DECLARATION

I, David McClelland, declare that this research report is my own unaided work. It is submitted in partial fulfilment of the requirements for the degree of Master of Commerce in Finance at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.

David McClelland

March 2014

ACKNOWLEDGEMENTS

I would like to thank the following people for their support in the completion of this report: My supervisor, Professor Christo Auret for giving me the space, encouragement and advice I required to complete this research; my mother, and personal editor, Angela McClelland for having the patience and detective skills to decipher, and correct, my writing; and my friends and colleagues Daniel Page, James Britten, and Yudhvir Seetheram for their useful critique and opposing views that kept me on my toes. Finally I would like to thank the many members of the Finance and Economics departments for their assistance, my friends for their support, and my girlfriend, Ling Ting, for her mathematical help and continued support through this process.

Definition of key terms and abbreviations

Abnormal return (Alpha): The return earned by a share or portfolio over and above that which is attributable to the risk loadings of a particular risk model.

B/M: The book value of equity divided by the corresponding market capitalisation.

B - S: Buy minus Sell: the buy portfolio is made up of the portfolio which theory and prior evidence suggests should earn a higher rate of return.

Beta: The sensitivity of a share to changes in the market (proxied by the J203T) portfolio.

Bid-ask spread: The highest limit order bid to buy a share minus the lowest limit order offer to sell on the order book.

Excess return: The return over and above the risk free rate (proxied by the 3 month Treasury Bill).

FF3F: The Fama and French Three Factor model.

Global financial crisis (GFC): The combined period encompassing the sub-prime mortgage crisis, the liquidity crisis, as well as the European sovereign debt crisis.

HML: The returns of high B/M quintile minus the low B/M quintile

JSE: The Johannesburg Stock Exchange

J203T: The JSE All Share total returns index

Liquidity: 'The ability to trade large quantities quickly at low cost with small price impact' (Liu, 2006, p. 1).

Liq CAPM: The Liquidity-Augmented CAPM

Liq_{12}: The return on shares classified as illiquid using the prior 12 months trading data minus the shares classified as liquid over the same period. The liquidity risk-premium used in the Liq CAPM.

Market capitalisation (size): The market price per share multiplied by the number of shares outstanding.

Momentum: Refers to price momentum.

Non-synchronous: Two share return series that do not trade at the same frequency and, therefore, have different intervals of informational updates in the form of price changes.

P-value: The probability of committing a Type-1 error, or the probability that the relationship that has been estimated does not exist.

Pre-ranking: Using only historical data to group shares.

Sharpe ratio: The excess return divided by the standard deviation.

SMB: The returns of the small market capitalisation quintile minus the returns of the big market capitalisation quintile.

Trading friction: Includes trading costs, taxes, bid-ask spreads. Can be approximated by the cost of a round-trip transaction (the cost of buying a security and then immediately selling the same security).

Turnover: Volume traded divided by the number shares in issue

Type-2 error: The probability of failing to detect a true relationship, or the probability of failing to reject the false null hypothesis (for our purposes the null hypothesis (H_0) states that a relationship is equal to zero).

ZDT: The number of zero daily trades over a particular period.

The Liquidity-Augmented CAPM:

Empirical evidence from the JSE

ABSTRACT

This study replicates the two-factor Liquidity-Augmented CAPM of Liu (2006) on the JSE over the period 1997–2011. To adjust any risk measures for the downward bias consistent with infrequently traded shares, Fowler and Rorke (1983)'s adjusted OLS coefficients are utilised. Liu's paper defines liquidity according to the relatively unexplored dimension of trading speed. Measured as a combination of the number of zero daily trades and turnover, this liquidity variable captures the ability of investors to move in and out of positions quickly. The two-factor model is tested against the ordinary CAPM as well as the Fama and French three-factor model in its ability to accurately explain the premiums previously documented with pre-sorted size, value, liquidity, and Beta portfolios. The results are supportive of the two-factor model. The study demonstrates that illiquidity is associated with higher levels of excess returns and that this relationship cannot be captured with the CAPM or the Fama and French three-factor models. Not only does the two-factor model accurately capture the cross-section of returns of liquidity sorted shares, it is also able to better explain the value premium than the other two models. There does not appear to be any size premium over the sample period, although the SMB risk factor is still able to explain some variation in the cross-section of returns. This indicates that although size does seem to proxy for some risk, it is not a priced risk. Most notably, shares sorted according to pre-ranking Beta generate the opposite relationship to what is predicted by the CAPM. There is a consistent and monotonically increasing premium from the high Beta portfolio to the low Beta portfolio. This anomaly cannot be explained by any of the three models.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

Within a rational investor paradigm, and in an environment of virtually instant computation and trade execution, it is natural to assume that arbitrage cannot survive for very long before it is traded away. Even 'risky arbitrage' (the oxymoronic nature of this term is noted, however arbitrage has evolved into a more colloquial definition encompassing any mispricing, even if the opportunity cannot realistically be secured with certainty) should not be able to survive unaccompanied by some structural competitive advantage. Thus, it seems reasonable that informational and allocative efficiency should exist, at least most of the time. However, there is a vast amount of literature, both locally and internationally, that show persistent and largely unexplained anomalies within the framework of many of the most popular asset pricing models. The most common of these anomalies are found when analysing excess abnormal returns across size (market capitalisation) sorted portfolios as well as value (typically a ratio of an accounting measure of value to price, such as book-to-market or earnings-to-price) sorted portfolios.

It is unlikely that the almost ubiquitous literature on the simple sorting procedures that result in superior risk-adjusted returns has gone unnoticed by the investor community. This leaves two potential explanations: a non-uniform investor population, each with their own comparative advantages and restrictions, and/or risk factors that are currently unaccounted for. The former proposition presents a major problem. If investors are materially different, then asset valuation could be investor specific. This would mean that prices become a function of not only the changing characteristics of the security, but also the changing preferences, constraints, influence, and composition of the investors that generate the demand for these securities. This problem is beyond the scope of this research but does offer a potential explanation to some of the unexpected results found in this study.

The focus of this research is to use a risk-factor explanation for common anomalies that have been documented on the Johannesburg Stock Exchange (JSE). One method for identifying common risk-factors that cross-sectionally explain returns is to compliment the asset pricing model, with which the anomalies were found, with risk factors derived directly from the characteristics which define the anomalous behaviour. Although seemingly circular, this methodology does not by design ensure the model's success. This is because the risk factors are not generated from the characteristics themselves but rather from the share's sensitivity to the performance of portfolios sorted according to that particular characteristic. This means that a share may have a high risk loading to the size risk factor because it is sensitive to the unobservable risk factor that causes small shares to do well, while not necessarily falling into the classification of small. Furthermore, small shares may still outperform large shares even within a risk model that uses the excess performance of small shares over large shares as an explanatory variable. If small shares did still outperform even after including a size risk factor, this would indicate that the risk measure is failing to fully capture the risk associated with small firms and a more direct measure of this risk is necessary.

Unanticipated innovations in the aggregate characteristic itself could be used as the risk factor series from which to calculate factor loadings. However, this measure could be highly noisy and more prone to measurement error. This is primarily because an estimation of the series' innovations requires the use of an estimation of expected changes in the aggregate characteristic.

Fama and French (1993) employed a mimicking portfolio methodology to successfully explain the abnormal returns attributed to small market capitalisation and high book-to-market firms. This is done by sorting portfolios at the beginning of each re-weighting period into deciles according to the risk characteristic in question (in their case market capitalisation and B/M ratios). The returns for each period are then calculated for the two extreme portfolios. From this the less risky portfolio's returns are subtracted from the risky portfolio's returns to create a risk-premium time series. This is then used as an explanatory variable to estimate the risk loading of a share (along with any other risk-factors such as the market). This study utilises the same methodology.

An alternative approach to identify risk-factors can be to approach the problem theoretically rather than empirically. The capital asset pricing model (CAPM) serves as the theoretical foundation for most asset pricing models. Thus, the source of anomalous returns behaviour is likely to stem from violations of its underlying assumptions. Despite the fact that all of the assumptions simplify reality to some extent, they are likely to impact prices differently depending on the extent to which they differ from reality, and the extent to which their representation of reality differs cross-sectionally.

The assumption of frictionless capital markets contains both of the aforementioned qualities. Markets have substantial friction. Even in the most active exchanges, transaction costs continue to be significant at high frequencies of trade. More important, however, is the difference in the degree of market frictions across shares listed on an exchange. If this cross-section of deviation from the assumption is correlated to a particular characteristic, and deviating from this assumption impacts prices, then it follows that sorting according to that characteristic would lead to an anomalous result under the paradigm of the model in question. Liquidity, defined by Liu (2006) as being able to trade large quantities quickly at low cost with small price impact, seems to embody this assumption. Thus, the degree of illiquidity of a share is likely to be a priced variable unaccounted for by the CAPM. If the previously documented firm characteristics producing the anomalies are related to liquidity then inclusion of liquidity as a variable in the risk model may remove some of the anomalies previously documented on the JSE.

The Liu (2006) Liquidity-Augmented CAPM (henceforth Liq-CAPM) adds an additional liquidity risk-factor to the traditional CAPM. This model is used to test a cross-section of JSE listed shares for the period 1997–2011. Although the period being studied is substantially shorter than in Liu (2006), it has the advantage of examining the model's performance over three years of extreme market uncertainty during the recent financial crisis. Furthermore, the application of this model to a developing country's stock exchange with substantially lower levels of liquidity will act to indicate the models international robustness.

Four key pre-sorting characteristics that are either theoretically or empirically associated with excess total returns will be tested. Namely: market capitalisation (size), book-to-market equity (B/M), market Beta (Beta), and the liquidity factor developed in Liu (2006), which uses a combination of days traded and trading volume (Liq_{12}). The study tests the model by reweighting the pre-sorted portfolios annually, for each characteristic, and regressing their returns against the two independent risk-factor series utilised by the Liq-CAPM. For completeness, all portfolios are also regressed using the traditional CAPM as well the Fama and French three factor model (FF3F model). Furthermore, steps are taken to ensure that the thin-trading that occurs across a large number of shares on the JSE is adjusted for where possible. Where adequate adjustments are not possible shares are excluded but this has been kept to a minimum. This allows the study to be robust to some of the least liquid securities while still maintaining the integrity of the data. The procedures used are outlined in detail in Section 4.

2 Literature review

The literature review is divided into four sections. The first two sections will discuss the literature concerning the theory of the capital asset pricing model and empirical studies of its performance from both an international and South African perspective respectively. These sections also outline the seminal research surrounding the discovery of pricing anomalies such as the small firm effect and the value effect. The third section of the literature review will analyse some of the adjustments that need to be applied when importing a methodology that was constructed for a more liquid exchange to a less liquid environment such as the JSE. The final section analyses the literature leading up to, and developing, liquidity as a risk factor. Specifically, this section provides the rationale and empirical literature demonstrating why liquidity is not only a priced variable, but that it is unable to be accurately captured using the CAPM.

2.1 The international literature

The CAPM, developed separately by Sharpe (1964), Lintner (1965a,b), and Mossin (1966), built upon the seminal research by Markowitz (1959) into diversification strategies and modern portfolio theory. Despite all of the restrictive assumptions, the CAPM was widely accepted in practice and academia, due to its simplicity and sound theoretical background. The model proposes that in a mean-variance optimising environment all market participants will choose to hold a combination of the risk-free asset and the corresponding tangency portfolio of risky assets. This means that the constituents of the tangent risky portfolio, under the models assumptions of rationality and frictionless markets (amongst others), should contain all risky assets in the investable universe. It can be inferred, therefore, that the only relevant risk of each asset is the portion of its total asset risk that contributes to the tangency portfolio's risk. In other words, all unique risk becomes irrelevant due to the investor's ability to diversify it away costlessly, without compromising the weighted-average relationship of expected excess returns. If the investable universe is then proxied by some market index, it follows that the only relevant risk is the covariance of each asset with the market. Standardising this by the aggregate risk of the market leads to the development of Beta, $\frac{Cov_{im}}{Var_m}$, where Cov_{im} is the covariance between the share and the market, and Var_m is the variance of the market index. The standardisation procedure allows an asset's a priori expected excess returns to be characterised by a regression where Beta is the coefficient to the independent variable of expected excess market returns. The CAPM is given below in Equation 1.

$$E[R_i] = \beta_i(E[R_m]) + \varepsilon_i \quad (1)$$

Where R_i and R_m are the share and the market excess return respectively, E is the expectations operator, β is the Beta relationship between the share and the market, and ε_i is a white noise error term. This relationship is intuitively appealing because $\frac{R_m}{Var_m}$ is the excess return per overall risk and Cov_{im} is the contribution of asset *i* to overall risk. The above equation leads to the posteriori regression of excess returns (Market Model), shown in Equation 2 below, and can be used to test the hypothesised one-factor model given above.

$$R_i = \alpha + \beta_i(R_m) + \varepsilon_i \quad (2)$$

Where R_i and R_m are the realised excess returns for the share and the market respectively over the sample period, α is the intercept term, and ε_i is an error term. The benefit of the Market Model given in Equation 2 is that if any alternative factors are necessary to explain excess share returns, then pre-sorting stocks according to that factor should reveal statistically significant, and monotonically changing, intercept values (α) (Jensen, 1968).

Basu (1983) documented excess abnormal positive (negative) returns associated with firms that exhibited high (low) earnings to price ratios (E/P). The effect of size (market capitalisation) on share returns was found by Banz (1981) and Reinganum (1981), whereby smaller market capitalisation firms were found to have positive abnormal excess returns to what would be expected from the CAPM. Studies by Stattman (1980) as well as Rosenberg, Reid, and Lanstein (1985) find that the ratio of the book value of equity to the market value of equity (B/M) is positively related to a share's excess returns. Bhandari (1988) finds that the higher a firms leverage, the higher the excess returns. This relationship was found to be pervasive to the inclusion of market Beta as well as size. In the above papers, shares are sorted into portfolios (usually deciles) according to the characteristic under examination. These portfolios are then reweighted periodically (usually annually) to ensure that any changes in the share's characteristic are accounted for over the sample period. The time series of each portfolio can then be regressed over the sample period using the market's excess return as the independent variable. If the intercept term of the regression is found to be significant at the extremes (highest or lowest of the sorting criteria), and relatively unidirectional in its magnitude, then there is evidence that the sorting criteria under examination is a priced risk variable unaccounted for by the CAPM.

Fama and French (1992), using data from all non-financial firms listed on the NYSE, NASDAQ, and Amex over the period 1962 to 1989, found that B/M and size combine to subsume the effects of E/P, leverage and market Beta. Fama and French (1993) identify five common risk factors that capture the cross-section of average excess stock and bond returns, three risk factors for stocks, and two risk factors for bonds. The stock risk factors that they identify are small minus big (SMB), high book-to-market minus low book-to-market (HML), and the market Beta. The two stock risk factors in addition to market Beta are derived using a mimicking portfolio technique, whereby portfolios are grouped into deciles based on market capitalisation (or book-to-market) and the explanatory variables are derived from the smallest market capitalisation (highest book-to-market) portfolio's return for a given month minus the largest market capitalisation (lowest book-to-market) portfolio's return. The market variable is simply the excess market return. Once these three independent explanatory variables are produced, a multifactor OLS regression is run to determine the sensitivities of each security to each of the three variables. Fama and French (1993) showed that their three factor model was able to capture the cross-sectional average returns of 25 portfolios pre-sorted simultaneously according to B/M and size far more accurately than the CAPM (only three out of the 25 portfolios have alphas in absolute value greater than 0.2% per month using the three factor model versus 15 out of 25 using the CAPM). This indicates that the majority of abnormal returns associated with size and value sorted portfolios can be explained by those shares' sensitivity to the aggregate performance of small firms and value firms.

2.2 The domestic literature

The effect of size and value on the cross-section of average JSE stock returns has been thoroughly tested. Van Rensburg and Robertson (2003a) illustrate that, over the period 1990 to 2000, share returns on the JSE are independently, and jointly, related to size and P/E. The authors find similar results to Fama and French (1992) in that small firms and value firms earn higher returns after taking into account their exposure to market risk. Surprisingly, the authors find that pre-ranking Beta is not only insufficient to accurately estimate future returns, but that it acts in the opposite manner from that proposed by the CAPM – low Beta firms earn slightly greater returns than high Beta firms. This anomalous finding could potentially be due to a downward bias in the pre-ranking Beta estimate of less frequently traded stocks (a point to which the authors concede due to the strong positive correlation between size and Beta). A further explanation for these surprising results is the high frequency of portfolio rebalancing, coupled with the lack of transaction cost implications of trade in the study. Stoll and Whaley (1983) showed that the small firm effect documented by Banz (1981) and Reinganum (1981) could be partially explained through the differences in the cross-section of transaction costs across a spectrum of market capitalisation ranked stocks. Specifically, the authors find that if transaction costs are incurred every month the returns to the smallest portfolio are negative. Due to the two-period nature of the transaction costs of an investment (they are only incurred at the purchase or sale of the asset), and the continuous nature of expected returns, the small firm effect does begin to materialise as the portfolio holding period is increased. The break-even point found in Stoll and Whaley (1983) was found to be four months.

Auret and Sinclaire (2006), using the same data and time period as Van Rensburg and Robertson (2003b), tested their results to the inclusion of a B/M variable. They showed that B/M was able to subsume the effects of both P/E and size when all three variables are run in a multivariate regression. However, the inclusion of the B/M variable was not found to improve upon the two-factor model (size and P/E) of Van Rensburg and Robertson (2003b) under the requirement of all variables being significant at the 5% level. Basiewicz and Auret (2009) also found that B/M was the superior indicator of value stocks. Using a data set from December 1989 to July 2005, which is complete with all known shares listed over the period, the authors attempted to test the persistence of the value and size effects to the inclusion of liquidity and transaction cost proxies.

The authors used various price filters to account for transaction cost implications, as per Bhardwaj and Brooks (1992), and used the twelve month moving average of volume per share as a liquidity filter mechanism, as per Hou and Moskowitz (2005). Further, the authors rebalance their portfolios annually (making trading costs less likely to impact the results) and use both equal-weighted as well as value-weighted portfolios. Their results show that the inclusion of a trading cost and liquidity filter reduce the magnitude and persistence of both the value and size effect but do not eliminate them. Furthermore, the value-weighted portfolios show consistently smaller premiums compared to their equally-weighted counterparts. Basiewicz and Auret (2010) use the same dataset to test the feasibility of the Fama and French (1993) three-factor model to explain the cross-section of excess returns on the JSE. They find that the model produces smaller and less significant pricing errors (intercept values or α) than both the CAPM and a two-factor APT model of Van Rensburg and Slaney (1997), which use the JSE Actuaries All Gold and Industrial Indices as proxies for their two-risk-factor APT model.

Despite the thorough investigation into size and value phenomena in South Africa, the literature on liquidity is scarce, particularly with respect to liquidity as a risk factor necessary in explaining the cross-section of equity returns. The focus has largely been on the effect that illiquid securities have on risk estimation due to the non-synchronous updating of security prices across the JSE. The following section discusses the international and domestic literature that deals with the measurement issues that most commonly accompany illiquid securities. This is followed by the final section of the literature review which analyses the research surrounding liquidity as a priced risk factor.

2.3 Literature on risk measurement issues

Differences in trading frequencies occur both across shares within an exchange as well as across exchanges. This creates a problem for cross-sectional tests of asset pricing phenomenon. Specifically, this is because the asset-pricing models contain measures of relationships between the share price series (dependant variable) and some independent variable series that are unlikely to have similar frequencies. This problem was first addressed by Scholes and Williams (1977). The authors showed that if a share price series trades (therefore updating the price returns series with new information that is likely to affect the underlying value) less frequently than the market proxy (independent variable), then the estimated relationship will be smaller than the true relationship between the two series' changes in underlying value. This problem occurs because the two series are assumed by the model to update information synchronously, and continuously, but in reality one series may lead or lag the other at different discrete intervals.

There has been substantial literature on the OLS estimation bias created by differences in the trading frequencies of a share and the risk-factor. Fischer (1966) is the first paper that highlighted the relevance of the non-synchronicity of returns in estimating the relationship between a market index and a share. The author showed that compared to the true unobservable relationship, the estimate would be downward bias in absolute value. Scholes and Williams (1977) proposed using a combination of contemporaneous, lead, and lag Betas as well as an autocorrelation coefficient, to adjust for the bias. This method was tested on daily data from the NYSE and Amex over the period 1963 to 1975 by grouping shares into quintiles based on their trading volume. The resulting portfolio aggregated-coefficient-Betas were shown to remove large amounts of the bias created by non-trading, illustrated by a narrowing of the distribution of the new Betas across the quintiles – compared to the ordinary OLS Betas.

The Scholes-Williams methodology was later generalised to include lead/lags up to order N by Cohen, Hawawini, Maier, Schwartz, and Whitcomb (1983). This more general form of the model allows lead/lags of higher orders to be included, dependant on the severity of the trading infrequencies with which the author is dealing. McClelland, Auret, and Wright (forthcoming) generated a simulated trading environment using an artificial cross-section of normally, independently, identically distributed shares. The authors then imposed randomised trading limitations with a probability of no trade ranging from 1/250 to 249/250. They demonstrated the relative strengths and weaknesses of the various Beta estimation techniques. The Cohen et al. (1983) aggregated coefficients methodology showed that the higher the number of lead/lag variables included, the greater the percentage of no trading that the model would be able to accurately estimate. However, as would be expected, the greater the number of lead/lag variables (n), the greater the standard error of the estimate. In fact, the standard error of the estimate becomes a function of n. A more data intensive, yet intuitively appealing, procedure of dealing with the non-synchronicity problem is to regress the two series only using data points where both variables have newly updated information in their prices (on days when both series trade). This method requires knowing the days of each trade for each security which is often not available. A Trade-to-Trade (henceforth TT) model that both eliminates the days in which one of the variables does not trade, as well as account for the higher order implications of lengthened periods between trades, was developed by Marsh (1979). This model has been shown to be robust to liquid markets (Dimson and Marsh, 1983) as well as illiquid exchanges (Bowie and Bradfield, 1993). The model not only requires trading date information but also requires reliable daily data to be available. Datasets appropriately adjusted for corporate actions on a daily basis are often not available, especially in less developed markets such as the JSE. This problem has caused some researchers to find less accurate but less data-intensive methods for estimating Beta.

Fowler and Rorke (1983) show that a simple adjustment to Beta can account for the downward bias in absolute value of any simple OLS regression caused by non-synchronousity between the two series. McClelland, Auret, and Wright (forthcoming) show that estimating the Beta against a frequently traded independent variable, through a cross-section of declining frequency dependant variables, generates a linear deterioration of the true underlying relationship. This is illustrated in Figure 1 below, taken directly from McClelland, Auret, and Wright (forthcoming).

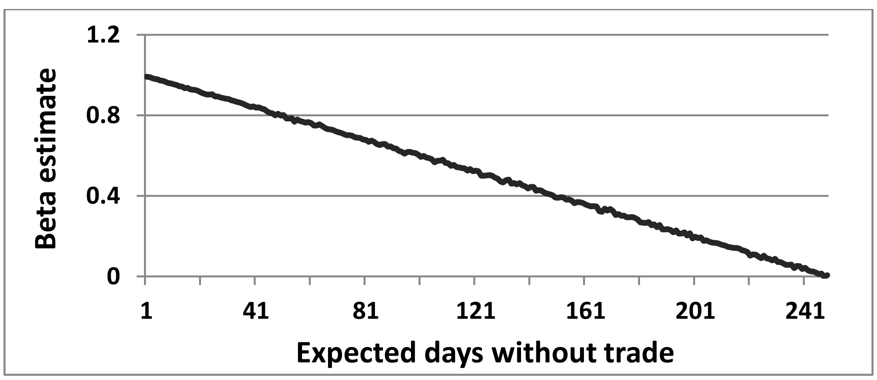

Figure 1. Arithmetic average of OLS

The figure shows the coefficient estimates of a regression of a frequently traded independent variable and a dependant variable with a cross-sectionally declining frequency. The X-axis represents the days not traded out of a potential of 250 and the Y-axis shows the coefficient estimate (the underlying true relationship was simulated with a coefficient of one). The simple solution, shown by Fowler and Rorke (1983), to a linear degeneration of the estimate is to simply multiply the OLS Beta by the inverse of percentage of days traded. This is shown by Equation 3 below.

$$\beta_{Adj} = \beta_{OLS} \frac{trading\ days}{days\ traded} \tag{3}$$

Where β_{Adj} is the new Beta estimate that has been adjusted for the downward bias, β_{OLS} is the unadjusted OLS estimate, $trading\ days$ is the total number of trading days in the estimation period, and days traded is the total number of days that the dependant variable actually traded. The authors tested the above methodology on both daily and monthly simulated returns series and found that the estimates remain bias-free up to about 90% of days without trade. However, the major problem associated with this approach is that by multiplying the original estimate by a factor of the inverse of the percentage of days traded, not only does the coefficient increase, but so does the standard error of the estimate. McClelland, Auret, and Wright (forthcoming) suggest that the solution to this problem is a liquidity filter based on the percentage of days traded, and the strictness of this filter should be informed by the universe of shares (and therefore number of shares per portfolio) at the researcher’s disposal.

The aforementioned solutions to risk estimation on illiquid markets fail to deal with two major problems associated with less liquid securities: transaction costs and liquidity risk. Unfortunately, this study fails to take transaction costs into account and this remains a major caveat to the research. Although this study is expressly concerned with addressing liquidity as a risk factor, the distinction between the cost and risk aspects of illiquidity are often highly inter-related and difficult to disentangle. Therefore, the literature concerning both costs and risks of liquidity will be reviewed together in the following section.

2.4 Transaction costs and liquidity risk literature

To fully appreciate the multiple dimensions of liquidity, as well as the difficulty in isolating its various forms, the following sections briefly outline some of the seminal work in market frictions and liquidity risk.

In Harry Demsetz's 1968 seminal paper 'The Cost of Transacting' the importance of understanding transaction costs and their effect on asset pricing and returns is outlined. Two main transaction costs are identified: brokerage fees and the bid-ask spread. Brokerage fees are relatively consistent within an exchange and can therefore be a predictable cost to investors regardless of their specific portfolio composition. However, the bid-ask spread can vary greatly among stocks. Due to the stock-specific nature of the spread, understanding its components and its effect on prices and returns is necessary in efficiently allocating capital to a portfolio of shares that will maximise investor specific constraints.

Although Demsetz (1968) highlighted a much neglected problem, the multitude of issues related to varying degrees of liquidity amongst shares remained unaccounted for in most asset pricing literature. It wasn't until the CAPM began to demonstrate its weaknesses in explaining the cross-sectional returns of shares sorted according to size, E/P, and B/M that researchers returned their focus to liquidity. The importance of liquidity in explaining CAPM anomalies stems from one of the underlying assumptions of the model: frictionless capital markets. The additional relevance of this assumption, over all of the other idealistic assumptions, is that it is not only different from reality; it is different from reality in varying degrees across shares listed on the exchange. Other assumptions (such as borrowing and lending at the risk-free rate; and rational, risk-averse and utility maximising investors with all information available at the same time to all investors) may differ in the extent to which they are accurate across exchanges but do not substantially differ within an exchange.

Amihud and Mendelson (1986) provide sound theoretical reasoning for two main propositions. Proposition 1 states that varying bid-ask spreads across securities leads to a clientele effect, insofar as stocks with higher spreads will be allocated in equilibrium to portfolios with longer holding periods. Proposition 2 claims that 'In equilibrium, the observed market (gross) return is an increasing and concave piecewise-linear function of the (relative) spread' (Amihud and Mendelson, 1986, p. 228).

The first proposition relies on the logical conclusion that the higher continuous returns associated with increased spread will compensate the investor more for the two-period costs of a high spread the longer the holding period. Longer holding period investors will, therefore, have a greater preference for portfolios with higher spreads (therefore higher returns) than shorter holding period investors. The compensation for the costs associated with higher spread stocks will be greater, the longer the holding period. A formal mathematical proof is provided that shows that in a profit maximising, rational and efficient market this proposition will hold. The second proposition is empirically tested and the following findings provide evidence for a priced liquidity measure. First, market-observed average returns are positively related to the relative spread. Second, the returns to investors net of transaction costs are positively related to the spread. Finally, due to the clientele effect established in Proposition 1, higher-spread stock returns are less spread sensitive and give rise to a concave return-spread relationship.

Although Amihud and Mendelson (1986) don't test whether the first proposition holds, other papers have found that holding periods and bid-ask spreads are in fact positively related. Atkins and Dyl (1997) directly test the relationship proposed by Amihud and Mendelson by testing the relationship between holding period (calculated as shares outstanding divided by trading volume) and bid-ask spread (calculated as the sum of daily relative bid-ask spreads divided by the number of trading days). They found a strong statistically significant relationship on both the NYSE and the NASDAQ. Interestingly, they found a stronger relationship on the NASDAQ which has lower level of liquidity. These findings are reassuring as this study will be applying the same tests on the JSE, which is characterised by far less liquidity than the NYSE or NASDAQ.

Constantinides (1986) formulates a two-asset intertemporal portfolio selection model with endogenous holding periods that incorporates proportional transaction costs and showed a strong positive trading volume – transaction cost relationship. The paper identifies holding periods of investors of a particular stock as being highly sensitive and positively related to transaction costs and their utility being insensitive to portfolio proportions. Constantinides, like Amihud and Mendelson (1986), concludes, therefore, that a clientele effect exists. He finds that this clientele effect is strong enough to eliminate any strong positive effect that transaction costs may have on returns, and therefore believes that they cannot explain anomalies such as the small firm effect identified by Banz (1981) and Reinganum (1981). The conjecture that a strong clientele effect mitigates the spread-returns effect is consistent with the research carried out by Amihud and Mendelson (1986). However, Amihud and Mendelson do not find that the relationship between returns and spread becomes negligible, or that it can be safely ignored. To the contrary, they find that it is sufficient to motivate firms to engage in liquidity enhancing behaviour.

The bid-ask spread is not a perfect measure of the liquidity of an asset and may inadvertently proxy for various other exposures to risks or costs other than liquidity. Huang and Stoll (1997) use a trade indicator model to identify and evaluate the three components of the bid-ask spread, namely: order processing, inventory holding costs, and adverse selection. The first two components comprise the majority of the spread, and are directly linked to the liquidity of the stock and the relative competitiveness of the exchange. Their sample consists of the 20 largest and most actively traded stocks on the NYSE in 1992, and the adjusted results show that the average order processing component made up 61.8%, the average inventory cost 28.7%, and the adverse selection component 9.6%. The adverse selection component may, however, be significantly larger when analysing smaller or less liquid firms, as shown by Seyhun (1986).

Chordia, Roll and Subramanyam (2000) explore aggregate daily liquidity using data from 1988 to 1998 on the NYSE. The authors analyse the liquidity using spreads plus market depth and use volume and the number of daily transactions as a measure of trading activity. Their findings indicate that liquidity reactions to updated information are highly asymmetrical: it plummets in down markets and rises gradually in up markets. Brennan and Subramanyam (1996) analyse market microstructure characteristics related to share liquidity and the impact they have on stock returns. They find that after controlling for the Fama and French risk factors, a significant inverse relationship exists between liquidity and share returns.

The presence of risks and costs requires different treatments for the investor, but both require compensation in the form of higher expected returns. Relative risk aversion of investors means that a firm's equity must provide a premium that will not only compensate the investor to a position with equal expected return (as is the case with liquidity costs under the assumption fixed holding periods), but must create a premium to incentivise holding a position with increased uncertainty.

Pastor and Stambaugh (2003) explore whether aggregate liquidity is a state variable that is priced in asset returns. The authors use a sample of daily data from 1966 to 1999 and control for size, value, market Beta, and momentum effects. Their findings indicate that returns are related cross-sectionally to sensitivities to aggregate liquidity, where high sensitivities are associated with high average returns. Furthermore, they find that up to half of the momentum strategy's abnormal returns can be attributed to sensitivities to aggregate liquidity.

Acharya and Pedersen (2005) solve for the liquidity risk of an asset using the expected liquidity and covariance of returns and liquidity with the market return and liquidity. They argue that liquidity varies over time for individual firms as well as for the market as a whole. Therefore, as with the logic underlying CAPM, the unique liquidity risk can be diversified away but the liquidity beta needs to have a premium assigned to it. This is the basis for their liquidity adjusted CAPM, and although self-admittedly simplistic, it does provide a framework with which to measure liquidity risk and account for any potential premium that should be assigned to variations in liquidity risk across various securities.

One of the major difficulties in estimating the impact that liquidity (both risk and cost) has on asset pricing is that liquidity is so difficult to define. The most easily observable measure of liquidity is the trading cost of an asset, which is measured as the spread between the bid and ask prices (Amihud and Mendelson, 1986). This measure essentially captures the liquidity cost of immediacy of transaction: how much of a premium would you have to pay (discount would you have to accept) for buying (selling) an asset immediately. Another very popular angle for measuring liquidity is trading quantity, measured using the turnover (Datar, Naik and Radcliffe, 1998). This measure allows the scalability of trading to be captured by observing the relative trading volume of a particular asset. Finally, the price reaction to trading volume is used by Amihud (2002) and Pastor and Stambaugh (2003) to measure the price impact dimension of liquidity. This captures the costs to investors of trading out their positions. In particular, this measure captures the bid-ask spread costs associated with larger transactions, indicating the relative depth and concentration of a securities order book.

Although the aforementioned literature all capture critical aspects of the impact of liquidity on asset prices, they do not provide a comprehensive measure for liquidity – one that simultaneously captures the multiple effects that liquidity can have on investors. Furthermore, the impact of trading speed had not been comprehensively explored. Trading speed is an important aspect of an assets' liquidity because it captures the ability of investors to move in and out of positions, limiting this ability can be very costly especially in periods of heightened uncertainty. The combination of the above mentioned dimensions of liquidity are likely to be partially redundant but combine to capture the essence of a generally accepted financial definition of liquidity: ‘The ability to trade large quantities quickly at low cost with small price impact’ (Liu, 2006, p. 1).

Liu (2006) developed a liquidity-augmented capital asset pricing model (henceforth Liq-CAPM) that attempted to capture all four of the above-mentioned liquidity dimensions (trading cost, trading speed, trading quantity, and price impact) in a single liquidity measure, and combine this with the market factor of the traditional CAPM. The two-factor model partially rectifies the theoretical limitations of the CAPM imposed by the simplifying assumption of perfect capital markets. Moreover, empirical analysis reveals that the model has a high level of cross-sectional explanatory power for stocks on the Amex, NYSE and NASDAQ for the period 1963–2003. Specifically, the two-factor model was able to explain abnormal return anomalies previously attributed to size, book-to-market, cash flow-to-price, earnings-to-price, dividend-to-price, and long-term contrarian premiums. To illustrate the robustness of these results the author performed the same empirical analysis over the period 1926–1963 and found results supporting the two-factor model.

The liquidity factor developed by Liu (2006) is defined as the standardised turnover-adjusted number of zero daily trading volumes over the prior twelve months. The number of zero daily trades is said to act as a proxy for the previously unexplored dimension of liquidity: the trading speed of a security. The idea of capturing trading speed through transaction data is, however, overly simplifying. This is because of the fact that for all tradable assets the true potential for trading speed is actually a function of the investor’s willingness to accept a given level of bid-ask spread. Any potential purchaser (seller) could accelerate her purchase (sale) speed by offering (bidding) an amount equal to the lowest ask (offer) price. The number of zero daily trades is, therefore, a measure of the frequency of trades at an acceptable cost. This rate could rise for two reasons: either the benefits of transaction change (which would be the case with both favourable and unfavourable new information), or the cost of the transaction declines (this could be the result of an increase in market participants).

One of the main downsides to zero daily trades is that they are a discrete series with clustering occurring at the most liquid portion of the market. This means that an additional variable is needed to distinguish between the levels of liquidity of two shares that traded every day over the past twelve months. To differentiate between shares that have the same number of zero daily trades, Liu (2006) then looks to turnover as a secondary sorting mechanism. This two-pass sorting system means that firms with a greater number of zero daily trades will always be classified as less liquid than a firm with fewer zero daily trades irrespective of turnover. However, if there are an equal number of zero daily trades then turnover acts as a tie-breaker.

This measure is shown to be highly correlated with the existing measures mentioned above, but also materially different. Analysis of U.S. equities showed that the stocks that the new measure identified as illiquid tended to be small, value, low-turnover, high bid-ask spread, and high return to volume stocks. A detailed description of the model parameters and construction will follow in the proceeding section.

3 Data description and variable construction

This section is divided into three parts. The first lists the data sources and adjustments that were made to the data as well as providing explanation, when warranted, for the use of a particular variable or measure where more than one possible measure is available. The second section lists and explains the inclusion and construction of each of the variables to be used in the empirical analysis. The third section mentions the limitations of the research as well as areas for improvement and development by future researchers.

3.1 The data

The total sample is composed of monthly data of all ordinary shares listed on the JSE from January 1992 to December 2011. However, a maximum of five years of historical data is utilised at each portfolio reweighting date, reducing the 20-year period to 15 years (180 monthly observations) of data to be used in the regression analysis. The Findata@Wits database was used to collect monthly total returns (adjusted for all dividends and corporate actions), number of shares outstanding, market capitalisation, and book-to-market ratios. The I-Net Bridge database was used for the market portfolio returns and the risk-free rate, proxied by the JSE All Share index total returns (J203T) and the three-month Treasury Bill rate (henceforth rfr) respectively. Daily volume data was kindly provided by the JSE. Total return data for each share was adjusted for dividends by adding the market value of any cash, special, or scrip dividend to the price at the trading day following the last day to register. Share consolidations and sub-divisions are adjusted for by multiplying the price at the execution date by the factor that would leave the shareholder's ownership stake prior to execution unchanged. Similarly, the shareholders of companies that go through an unbundling are assumed to sell their stake in the newly unbundled company immediately at the going market price, this amount is then treated as a dividend. For simplicity, delisted companies are assumed to provide the investor with a zero return for that share for the remainder of the holding period. New listings are only considered for each sorting procedure when the minimum required pre-ranking sorting information is available. The least amount of pre-ranking data needed is a year.

To avoid look-ahead bias it is assumed that investors do not have access to final and interim results for three months after their release. Each book value is brought forward by three months, following Basu (1983), using the price returns of the share in question and utilised as an updated book value only at that time. For the three months after the release of new information the previous periods book values are utilised. To deal with any complications arising from name changes and/or ticker code changes, a first true code (FTC) system is used. This system allocates the ticker code to each share that lists if the code is not currently used as an FTC. If the code is already taken, then a 0 is added to the end of the code to identify it as unique; if that code is already taken, a 1 is added, and so on. When a company delists its FTC remains unavailable for new listings. If a company name or ticker code changes it will still be able to be identified by its unique FTC, guaranteeing that no share is incorrectly shown as delisted when it simply changed names or code. Furthermore, every share suspension and delisting has been recorded to ensure that the database contains no survivorship bias.

3.2 The variables

The variables used in this study consist of size, B/M, Turnover, Total Excess Return, Liq, and pre-ranking Beta. The size variable is simply the market capitalisation of the previous month's close. The B/M variable is calculated as the median book-to-market ratio of the previous 12 months. This is done because of the non-normality of the distribution of B/M ratios over time and due to the fact that this sorting mechanism provides the highest absolute and abnormal returns. Turnover is calculated as the total volume traded over the prior twelve months divided by the number of shares in issue. If the number of shares in issue changes over a 12-month period, turnover is calculated for each period with a consistent number of shares in issue and then averaged with a weighting allocated to each sub-period according to the proportion of a year attributed to it. Total excess returns for all shares, as well as the market portfolio proxy, are calculated by subtracting the monthly risk-free rate from each month's total returns. The Liq variable is constructed identically to the Liu (2006) LM12 variable (only the LM12 variable is tested, as this proved to be the best predictor of future returns in the Liu (2006) study of U.S security returns) which is constructed as follows:

$$\text{Liq} = [\text{ZDT}_{12} + \frac{1/(Turnover)}{Deflator}] \text{ x } \frac{252}{NoTD} \quad (4)$$

Where ZDT_{12} is the number of zero daily volumes over the prior 12 months, turnover is calculated as the total volume traded over the period divided by the average number of shares in issue, NoTD is the total number of trading days over the past 12 months, and the deflator is chosen such that: $0<\frac{1/(Turnover)}{Deflator}<1$

The final term in the equation is necessary to standardise the number of trading days from year-to-year to make the liquidity measure comparable over time. The turnover term acts as a tie-breaker between securities that have the same number of zero daily trading volumes over the past year, where a higher turnover will be consistent with a lower Liq value. As can be seen from the above equation, higher values of Liq indicate lower levels of liquidity, and vice versa.

Pre-ranking Betas are calculated using the previous 36 to 60 months total returns data depending on the amount of time that the share has been listed. This is done so as not to restrict the sample too severely from newly listed shares, but still allows more accurate estimates when longer periods of share returns are available. The calculation of pre-ranking Beta differs rather substantially from Liu (2006) due to the different market microstructure characteristics of the JSE compared to the NYSE, Amex, or NASDAQ. Specifically, the cross-sectional range of trading frequencies found on the JSE creates a severe systematic bias of any ordinary OLS estimate of Beta. Following the methodological recommendations of McClelland, Auret, and Wright (forthcoming), and due to the monthly data frequency of this study, the sample is restricted to shares that have no more than an average of 150 days without trade per year over the beta estimation period. The Adjusted OLS is used to calculate individual beta estimates. The Adjusted OLS is calculated according to Fowler and Rorke (1983) as follows:

$$\beta_{Adj\ OLS} = \beta_{OLS} \frac{Trading\ days}{Trading\ days - ZDT} \quad (5)$$

Where ZDT is the number of zero daily trades over the Beta estimation period, $Trading\ days$ is the amount of possible trading days over the same period, and β_{OLS} is the ordinary OLS Beta estimate. The sample is restricted to shares that have no more than 150 days without trade. This is because, as shown in McClelland, Auret, and Wright (forthcoming), at this level of trading infrequency the standard error of the Beta estimate is approximately double what it would be using ordinary OLS. Given the cross-sectional sample size, the number of shares per pre-sorting portfolio ranges from 32–88, which will allow stable and sufficiently accurate estimates of portfolio Betas with the aforementioned increase in the estimate's standard error. The Fowler and Rorke (1982) adjusted OLS is used because it is shown to remain an unbiased estimator for shares that have fewer than 150 days without trade per year.

Finally, the OLS regression used to calculate β_{OLS} forces the intercept through the origin. This is done to calculate a pre-ranking Beta under the full assumptions of the CAPM (in other words, the slope is calculated assuming that Jensen's alpha is equal to zero).

3.3 Limitations

There are a number of limitations to this study which leaves open interesting avenues of investigation for future researchers. The most prominent of this study's shortcomings is that it does not address the problem of separating the risks of liquidity with their associated costs of trade. This separation is both difficult (due to the multiple dimensions of trading costs such as bid-offer spreads and price impacts, as well as the more subtle costs of trade delays) and imperative to a full understanding of asset pricing. The reason that the risks and costs should not be analysed collectively is because of the different effects that they will have on investors. For example the solution to cross-sectional differences in trading costs alone leads to optimal portions of the investable universe becoming a function of investment horizon (Amihud and Mendelson, 1986), which in turn leads to the return premium being a function of the relative distribution of different holding period investors and the trading cost characteristics of the supply of securities available. Liquidity risks, however, will be priced according to the general risk aversion of investors, which is a function of aggregate uncertainty. The ability to separate these two dimensions of liquidity should lead to a more accurate measurement of the true impact of liquidity risk.

A further limitation is that this study does not cover a comprehensive list of potential risk factors. An expansion of this study to include the likes of momentum (Jagadeesh and Titman, 1993) and long-term reversal (DeBondt and Thaler, 1985) would certainly provide a greater understanding of the strengths and limitations of the two-factor model to accurately capture priced risk factors. An analysis of comparative power of the Carhart model to the two-factor model would be especially enlightening given the recent evidence of momentum effects on the JSE (Page, Britten, and Auret, 2013).

In summary, the data collection and variable construction is designed to ensure as much consistency as possible with the Liu (2006) paper, upon which this study is based, while adapting the research, where necessary, to the nuanced trading environment of the JSE. The adaptations are designed primarily to deal with the smaller universe of stocks and greater range of trading frequency found on the JSE. Further adaptations are explained in Section 4, which details the methodology of the paper.

4. Methodology

The theoretical framework and methodology of this paper largely follow the Liu (2006) study. Whenever this study deviates materially from the methodology employed by Liu it is clearly discussed why, and to what extent, the deviation or omission occurs. Broadly speaking, Liu (2006) tested the performance of the Liq-CAPM against the CAPM and FF3F model by sorting portfolios according to a variety of characteristics that have been associated with abnormal returns. He then regressed each of these pre-sorted portfolio returns against each of the various models' factors. If this process leads to a significant intercept term for any of the pre-sorted portfolios then the risk model in question has failed to explain the returns premium associated with that particular pre-sorting characteristic. The main differences in the methodologies and techniques employed have to do with the characteristics of the exchanges under examination. The U.S exchanges that were analysed by Liu (2006) have a far greater cross-section of stocks at any given time and those stocks, on aggregate, trade at a far higher frequency than shares on the JSE.

4.1 Portfolio formation

To isolate the various systematic factors from idiosyncratic share return variations, portfolios are formed according to size, B/M, Liq and pre-ranking Beta. Portfolios are re-weighted at the beginning of January each year. Annual re-weighting was chosen over the more frequent monthly holding period of Van Rensburg and Robertson (2003a) to avoid exacerbating the problem of cross-sectional differences in trading costs. This is especially important for a study that is focussed on the pricing of illiquid securities as the sorting procedure will introduce variations in the trading cost implications of each portfolio. Large differences in the trading costs of different strategies may cause premiums to be found where practical considerations would make trading at that particular periodicity unrealistic. This is more relevant due to the large cross-section being examined in this study (limiting the sample to shares that have fewer than 150 zero daily trades is a relatively lenient trading filter that is applied to avoid measurement bias but does not necessarily address trading cost implications). A further benefit of annual reweighting is that the Liq variable requires a one-year look-back period and, therefore, data points will neither be ignored nor overlap each other. Additional care needs to be taken whenever equally-weighted portfolios are used and the data frequency does not match up to the holding period frequency. This is because equally weighting the total returns of the portfolio will always give a contrarian tilt to returns after the first month of reweighting. If share A goes up and share B goes down in the first month then if we reweight the returns in the following month it will assume we sold a small portion of the previous winner (A) and bought a small portion of the previous loser (B). The process used to correctly estimate the monthly returns is outlined in detail in section 4.1.4.

If a share delists within the one year holding period, the returns for the remaining months are assumed to be zero. This is a conservative and slightly downward bias if the shareholder could invest at least in the risk-free security for the remainder of the period. The sole reason for the choice of a zero return being assigned to shares that delist is due to computational simplicity. The author believes that this caveat is minor and does not materially change any of the results or conclusions that are drawn from the study.

4.1.1 Sorting of portfolios

Portfolios are sorted into quintiles from smallest (S) to largest (L) (the other three portfolios are labelled 2, 3, and 4, with 2 being the smallest of the three) for each of the four variables being analysed: pre-ranking Beta, size, B/M and Liq. The use of quintiles rather than the deciles employed by Liu (2006) or the median split employed by Auret and Cline (2011) was chosen to maximise cross-sectional power of the tests without compromising the stability and reliability of the portfolios. The Liu (2006) study had a far greater cross-section of shares to analyse (NYSE, Amex and NASDAQ) and the Auret and Cline (2011) study had a far stricter liquidity and trading cost filter and, therefore, a much smaller sample. A further digression from the Liu (2006) methodology is the choice to not run multiple characteristic independent sorts. This, too, is done to accommodate the smaller cross-sectional sample of this study and avoid the econometric bias to which this particular research method lends itself if concurrent portfolio sizes differ and are relatively small. Specifically, a research method that examines the significance of the intercept term of an OLS regression is highly susceptible to false positives if one portfolio has a greater likelihood of insufficiently diversifying the unique risks of its constituent shares. A two or three-way independent sort creates a problem in a number of ways. First, changing the number of shares per portfolio from one analysis to another makes the results potentially incomparable (granted this is only a problem if portfolio sizes are sufficiently small, less than 20, or if a value-weighted portfolio formation is implemented). Second, independently sorted portfolios do not guarantee consistent portfolio sizes. In fact, if there is any correlation between the two or three sorting characteristics, portfolio sizes will differ significantly making a comparison across portfolios potentially misleading. Third, preliminary tests have revealed that certain risk characteristics' premiums are not evenly spread across the market making a median split insufficient to isolate their effect. Unfortunately, anything greater than a median split would result in a minimum of nine portfolios and portfolio sizes would be too small (which goes back to point one). This limitation could potentially be reduced by allowing more shares into the sample by making a more lenient trading filter. For this to work a more sophisticated adjustment process would be needed to calculate the risk factor coefficients, such as the Trade-to-Trade model of Dimson and Marsh (1983). Unfortunately, this model requires the use of daily data, which was not available to the author in a format already adjusted for corporate actions.

The sorting procedure is consistent and maintains relatively large portfolio sizes (all greater than 30) reducing the probability of idiosyncratic factors affecting the results.

4.1.2 Exclusions

From the original sample of every ordinary share on the JSE from the period January 1992 to December 2011, several exclusion criteria were implemented. The exclusion criteria are analysed on an annual basis so that no share is permanently excluded unless it fails to satisfy the following criteria consistently.

As mentioned previously, to avoid systematic bias of regression coefficients, any share that has more than an annual average of 150 days without trade is excluded from the sample. Due to the data requirements of the different variables, the following constitute additional reasons for exclusion from a particular portfolio sorting process: any share that has not been listed for a minimum of 36 months is not considered during portfolio formation for pre-ranking Beta portfolios; any share with a negative median B/M ratio over the previous 12 months, or in the financial sector, is excluded from B/M sorted portfolios.

These procedures leave portfolio sizes that range from 36 to 84. The differences in portfolio size should not influence the results as even the smallest portfolio is sufficiently large to eliminate any idiosyncratic factors. This is particularly true due to the equal weighting process which does not cause the same concentration problems as a value-weighted index.

4.1.3 Independent variables

The market risk premium (R_M) is calculated as the total return of the JSE All Share (J203T) minus the three-month Treasury Bill rate (rfr). The other independent variables used in the multifactor risk models being analysed (Fama and French Three Factor Model and the Liquidity-Augmented CAPM) are calculated indirectly using the mimicking portfolio method as described in Fama and French (1993) and are calculated as follows:

Note that S simply refers to the smallest portfolio and L to the largest portfolio.

Liq_{12}: The total returns of the illiquid portfolio (L) minus the total returns of the liquid portfolio (S).

SMB: The total returns of the small portfolio (S) minus the total returns of the big portfolio (L).

HML: The total returns of the high B/M portfolio (L) minus the total returns of the low B/M portfolio (S).

4.1.4 Equally-weighted portfolios

Equally-weighted portfolios are used instead of value-weighted portfolios primarily because the effects of value, size, and liquidity tend to manifest themselves in smaller shares. If portfolios are value weighted, there is a good chance the effects of these risks will be drowned out by the overwhelming effects of larger shares: specifically because the cross-sectional difference in market capitalisation of shares listed on the JSE is so large. Furthermore, the high levels of concentration in the resource sector (documented by Kruger and Van Rensburg, 2008) are more apparent when weighting shares by market capitalisation rather than number of listings. Finally, this study analyses the Jensen's Alpha of portfolios sorted according to a variety of characteristics (liquidity, value, size, and pre-ranking Beta). This makes the results highly sensitive to any remaining idiosyncratic risk in each portfolio. Value weighting will increase the concentration of larger shares and therefore make the number of shares necessary to fully diversify the portfolio far greater than our quintile sorting procedure can sustain. The calculation for portfolio total returns for an equally-weighted portfolio that has a holding period greater than the frequency of returns data is described below.

The calculation for equally-weighted portfolio returns will differ from the conventions of equally-weighted indices' calculations because of the difference in frequency of the returns we would like to capture (monthly) and the intervals between re-weightings (annually). Equally-weighted indices are simply calculated as the arithmetic average of the constituent's returns, essentially keeping the weightings of all shares equal at each point in time. However, when the holding period is greater than the frequency of the portfolio returns that we would like to capture, it is necessary to account for a change in weights (based on the prior returns since the last portfolio reweighting) for every period other than the first month of each annual period. This is done by first creating a weighting variable for each share (Wi) that takes the value of one at each January and then grows according to the below equation (where each month is represented by their chronological order in the calendar year, 1=January to 12=December):

$$W_{i+1} = 1 \times (1 + R_i), \tag{6}$$

where R_i is the total return month i.

Each total return is then multiplied by its respective weight to create a weight-adjusted return (A_i):

$$A_i = W_i \times R_i. \tag{7}$$

The total portfolio return, sorted according to a particular characteristic x, P_x is then simply the sum of all relevant weight-adjusted returns A_i that have characteristic x, divided by the sum of the shares weights (W_i) that have characteristic x:

$$P_x = \sum_{i=x} A_i \div \sum_{i=x} W_i, \tag{8}$$

where A_x and W_x represent the weight-adjusted return and share weight, respectively, for shares that exhibit characteristic x.

The above methodology allows for an adjustment in the weight according to the relative performance of each share in comparison to the overall portfolio within each holding period. This takes into account that shares that perform above (below) average will necessarily have a greater (smaller) weighting in subsequent periods but still equal weights each portfolio at each portfolio formation period. This procedure also largely avoids the built-in contrarian effects that equally-weighted portfolios create. Specifically, a truly equally-weighted portfolio that rebalances its share weights monthly would, by design, sell a certain portion of the previous month's winners and buy a portion of the previous month's losers in order to keep weightings equal. On a monthly frequency this could allow market timing effects to affect the results.

4.2 Preliminary examination

To establish whether liquidity (Liq) is a priced risk factor, this study first examines whether illiquid firms do, in fact, earn higher returns than liquid firms. To do this, portfolios are sorted into quintiles ranging from most liquid (S) to least liquid (L) and reweighted every year. This creates a single return series for each of the five portfolios over the 15-year period. A graphical analysis, T-test, as well as some descriptive statistics are then analysed to see whether the illiquid portfolio outperforms the liquid portfolio; whether the liquidity effect is monotonic across the range of liquidity; and whether or not the supposed risk associated with liquidity manifests itself during the global financial crisis.

The initial hypothesis of a liquidity risk will take the following form:

$$H_0: E(r_B) = E(r_s), \tag{9}$$

$$H_1: E(r_B) > E(r_s). \tag{10}$$

Where r_B is the returns of the illiquid portfolio (buy portfolio), r_s is the returns of the liquid portfolio (sell portfolio), and E is the expectations operator. The shortcoming of this approach is that it does not identify whether liquidity risk is simply capturing the individual or joint effects of the well-documented size and value effects. To add robustness to the preliminary examination, a characteristic-adjusted evaluation of the effect of liquidity on stock returns is undertaken.

4.2.1 Characteristic-adjustment

For completeness, the potential effects of size and value will be taken into account. To adjust for size or B/M individually, each share's total return at every month over the 15-year period is adjusted to remove the return that can be attributed to its own size, or B/M, characteristic. This is done by first identifying the size, or B/M, quintile that each share is a part of and then subtracting the return of that portfolio (excluding the return of the share under examination). This is done for every share to create new adjusted total return series ($\widehat{R_{\iota\jmath q}}$) as shown by Equation 11 below:

$$\widehat{R_{\iota\jmath q}} = R_{ijq} - \overline{R_{(\iota'-\iota)\jmath q}}, \tag{11}$$

where R_{ijq} represents the return for share i that forms part of quintile q when sorted according to characteristic j, and j represents either the size or value characteristic of share i, i' is used to refer to all shares in portfolio q sorted according to characteristic j, such that $\overline{R_{(\iota'-\iota)\jmath q}}$ represents the average return for all shares in quintile q sorted according to characteristic j excluding share i.

The adjusted total return series can then be sorted according to liquidity, as described above, resulting in size and value characteristic-adjusted liquidity portfolios. These portfolios are then re-evaluated to identify whether liquidity is a priced risk that is not captured by the size and value effects.

To test whether the effect of liquidity is robust to the joint effect of size and value, a simultaneous sorting procedure is utilised. To do this, the size and B/M characteristics of each share need to be individually identified and portfolios created where the two characteristics appropriately coincide. This is done to maintain the integrity of each characteristic and the effect it has on returns. For example, the cross-sectional market capitalisation of shares within a particular B/M portfolio may be relatively concentrated, which will not allow a true size premium to be extracted if shares are first sorted according to value. Likewise, a true value premium may not be found if shares are first sorted according size.

In contrast to the two-step sorting procedure described above, which sorts portfolios according to a single characteristic and then applies a secondary sort within each portfolio according to the other characteristic, the simultaneous sorting procedure does not guarantee a similar number of shares within each dual sorted portfolio. In fact, due to the correlations between size and value, an equal distribution of shares is unlikely. For this reason, only three portfolios are formed to adjust for joint characteristic performance. This is done to ensure that a sufficient (greater than 30) number of shares is present in each of the characteristic's buy, hold, and sell portfolios. The 'buy' portfolio is constructed as all four combinations of the two smallest market capitalisation sorted quintile portfolios and the two highest B/M sorted quintile portfolios. The 'sell' portfolio is constructed as all four combinations of the two largest market capitalisation sorted quintile portfolios and the two lowest B/M sorted quintile portfolios. The 'hold' portfolio is constructed from the remaining 17 possible combinations of the size and value sorted quintile portfolios. The 'hold' portfolio offers little economic significance because it contains a variety of combinations of size and value, the effects of which are beyond the relevance of this procedure. The 'hold' portfolio can, however, be thought of as a 'normal returns' adjustment, one which is used to adjust shares that are not small and value, or big and growth, using the average returns of similarly non-extreme characteristic shares. Each share that would form part of each of these three portfolios has the average return of that portfolio (excluding the share in question) subtracted from its return for the same month. This is repeated for every month and every share over the 15-year period to create a total return series that is adjusted for the joint effects of size and value. Shares are then sorted according to their liquidity into quintiles to see whether or not the effects of liquidity are pervasive to the joint characteristic adjustment of size and B/M. Table 1 below illustrates the simultaneous sorting procedure that is applied.

Table 1. Dual sorted portfolio procedure

	Growth				**Value**
B/M	A	B	C	D	E
	Small				**Big**
Market cap	1	2	3	4	5
Buy	D1	D2	E1	E2	
Sell	A4	A5	B4	B5	
Hold	A1	A2	A3	B1	B2
	B3	C1	C2	C3	C4
	C5	D3	D4	D5	E3
	E4	E5			

The process for calculating the new dual adjusted share return series ($\widehat{R_{\imath svq}}$) is shown in Equation 12 below:

$$\widehat{R_{\imath svq}} = R_{isvq} - \overline{R_{(\imath' - \imath)svq}}, \tag{12}$$

where R_{isvq} is the return for share i that forms part of portfolio q when sorted according to characteristic s (size) and v (value) simultaneously, $\overline{R_{(\imath' - \imath)svq}}$ is the average return of all shares that make up portfolio svq excluding share i.

4.3 The three risk models

The two main hypotheses of this paper are as follows: liquidity risk cannot accurately be captured by either the Fama and French Three Factor Model or the CAPM; and that the Liquidity Augmented CAPM of Liu (2006) provides greater explanatory power of B/M, size, liquidity and market risk than either the CAPM or the Fama and French Three Factor Model. The equations below describe each of the three models.

CAPM: $$r_{it} - r_{ft} = \alpha_i + \beta_i (r_{mt} - r_{ft}) + \varepsilon_{it} \tag{13}$$

FF 3 Factor Model: $$r_{it} - r_{ft} = \alpha_i + \beta_i (r_{mt} - r_{ft}) + s_i SMB_t + h_i HML_t + \varepsilon_{it} \tag{14}$$

Liq-Aug CAPM: $$r_{it} - r_{ft} = \alpha_i + \beta_i (r_{mt} - r_{ft}) + \beta_{l,i} LIQ_t + \varepsilon_{it} \tag{15}$$

Where r_{it}, r_{ft}, and r_{mt} are the returns for share i, the risk free asset, and the market portfolio at time t respectively and α_i is the Jensen's alpha for share i. β_i, s_i, h_i, and $\beta_{l,i}$ are the risk loadings for share i for the market, size (SMB), value (HML), and liquidity (LIQ) risk factors.

To test these two hypotheses, pre-sorted portfolios' total returns, distinguished by their varying degrees of the different risk characteristics, are regressed against the independent variables of each risk model. The basic premise of the test is that if the premium of a risk characteristic is not captured by a risk model then the unaccounted for premium will present itself in the form of a significant intercept, or alpha. This method of identifying systematic abnormal returns in a particular risk model framework was introduced by Jensen (1968).

To account for the problem of biased estimators, caused by cross-sectional discrepancies in trading frequencies, the Adjusted OLS method described earlier and prescribed by McClelland, Auret, and Wright (forthcoming) will be employed. A number of additional difficulties arise when trying to apply this method to portfolios rather than individual shares, specifically when the test statistic relies on an analysis of the intercept rather than the coefficient. These problems are individually addressed below. All of the following adjustment processes are supplementary to the Liu (2006) methodology. They are necessary in this study and not in Liu's because of the market microstructure differences. Particularly, cross-sectional differences in trading frequencies are a minor problem on the U.S stock exchanges in comparison to the JSE.

4.3.1 Portfolio adjustment for various trading frequencies

To calculate the percentage of zero daily trades of a portfolio over the regression period (15 years), the following steps were taken: To calculate the average number of zero daily trades for a given time period ($\overline{ZDT_{pt}}$), the number of zero daily trades (ZDT) of each share (i) in the portfolio (p) are summed and then divided by the number of shares (N) in that portfolio for the particular month. This process is shown by Equation 16. The average zero daily trades are then summed over the entire period (T) and divided by the potential trading days during the 15 year period (TD) to calculate the percentage of zero daily trades ($\%ZDT_p$). This can be seen in Equation 17. To then arrive at the Beta adjustment factor (AF_β), I simply divided one by one minus the percentage of zero daily trades. This is equivalent to dividing the potential trading days by the actual days traded. This Beta adjustment factor is then multiplied by the OLS coefficient (B_{ols}) to arrive at the Adjusted OLS estimate (B_{adj}). The final two procedures are given by Equation 18 and 19.

$$\overline{ZDT_{pt}} = (\textstyle\sum_{i=1}^{N} ZDT_{itp})/N \quad (16)$$

$$\%ZDT_p = (\textstyle\sum_{t=1}^{T} \overline{ZDT_{pt}})/TD \quad (17)$$

$$AF_\beta = 1/(1 - \%ZDT_p) \quad (18)$$

$$B_{adj} = AF_{\beta} \times B_{ols} \tag{19}$$

4.3.2 The adjusted statistics

The major statistical problem that arises when using the adjusted Beta of Fowler and Rorke (1982) is that by increasing the slope coefficient by a factor determined by the ratio of trading days to days traded impacts several other necessary statistics. The standard error of the estimate will be increased by the same factor as the slope coefficient and the intercept term will change to fit the new coefficient. This sub-section explains the adjustments that are applied to slope coefficents p-value, the intercept term estimate, as well as the intercept term's p-value.

The difficulty of accurately estimating the p-values of adjusted coefficients is fortuitously avoided because the two necessary adjustments cancel each other out. Calculating p-values from first principles illustrates that the coefficient is divided by the standard error of the estimate. Fortunately, the adjustment process for both the coefficient and the standard error is common, and thus cancels.

To estimate the adjusted intercept, adjusted residuals are first calculated by subtracting the zero-alpha expected returns for each month (using the adjusted coefficient and no intercept) from the actual returns. Because residuals must by nature have a mean of zero, the mean of these zero-alpha residuals is the adjusted intercept.

Finally, to estimate the significance level of the adjusted intercept, the p-value of the mean of the zero-alpha residuals (adjusted intercept) is calculated with an additional degree of freedom removed for each independent variable in the particular risk model.

4.4 Additional statistics for informal portfolio evaluation

Additional statistics are analysed to better understand the nature of the various portfolio time series. The statistics are derived for each portfolio (shown in Tables 3a, 4a, 5a, and 6a), and are briefly listed and explained below.

Primary descriptive statistics: These include the average total excess return and standard deviation of the total excess return; the minimum and maximum monthly total excess return; the minimum, maximum, and average number of shares in a portfolio; as well as a Sharpe ratio and an adjusted R-squared.

Liquidity profile statistics: To gain an understanding of potential trading cost implications, and the effect that the coefficient adjustment had on the significance levels, the percentage of days not traded – as well as the adjustment factor for each portfolio – is calculated. These calculations are shown in Equation 5.

Portfolio mobility: To understand the appropriateness of the annual holding period (over and above the potential practical problem of trading costs mentioned earlier), the percentage of shares that migrate out of a portfolio each year, and every three years, is calculated. High levels of migration would indicate that a more frequent portfolio sort could be appropriate to better isolate the risk factor in question. High levels of migration over these periods may further provide insight into the extent that trading costs (loosely proxied by the liquidity profile statistics) may distort the alpha calculated from a realisable alpha in practice.

Portfolio consistency: To analyse whether outperformance is caused by extreme returns from a small number of shares or consistently by most shares, the percentage of shares that outperformed the risk free rate over the annual holding period is calculated. To check whether the portfolio alpha is derived from a small number of periods (likely to be period specific) rather than a relatively consistent outperformance, the percentage of months and years in which the portfolio outperformed the market is calculated.

Finally, a graphical representation of a R1 investment in each of the five quintiles is illustrated, and contrasted, to the performance of the overall market (J203T). This is repeated for each risk category: pre-ranking Beta, size, B/M, and Liq. The sample is also split into pre-global-financial-crisis (1997–2006) and post-global financial crisis (2007–2011). This break-point is designed to analyse the effects of an exchange-wide shock to risk-appetite. The “buy” risk categorisations are expected to underperform immediately post-global-financial-crisis and outperform consistently better during the pre-global financial crisis period. It is important to keep in mind that the equal weighting of the portfolios increases the weighting of small firms relative to the value-weighted market proxy (J203T). This indirectly creates a size effect on all equal weighted portfolios when they're contrasted to the market. If the small portfolios underperform post-global financial crisis, it can be expected that most, if not all, equally-weighted portfolios underperform the market during the same time period. The converse is equally true pre-global financial crisis.

The methodology employed in this study is designed to translate the methods employed by Liu (2006) onto the JSE. This is done by adjusting certain procedures to match the smaller and less frequently traded JSE. Unfortunately, this does come at the cost of a smaller sample, both cross-sectionally and over time, as well as less powerful test statistics. Despite these drawbacks Section 5 is able to highlight some compelling evidence that liquidity is a priced risk factor and that the Liq-Augmented CAPM is better equipped to explain the cross-section of returns on the JSE.

5 Results

The results are divided into two major sections. The first section (5.1) examines the outcomes of the preliminary tests. These are intended to justify the more rigorous statistical analysis that follows (Section 5.2). Section 5.1 displays the compounded performance of each of the sorting characteristics over the entire sample period, as well as two sub-periods 1997–2006 and 2007–2011. The sub-periods are utilised to examine the performance of each characteristic portfolio both pre- and post- global financial crisis. Section 5.2 tabulates both the descriptive statistics for each characteristic sorted portfolio as well as the results of the various regression analyses. The descriptive statistics provide a greater insight into the results of the regressions. For instance, if a share is displaying significant abnormal returns, are those returns the result of a small number of shares within the portfolio or are the returns more uniformly above average?

5.1 Preliminary results

Figures 2 to 4 show the results of a R1 investment in each of the liquidity quintiles over the period January 1997 to December 2011, as well as pre- and post- global financial crisis sub-periods (1997–2007 and 2008–2011). This rudimentary analysis is necessary to check whether further investigation into liquidity as a risk factor is warranted.

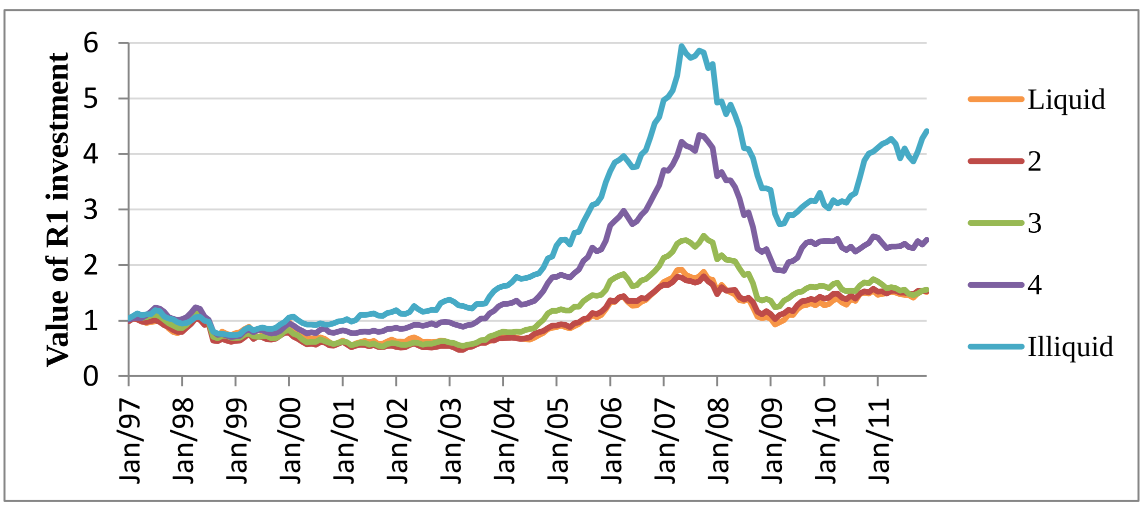

Figure 2 - Liquidity sorted quintiles

Figure 2 shows a monotonic increase in geometric total excess returns as portfolio liquidity decreases. Furthermore, it seems as if liquidity risk is virtually absent from the most liquid 40% to 60% of the market (note that the two most liquid quintiles are drawn from a population that has already excluded shares that have more than 150 days without trade and the true proportion of the market captured in these two quintiles is actually smaller than 40%).

To see whether the risk premium rises with global uncertainty, the 15 year period is broken up into two sub-periods. The 11 years running up to the crisis, seen as a relative bull market with declining global risk aversion, and the four years including the crisis, seen as a period of global shock and a sudden increase in global risk aversion. The periods span from January 1997 to December 2007 and from January 2008 to December 2011.

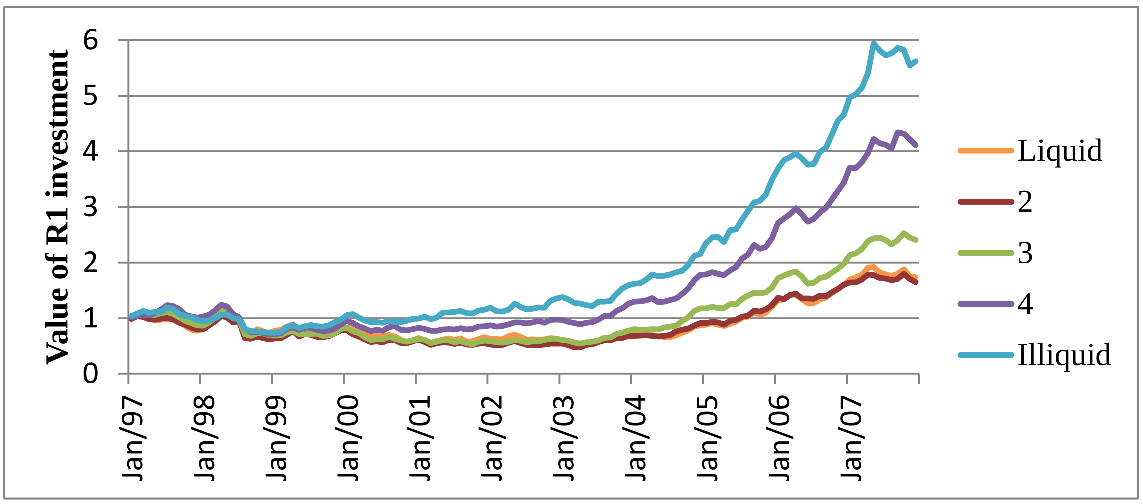

Figure 3 - Liquidity sort: pre-GFC

The risk sorted portfolios in Figure 3 conform to the manner in which one would expect a risk characteristic to perform in a period absent of any structural shocks to risk aversion. The riskier portfolios outperform their less risky counterparts over the 11 year period, with the exception of the two least risky quintiles which performed similarly. It must be reiterated that this analysis does not take into account the higher transaction costs (in the form of wider bid-offer spreads) that are also associated with less liquid shares. However, due to the relatively long holding period of the study (one year) this should not materially affect the results.

Figure 4 displays the performance in the latter period, which contains the structural shock to global risk aversion in the form of the global financial crisis. The second sub-period shows that, as would be expected, the more risky portfolios, in general, underperform their less risky counterparts. Interestingly, the second most liquid quintile seems to be less affected than the other four portfolios.

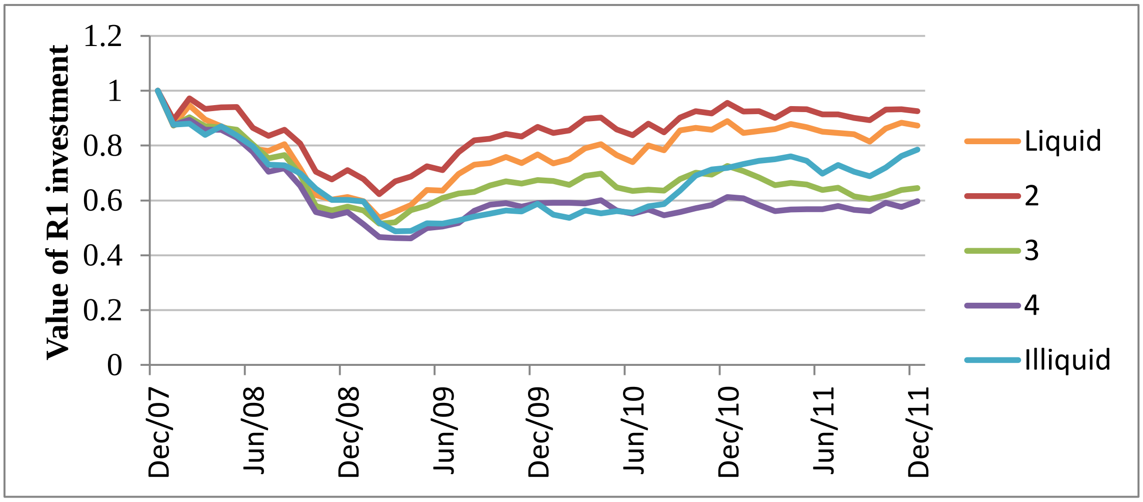

Figure 4 - Liquidity sort: post-GFC

Figures 2 to 4 provide a picture that is generally in line with a risk factor. However, the possibility still remains that this risk characteristic may be captured either partially or wholly by well-documented risk characteristics such as size or value. To ascertain whether or not the liquidity sorted portfolios are capturing other well-established risks rather than an independent liquidity risk, individual share returns are adjusted for their B/M and market capitalisation characteristics, both individually and jointly. Portfolios are then once more sorted into quintiles according to their liquidity characteristics to check whether the risk is independent and robust to the size and value risk factors. The adjustment procedure is outlined in Equation 11 in Section 4.2.1. The results are shown in Figure 5, 6 and 7.

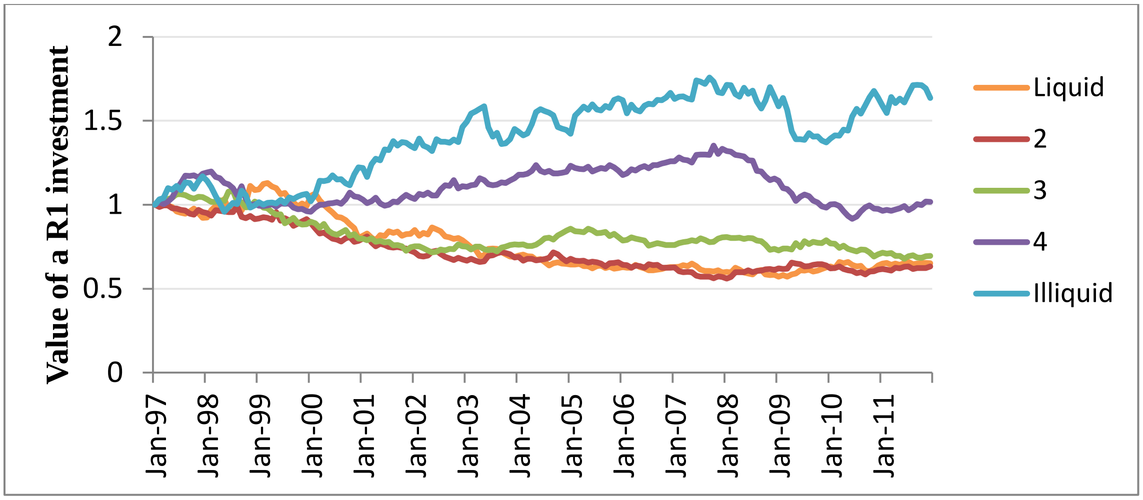

Figure 5 - Value adjusted liquidity sort

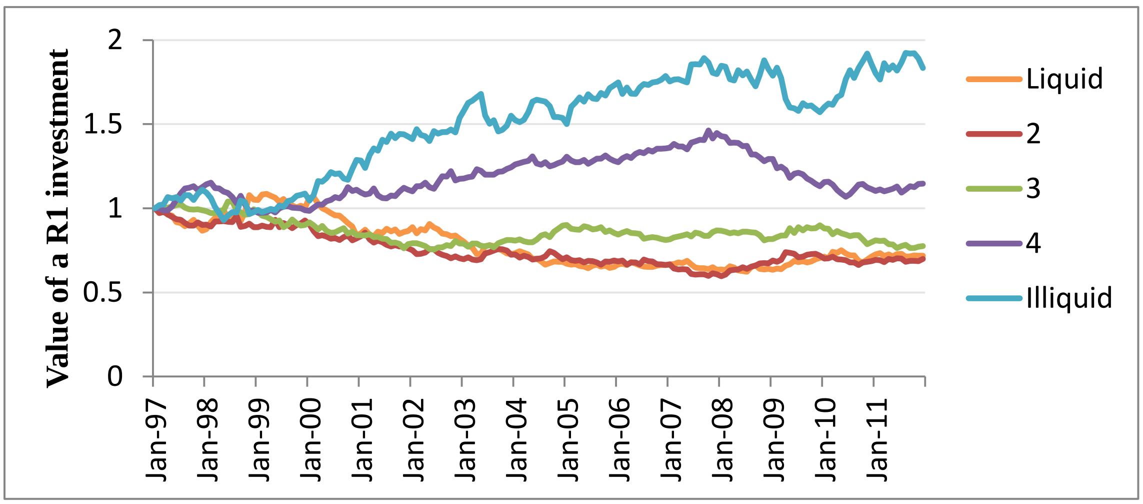

Figure 6 - Size adjusted liquidity sort

Figures 5 and 6 above show that the liquidity risk is pervasive after adjusting total excess returns for the value or size quintile to which each share belongs. The two least liquid quintiles continue to substantially outperform their more liquid counterparts and the two most liquid quintiles remain largely indistinguishable.

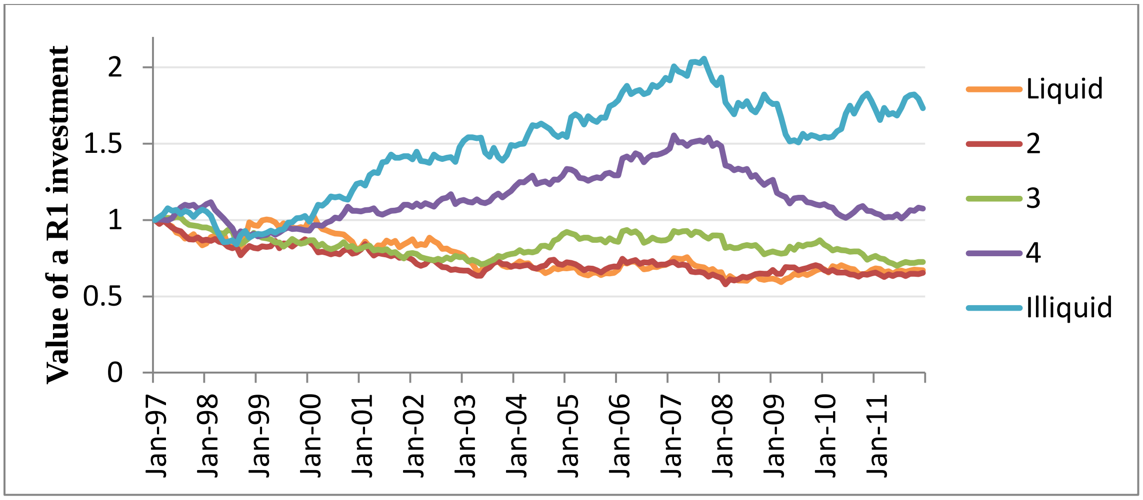

Figure 7 – Dual adjusted liquidity sort

A dual adjustment for the joint effects of size and value also fails to account for the higher total excess returns attributed to the less liquid portfolios. This is shown in Figure 7 above and the adjustment procedure is outlined in Equation 12 in Section 4.2.1. Through all three adjustment processes the effect of sorting portfolios according to the liquidity measure of Liu (2006) creates distinct return characteristics consistent with a priced risk variable. The adjustment processes seem to have little or no effect on the returns premiums across liquidity sorted shares. Furthermore, a student's t-test of the two extreme liquidity portfolios (the most liquid and least liquid quintiles), assuming independent samples with unequal variances, is significant at the 10% level for the unadjusted, both single-factor-adjusted, as well as dual-adjusted time series. The results are shown in Table 2 below.

Table 2. T-test of difference in mean between most and least liquid quintile

Quintile series	P-value
No adjustments	0.065
Size adjusted	0.059
Value adjusted	0.061
Size and value adjusted	0.062

From the results of the preliminary analysis it seems likely that liquidity is a priced risk variable that is separate from other characteristics such as size and value. The geometric series from both unadjusted and adjusted portfolio sorts show monotonically rising returns as the liquidity of the portfolio declines. Furthermore, the liquidity portfolios, for the most part, seem to underperform during shocks to aggregate risk aversion such as the global financial crisis. This is what would be expected from a characteristic that is capturing a systematic risk factor.

Section 5.2 applies more rigorous statistical analysis to determine whether or not this risk factor is captured by either the CAPM or the FF3F model, and whether Liu's two-factor Liq CAPM does a better job at explaining the cross-section of returns. Furthermore, the Liq CAPM is tested using portfolios pre-sorted according to size, value, and pre-ranking Beta to see whether the model can explain some of the more traditional risk characteristics.

5.2 The risk model results

The CAPM, Fama and French Three Factor Model (FF3F), and the Liquidity-Augmented CAPM (Liq-CAPM) are tested on quintiles sorted according to market capitalisation, book-to-market ratios, liquidity, and pre-ranking Beta. The portfolio statistics are summarised in Tables 3a, 4a, 5a, and 6a and the regression statistics are summarised in Tables 3b, 4b, 5b, and 6b below. A description of the methods utilised in the 'a' and 'b' tables can be found in Section 4.4 and Section 4.3 respectively. The B – S column in each of the 'b' tables refers to a 'buy' (B) minus 'sell' (S) portfolio which represents the portfolio that is expected to earn the highest total return minus the portfolio which is expected to earn the lowest total return. Specifically it is high Beta minus low Beta, small minus big market capitalisation, high B/M minus low B/M, and illiquid minus liquid for each of the four risk categorisations respectively. The B – S portfolios aid in concentrating Jensen's Alpha for ease of identification when the abnormal returns of the quintiles increase or decrease in magnitude in a linear fashion. If the relationship across the quintiles is non-linear (such that the minimum or maximum alpha is not found in either of the extreme portfolios), or if the alphas of the extreme portfolios have the same sign (both negative or positive) then the B – S column may not correctly isolate the unaccounted for risk premium in question and the individual quintiles may serve as a better test for unidentified risks.

Furthermore, it is important when analysing the results to keep in mind that the power of the tests is unfortunately not consistent across portfolios within a risk classification, nor is the degree of inconsistency within a risk classification comparable across the different classifications. This is due to the adjustment mechanism utilised to remove the systematic bias of the coefficients. The procedure multiplies the ordinary OLS coefficient by the total days traded divided by the total trading days over the 15-year period. This mechanical adjustment is done to remove any systematic bias of the coefficient but it also magnifies the standard error of the estimate, which in turn decreases the t-statistic for both the coefficients as well as the intercept term. This makes identifying the significant Jensen's alphas (which signify uncaptured risk) more difficult and increases the probability of committing a Type 2 error. The degree to which this will affect a portfolio can be seen by the percentage of days not traded in each of the 'a' tables. The greater the percentage of days not traded, the greater the probability of a Type 2 error. For the B – S column, which is not shown in any of the 'a' tables, a simple average of the two extreme portfolios percentage of days not traded is indicative of the decrease in power of the portfolio.

Table 3a. Pre-ranking Beta sorted portfolio statistics

Portfolio	Low Beta	2	3	4	High Beta
Monthly excess return	1.14%	0.83%	0.59%	0.51%	0.44%
Standard deviation	4.70%	4.95%	5.35%	5.55%	6.03%
Minimum	-24.47%	-29.60%	-26.64%	-26.41%	-23.93%
Maximum	12.53%	13.02%	12.95%	15.85%	13.95%
Sharpe ratio	0.24	0.17	0.11	0.09	0.07
Number of shares	44	43	43	43	44
% of days not traded	15.80%	9.27%	10.10%	11.14%	15.86%
% that remain in portfolio next year	59.59%	45.97%	40.77%	43.71%	60.90%
% that remain in portfolio for next 3 years	24.43%	11.99%	8.69%	9.31%	24.75%
% of winners per portfolio per year	65.29%	61.09%	55.40%	51.52%	52.31%
% months outperformed market	57.22%	53.89%	51.11%	51.67%	53.89%
% of J203 value 2011	321.15%	177.66%	111.06%	94.33%	80.00%
Range of returns	37.00%	42.62%	39.59%	42.25%	37.88%
% years outperformed market	66.67%	46.67%	53.33%	46.67%	53.33%

Table 3a shows the descriptive statistics for portfolios sorted according to pre-ranking Beta and presents an interesting problem for the use of the CAPM on the JSE. There are monotonically decreasing average monthly excess returns over the 15-year sample period as you move from the lowest pre-ranking Beta quintile to the highest, from 1.14% to 0.44% per month. This effect is even greater on a risk-adjusted (Sharpe ratio) basis with 0.24 for the low Beta portfolio compared to 0.07 for the high Beta portfolio. This relationship is the inverse of what is predicted by the CAPM. The use of the Beta adjustment mechanism means that it is unlikely that low Beta shares are simply capturing low liquidity shares due to the measurement error that is introduced by the share series trading less frequently. The trading characteristics of the portfolios, seen as the percentage of days not traded, show that the low Beta portfolio trades about as often as the high Beta portfolio and almost as much as the three middle portfolios. This demonstrates that the phenomenon cannot be explained by the measurement bias introduced by thin-trading.

The low Beta portfolio also seems to be the most consistent performer, both on a constituent basis whereby an average of 65% of the low Beta portfolio's shares outperformed the market, as well as on a time series basis, whereby approximately 57% of the months the portfolio outperformed the market.

It must be noted that, when comparing the performance of the portfolios to the market, one must keep in mind that the market is value-weighted whereas the portfolios are equal weighted. This means that the market does not always fit as an aggregate of the five portfolios because the market is relatively overweight in large market capitalisation stocks. Thus one would find that when small firms outperform (underperform) large firms, the majority of the equally-weighted quintiles will outperform (underperform) the market. Furthermore, due to the long period over which the descriptive statistics are calculated, some of the results may appear contradictory which warranted further examination. For example, the high Beta portfolio achieved the second highest percentage of years outperforming the market, had a low minimum return, a high maximum return and yet it had the lowest value as a percentage of the market at the end of the period. This was found to be due to the high monthly standard deviation of returns which increases the differences between arithmetic and geometric measures of performance.

From Table 3a we can see that the percentage of zero daily trades for the low Beta portfolio is approximately the same as the high Beta portfolio, both of which seem to be slightly higher (less frequently traded) than the middle three portfolios. At first sight this makes the monotonic decline of returns across pre-ranking Beta sorted portfolios unlikely to be fully explained by the addition of liquidity risk to the ordinary CAPM framework. This is confirmed in the regression results in Table 3b where the CAPM, FF3F, and Liq-CAPM all fail to explain the abnormal returns of the lowest and, to a certain extent, the second lowest Beta quintile. The lowest Beta quintile exhibits positive and economically significant monthly alphas of 82bp, 109bp, and 81bp for the CAPM, FF3F and Liq-CAPM respectively. All of these alphas are significant at the 1% level. The second lowest Beta portfolio also shows positive and economically significant alphas (especially using the FF3F model) which are all significant at the 10% level (significant at the 5% level for the FF3F model).

Sorting portfolios according to pre-ranking Beta, by design, demonstrates increasing market risk as you move from the low Beta to high Beta portfolios. For all three risk models, the market risk coefficients seem to be largely consistent and in line with their backward looking Beta characteristics. Liquidity risk (Liq) does not show a clear relationship to the pre-ranking Beta portfolios, but it does show that the low Beta portfolio is the only quintile to have a positive risk exposure to liquidity. However, this may be an econometric issue rather than a true economic relationship. Although every measure was taken to avoid the downward bias problem of thin trading, the increased magnitude of standard errors due to the Beta adjustment procedure may not have been fully compensated for by portfolio sizes of approximately 44 shares. However, it must be noted that this is unlikely considering the almost even level of thin trading across portfolios, 9.27%–15.86% of days not traded illustrated in Table 3a.

Considering the two non-market risks of the FF3F model, it can be seen in Table 3b that there seems to be no real relationship between pre-ranking Beta portfolios and the exposure to HML (value) risk but that there seems to be an almost significant relationship to SMB risk (size) at the two extremes of pre-ranking Beta sorted portfolios. The second highest Beta portfolio has a marginally significant exposure to size risk at the 10% level. Looking at the Liq CAPM model, Portfolio 3 is the portfolio that has a significant exposure to liquidity risk at the 5% level. Not surprisingly, because this portfolio was actually found to have traded relatively frequently during the estimation period, that exposure is negative. However, these risk exposures are mostly insignificant. Looking at the adjusted R-squared statistics for the three models, all of the models seem to explain approximately equal proportions of the variability in the quintiles over the 15 year period, ranging from a low of 35% in the low Beta portfolio to a high of 62% for the middle portfolios. The positive abnormal returns to low Beta stocks cannot be explained through any of the three models under examination. This phenomenon should be an interesting area for future research.

Table 3b. Pre-ranking Beta sorted regression results

			Low Beta	2	3	4	High Beta	B - S
Adjusted Betas	CAPM	Market	0.56	0.66	0.77	0.78	0.81	0.25
	FF3F	HML	0.08	0.05	-0.05	0.03	0.05	-0.03
		SMB	0.10	0.04	0.03	0.10	0.12	0.02
		Market	0.60	0.68	0.77	0.81	0.85	0.25
	Liq-Aug CAPM	Market	0.56	0.64	0.73	0.74	0.81	0.24
		Liq	0.01	-0.04	-0.14	-0.12	-0.03	-0.03
Adjusted Beta's P-Values	CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.00
	FF3F	HML	0.40	0.55	0.53	0.76	0.65	0.81
		SMB	0.14	0.43	0.59	0.09	0.13	0.82
		Market	0.00	0.00	0.00	0.00	0.00	0.00
	Liq-Aug CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.01
		Liq	0.92	0.51	0.03	0.11	0.80	0.75
Adjusted Alphas %	CAPM		0.82%	0.45%	0.14%	0.06%	-0.02%	-0.84%
	FF3F		1.09%	0.80%	0.62%	0.49%	0.41%	-0.68%
	Liq-Aug CAPM		0.81%	0.48%	0.25%	0.15%	0.00%	-0.82%
Adjusted Alpha's P-Values	CAPM		0.00	0.08	0.57	0.83	0.94	0.02
	FF3F		0.00	0.03	0.14	0.23	0.35	0.06
	Liq-Aug CAPM		0.00	0.06	0.30	0.58	0.99	0.02
Adjusted R squared	CAPM		0.36	0.53	0.61	0.56	0.46	0.06
	FF3F		0.36	0.52	0.60	0.56	0.45	0.05
	Liq-Aug CAPM		0.35	0.53	0.62	0.57	0.45	0.06

Table 4a. Market capitalisation sorted portfolio statistics

Portfolio	Small	2	3	4	Big
Monthly excess return	0.46%	0.45%	0.52%	0.49%	0.51%
Standard deviation	5.66%	4.95%	4.92%	5.07%	5.60%
Minimum	-26.30%	-24.89%	-26.33%	-26.65%	-27.09%
Maximum	17.87%	9.87%	12.69%	11.74%	15.26%
Sharpe ratio	0.08	0.09	0.11	0.10	0.09
Number of shares	62.73	62.40	62.40	62.40	63.07
% of days not traded	35.51%	20.14%	13.78%	5.40%	1.00%
% that remain in portfolio next year	61.48%	55.71%	57.68%	66.67%	84.68%
% that remain in portfolio for next 3 years	21.81%	16.70%	19.75%	26.65%	63.01%
% of winners per portfolio per year	44.23%	55.28%	60.13%	60.84%	53.45%
% months outperformed market	47.22%	51.67%	50.56%	51.11%	49.44%
% of J203 value 2011	86.63%	89.79%	102.33%	95.50%	94.49%
Range of returns	44.17%	34.76%	39.02%	38.40%	42.35%
% years outperformed market	60.00%	53.33%	46.67%	46.67%	53.33%

The portfolio statistics and regression outputs of sorting portfolios according to size are shown in Tables 4a and 4b respectively. Table 4a doesn't reveal any strong signs of a cross sectional difference in either of the first two moments of the distribution of monthly excess returns. This result is contrary to Van Rensburg and Robertson (2003a, 2003b) that finds a strong size effect which is independent of value. The main reason for the difference in results is probably due to the different sample periods. It was found by analysing the period 1992–1997, a period which forms a portfolio formation period in this study but does not make up part of the testing sample period, that a very strong size premium existed. This premium seems to completely disappear after 1997. An additional reason may be the choice of reweighting periods. The annual reweighting used in this study may be too long to capture the premiums associated with smaller shares.

There seems to be a strong relationship between size and trading frequency, as one would expect, demonstrated by the increasing percentage of days not traded as the portfolio size decreases. Furthermore, the larger quintiles appear far more stable than the smaller groups with 63% of shares remaining in the large portfolio over an average three year-period, whereas only 21.8% remain over the same period for the small portfolio. The percentage of market value at the end of 2011 shows that four out of the five portfolios underperformed the market over the 15-year period (albeit slightly). As mentioned earlier, this is possibly because of the difference in equally-weighted and value-weighted portfolios and their sensitivity to the relative performance of small firms versus large firms over the sample period.

Looking at the adjusted alphas' p-values in Table 4b we can see that the small firm anomaly is absent from the time period under examination where even the ordinary CAPM is able to fully explain the cross-section of size-sorted excess returns. All three risk models fail to reject the null hypothesis of a zero intercept term (Jensen's Alpha) for all five quintiles at the 10% significance level. The adjusted R-squared statistics for the three models reveal that the FF3F model seems to explain the greatest amount of variation in the two smallest portfolios (28% and 39%) but the Liq-CAPM explains more in the two larger portfolios (57% and 81%). The liquidity risk factor is significant at the 5% level for three out of the five portfolios. Specifically, it shows that the smallest portfolio has a significantly positive exposure to liquidity risk and the two biggest portfolios have a significantly negative exposure to liquidity risk. This would be expected given the trading characteristics of the five portfolios. Surprisingly, only the two smallest of the five portfolios showed significant exposure at the 5% level to the SMB risk factor of the FF3F model, although portfolio three marginally misses the cut-off and is significant at the 10% level. Size, liquidity, and value showed significant coefficients at the 1% level for the B–S portfolio. Only the biggest quintile portfolio showed an exposure to the HML risk factor, which was a significantly negative exposure. Most surprisingly, however, was that the CAPM showed the least statistically, and economically, significant intercept term. This indicates that the significant exposures to certain risk factors, size does not seem to be a priced risk factor, offering a premium for taking on higher levels of systematic risk.

Table 4b. Market capitalisation sorted regression results

			Small	2	3	4	Big	B - S
Adjusted Betas	CAPM	Market	0.81	0.65	0.67	0.65	0.82	-0.01
	FF3F	HML	0.25	0.00	-0.05	-0.08	-0.13	0.38
		SMB	0.50	0.20	0.11	0.05	-0.03	0.53
		Market	0.99	0.70	0.69	0.65	0.79	0.20
	Liq-Aug CAPM	Market	0.92	0.66	0.66	0.60	0.75	0.17
		Liq	0.35	0.02	-0.05	-0.18	-0.23	0.58
Adjusted Beta's P-Values	CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.00
	FF3F	HML	0.19	1.00	0.59	0.27	0.01	0.00
		SMB	0.00	0.01	0.06	0.33	0.38	0.00
		Market	0.00	0.00	0.00	0.00	0.00	0.01
	Liq-Aug CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.02
		Liq	0.05	0.84	0.47	0.00	0.00	0.00
Adjusted Alphas %	CAPM		0.00%	0.07%	0.13%	0.11%	0.04%	-0.04%
	FF3F		0.31%	0.44%	0.54%	0.53%	0.59%	-0.28%
	Liq-Aug CAPM		-0.27%	0.06%	0.17%	0.25%	0.22%	-0.49%
Adjusted Alpha's P-Values	CAPM		0.99	0.80	0.61	0.66	0.85	0.91
	FF3F		0.41	0.23	0.16	0.19	0.21	0.53
	Liq-Aug CAPM		0.46	0.84	0.50	0.31	0.23	0.13
Adjusted R squared	CAPM		0.22	0.38	0.50	0.55	0.78	0.00
	FF3F		0.28	0.39	0.49	0.54	0.78	0.25
	Liq-Aug CAPM		0.24	0.37	0.50	0.57	0.81	0.18

Tables 5a and 5b present strong evidence for a value effect. Table 5a shows that average excess monthly returns increase almost monotonically as you move from extreme growth to extreme value portfolios (the exception is portfolio 4 which actually outperforms extreme value by having the highest returns and the lowest risk). The Sharpe ratios show that this cannot be explained by the standard deviations of the portfolios which are relatively uniform across the B/M categories. Looking at the percentage of days not traded; there is evidence that the relative lack of liquidity in the more value portfolios could play a role in their outperformance as there does seem to be a consistent deterioration of liquidity as the portfolios move from growth to value (11.5% of days not traded in growth to 27% of days not traded in value). This is confirmed in the regression outputs in Table 4b whereby the quintiles show increasing sensitivity to the liquidity risk premium (Liq) as we move from growth to value. Looking at the percentage of months that each portfolio outperformed the J203T, and the overall performance relative to the J203T, it is once again shown that portfolio 4 is superior not only to portfolios 1, 2, and 3, but also the highest B/M portfolio. The performance of portfolio 4 is shown to also have the smallest range of returns over the sample period.

Table 5a. Book-to-market sorted portfolio statistics

Portfolio	Growth	2	3	4	Value
Monthly excess return	0.27%	0.32%	0.45%	0.85%	0.82%
Standard deviation	5.61%	5.07%	5.04%	4.91%	5.40%
Minimum	-33.45%	-26.36%	-28.01%	-21.73%	-24.29%
Maximum	14.57%	11.89%	12.00%	11.55%	18.62%
Sharpe ratio	0.05	0.06	0.09	0.17	0.15
Number of shares	54.67	55.27	55.27	55.27	62.93
% of days not traded	11.45%	11.24%	12.71%	17.19%	26.98%
% that remain in portfolio next year	47.69%	32.59%	32.68%	32.73%	47.41%
% that remain in portfolio for next 3 years	22.91%	11.21%	9.43%	6.50%	19.90%
% of winners per portfolio per year	51.57%	53.62%	54.05%	60.15%	53.93%
% months outperformed market	51.11%	50.56%	48.33%	60.00%	53.89%
% of J203 value 2011	60.80%	70.55%	89.78%	185.97%	170.37%
Range of returns	48.02%	38.25%	40.01%	33.28%	42.91%
% years outperformed market	46.67%	40.00%	33.33%	60.00%	53.33%

Looking at Table 5b, the coefficients to HML behave as would be expected with the sensitivity to the value risk factor increasing as the B/M of the portfolio increases. The coefficients are significantly negative at the 5% level for the two lowest B/M portfolios and significantly positive at the 5% level for the highest B/M portfolio. As mentioned above the exposure to liquidity risk increases monotonically as the B/M of the portfolios increase. These coefficients are found to be significant at the 10% level for the lowest three B/M portfolios (growth portfolios) but insignificant for the two value portfolios. The B – S portfolio coefficients are significantly positive at the 1% level for both Liq as well as HML risk factors. The size and market risk factors show an inconsistent relationship across portfolios.

Table 5b. Book-to-market sorted regression results

			Growth	2	3	4	Value	B - S
Adjusted Betas	CAPM	Market	0.79	0.74	0.71	0.69	0.74	-0.05
	FF3F	HML	-0.40	-0.16	-0.06	0.06	0.64	1.04
		SMB	0.12	0.10	0.08	0.08	0.20	0.08
		Market	0.74	0.74	0.72	0.72	0.91	0.17
	Liq-Aug CAPM	Market	0.74	0.71	0.67	0.67	0.78	0.04
		Liq	-0.17	-0.13	-0.12	-0.07	0.12	0.30
Adjusted Beta's P-Values	CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.02
	FF3F	HML	0.00	0.04	0.46	0.57	0.00	0.00
		SMB	0.04	0.05	0.16	0.25	0.02	0.24
		Market	0.00	0.00	0.00	0.00	0.00	0.84
	Liq-Aug CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.21
		Liq	0.02	0.05	0.10	0.41	0.34	0.00
Adjusted Alphas %	CAPM		-0.18%	-0.11%	0.05%	0.45%	0.40%	0.58%
	FF3F		0.50%	0.41%	0.49%	0.82%	0.45%	-0.04%
	Liq-Aug CAPM		-0.05%	-0.01%	0.14%	0.51%	0.30%	0.35%
Adjusted Alpha's P-Values	CAPM		0.52	0.65	0.86	0.09	0.23	0.07
	FF3F		0.38	0.35	0.22	0.02	0.16	0.94
	Liq-Aug CAPM		0.87	0.97	0.59	0.06	0.37	0.26
Adjusted R squared	CAPM		0.56	0.62	0.54	0.4724	0.32	0.02
	FF3F		0.61	0.62	0.53	0.4644	0.46	0.65
	Liq-Aug CAPM		0.57	0.63	0.55	0.4722	0.32	0.08

Positive and significant alphas are found in the fourth portfolio (second most value portfolio). Surprisingly, the highest (0.82% per month) and most significant (p-value = 0.02) coefficient is found using the FF3F model. The CAPM and Liq-CAPM both show significant alphas at the 10% level with p-values of 0.09 and 0.06 respectively. This seems to indicate that the CAPM and Liq-CAPM have lower and less significant levels of unexplained returns when sorting according to the B/M. This is partially in line with the results of Liu (2006), which also found that the Liq CAPM was better at explaining the value premium than the FF3F model. The adjusted R-squares show slightly higher explanatory power for the FF3F model in the two extreme portfolios. Furthermore, despite the more significant intercept term found using the FF3F model, this model also provided the smallest range of alpha values across the B/M sorted portfolios. Overall, however, it seems as if the Liq-CAPM does at least as well if not better at explaining the value premium as the CAPM and FF3F model.

The final quintile formation procedure uses the liquidity risk measure proposed by Liu (2006) as the pre-sorting characteristic. The portfolio statistics and regression outputs are illustrated in Tables 6a and 6b respectively. The average percentage of days not traded ranges from 0.92% for the most liquid quintile to 42.34% for the least liquid quintile. It is clear that investors are being compensated for holding less liquid securities shown by the steadily increasing monthly excess returns: 0.34% in the liquid portfolio to 0.87% per month in the illiquid portfolio. The positive relationship between liquidity and standard deviations is most likely an econometric phenomenon due to infrequent trading and cannot be interpreted as declining total risk as liquidity declines.

The illiquid portfolio outperformance is relatively consistent over time (whereby the portfolio outperforms the market over 55% of the months or 67% of the years under examination) but relatively inconsistent with respect to performance within the portfolio (whereby only 54.5% of shares have positive excess returns over any given month, the second lowest portfolio with respect to this statistic), indicating that overall performance is being driven by a relatively small group of shares that have shown disproportionately good performance in any given month. A high number of shares would thus be necessary to keep portfolio performance consistent.

Surprisingly the range of returns is relatively low for illiquid shares. This is both due to a low maximum return, as well as a relatively high minimum return. These characteristics may be the result of the investor that less liquid securities attract; a longer optimal holding period, less myopic, fundamental investors and fewer traders who enjoy the freedom of switching in and out of positions at a moment’s notice. If the characteristics of the investors are driving the stock price behaviour across the liquidity spectrum, then there may be room for a segmented market analysis of share returns. This is beyond the scope of this study but the author believes it offers an interesting avenue for future research.

Table 6a. Liquidity sorted portfolio statistics

Portfolio	Liquid	2	3	4	Illiquid
Monthly excess return	0.34%	0.30%	0.35%	0.57%	0.87%
Standard deviation	6.02%	5.47%	5.21%	4.82%	4.58%
Minimum	-29.72%	-32.19%	-26.86%	-20.82%	-18.04%
Maximum	15.38%	10.36%	11.15%	11.34%	10.80%
Sharpe ratio	0.06	0.05	0.07	0.12	0.19
Number of shares	62.07	59.67	62.87	62.27	62.67
% of days not traded	0.92%	2.03%	8.67%	21.32%	42.34%
% that remain in portfolio next year	68.20%	56.78%	47.67%	47.06%	50.90%
% that remain in portfolio for next 3 years	42.65%	17.36%	8.39%	9.47%	8.35%
% of winners per portfolio per year	53.35%	55.23%	55.11%	56.49%	54.50%
% months outperformed market	47.22%	44.44%	50.56%	53.89%	55.00%
% of J203 value 2011	66.42%	65.05%	73.31%	114.10%	198.37%
Range of returns	45.10%	42.55%	38.01%	32.16%	28.84%
% years outperformed market	46.67%	33.33%	40.00%	46.67%	66.67%

Table 6b. Liquidity sorted regression results

			Liquid	2	3	4	Illiquid	B - S
Adjusted Betas	CAPM	Market	0.81	0.74	0.66	0.68	0.66	-0.15
	FF3F	HML	-0.11	-0.07	-0.01	-0.01	0.19	0.30
		SMB	0.03	0.06	0.17	0.24	0.29	0.26
		Market	0.80	0.74	0.70	0.73	0.77	-0.03
	Liq-Aug CAPM	Market	0.72	0.69	0.65	0.69	0.80	0.08
		Liq	-0.32	-0.17	-0.04	0.06	0.46	0.78
Adjusted Beta's P-Values	CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.00
	FF3F	HML	0.10	0.27	0.86	0.93	0.37	0.02
		SMB	0.53	0.20	0.00	0.00	0.05	0.03
		Market	0.00	0.00	0.00	0.00	0.00	0.00
	Liq-Aug CAPM	Market	0.00	0.00	0.00	0.00	0.00	0.00
		Liq	0.00	0.00	0.60	0.50	0.01	0.00
Adjusted Alphas %	CAPM		-0.13%	-0.13%	-0.03%	0.18%	0.49%	0.61%
	FF3F		0.41%	0.34%	0.35%	0.57%	0.75%	0.34%
	Liq-Aug CAPM		0.13%	0.01%	0.00%	0.13%	0.13%	0.00%
Adjusted Alpha's P-Values	CAPM		0.63	0.60	0.91	0.51	0.13	0.05
	FF3F		0.41	0.44	0.37	0.12	0.02	0.44
	Liq-Aug CAPM		0.60	0.98	1.00	0.63	0.69	1.00
Adjusted R squared	CAPM		0.66	0.65	0.49	0.41	0.11	0.18
	FF3F		0.66	0.65	0.50	0.44	0.10	0.22
	Liq-Aug CAPM		0.72	0.67	0.49	0.41	0.14	0.48

Table 6b shows the regression outputs of the three risk models under examination. Market risk does not seem to show any consistent discernible pattern across liquidity quintiles using the three models, except using the CAPM model whereby the most liquid quintiles appear to have higher levels of market risk. This is promising as it demonstrates that the Beta adjustment procedure of Fowler and Rorke (1982) has removed at least large-scale downward bias from the illiquid portfolios. There does seem to be a slight relationship between liquidity and market risk.

It should be noted that all of the portfolios from Tables 3b, 4b, 5b, and 6b have average market risk that is less than 1, which seems unlikely. This is only possible due to the equal weightings of the portfolios versus the value weighting of the market, but it demonstrates that smaller firms (which will have a higher weighting using equal weighting) will have lower than unity Betas. More accurately, given the high levels of concentration in the market, very high market capitalisation shares have, on average, higher than unity Betas. Value weighting the index then arrives at unity due to a small number of shares with very high Betas, whereas equal weighting the index arrives at less than unity Beta estimates.

The SMB risk factor does seem to capture the some liquidity risk. This can be seen by the rising SMB coefficients as you move from the most liquid to least liquid quintiles, the coefficient is significant at the 5% level for the three least liquid portfolios. The HML risk factor also shows a negative relationship to liquidity, increasing consistently from a negative factor of -0.11 to 0.19 from the most liquid to least liquid quintile. Furthermore, although the HML factor is insignificant for each of the quintiles, SMB and HML are significant at the 5% level for the B–S portfolio. This may indicate that size and value combine to capture liquidity risk. The Liq risk variable is significant at the 1% level for the two most liquid and the least liquid quintiles.

The ability of the three models to capture the return associated with liquidity risk differs quite substantially. Although FF3F model factors of SMB and HML did seem to capture some liquidity risk, the model is unable to explain the returns of the least liquid quintile. This is shown by the economically significant 0.75% excess monthly return which is statistically significant at the 5% level (p-value = 0.02). In fact the FF3F model fares even worse than the ordinary CAPM in capturing liquidity risk. The CAPM does not show any significant alphas for each of the quintiles, however it is unable to fully capture the excess abnormal return associated with the B –

S portfolio. Here the alpha using the CAPM is economically significant at 0.61% per month and statistically significant at the 5% level. The Liq-CAPM is found to be superior to both the FF3F as well as the ordinary CAPM in explaining liquidity risk. The model consistently finds the lowest (in absolute value) alphas, all of which are insignificant at the 10% level. Furthermore the Liq-CAPM shows the greatest explanatory power for three out of five quintiles using the adjusted R-squared.

Overall the liquidity augmented CAPM of Liu (2006) performs relatively well compared to its most prominent competition. Specifically, the two-factor model performs the best out of the three models when sorting portfolios according to Beta (albeit very poorly) or liquidity; all three models are adequate to explain size, as a premium doesn't seem to exist; the only pre-sorting variable in which the two-factor model could be viewed as inferior is B/M, in this case the CAPM better explains the returns in each portfolio as evidenced by the smallest and least significant alpha in portfolio four, however the CAPM does show a significant alpha for the B – S portfolio where the two-factor model does not.

6 Conclusion

Section 6 summarises the problem statement and findings as well as outlining the study's limitations and recommendations for future research.

This study can be summarised in two sections: the establishment of liquidity as a priced risk on the JSE that is independent of value and size, and the testing of the Liquidity-Augmented CAPM of Liu (2006) which incorporates this risk factor into a traditional CAPM.

The establishment of liquidity as a priced risk factor was achieved through an annual pre-sorting of shares into quintiles according to a primary sort based on their number of zero daily trades over the past 12 months, and then to a secondary sort according to their turnover. This sorting procedure created almost monotically increasing excess returns from the most liquid to least liquid quintiles. The two most liquid quintiles were found to be relatively indistinguishable demonstrating that the liquidity measure proposed by Liu (2006) may be less sensitive to changes in the liquidity of more liquid shares. Overall, however, the excess return characteristics seemed to demonstrate that the pre-sorting mechanism did, in fact, proxy for some risk factor.

This sample period was then divided into pre- and post- global financial crisis years and, for the most part, the less liquid quintiles outperformed during the period of low risk aversion and under-performed during the period of heightened risk-aversion. The exception was the least liquid quintile, which demonstrated more stability than the other quintiles. This may be a consequence of the liquidity characteristics of a share attracting longer holding period investors, as per Amihud and Mendelson (1986), which induces a price delay to changes in underlying value. Apart from the least liquid quintile, the return characteristics in the two sub-periods are in line with a priced risk factor.

Finally, to check whether the cross-section of returns associated with illiquid shares is independent of the well-documented size and value effects, each share's returns are adjusted for size and value. This is done by subtracting the returns associated with each share's particular size and value quintile from the share's return. The cross-sectional returns of pre-sorting shares according to liquidity are robust to adjustments made for size and value, both separately as well as jointly.

This study then analysed the power of the Liquidity-Augmented CAPM of Liu (2006) in explaining the cross-section of JSE-listed securities pre-sorted according to size, B/M, Beta, and liquidity. The results illustrated an improvement over the FF3F model as well as the traditional CAPM. The superiority of the model was most evident when pre-sorting according to liquidity. The FF3F and the CAPM showed highly significant intercept values for the illiquid portfolio and the B–S portfolio respectively, while the Liq-CAPM generated no noticeable differences in intercept values across the portfolios and none of any statistical significance. The pre-sorting characteristics of size and value, from which the FF3F model's risk factors are derived, yielded surprising results. First, small firms seemed to show no real outperformance over the sample period on an absolute returns basis. In fact, using any of the three models showed that higher market capitalisation firms outperformed smaller market capitalisation firms on a risk adjusted basis. However, these results were statistically insignificant. It also appeared that although there was no premium associated with smaller shares, SMB as a risk factor did seem to capture some of the cross-sectional returns. This may indicate that although size is a risk proxy, it is not a priced risk. B/M sorted firms presented a cross-section of variation of absolute returns that we would expect from the results of previous literature. However, the abnormal performance seemed to concentrate in the fourth highest B/M quintile rather than the highest B/M quintile which was unusual. This quintile's outperformance was unable to be fully explained by any of the three models at a 10% level of significance. Surprisingly, the FF3F model demonstrated the most economically and statistically significant intercept terms: demonstrating that, even though the two-factor model of Liu (2006) may not fully explain the risk which is proxied by B/M, it explains the cross-section of returns at least as well, if not better, than the FF3F model.

The most interesting result of this study is undoubtedly the inexplicable outperformance of low Beta shares. This result demonstrated not only a failure of the CAPM, but also the FF3F model and the Liq-CAPM in explaining the performance of shares pre-sorted according to their sensitivity to the overall market. The relationship found was in direct contradiction of a solely mean-variance optimising investor population. The monotonic nature of the inverse relationship between Beta and returns suggested that there may have been an unaccounted for risk associated with insensitivity to market performance. Evidence that these findings are unlikely to have been driven by outliers was demonstrated by the fact that the low Beta portfolio outperformed the market more often than any other pre-sorting characteristic, and consisted of the greatest percentage of shares that increased in value per year of any portfolio. This showed both the time series and internal consistency of the anomalous finding.

There are a number of limitations to this research, most of which are due to data availability. For instance, the availability of daily data would have allowed the risk coefficients to be estimated using the Trade-to-Trade model of Marsh (1979), which is shown to have minimal bias in the first two moments of the distribution of the coefficient estimates. This would allow a more inclusive sample with a greater range of trading frequency. An additional limitation due to data availability is the comparison against a wider variety of liquidity measures such as the bid-ask spread and the price impact of trade. An analysis of Liu's liquidity measure against a variety of alternative measures would provide justification for a particular liquidity measure rather than simply a justification for the inclusion of liquidity.

Apart from data availability constraints, a number of additional limitations to this research exist. Primarily, the research would benefit from testing the two-factor model using a greater variety of pre-sorting characteristics. This could include various other proxies for value such as price to earnings and cash flow to price, as well as other measures for documented anomalies such as price momentum and long term reversal.

The aforementioned limitations provide interesting areas for future research. Additionally, the abnormal returns found in low pre-ranking Beta portfolios present a persistent and unexplained anomaly that warrants further investigation. This anomaly presents an interesting opportunity for academics and practitioners alike due to its peculiar performance characteristics. Specifically, the low Beta portfolio was found to have the lowest risk, highest absolute return, and greatest level of overall consistency throughout the sample period. None of the three risk models examined in this study could explain the returns in either of the two lowest Beta portfolios.

7 Reference list

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31-56.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid ask spread. *Journal of Financial Economics*, *17*(2) 223-249.

Armitage, S., & Bizeszczynski, J. (2011). Heteroscedasticity and interval effects in estimating beta, UK evidence. *University of Edinburgh Business School, unpublished*, 1-37.

Atkins, A., & Dyl, E. (1997). Transactions costs and holding periods for common stocks. *The Journal of Finance*, *52*(1), 309-325.

Auret, C., & Sinclaire, R. (2006). Book-to-market ratio and returns on the JSE. *Investment Analysts Journal*, *63*(1), 31-38.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3-18.

Basiewicz, P., & Auret, C. (2009). Another look at the cross-section of average returns on the JSE. *Investment Analysts Journal*, *69*(1),23-38.

Basiewicz, P., & Auret, C. (2010). Feasibility of the Fama and French three factor model in explaining returns on the JSE. *Investment Analysts Journal, 71*(1), 13-25.

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks. *Journal of Financial Economics*, *12*(1), 129-156.

Bhandari, L. (1988). Debt/equity ratio and expected common stock. *Journal of Finance, 43*(2), 507-528.

Bhardwaj, R., & Brooks, L. (1992). The January anomaly: effects of low share price, transaction costs, and bid-ask bias. *Journal of Finance, 47*(2), 553-575.

Bowie, D. C., & Bradfield, D. J. (1993). A review of systematic risk estimation on the JSE. *De Ratione*, *7*(1), 6-22.

Cohen, K., Hawawini, G., Maier, S., Schwartz, R., & Whitcomb, D. (1983). Friction in the trading process and the estimation of systematic risk. (North-Holland, Ed.) *Journal of Financial Economics*, *12*(2), 263-278.

Constantinides, G. M. (1986). Capital market equilibrium with transaction costs. *The Journal of Political Economy*, *94*(4), 842-862.

Datar, V., Naik, N., & Radcliffe, R. (1998). Liquidity and asset returns: an alternative test. *Journal of Financial Markets*, *1*(2) 203-220.

DeBondt, W. F., & Thaler, R. (1985). Does the stock market overreact? *Journal of Finance, 40*(3), 793-805.

Demsetz, H. (1968). The cost of transacting. *The Quarterly Journal of Economics*, *82*(1), 33-53.

Dimson, E., & Marsh, P. (1983). The stability of UK risk measures and the probem of thin trading. *The Journal of Finance, 38*(3), 753-783.

Fama, E., & French, K. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*(1), 3-25.

Fama, E., & French, K. (1992). The cross-section of expected stock returns. *The Journal of Finance, 48*(1), 26-32.

Fama, E., & French, K. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Fama, E., & French, K. (1996). Multifactor explanations to asset pricing anomalies. *Journal of Finance*, *51*(1), 55-84.

Fama, E., & French, K. (2004). The Capital Asset Pricing Model: theory and evidence. *The Journal of Economic Perspectives, 18*(3), 25-46.

Fischer, L. (1966). Some new stock market indexes. *Journal of Business, 39*(1), 191-225.

Fowler, D., & Rorke, C. (1983). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, *12*(2), 279-283.

Hou, K., & Moskowitz, T. (2005). Market frictions, price delay, and the cross-section of expected returns. *The Review of Financial Studies, 18*(3), 981-1020.

Jagadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, *48*(1), 65-91.

Jensen, M. (1968). The performance of mutual funds in the period 1945-1964. *Journal of Finance, 23*(2), 389-416.

Kruger, P., & van Rensburg, P. (2008). Evaluating and constructing equity benchmarks in the South African portfolio management context. *Investment Analysts Journal, 67*(1), 5-18.

Lintner, J. (1965). Security prices, risk and the maximal gains from diversification. *The Journal of Finance, 20*(1), 587-616.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13-37.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics, 82*(1), 631-671.

Markowitz, H. (1959). *Portfolio Selection: efficient diversification of investments.* New York: Yale University Press.

Marsh, P. (1979). Equity rights issues and the efficiency of the UK stock market. *The Journal of Finance, 34*(1), 839-862.

McClelland, D., Auret, C., & Wright, T. (forthcoming). Thin-trading and Beta estimation: results from a simulated environment. *Journal for Studies in Economics and Econometrics*.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768-783.

Page, D., Britten, J., & Auret, C. (2013). Momentum and liquidity on the Johannesburg Stock Exchange. *International Journal of Economics and Finance Studies*, *5*(1), 56-73.

Pastor, L., & Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3) 642-685.

Reinganum, M. (1981). Misspecification of capital asset pricing: empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics, 9*(1), 19-46.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9-16.

Scholes, M., & Williams, J. (1977). Estimating beta from nonsynchronous data. *Journal of Financial Economics, 5*(3), 309-327.

Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442.

Stattman, D. (1980). *Book values and stock returns.* Chicago, IL: University of Chicago.

Stoll, H., & Whaley, R. (1983). Transaction costs and the small firm effect. *Journal of Financial Economics, 12*(1), 57-79.

Van Rensburg, P., & Robertson, M. (2003). Size, price-to-earnings and beta on the JSE Securities Exchange. *Investment Analysts Journal*, *58*(1), 1-11.

Van Rensburg, P., & Robertson, M. (2003). Style characteristics and the cross-section of JSE returns. *Investment Analysts Journal*, *57*(1), 7-15.

Van Rensburg, P., & Slaney, K. (1997). Market segmentation on the Johannesburg Stock Exchange. *Journal for Studies in Economics and Econometrics*, *23*(3), 1-23.