1.27	0.98	1.18	1.02	1.01	0.94
Cap-9	0.95	0.98	0.88	1.02	1.14	1.07	1.23	0.94	0.82	0.88	0.59
Large-Cap	0.89	1.01	0.93	1.10	0.94	0.93	0.89	1.03	0.71	0.74	0.56

A series of 2x2 matrices were constructed to compare both 'market capitalisation and beta' and 'market capitalisation and the book to market ratio'. Table 3.3 was taken from Fama and French (1992:434) and illustrates the disentangled effects of beta and market capitalisation. In this matrix 10 portfolios were constructed at the beginning of each year based on 10 deciles of market capitalisation. Each of these 10 portfolios were then subdivided into a further 10 based on pre-ranked betas calculated over a two to five year period prior to 1963 (see section 3.2.2.1). The rationale behind this methodology is that it enables separation of the size and beta effects which were previously shown to be highly correlated (Chan and Chen 1988:317). This methodology resulted in the calculation of the average return of 100 portfolios over the study period which is displayed in the 2x2 matrix. By examining the rows (increasing market capitalisation for various deciles of beta) and columns (increasing betas for various deciles of market capitalisation) a disentangled view of both the market capitalisation effect and beta is obtained.

The same strategy was followed for the market capitalisation – book to market ratio matrix. Multiple regression analysis showed that the correlation between market capitalisation and returns as well as book to market ratio and returns held independently, but beta had no explanatory power when it was corrected for market capitalisation. Fama and French (1992:444) offered 'company distress' as a possible risk factor captured by the book to market equity ratio but no explanation is offered for the market capitalisation effect.

Interestingly the natural logarithm of both the market capitalisation effect and the book to market equity effect was used in the multiple regression which suggests that these two factors could also be proxies for liquidity (see section 3.3.4).

3.3.2 The Amihud, Christensen and Mendelson (1992) model

Following the portfolio formation technique of Fama and French (1992:430-432), Amihud, Christensen and Mendelson (1992:1-16) tested the significance of beta, market capitalisation and the residual standard deviation of the regressed portfolios (a proxy for firm specific risk) for the period from 1953 to 1990. They criticised the regression methodology of Fama and French in that it did not take into consideration cross-sectional correlation and heteroscedasticity in portfolio returns. By using more sophisticated regression techniques (joint pooled cross-section and time-series estimation as well as generalised least square regression) first suggested by Amihud and Mendelson (1986:233-237) they show that all three of the tested independent variables have statistically significant coefficients. Thus although the beta expected return relation was not exact, it did show at least some predictive power.

3.3.3 The Roll and Ross proposition

Roll and Ross (1994:101-120) also criticised the rejection of beta by Fama and French (1992:427-464). Using a simulation technique they show that when the market index proxy is not on the efficient frontier (even if it is only 22 basis points below the efficient frontier) it can produce errors in the beta estimations that lead to a cross-sectional regression between beta and the portfolio return that is virtually zero. They conclude that since the true market portfolio cannot be observed, CAPM is of little use in analysing stock returns. Roll and Ross also comment on the use of generalised least squares as used by Amihud, Christensen and Mendelson (1992:2-5). Using this technique the cross-sectional relation between beta and expected returns becomes more robust and will produce a positive relation as long as the market portfolio proxy is not grossly inefficient. This helps to explain why Amihud, Christensen and Mendelson (1992:15) found a statistically significant positive correlation when Fama and French (1992:447) did not.

3.3.4 The Amihud and Mendelson (1986 and 1989) model with liquidity incorporated

Following their theoretical development of liquidity theory, Amihud and Mendelson (1986:231-246) tested the use of the bid-ask spread as a proxy for liquidity in a multifactor model that included beta, the bid-ask spread and market capitalisation. They found that market capitalisation had no explanatory power when the bid-ask spread was included as an independent variable. They also repeated the study with an expanded model that also included residual deviation as a proxy for firm specific risk (Amihud & Mendelson, 1989:479-485). Residual deviation had predictive power when ordinary least squares regression was used, but not with generalised least squares regression. Their results suggest that market capitalisation may have had predictive power, because it serves as a proxy for liquidity.

3.3.5 The Fama and French three factor model

Davis, Fama and French (2000:389-405) expanded on the initial work of Fama and French (1992:427-464) and used beta, book to market equity ratio, and firm size as regression factors in what has become known as the Fama and French three factor model (so called due to the very high R^2 that was achieved with these three factors). Their interest was mainly in establishing the book to market equity ratio as an explanatory factor. It is argued that a book to market equity ratio may capture additional risk not captured by beta, may be a proxy for investor overreaction to corporations with strong fundamentals, or may indicate investor preference towards strong firms.

3.3.6 The conditional CAPM model of Jagannathan and Wang (1996)

Jagannathan and Wang (1996:3-37) addressed two problems of the Fama and French three factor model. Firstly they allowed the market risk premium and betas to vary over time and thus prevent market capitalisation and book to market equity ratio to capture risk not captured by a static beta and MRP. They also included return on human capital in the regression equation in an attempt to improve on the possibly poor ability of the market weighted portfolio to capture the true market portfolio as suggested by Roll and Ross (1994:101-102). They show that with these inclusions, market capitalisation and book to market equity ratio have no further predictive power and hence may have captured risk due to its correlation to the true market portfolio and true portfolio beta.

3.3.7 Chen and Kan criticism of the Amihud and Mendelson model with liquidity incorporated

The criticism of Chen and Kan (1995:1-12) is along the same line as the criticism of Jagannathan and Wang (1996:3-37) on the work of Davis, Fama and French (2000), namely that when beta is assumed to be constant when it isn't, factors that are correlated to beta (the relative spread in this case) will have explanatory power because it captures risk not captured by the static beta. They show that when a different methodology is used, the spread-return relationship largely disappears. Therefore, their study seems to support the suggestion of Constantinides (1986:843), namely that the spread-return relationship is at best, marginal.

3.4 Conclusion

The poor performance of beta in empirical studies poses a dilemma for the market researcher. The *ex ante* true market portfolio cannot be observed and similarly the calculated betas are only estimates of the true betas. These two problems create a dilemma when other explanatory independent variables are uncovered as there is no way of knowing whether the variables are truly explanatory or whether they are just correlated to the true beta and market portfolio. As illustrated in the study of Roll and Ross (1994:101-120), even slight errors in the estimation of these variables can severely distort the cross-sectional relation between expected returns and beta. Therefore there is no empirical way of testing the mean-variance assumption that underlies CAPM and no way of knowing whether or not beta captures all risk. Contrary to the uncertainty associated with single factor (beta) models the three factor model of Davis, Fama and French (2002:395-397) provides excellent cross sectional correlation ($R^2 \approx 0.91$). It is important to note that even though this model has excellent predictive capability, it gives no clue as to what systematic factors the firm size and book to market ratio are proxies for or, (and this is even less plausible) why investors should earn a riskless return in addition to what is predicted by CAPM for selecting portfolios of smaller corporations with higher book to market ratios. Bodie, Kane and Marcus (2005: 431) give the following possible explanations:

I. The size and book to market ratios capture risk that is not explained by CAPM. This could be due to shortcomings in the theory, or to the problems associated with the empirical testing of CAPM.

II. Irrational investor preference for large company and low book to market shares causes mis-pricing in markets.

Another possible explanation that is not considered in Bodie, Kane and Marcus (2005:431) is that firm size and book to market value may be correlated to violations in the assumptions of CAPM, i.e. liquidity premiums. This idea is supported by the study of Amihud and Mendelson (1989:479-485) and is more satisfying from a theoretical perspective, since real markets notably do not abide by the liquidity assumption of CAPM. It would thus seem prudent, from a theoretical perspective at least, to include at least beta and liquidity in all multi-factor models, even though Chen and Kan (1995:11-12) could find no evidence of a liquidity premium when beta estimation was optimised.

CHAPTER 4: EMPIRICAL EVIDENCE ON LIQUIDITY PREMIUMS

4.1 Introduction

This chapter provides an overview of the literature on empirical evidence of liquidity premiums. After a discussion on the measurement of liquidity, the remainder of the chapter is subdivided into two parts. The first part focuses on the spread-return relationship (proposition two of Amihud and Mendelson) and the second on the clientele effect (proposition one). The studies are discussed in terms of focus, methodology and results. The aim of this chapter is to provide a critical appraisal of landmark articles that contributed to the current understanding of liquidity premiums. From this analysis flows the formulation of the hypothesis of this study. The chapter is not intended to be a comprehensive overview of all the literature on the subject.

4.2 Measurement of liquidity

The model of Amihud and Mendelson (1986:225-231) assumes perfect agreement between the bid-ask spread and liquidity. Under this assumption the bid-ask spread captures all liquidity and no other factors influence the bid-ask spread. This assumption is not well accepted amongst researchers and numerous other parameters of liquidity have been proposed. A total of 68 measures of liquidity have been recorded in the literature (Aitken & Winn, 1977 in Aitken & Comerton-Forde, 2003: 46). Aitken and Comerton-Forde (2003:46-51) classify liquidity measures into two categories, namely trade- and order-based measures. Under trade-based measures the following four are identified:

I. Trading value: This is the absolute value of trade per time unit stated in monetary terms.

II. Trading volume: The absolute number of shares traded per unit of time. For the analysis of the data both the absolute and the relative trading volume, which is the absolute volume expressed as a fraction of the total number of shares in issue, were used.

III. Frequency/count: The number of transactions per unit of time, regardless of the size of the transaction.

IV. Turnover ratio: Trading value as a fraction of the total market capitalisation.

According to them, trade measures were used in earlier research due to the simplicity of the calculation and the availability of these measures. These are however, problematic because they reflect *ex post* market conditions rather than *ex ante*. Furthermore their research indicates that these measures reflect liquidity poorly in an *ex ante* sense which suggests that liquidity is not stable over time (at least not in their study population) and that past liquidity is a poor indicator of future expectations.

Order-based measures give a better indication of what the market expects the liquidity to be. The following order-based measures are examined:

I. The relative bid-ask spread: this is the absolute value of the bid-ask spread divided by the midpoint price. The calculation ensures that the spread is comparative across differently prices shares. This measure reflects only the best limit orders which are not necessarily the spread of all executed orders. This measure thus in a sense reflects the liquidity cost for small investors, but underestimates the cost for large investors.

II. Relative order depth: This reflects the number of shares in the order book divided by the number of shares in issue. It tries to address the above-mentioned problem. The measure however has problems of its own. As the orders moves away from the best bid-ask spread the probability of execution decreases. This would suggest that not all orders in the book are equally important, and that a weighting system should probably be used.

III. Aitken and Comerton-Forde (2003:56-57) suggest a new liquidity measure which is the main contribution of this study. The quoted distance that a limit order is removed from the best bid-ask spread, expressed as a percentage, is calculated (the order value) and all the distances are grouped into price bands. The probability of execution given a specific price band is then established (order weight). The weighted order value is then calculated as follows:

$$Weighted\ order\ value = \sqrt{\sum (AOV * AOW) * \sum (BOV * BOW)} \quad \text{.................4.1}$$

Where AOV and AOW are the ask order value and weight, and BOV and BOW are the bid order value and weight. They give no indication why the weighted order value was calculated in this way instead of using the arithmetic average which would be more consistent with the way in which bid-ask spreads are usually calculated. They do however, report that the use of the arithmetic average produces results consistent with the calculation shown above.

Aitken and Comerton-Forde (2003:49-57) examined the performance of these common measures of liquidity on the Jakarta stock exchange during the liquidity crisis experienced in Asia in the late 1990s. This provided an opportunity to study the effects of a change in liquidity on liquidity proxies. They examined the period from 1 June 1996 to 28 August 1998 by grouping 178 out of a possible 221 securities into four quartiles based on trading value and assumed that the 14th of August 1997, when the rupah was floated, was the start of the liquidity crisis. The difference and significance of the pre- and post-liquidity crisis period were calculated for each of the measures listed above, across each of four quartiles and collectively calculated with a t-test.

Prior to significance tests Aitken and Comerton-Forde (2003:51) calculated the correlation between all of measures mentioned above. Strong correlations, in excess of 0.8 were observed between three of the trade based measures namely; volume, value and frequency. The correlation between the relative volume and the other measures were in excess of 0.4. Only trading volume was further considered, as it can be deducted from the good correlations that the other parameters will give similar values. Poor to no correlation was observed between trade- and order-based measures and even between order measures. This suggests the measure of liquidity influences the results of empirical research.

The results and interpretation of the Aitken and Comerton-Forde (2003:51-57) study for the different measures are as follows:

I. Trade-based measures: Trading volume increased by 51% in the post crisis group. This is in contrast to the expectation that a decrease in liquidity will lead to a decrease in trading volume. Buyer initiated trading fell across all quartiles while seller initiated trading increased across all percentiles. The change increased with increasing illiquidity. This suggests that investors were more aggressive to exit shares and that this was the greatest in the most illiquid quartile. Since the expected relationship between liquidity and trading volume does not hold, they conclude that trading volume is not an effective measure of liquidity.

II. Relative bid-ask spread: In both the pre- and post-crisis groups the bid-ask spread increases with decrease in trading value. The bid-ask spread is also significantly higher in all of the post-crisis groups. The data seems to suggest that the liquidity crisis affects the liquid shares more than the illiquid ones because the percentage change decreases across increasingly illiquid quartiles. Aitken and Comerton-Forde (2003:53) offer the increase in the number undefined observations in the higher quartiles, as a possible explanation for this observation[7]. The relationship between the relative bid-ask spread and liquidity holds, which would suggest that it is a satisfactory measure of liquidity.

III. Order depth: With the exception of quartile 3 the order depths increased on both the bid and ask side of the order book. This is again in contrast to expectation. Further analysis, however reveals that orders were placed further away from the best bid-ask spread in the post-crisis period and thus had less chance of execution. Thus while order depth may not be an effective measure of liquidity, a measure that takes into account the probability of order execution may well be effective. This observation led Aitken and Comerton-Forde (2003:56-57) to the development of the weighted order value measure.

IV. Weighted order value measure: The new measure was also empirically tested. The results are consistent with expectations. The weighted order value decreases in all the quartiles. The results also show a progressively greater decline in the least liquid stocks which is again according to expectations. Recall that this was not the case when the bid-ask spread was used as a measure.

The study by Aitken and Comerton-Forde (2003:45-59) gives some important insights into the measures of liquidity. It specifically illustrates that trade-based measures cannot be used as measures of liquidity. It further makes the point that the best bid-ask spread may not be a reflection of the true illiquidity cost. The following critique and/or observations should however also be considered:

I. The authors postulate that the performance of the measures in a known liquidity crisis will be a good reflection of their performance in general. This seems reasonable but there is however, no way of knowing for sure that the performance will be the same under conditions not as extreme as that reported by this study. Market drivers may well be less rational during a crisis than during periods of stability. The uncertainty here is whether or not the results of Aitken and Comerton-Forde (2003:45-59) can be generalised to fairly stable markets?

II. The sample is partitioned into for quartiles of liquidity, with the first being the most liquid and the fourth the least. The partitioning is based on trading value. As discussed by the authors themselves trading value is an ex post measure of liquidity. Such *ex post* measures of liquidity may not be a good predictor of *ex ante*

---

[7] Whenever only a bid or ask price was quoted the data point was deleted from the sample.

expectations. In fact, the poor performance of the trade measures clearly illustrates that past performances are in fact not good indicators of future performance. If this was the case one would expect very good performance from the trade-based measures, especially since they are so strongly correlated. The impact of this observation is that the standard created by the classification system may have resulted in classification bias which would influence the performance of the order-based methods. On the upside, the bid-ask spread which is arguably the most accepted measure of liquidity, and which has performed fairly well suggests that the classification may be satisfactory.

III. The study design of Aitken and Comerton-Forde (2003:45-59) shows to what extent the measures and liquidity change simultaneously. It does not examine the predictive power of the measures which is what is actually required. In order to accomplish this, the ability of a measure in one period to predict changes in liquidity in the following period, should be investigated. Thus even though bid-ask spread seems to be a good indicator of current liquidity, it does not necessarily reflect future liquidity.

IV. The proposed weighted order value has to be considered with caution. The calculation does not account for trading within the bid-ask spread. It is thus likely that the measure will overestimate illiquidity. Furthermore, the bulk of the orders (81% of bids and 63% of asks in the pre-crisis group) are placed within 5% of the best quoted spread and there is a very strong inverse relationship between the percentage that an order is placed from the best spread, and the probability of execution. The point is that the bulk of executed orders are placed very close to the bid-ask spread. One has to consider what the contribution of the new measure is to the total capturing of illiquidity, especially in the light of other, possibly larger contributors which are not accounted for, like brokerage fees and tax considerations. The new measure is laborious to calculate and will certainly limit studies that look at longer periods. To the credit of the new method, recall that its results are slightly more consistent with expectations across the quartiles than the bid-ask spread.

4.3 Empirical evidence on proposition 2: The spread-return relationship

Recall from Chapter 3 that Amihud and Mendelson (1986:231-246) empirically tested their own theory on the New York Stock Exchange over the period of 1960-1979. They followed the methodology (described in section 3.2.2.1) of estimating portfolio betas over a period prior to the testing period. The portfolio formation process takes place before the portfolio beta estimation period and is based on grouping securities according to estimated betas into portfolios. Even with ordinary least square regression, as described in section 3.2.1, they found both beta and the bid-ask spread to be highly significant coefficients when only these two independent variables and dummy variables for the 20 year test period were included in the regression model. In addition to this, the market risk premium as calculated by the model, differed with only 0.001% from the average excess return on the market portfolio proxy. This provides empirical evidence of the fifth testable relationship of the mean variance efficiency hypothesis as outlined in section 3.1. They subsequently re-calculated the model using covariance analysis and pooling of cross-sectional and time series data as well as generalised least squares regression. The results confirmed the model hypothesis with a positive cross-sectional relationship between beta and portfolio return, and a concave piecewise linear relation between bid-ask spreads and returns. Neither market capitalisation, nor the natural logarithm of market capitalisation had any predictive power when beta and the bid-ask spread were included in the model.

Chen and Kan (1995:2-15) tested the same data set as used by Amihud and Mendelson (1986:231-246) with different methodologies. Their critique is based on the findings of Miller and Scholes (1982:1118-1141) in which the inverse of price was found to be proxying for estimation errors in beta. When either beta is estimated with errors, or when a change in the market risk premium is unaccounted for, a variable like the bid-ask spread (that may serve as a proxy for the inverse of price) can be expected to have predictive power. In the Amihud and Mendelson model (1986:231-246) the market risk premium is held constant in the second pass regression by lumping all time periods into the same regression which makes this model particularly prone to this kind of effect. Chen and Kan (1995:6-8) show that when the Fama &Macbeth (1973:614-618) approach[8] is used the predictive power of the bid-ask spread becomes largely insignificant. They postulate that the reason for this may be that even though the Fama and Macbeth approach also suffers from the same beta estimation problem, the market risk premium is estimated continuously over the test period, which allows for a fluctuating premium, and thus less chance of cross-sectional estimation error between beta and return. They also use a number of more sophisticated models with similar outcomes.

Marshall and Young (2003:173-178) examined the ability of the relative bid-ask spread, the turnover rate and the amortised spread to predict returns by using seemingly unrelated regressions (SUR) and cross-sectionally correlated and timewise autoregression (CSCTA)

---

[8] Fama and Macbeth estimated second pass regression equations for each month which allows the MRP to fluctuate from month to month. The average of the regression and standard deviation was then subsequently used in tests of significance.

as the statistical methodology. Each of the defined liquidity measures is examined in turn using both of the statistical methods. Other independent variables that were also included in various combinations were beta, market capitalisation and dummy variables for January market capitalisation and January liquidity. The methodology used allows for the simultaneous estimation of the coefficients for beta, market capitalisation and the liquidity measure. While this is advantageous when the aim is to prove or disprove various theories, it gives no indication of the ability of past variables to predict future returns. The advantages and drawbacks of the two statistical methodologies were explained in section 3.2.2.4. The definition of the relative bid-ask spread and turnover ratio is the same as in Aitken and Comerton-Forde (2003:50-51). The amortised spread is similar to the turnover ratio, except that the numerator now becomes the sum of the number of trades, times the effective bid-ask spread of the trading day. Based on the coefficient of determination, the bid-ask spread and turnover ratio gives similar results and were generally more significant in the SUR than CSCTA models. The amortised spread was slightly less effective.

The study also gives some important information on the effect of including/excluding beta. In the SUR model, spread is only significant when beta is included, but against all expectations the coefficient is negative. When beta is excluded the coefficient becomes positive, but is insignificant[9]. The CSCTA model follows the same pattern, although the spread is never significant. Overall the slope of the beta coefficient is more than 30 times greater than the slope of the spread, suggesting that if spread-return relationship is present then it has at most, a small effect. Overall beta itself, was highly significant in all models. Market capitalisation was only significant when beta was excluded from the model, and then highly so[10]. This study suggests that if a spread effect is present it is a small effect, and that market capitalisation is only significant when beta is excluded, or not estimated with the required accuracy.

4.4 Empirical evidence on proposition 1: The clientele effect

While the spread-return relationship has undergone extensive empirical testing, the clientele effect has received little attention. As discussed in section 1.3 implicit liquidity costs are

[9] The alternative hypothesis should not necessarily be rejected in favour of the nil hypothesis as the study may not have had enough power to detect a difference (see footnote 22).

[10] This suggests that market capitalisation may be a significant variable due to inaccurate estimation of beta, especially in models where beta is estimated from past performance. Although this effect is also to some extent present in the various measures of liquidity, it is of far lesser magnitude.

generated due to market imperfections. The bid-ask spread has been shown (see sections 4.2 and 4.3) to be a reasonable proxy for this liquidity cost, and was used by Atkins and Dyl (1997:309-324) as the independent variable in a study of the clientele effect[11]. Amihud and Mendelson (1986:225-231) argued that the clientele effect is achieved through mechanisms that drive investor indifference between assets. This behaviour forms the core principle of asset pricing theory which states that if an asset is under priced (overpriced) a riskless profit can be locked in by taking a long position in the under priced asset (short position in the overpriced asset) and a simultaneous short position (long position) in an asset that is in all ways similar to the incorrectly priced asset, except that it is priced correctly. It is thus plausible to argue that the bid-ask spread is the causative factor, as an increase in the spread necessitates an increase in the holding period as argued in Chapter 2. To the contrary however, an increase in the trading volume is just as likely to decrease the bid-ask spread due to competitive market forces. This type of behaviour ensures constant equilibrium in asset prices. Equilibrium systems rely on feedback mechanisms which makes it difficult to determine the direction of causality. One could thus expect that it will be difficult to determine causality in the clientele effect. This argument is supported by the research of Atkins and Dyl (1997:322-324) which used the Granger causality test[12] to determine the direction of causality. The results of their research are discussed next.

Atkins and Dyl (1997:309-324) looked at holding periods of shares on the NASDAQ and the NYSE for the periods of 1984 to 1991 and 1976 to 1989 respectively. They acknowledge that the bid- ask spread may not be the only factor that determines the holding period and identify the following three factors which could influence it:

I. Investor beliefs: Larger firms are more often considered by investors to be of so-called investment grade quality, which may result in longer holding periods.

II. Information availability: Information on corporations which are frequently followed by many analysts tends to converge over time. The convergence of information will cause investors to have the same outlook on the corporation which will decrease the trading volume. This is in contrast to limited information corporations where the divergence or absence of information leads to many outlooks and a higher trade volume.

---

[11] Regression cannot determine causality although the independent variable is often assumed to be the causative factor.

[12] The Granger test states that past values of the causative variable should be helpful to predict future values of the predicted variable. This is done by regressing the causative variable of a previous period against the current period of the predicted variable. If Granger causality exists in both directions then the conclusion is that a feedback system exists (Ramanathan 2002: 476).

III. Risk and stability: Corporations that are less risky, with greater stability over time require less portfolio rebalancing which decreases trading volume.

Atkins and Dyl (1997:312) used the bid-ask spread, the market capitalisation and the daily variance as proxies for these factors in order to investigate holding period behaviour. The regression equation was as follows:

$$Hp_i = \alpha + \gamma_1 S_i + \gamma_2 Cap_i + \gamma_3 Var_i + \sum_{j=1}^{k} \gamma_4 D_j + \varepsilon_i \quad \text{.....4.2}$$

Where:

Hp refers to the average holding period of security i calculated as shares in issue divided by trading volume,

S is the relative bid-ask spread,

Cap is the market capitalisation,

Var is daily variance on returns and

D_j is the dummy variables for each of the years $j=1$ to k tested.

All values, except for the dummy variables were log-transformed to improve on the skewness of the data. It is not stated in the article why the variance was used as opposed to the more traditional beta. Perhaps the authors were concerned about the accuracy of beta estimations of individual securities. Variance does not require estimation, but consists of both systematic and non-systematic variance of which the last is unlikely to be correlated to the holding period, as it represents firm specific events that are not predictable. This will make estimates of the coefficients less precise. A further criticism is that daily returns were used in the calculation of the standard deviation. This will make some of the least liquid shares, which do not trade on a daily basis, seem less volatile. The authors compensate for this to some extent by excluding some of the least liquid shares[13].

As discussed earlier the bid-ask spread and holding period are likely to be simultaneously determined during any period. Atkins and Dyl (1997:313) used two-stage least square regression to circumvent this problem. For the first-pass regression they estimated the spread of period T from the bid-ask spreads of period T-1. The equation is as follows:

---

[13] The shares of corporations with a trading volume of less than 25 000 shares per year, with less than 500 000 shares outstanding and with a bid-ask spread of less than 0.2% or greater than 50% were excluded.

$$S_{iT} = \alpha + \gamma_1 S_{iT-1} + \gamma_2 Cap_{iT} + \gamma_3 Var_{iT} + \sum_{j=1}^{k} \gamma_4 D_j + \varepsilon_i \qquad \text{.......................4.3}$$

This spread was then used equation 4.2 which is second-pass regression equation. All of the coefficients on both the NASDAQ and the NYSE data were highly significant. However the R^2 were only 0.43 and 0.27 respectively. Thus a large part of the variance in holding periods remains unaccounted for.

Atkins and Dyl (1997:322-323) also tested for causality between the spread and the holding period and found causality in both directions and thus concluded that a feedback effect exists.

4.5 Conclusion

To summarise, the bid-ask spread seems to be a reasonably good measure of liquidity. The proposed weighted order value measure of Aitken and Comerton-Forde (2003:56-57) could add additional strength, but is laborious and should probably be confirmed in other study populations. While the bid-ask spread may be an effective proxy, some have advocated that the inability to accurately measure beta becomes an important constraint when liquidity is empirically investigated. Proponents of this view argue that high correlation between the spread and the true beta cause the spread to capture risk not captured by the estimated beta. This concern also applies to the clientele effect, although CAPM does not advocate that risk influences the holding period[14] and therefore the concern is of lesser importance, from a theoretical perspective at least.

[14] One of the assumptions of CAPM is actually that all investors have the same identical holding period (Bodie, Kane & Marcus, 2005: 282).

CHAPTER 5: RESEARCH HYPOTHESIS AND METHODOLOGY

5.1 Introduction

This chapter provides a description of the methodology that was used in the investigation of the spread-return relationship and clientele effect as proposed by Amihud and Mendelson (1986)[15]. The data on 416 securities traded on the Johannesburg Stock Exchange, obtained from I-Net Bridge, is summarised in Table 5.1

Table 5.1 – Summary of the data provided by I-Net Bridge

Data Type	Period	Sampling interval
Closing bid prices	2002/01/01 - 2007/08/30	Daily
Closing ask prices	2002/01/01 - 2007/08/30	Daily
Closing share prices	1996/12/31 - 2007/07/31	Monthly
Trading volume	1997/01/01 - 2007/07/31	Monthly
Shares in issue	1997/01/01 - 2007/07/31	Monthly
All share index (J203)	1997/01/01 - 2007/07/31	Monthly
Money market index (GMC1)	1997/01/01 - 2007/07/31	Monthly

The all share index (ALSI) was used as a proxy for the market portfolio. The data does not include de-listed corporations and dividends. All closing prices were adjusted for share splits. This chapter provides a detailed description of firstly the methodology that was used for the investigation into the spread-return relationship, and secondly for the clientele effect.

5.2 General remarks on the methodology

In general the methodology of Fama and Macbeth (1973), as described in section 3.2.2.1 was used for the investigation of the spread-return relationship. It is simple and can easily be incorporated into a spreadsheet model. The model allows for beta to change from period to period, which is contrary to the more complicated procedure of pooling of cross-sectional

[15] From here onwards page numbers are generally not supplied (unless of specific importance) because the related literature has been discussed in detail in chapters 2-4.

and time series data. The disadvantage of the methodology is that the estimated betas can be rather inaccurate, in which case other independent variables may proxy for errors in the beta estimations as opposed to being true explanatory variables. In the estimation of security betas the aggregated coefficient method of Dimson (1979) was used. This allows for more accurate estimation of the betas of illiquid portfolios. Chen and Kan (1995) did not find a significant spread-return effect using the same data as Amihud and Mendelson (1986), but with the methodology of Fama and Macbeth (1973) as discussed in section 3.3.7, although the aggregated coefficient method was not used.

5.3 The relative spread

The relative spread for a security in a specific month was calculated as

$$S = \frac{\frac{\sum_{n=1}^{k}(Ask_n - Bid_n)}{k}}{(Close_{end} + Close_{begin})/2} \qquad \text{5.1}$$

Where k is the number of trading days for which both closing ask and bid prices were available, and $Close_{end}$ and $Close_{begin}$ are the closing share prices at the beginning and end of the relevant month[16].

5.4 The first pass regression: The calculation of security beta

The betas of all securities were calculated from the data of the previous 60 months.

The return of all securities as well as the ALSI and money market index in period T were calculated as $\ln(Close_T/Close_{T-1})$. Every month the beta is recalculated from the previous 60 months. This is known as rolling betas. In order for the beta to be calculated the security has to have recorded returns available for at least 24 of the 60 months, as a recorded return is only valid when trade in the security has taken place during the specific month. This ensures, from a theoretical point at least, that all systematic risk is captured by regressing the security return against the current and previous month's return on the ALSI. Thus the first pass regression, recalculated every month for the months January 2002 to June 2007 for security i is:

[16] Actually the closing prices of the last day of the previous month were used as an estimation of the closing price of the first day of the current month.

$$r_{i_T} - rf_T = \beta_1\left(r_{m_T} - rf_T\right) + \beta_2\left(r_{m_{T-1}} - rf_{T-1}\right) + \varepsilon_i \text{ ....5.2}$$

and so the security return is regressed against the ALSI return and the ALSI return of the previous month. In accordance with the methodology of Dimson (1979), the security aggregate beta is then the sum of β_1 and β_2.

In order to investigate the relevance of this methodology the β_1 and β_2 were regressed against the relative volume (RV), i.e. the monthly trading volume divided by the number of shares in issue. The regression equation is as follows:

$$\ln(RV_{AVi}) = \alpha + \phi_1\beta_{1AVi} + \phi_2\beta_{2AVi} \text{ ....5.3}$$

The subscript AVi indicates that the average of the relative volume, β_1 and β_2 for the ith security over the period from January 2002 to June 2007 were used. The relative volume data showed right skewness and were therefore log-transformed (natural logarithm). Under the nil hypotheses β_1 and β_2 should not be influenced by trading volume and should have coefficients (ϕ_1 and ϕ_2) of zero. Under the hypothesis of Dimson (1979), the lower the trading volume of a security, the larger the portion of the systematic risk captured by β_2 and thus β_2 should be negatively correlated to the relative trading volume. To the contrary the higher the trading volume of a security, the larger the portion of the systematic risk captured by β_1, and thus β_1 should be positively correlated to the trading volume. Thus, in order to accept the hypothesis of Dimson ϕ_1 should be positive and ϕ_2 should be negative.

5.5 The portfolio formation process

The two methodologies that have been suggested for the portfolio formation process were described in Section 3.2.2.1. The methodology of Fama and Macbeth (1973) has serious limitations when applied to studies of liquidity, because it requires a large number of historical data which is typically not available for the most illiquid securities. The methodology of Chan and Chen (1988) which circumvents this problem by using the market capitalisation (a variable that is highly correlated with beta) was used to allocate securities into portfolios. Unfortunately the methodology is not as effective on the JSE, probably due to a number of high market capitalisation, high beta mining corporations. The effect of this is a poorer spread of betas for the second pass regression.

In more detail the portfolio formation process was as follows:

I. In order for a security to be eligible for allocation to a portfolio it had to have data on the following:

(a) Market capitalisation and relative spread for the current and previous month. The previous month's data is used in the portfolio formation process and the current month in the second pass regression.

(b) Beta for the current month where beta was calculated as described in Section 4.2.3.

(c) Return, shares in issue and volume traded of the following month which are used in the second pass regression. The shares in issue data and the volume traded data were used to calculate the average holding period as: HP = SII / Volume

II. All securities with negative betas were kept separately in one portfolio. This ensures the widest range of betas.

III. The remaining eligible securities were allocated to one of six groups based on the market capitalisation of the previous month. Whenever the number of securities was not exactly dividable by six, the remaining securities were allocated to the lower market capitalisation groups. Thus a low market capitalisation group may have had up to one more security than a high market capitalisation group.

IV. Each of the 6 groups was then divided into a high and low relative spread portfolio. Thus, including the negative beta portfolio, a total of 13 portfolios result.

V. For each of the 13 portfolios the portfolio aggregate beta, relative spread, market capitalisation and return and holding period of the following month were calculated as the average of the individual securities for all the months ranging from February 2002 to June 2007.

5.6 The consistency of beta

As already mentioned, the methodology of Fama and Macbeth (1973) is less efficient than the more complex regression procedure of pooling of cross-sectional and time series data and in addition, runs the risk that other explanatory variables that are correlated to the true beta may serve as proxies for errors in the estimated beta. The more complex regression procedures may however, suffer from the same problem if beta is not constant over time as assumed by these models. In order to test the consistency of beta the portfolio betas were regressed against successive time periods (months) covering the period from February 2002 to June 2007. The slopes of the thirteen portfolio betas were recorded and evaluated for significance.

5.7 The second pass regression: The spread-return relationship

The regression equation of the second pass regression is as follows:

$$r_{pT+1} = \alpha + \gamma_1 \beta_{pT} + \gamma_2 \ln(Cap_{pT}) + \gamma_3 \ln(S_{pT}) + \varepsilon_{pT} \quad \text{.........................5.4}$$

where the subscript pT refers to the p[th] portfolio (p = 1 to 13) in month T (T = February 2002 to June 2007). The portfolio beta, market capitalisation and relative spread of the current month are regressed against the portfolio return of the next month. This approach ensures that the second pass regression output at the end of the period is based on the information that was available at the beginning of the period. The model allows for the portfolio betas to fluctuate, but the MRP (γ_1) is assumed to be constant.

5.8 Hypothesis testing: The spread-return relationship

The second pass regression portrays the relationship between return and beta. This is also the relationship that is portrayed by the SML in the CAPM, except that in the CAPM the return is the expected return in a market where all investors has the same economic outlook (Bodie, Kane & Marcus, 2005: 290). If markets behave according to theoretical expectations one could expect, at least on average, that the second pass regression will be a reasonable proxy for the SML. The characteristics of the SML, and by implication, the second pass regression, are that slope equal the MRP and the intercept equal the risk free rate. The MRP and money market rate can also be calculated as the average return on the market portfolio (for which the excess return on the ALSI serves as a proxy) and money market index respectively. Thus the average return on the ALSI should equal γ_1 of the second pass regression and the money market index should equal α. In addition to this γ_2 should be positive and γ_3 should equal zero. The cumulative return on the two indexes (the money market index and the excess on the ALSI) for the period T_{begin} to T_{end} can be calculated as:

$$cr = \prod_{k=Tbegin}^{Tend} \left(1 + r_{I_k}\right) - 1 \quad \text{..5.5}$$

where cr is the cumulative return on either the money market index or the excess on the ALSI and r_{Ik} is the return on the index in the k[th] month. The average index can then be calculated as:

$$r_{ave} = \sqrt[n]{cr+1} - 1 \quad \text{5.6}$$

and so, combining equation 5.5 and 5.6 gives:

$$r_{ave} = \sqrt[n]{\prod_{k=T_{begin}}^{T_{end}} \left(1 + r_{I_k}\right)} - 1 \quad \text{5.7}$$

The 95% confidence intervals of the estimated coefficients were used to compare each of the estimated coefficients to the expected value.

5.9 The second pass regression: The clientele effect

As in the investigation of the spread-return relationship, the estimated beta, market capitalisation and spread were regressed against the holding period of the following month. The regression equation is as follows:

$$\ln\left(HP_{pT+1}\right) = \alpha + \theta_1\beta_{pT} + \theta_2\ln\left(Cap_{pT}\right) + \theta_3\ln\left(S_{pT}\right) + \varepsilon_{pT} \quad \text{5.8}$$

Atkins and Dyl (1997) gives evidence that the relationship between spread and holding period cannot be estimated straight away, because the two are likely to be simultaneously determined. In order to circumvent this problem, they use a two stage regression approach where the spread of the current period is determined from the spread of the previous year, as well as the variance and market capitalisation of the previous month. This approach however requires that portfolios are kept constant which does not fit the methodology of this study. Thus the results of the suggested regression cannot be used to assess causality, it only indicates to what extent the holding period of the following month can be predicted from the relative spread of the current month.

Atkins and Dyl (1997) found significantly positive coefficients for both spread and market capitalisation and a significantly negative coefficient for security variance. The coefficients supported their hypothesis outlined in section 4.4. Atkins and Dyl (1997) investigated variance as opposed to beta, but their proposed argument, namely that high volatility (and thus high variance and high beta) drives high trading volume should generally also be applicable to beta. For this reason the expected coefficient of beta is negative or stated differently, higher betas should lead to shorter holding periods. Since there is no theory that indicates what the actual values of the expected coefficients should be, they were tested firstly for being significantly different from zero, and secondly for similarity in direction to the direction found in the Atkins and Dyl (1997) study.

5.10 Conclusion

In order for the study to support the Amihud and Mendelson (1986) model, both the clientele and the bid-ask spread should hold. The model proposed by Constantinides requires only that there be a positive correlation between the relative spread and the holding period, with the relative spread-return relationship assuming a minor role. The selected methodology has previously been employed in studies of the relative spread-return relationship which generally did not have a positive outcome, and so is not biased towards a positive outcome. If the outcome does support the model, then it is likely to be due to better estimation of the betas by way of the aggregated beta coefficient method.

CHAPTER 6: RESULTS AND DISCUSSION

6.1 Introduction

In this chapter the results of the analysed data is presented. The implications of the results are discussed and various limitations are also pointed out.

6.2 The relevance of using the aggregated beta

Descriptive statistics of the average of the rolling $beta_1$'s and $beta_2$'s for each of the qualifying securities is presented in Table 6.1.

Table 6.1 – Descriptive statistics of averaged rolling betas

	$Beta_1$	$Beta_2$
Number of qualifying securities	296	296
Mean	0.52	0.14
Median	0.53	0.09
Range	-1.65 – 2.51	-2.19 – 3.14
Interquartile range	0.24 – 0.77	-0.05 – 0.28

As proposed by Dimson (1979) the aggregated beta calculation method has a greater effect on the less liquid securities. The result of the regression of the relative trading volume against β_1 and β_1 is presented in Table 6.2 and the scatter plot in Figure 6.1. The slopes of both β_1 and β_2 are significantly different from zero and the direction is as predicted in section 4.2.3. This is confirmation of the importance of using the aggregated beta in studies that include shares with low trading volume as first proposed by Dimson (1979).

Table 6.2 – Regression statistics of the average β_1 and β_2 against the relative trading volume average for the period January 2002 to June 2006

Overall significance of the model

	Sum of Squares	Degrees of freedom	Mean Squares	F ratio	p-Value
Regression	26	2	13.1	12.1	<0.0001
Residual	318	293	1.1		
Total	344	295			
R^2	7.6%				

Significance of the model coefficients

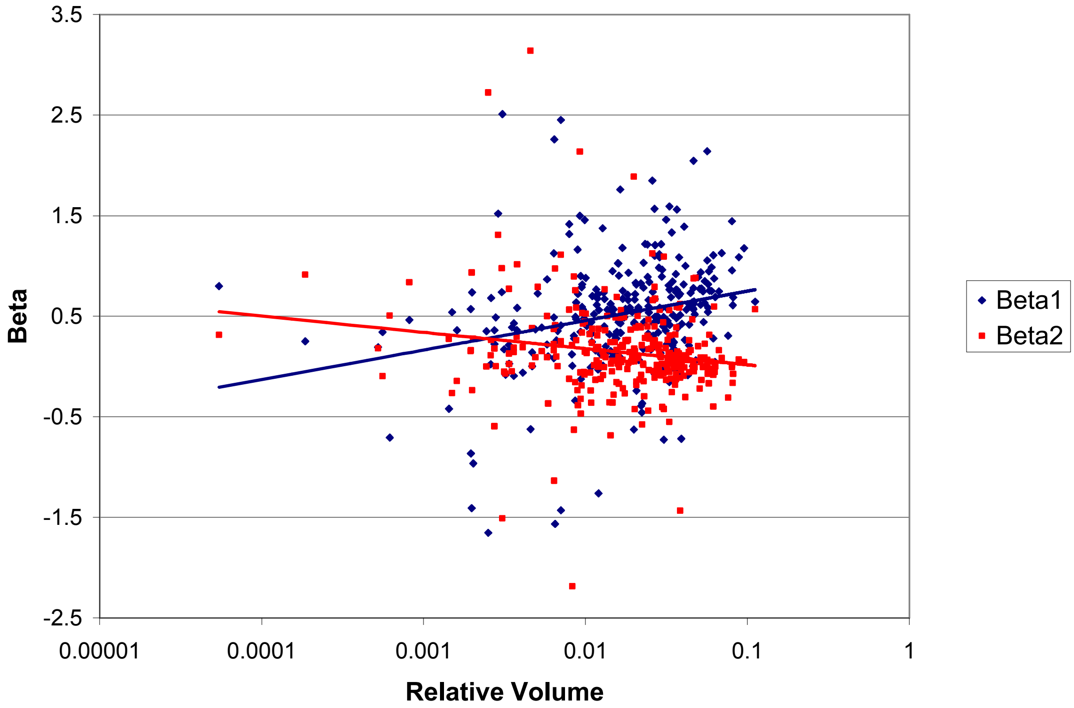
	β_2	B_1	Intercept
Coefficient	-0.28	0.43	-4.29
Standard Error	0.13	0.11	0.09
t-Value	2.20	3.98	49.05
p-Value[17]	0.01	<0.0001	<0.0001

Part Correlations

	β_2	B_1	Relative Volume
B_2	1.00		
B_1	-0.16	1.00	
Relative Volume	-0.16	0.25	1.00

[17] These p-values are for a one tail t-test, since prior expectation of the direction of the slopes existed.

Figure 6.1 – Scatter diagram of the natural logarithm of β_1 and β_2 against the relative volume traded



6.3 The portfolio characteristics

An excerpt of the 13 portfolios' aggregate betas, relative spreads, market capitalisations, returns and relative holding period from February 2002 to June 2007 are presented in Tables 6.3, 6.4 and 6.5 respectively. Recall that the portfolios are reformed every month from the previous month's market capitalisation and relative spread for all securities with positive betas. Securities with negative betas are grouped into a separate portfolio. The security data was then used to calculate the portfolios average beta, market capitalisation and relative spread, as well as the average return and relative holding period in the following month. From Table 6.3 it is evident that the strategy produced a fairly wide range of betas stretching from -0.42 to 1.08[18]. The correlation between the natural logarithm of the market capitalisation and beta is -0.19 and is not nearly as high as reported by Chan and Chen (1988) which was generally above -0.90. This is largely due to the increase in beta in the 6th group which is likely due to the presence of a number of large mining corporations that

[18] See Section 3.2.2.1 for a discussion of why a wide range and even spread of the independent variables of the second pass regression is required.

typically have fairly large betas. High relative spread portfolios consistently had higher betas than low spread portfolios, which suggests that spread may also serve as a proxy for beta in the portfolio formation process.

Table 6.3 – Aggregate beta of the 13 portfolios

Portfolio*	Market Cap group	Spread group	28-02-2002	31-05-2007	30-06-2007	Average
1	1	High	1.24	1.05	1.15	1.08
2	1	Low	1.16	1.05	1.05	1.08
3	2	High	1.25	1.23	0.92	0.95
4	2	Low	1.00	0.74	0.96	0.84
5	3	High	1.10	0.80	0.70	0.65
6	3	Low	0.75	0.47	0.48	0.60
7	4	High	1.04	0.79	0.77	0.75
8	4	Low	1.06	0.55	0.66	0.68
9	5	High	1.04	0.83	0.84	0.72
10	5	Low	1.00	0.68	0.63	0.67
11	6	High	0.86	0.96	0.96	0.85
12	6	Low	0.77	0.89	0.90	0.80
13	-	-	-0.29	-0.50	-0.51	-0.42

*Portfolios were formed based on 6 groups of increasing market capitalisation (1-6) with each of the 6 groups divided into a high and low relative spread group. The 13th portfolio was formed from all securities with negative betas

Table 6.4 and Table 6.5 demonstrate that a wide range of both relative spread and market capitalisation results from the portfolio formation process. This is expected as both the spread and the market capitalisation of prior periods were used in the portfolio formation process. The procedure is valid as actual values (as opposed to estimates of beta) were used and thus the procedure does not suffer from the regression problem discussed in section 3.2.2.1.

Table 6.4 – Relative spread (in percentage) of the 13 portfolios

Portfolio	Market Cap group	Spread group	28-02-2002	31-05-2007	30-06-2007	Average
1	1	High	17.2	9.7	6.5	14.0
2	1	Low	45.7	25.3	31.3	52.7
3	2	High	8.2	2.3	2.1	5.2
4	2	Low	15.2	8.7	10.7	12.0
5	3	High	3.8	1.6	1.6	3.0
6	3	Low	11.3	3.3	3.2	5.7
7	4	High	2.5	0.8	0.9	1.9
8	4	Low	5.5	2.3	2.3	4.2
9	5	High	1.4	0.5	0.6	1.1
10	5	Low	2.5	1.5	1.7	2.6
11	6	High	0.6	0.3	0.3	0.5
12	6	Low	1.0	0.5	0.5	1.0
13	-	-	33.6	12.6	11.9	35.6

Table 6.5 – Market capitalisation (in billions) of the 13 portfolios

Portfolio	Market Cap group	Spread group	28-02-2002	31-05-2007	30-06-2007	Average
1	1	High	2.6	12.0	12.6	5.6
2	1	Low	1.3	5.7	5.8	2.6
3	2	High	10.0	65.3	68.8	28.3
4	2	Low	7.3	56.3	57.4	21.6
5	3	High	34.3	183.5	200.2	83.8
6	3	Low	25.7	187.0	169.1	77.8
7	4	High	94.7	492.4	481.1	212.1
8	4	Low	69.3	408.2	399.5	189.0
9	5	High	278.1	1240.0	1208.8	666.9
10	5	Low	251.0	1009.7	972.0	522.8
11	6	High	6458.3	14948.5	15448.3	7722.8
12	6	Low	1934.3	4730.2	4083.0	3025.4
13	-	-	55.9	45.8	45.9	92.9

Table 6.6 illustrates that there is generally a trend towards higher returns as market capitalisation decreases and spread increases although it is not a consistent finding.

Table 6.6 – Returns (in percentage) in the following period of the 13 portfolios[19]

Portfolio	Market Cap group	Spread group	28-02-2002		31-05-2007	30-06-2007	Average
1	1	High	0.1		2.3	0.3	2.2
2	1	Low	-6.5		3.3	-15.3	3.3
3	2	High	-3.3		-2.3	-3.8	2.6
4	2	Low	-6.4		-2.5	1.2	1.3
5	3	High	0.5		0.1	-3.6	2.3
6	3	Low	-3.5		-2.3	-1.5	2.2
7	4	High	-1.4		-2.4	-0.8	1.9
8	4	Low	-1.3		-1.5	-0.1	1.9
9	5	High	0.6		-3.8	0.4	2.2
10	5	Low	-0.5		-3.5	-2.1	1.7
11	6	High	1.2		-1.5	0.1	1.1
12	6	Low	-3.0		-3.3	-1.9	1.5
13	-	-	16.3		2.2	-3.7	3.0

Contrary to the return the relative holding period demonstrates a very close correlation to market capitalisation and relative spread as is evident in Table 6.7.

[19] The return of the following month is recorded under the current month. Thus when the data of 28-02-2002 is fed into the second pass regression it is the beta, market capitalisation and relative spread of a portfolio for the current month (28-02-2002) as well as the return or holding period of the following month (31-03-2002).

Table 6.7 – Relative holding period in the following period of the 13 portfolios

Portfolio	Market Cap group	Spread group	28-02-2002	31-05-2007	30-06-2007	Average
1	1	High	1118	356	155	9030
2	1	Low	9588	1741	791	129723[20]
3	2	High	249	50	31	249
4	2	Low	990	1168	2079	5500
5	3	High	164	55	46	146
6	3	Low	1056	537	233	2290
7	4	High	193	26	24	88
8	4	Low	473	112	110	996
9	5	High	36	25	19	50
10	5	Low	77	76	85	1004
11	6	High	32	31	22	25
12	6	Low	43	33	41	53
13	-	-	901	275	1040	1990

6.4 The consistency of beta

Table 6.8 summarises the results of the regression to the rolling portfolio betas against successive time periods. Out of the 13 portfolios, 10 had significant coefficients if significance is set at $p<0.05$. The portfolio with the steepest slope/coefficient was that of portfolio 6 where the coefficient translates into an average period to period decrease in beta of 1%. It would thus appear that beta is not constant over time, and that even though pooling of cross-sectional and time series data regression techniques may yield better estimates of beta other variables, may still proxy for errors in the estimation due to the non-fixed nature of the portfolio betas.

[20] This very high holding period is due to a single far outlier in one of the securities. After exclusion of the outlier the value is 2859 months. The outlier was also excluded in the second pass regression.

Table 6.8 – Regression of the successive time periods against portfolio betas

Overall significance

Portfolio	Slope (x10-3)	Standard error (x10-4)	t-Value	p-Value	Intercept	Correlation
1	-1.57	9.36	1.67	0.10	1.13	0.21
2	-2.73	10.1	2.71	0.01	1.16	0.32
3	-1.23	9.15	0.13	0.89	0.95	0.02
4	-1.71	7.23	2.36	0.02	0.89	0.29
5	-6.37	9.23	6.91	$2.9*10^{-9}$	0.85	0.66
6	-7.79	6.99	11.14	$1.5*10^{-16}$	0.84	0.81
7	-4.17	7.62	5.48	$8.0*10^{-7}$	0.89	0.57
8	-7.73	7.73	10.01	$1.2*10^{-14}$	0.93	0.78
9	-4.69	7.56	6.20	$4.8*10^{-8}$	0.87	0.62
10	-6.15	6.86	8.96	$7.6*10^{-13}$	0.87	0.75
11	0.91	4.07	2.24	0.03	0.82	0.27
12	0.96	4.58	2.11	0.04	0.77	0.26
13	-0.15	3.69	0.41	0.69	-0.41	0.05

6.5 The spread-return relationship

The results of the second pass regression in which the aggregated beta, relative spread and market capitalisation of all 13 portfolios are regressed against the excess return of the following period, over the period ranging from February 2002 to June 2007 is presented in Table 6.9.

Table 6.9 – Second pass regression statistics for the period February 2002 to June 2007

Overall significance of the model

	Sum of Squares	Degrees of freedom	Mean Squares	F ratio	p-Value
Regression	0.025	3	0.0082	3.28	0.02
Residual	2.1	841	0.0025		
Total	2.124	844			
R^2	1.16%				

Significance of the model coefficients

	Relative Spread[21]	Market Capitalisation	Aggregate Beta	Intercept
Predicted coefficient[22]	Positive	0	0.66%	0.76%
Coefficient	-0.57%	-0.46%	-1.54%	3.53%
Standard Error	0.34%	0.20%	0.58%	0.55%
t-Value	1.65	2.30	2.67	6.38
p-Value	0.10	0.02	0.008	<0.0001
95% Confidence interval	-1.24% - 0.11%	-0.85% - -0.07%	-2.66% - -0.41%	2.45% - 4.62%
Power	93%	-	-	-

Part Correlations

	Relative Spread	Market Capitalisation	Aggregate Beta	Excess Return
Relative Spread	1.00			
Market Capitalisation	-0.88	1.00		
Aggregate Beta	-0.12	-0.19	1.00	
Excess Return	0.05	-0.06	-0.06	1

[21] The natural logarithm of the relative spread and market capitalisation produced smaller p-values than the untransformed spread. In addition to this the R^2 also increased when the data was transformed.

[22] The predicted coefficient is calculated according to the methodology discussed in Section 4.2.6

Overall the regression model is statistically significant but the R^2 is only 1.16%[23]. The intercept is significantly different from zero but the 95% confidence interval does not overlap the predicted intercept, and thus it does not behave as expected from theory. The relative spread was not significantly different from zero[24] and all other coefficients (although significantly different from zero) were significantly different from the predicted coefficient judged by the 95% confidence intervals. Thus the relative spread had no predictive power and beta were the opposite of theoretical expectations. Only market capitalisation behaved as proposed by the Fama and French three factor model (Davis, Fama & French 2000) which is contrary to the expectations of this study. A high correlation exists between market capitalisation and relative spread, which is as expected. The correlation between beta and both relative spread and market capitalisation were in excess of 0.1 which should alert one to the risk of both market capitalisation and relative spread serving as proxies for the true beta when the estimated beta is inaccurate. The correlation between excess return and all the independent variables were poor.

A graph of the cumulative returns[25] on the 13 portfolios and the ALSI is presented in Figure 6.2. The average ranked beta of each portfolio is tabulated against the ranked total cumulative return in Table 6.10. The table clearly illustrates that the average portfolio beta had very little ability to predict the cumulative portfolio return.

---

[23] A p value of less than 0.05, i.e. a less than 5% probability that the result is due to statistical fluctuation is usually considered significant. The p-value of 0.02 should be interpreted as a 1 in 50 probability that the effect occurred simply through statistical fluctuation. The R^2 of 1.16% indicates that only about 1% of the variation in the return is explained by the three independent variables which should be interpreted as poor predictive model capability.

[24] When a significant difference is not detected one cannot simply conclude equivalence to the 0 hypothesis as it may be that the study lacks sufficient power, i.e. the study was not large enough to show a difference. In order to calculate the study power a judgment call often has to be made what a significant difference would be. For this study it was arbitrarily decided that a significant difference has to be at least one 10^{th} of the expected beta coefficient (γ_1). In other words, a slope of at least one 10^{th} of the SML is judged to be of practical importance. A power of 80% or more is usually considered sufficient

[25] The cumulative returns are based on a one Rand investment in the starting period.

Figure 6.2 – Cumulative return of the ten portfolios and the ALSI

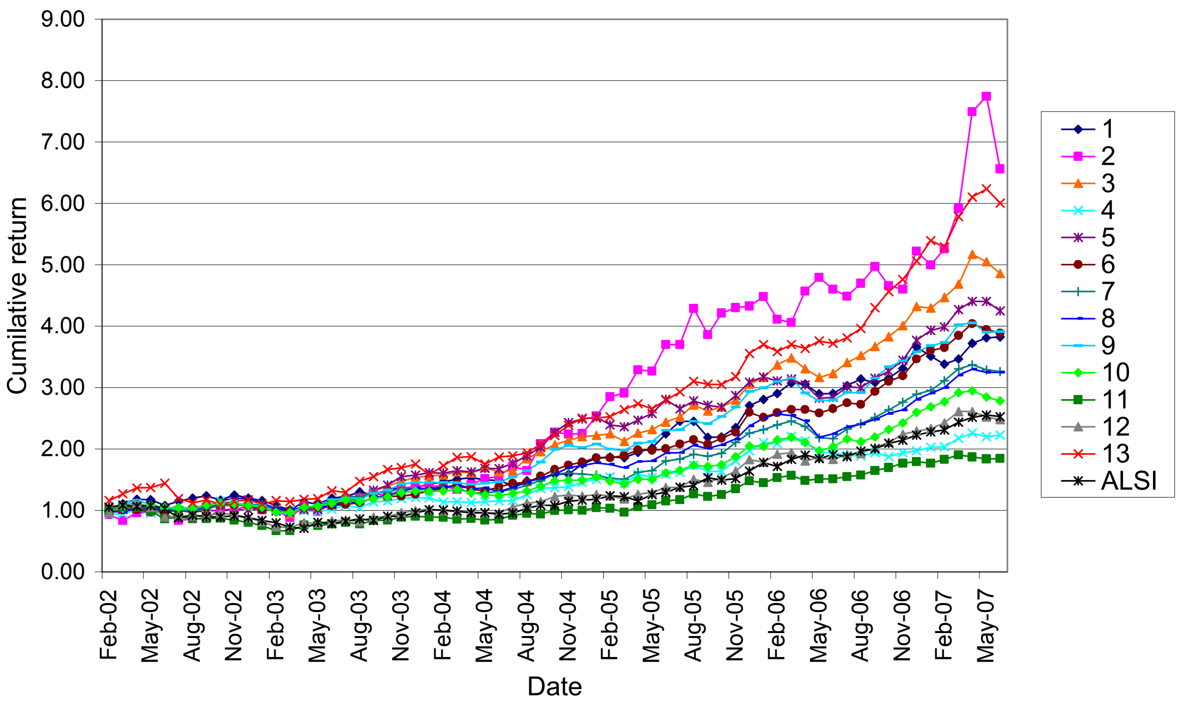


Table 6.10 – Ranked portfolio betas compared to ranked cumulative returns

Portfolio Beta	Ranked portfolios based on beta	Ranked portfolios based on cumilative return	Portfolio cumilative return
1.08	2	2	556%
1.08	1	13	501%
1.00	ALSI	3	386%
0.95	3	5	325%
0.85	11	9	291%
0.84	4	6	289%
0.80	12	1	282%
0.75	7	7	226%
0.72	9	8	225%
0.68	8	10	178%
0.67	10	ALSI	153%
0.65	5	12	147%
0.60	6	4	123%
-0.42	13	11	84%

From Figure 6.2 it can be deducted that the market was fairly flat during the first 18 months. Since CAPM is based on expected returns and since expected returns are limited to values greater than the risk-free rate, it may be of importance to evaluate the model over a period of generally positive market growth. In order to investigate this, the regression coefficients were again estimated for the period from July 2003 to June 2007, thus excluding the first 18 months. The statistics of this regression are presented in Table 6.11. These results are quite different from the previous results. The model is marginally insignificant with a R^2 of just less than 1%. The behaviour of the intercept is similar to the previous regression. The relative spread is again not statistically different from zero, but now the study lacks sufficient power to conclude equivalence. The aggregate beta is not significantly different from zero and its confidence interval does not overlap the predicted value. The market capitalisation now behaves as predicted by this study. The correlations are similar to those of the previous regression.

Table 6.11 – Second pass regression statistics for the period July 2003 to June 2007

Overall significance of the model

	Sum of Squares	Degrees of freedom	Mean Squares	F ratio	p-Value
Regression	0.013	3	0.0044	2.07	0.10
Residual	1.305	620	0.0021		
Total	1.318	623			
R2	0.99%				

Significance of the model coefficients

	Relative Spread	Market Capitalisation	Aggregate Beta	Intercept
Predicted coefficient	Positive	0	1.75%	0.68%
Coefficient	0.09%	-0.15%	-0.37%	3.93%
Standard Error	0.38%	0.22%	0.61%	0.60%
t-Value	0.24	0.71	0.61	6.61
p-Value	0.81	0.48	0.54	<0.0001
95% Confidence interval	-0.65% - 0.83%	-0.58 - 0.27%	-1.57% - 0.82%	2.76% - 5.10%
Power	18%	87%	99%	-

Part Correlations

	Relative Spread	Market Capitalisation	Aggregate Beta	Excess Return
Relative Spread	1.00			
Market Capitalisation	-0.89	1.00		
Aggregate Beta	-0.13	-0.15	1.00	
Excess Return	0.10	-0.09	-0.02	1

6.6 The clientele effect

The results of the second pass regression in which the aggregated beta, relative spread and market capitalisation of all 13 portfolios were regressed against the natural logarithm of the relative holding period of the following month, over the period ranging from February 2002 to June 2007, is presented in Table 6.12. The overall model is highly significant. Of the three coefficients the relative spread is by far the most significant. This is similar to the findings of Atkins and Dyl (1997). This is followed by beta which is marginally significant at P=0.03 although the direction is not as expected. Market capitalisation was not statistically different from zero and equivalence to zero can be concluded based on the high power. This is contrary to what Atkins and Dyl (1997) found. A scatter diagram that illustrates the strong correlation between the relative spread and the holding period is presented in Figure 6.3.

Table 6.12 – Second pass regression statistics for the period Feb 2002 to June 2007

Overall significance of the model

	Sum of Squares	Degrees of freedom	Mean Squares	F ratio	p-Value
Regression	1.56*103	3	5.21*102	568	<0.0001
Residual	7.73*102	841	9.19*10-1		
Total	2.34*103	844			
R2	67%				

Significance of the model coefficients

	Relative Spread[26]	Market Capitalisation	Aggregate Beta	Intercept
Predicted coefficient[27]	Positive	Positive	Negative	Positive
Coefficient	1.14	7.00*10-2	0.23	8.57
Standard Error	0.07	3.84*10-2	0.11	0.11
t-Value	17.35	1.82	2.12	80.63
p-Value	<0.0001	0.07	0.03	<0.0001
95% Confidence interval	1.01 – 1.27	-5.53*10-3 – 1.45*10-1	0.02 – 0.45	8.36 – 8.78
Power	-	93%	-	-

Part Correlations

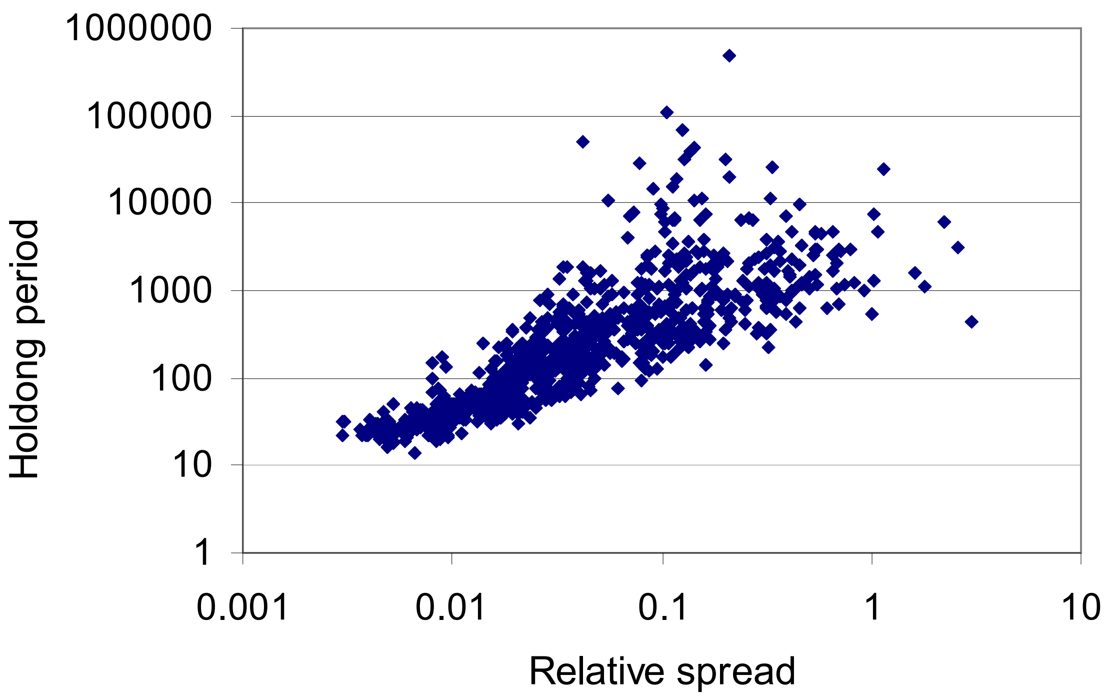
---

[26] The overall model had greater significance when the natural logarithm of the relative spread produced was used as opposed to the untransformed spread. In addition to this the R^2 increased from 56% to 67% when the transformed spread was used which confirms the appropriateness of log-transforming the relative spread data.

[27] The predicted coefficients is according to the research of Atkins and Dyl (1997) as discussed in Section 4.2.6

	Relative Spread	Market Capitalisation	Aggregate Beta	Excess Return
Relative Spread	1			
Market Capitalisation	-0.88	1		
Aggregate Beta	-0.12	-0.19	1	
Holding period	0.82	-0.71	-0.074	1

Table 6.13 – Scatter diagram of the relative spread and relative holding period



6.7 Conclusion

The aggregated beta, a phenomenon due to non-synchronised trading and first suggested by Dimson (1979), was empirically confirmed to be present on the JSE. In addition beta changes over time which suggest that statistical models that assume that beta is constant may not be valid. Significant empirical evidence was not found for a relation between return and either beta, market capitalisation or relative spread although strong evidence is present for the clientele effect.

CHAPTER 7: CONCLUSION AND RECOMMENDATIONS

This study set out to investigate relevance of the spread-return relationship and the clientele effect as proposed by Amihud and Mendelson (1986) on the JSE. This is theoretically sound because it addresses one of the assumptions of CAPM that is in violation of real markets, for example, the liquidity assumption. Unfortunately these relations cannot be studied in isolation because the omission of important independent variables in a regression equation leads to biased estimates of the regression coefficients, especially when the omitted variable is correlated to the included variables (Ramanathan, 2002:166). Strong theoretical evidence exists for the inclusion of beta although there is varying empirical evidence. To the contrary the empirical evidence for the inclusion of market capitalisation is generally strong with less sound theoretical support.

The inclusion of beta is particularly problematic because it has to be estimated. Roll and Ross (1994) showed that even small errors in the estimates can destroy the cross-sectional relation between beta and excess return. This creates two kinds of problems. Firstly the estimates of other coefficients become inaccurate, and secondly coefficients that are correlated to the true beta may now serve as proxies for the true beta, which would make them seem significant when in fact they are not. In order to improve on the estimates of beta the following two methodologies were used:

I. The data was examined for non-synchronous trading which was found to be present. Beta was subsequently calculated with the aggregated coefficient method of Dimson (1979) which minimises the errors in the estimates.

II. Portfolios, as opposed to individual securities were used, which largely eliminates variance due to non-systematic risk and improves the beta estimates. The methodology of Chan and Chen (1988) was used in the portfolio formation process. Unfortunately the use of market capitalisation was a less effective beta proxy than in other studies on empirical evidence for expanded CAPM models which leads a less effective beta spread in the second pass regression (see section 3.2.2.1). This may be due to the large mining sector in the South African economy.

Pooling of cross-sectional and time series data regression allows for greater accuracy in the estimation of beta but assumes that beta is constant for the duration of the study. This study gives evidence that beta is not constant, and thus suggests that these advanced regression techniques may not be more accurate than the two stage regression procedure described in section 3.2.2.1. The inability to measure beta accurately places great limits on empirical market research.

No convincing conclusions could be formulated for a cross-sectional correlation (as hypothesised) between either beta, relative spread or market capitalisation and excess return. The study also lacked sufficient power at times. Thus although the existence of a positive spread-return relationship cannot be rejected outright, it seems to have a minor effect size at most.

In contrast to the spread-return relationship, strong evidence for the clientele effect is shown. This is more in line with the propositions of Constantinides (1986) which was not really the main focus of this study. Only a marginally significant relation between the holding period and beta and a marginally insignificant relation between the holding period and the market capitalisation were shown, which gives evidence that the spread is by far the most important factor in the holding period with good predictive power given the model R^2 of 67%.

As already mentioned the inability to measure beta accurately places great restraints on empirical market research. The accuracy is to a large extent dependant on the quality of the data. At least part of the inability to measure portfolio betas accurately can be ascribed to exclusion of dividends in security data sets as well as survivorship bias.[28] Even the relative spread may have been influenced by the data quality. The closing spreads may not be an accurate reflection of the average daily spread due to increased volatility prior to market closure. Establishing a centralised data base with high quality data and easy access for market researchers will definitely promote empirical market research in South Africa.

Since no clear evidence of a liquidity premium on the JSE could be found, foreign investors should prefer more liquid markets as there appears to be no incentive to invest in less liquid markets (assuming that the findings of this study holds for other emerging markets as well). This may be of importance to exchange regulators and government as it suggests that measures that lower transaction costs will lead to increased foreign portfolio inflow.

Future research into liquidity should take the following into account:

I. Market capitalisation as a proxy for beta was less effective in the portfolio formation process then in previous research. This may be due to the large mining industry in

[28] Survivorship bias may be of particular importance when portfolios are used. If the data set does not include delisted securities (of which at least some will have delisted due to insolvency) then the performance of higher risk portfolios will be biased upwards.

South Africa which typically has high betas. Bid-ask spread may be a preferable proxy.

II. The five year study period in this study was probably not sufficient based on the power calculations. A study period of at least ten years is suggested, although it has to be balanced against data accuracy.

III. Improving on the data accuracy, especially with regard to delistings and dividend may be beneficial.

REFERENCES

Aitken, M. & Comerton-Forde, c. 2003. How should liquidity be measured? *Pacific-basin Finance Journal,* 11(1): 45-59

Amihud, Y., Christensen, B.J. & Mendelson, H. 1992. Further evidence on the risk-return relationship. *Research Paper no. 1248,* Graduate School of Business, Stanford University.

Amihud, Y. & Mendelson, H. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics,* 17 (2):223-49

Amihud, Y. & Mendelson, H. 1989. The effects of Beta, Bid-Ask spread, residual risk and size on stock returns. *Journal of Finance,* 44(2): 479-86

Amihud, Y & Mendelson, H. 1991. Liquidity, asset prices and financial policy. *Financial Analysts Journal,* 47(6): 56-66

Atkins, A.B & Dyl, E.A.1997. Transaction costs and the holding periods for common stocks. *The Journal of Finance,* 52(1): 309-25.

Black, F., Jensen, M.C. & Scholes, M. The capital asset pricing model: Some empirical tests. In: Jensen, M.C. 1972. *Studies in the theory of capital markets.* New York. Praeger publishers.

Blume, M.E. 1970. Portfolio theory: A step towards its practical implication. *Journal of Business*, 43(2): 152-73

Bodie, Z., Kane, A & Marcus, A.Z. 2005. *Investments.* Sixth Edition. New York: McGraw-Hill companies Inc.

Chan, K.C. & Chen N. 1988. An unconditional asset pricing tests and the role of firm size as an instrumental variable for risk. *Journal of Finance,* 43(2): 309-25

Chen, N. & Kan, R. 1995. Expected return and the bid ask spread. Working paper. University of California.

Constantinides, G.M. 1986. Capital market equilibrium with transaction costs. *Journal of Political Economy,* 94(4): 842-62

Davis, J.L., Fama, E.F. & French, K.R. 2000. Characteristics, covariances and average returns: 1929 to 1997. *The Journal of Finance,* 55(1): 389-406

Dimson, E. 1979. Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics,* 7(2): 197-226

Fama, E.F & French, K.R. 1992. The cross-section of expected stock returns. *Journal of Finance,* 47(2): 427-65

Fama, E. & MacBeth, J. 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy,* 81(3):607-36

Lintner, J. 1965. Security prices, risk, and maximal gains form diversification. *Journal of Finance,* 20(4); 587-615

Markowitz, H. 1952. Portfolio selection. *Journal of Finance,* 7(1): 77-91

Marshall, B.R. & Young, M. 2003. Liquidity and stock returns in pure order driven markets: evidence form the Australian stock market. *International Review of Financial Analysis,* 12(2):173-88

Miller, M.H. & Scholes, M. Rate of return in relation to risk: A re-examination of some recent findings. In: Jensen, M.C. 1972. *Studies in the theory of capital markets.* New York. Praeger publishers.

Miller, M.H. & Scholes, MF. 1982. Dividends and taxes: empirical evidence. *Journal of Political Economy*, 90(6):1118-1141

Ramanathan, R. 2002. *Introductory econometrics with applications*. Fifth edition. Ohio: South-Western publishers.

Roll, R. 1977. A Critique of the asset pricing theory's tests: Part I: On past and potential testability of the theory. *Journal of Financial Economics,* 4(2): 129-76

Roll, R. & Ross, S.A. 1994. On the cross-sectional relation between expected returns and beta. *Journal of Finance,* 49(1): 101-21

Sharp, WF. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance,* 19(3): 425-42