

Market Betas on the JSE:
Factor Selection, Estimation and Empirical Evaluation

James Laird-Smith

LRDJAM002

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In full fulfilment of the requirements for the degree

Master of Commerce specialising in Finance

Faculty of Commerce

UNIVERSITY OF CAPE TOWN

5th June 2017

Supervisor: Associate Professor Kanshukan Rajaratnam

Department of Finance and Tax

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

PLAGIARISM DECLARATION

COMPULSORY DECLARATION:

1. This dissertation has been submitted to the Turnitin module (or equivalent similarity and originality checking software) and I confirm that my supervisor has seen my report and any concerns revealed by such have been resolved with my supervisor.
2. I certify that I have received Ethics approval (if applicable) from the Commerce Ethics Committee
3. This work has not been previously submitted in whole, or in part, for the award of any degree in this or any other university. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, and has been cited and referenced.

Signed by candidate

Abstract

This paper examines the nature and significance of market betas on the Johannesburg Stock Exchange (JSE).

The identity of market betas is determined by means of Principal Component Analysis (PCA) performed on the returns of the *FTSE/JSE Africa Index Series*. A scree test shows two factors necessary for inclusion in the appropriate Arbitrage Pricing Theory (APT) model. Based on the promax rotated factor loadings, it is argued that the *Financials* (J580) and *Basic Materials* (J510) indices ought to be used as the appropriate observable index proxies for the first and second factors respectively.

Regarding the estimation of beta, this paper makes the case for the use of Reduced Major Axis (RMA) regression over the traditional Ordinary Least Squares (OLS) approach. A number of characteristics are assessed when arriving at this conclusion. Importantly, it is shown that the traditional OLS regression method chronically underestimates the magnitude of the beta parameter whereas RMA regression does not. In addition, it is shown that, while OLS beta values are more stable in absolute terms than RMA beta values, the RMA values are more stable when adjusted for their magnitude.

This paper does not make use of a thin trading filter to narrow the sample of stocks for empirical evaluation. Instead, an examination is made of the significance of beta values at the point at which they are estimated. This is accomplished by means of a rolling window of regressions. It is shown that, while most stocks do exhibit betas which are consistently significant over their listing period, many stocks do not. Some stock returns result in almost no significant beta values while some others exhibit beta values which are significant for only a portion of their listing period. It is shown that a median beta p-value value of 5% is an appropriate 'significance filter' for limiting the sample of stocks to only those significant for the majority of their listing period.

Using only these stocks, an empirical evaluation of beta is conducted using portfolios sorted on both OLS and RMA beta values. It is found that neither beta measure explains the cross-section of returns in the case of resource stocks. However, in the case of non-resource stocks the results show a clear divergence between the methods. In the case of OLS sorted portfolios, the results show a negative relationship between beta and returns. This surprising and counterintuitive result has also been arrived at by other researchers and is the opposite of what the APT would predict. However, in the case of RMA sorted portfolios, this pattern reverses itself, showing a positive relationship between beta and returns. For some holding periods, this is shown to be significant, providing evidence in support of the APT. As a result it is demonstrated that OLS regression not only underestimates the magnitude of beta, but that it distorts the results of empirical tests. On this basis it is argued that RMA regression ought to replace OLS regression as the preferred method of beta estimation for the JSE.

Contents

1	Introduction	1
2	Literature Review	4
2.1	Background	4
2.1.1	The Capital Asset Pricing Model (CAPM)	5
2.1.2	Arbitrage Pricing Theory (APT)	8
2.2	Factor Selection	9
2.2.1	Approaches	10
2.2.2	Application on the Johannesburg Stock Exchange (JSE)	11
2.3	Beta Estimation	15
2.3.1	Measurement Considerations	15
2.3.2	Correcting for Bias Estimates & Thin Trading	16
2.3.3	Estimation Method	17
2.4	Empirical Evaluation	21
2.4.1	Regression-Based Methods	21
2.4.2	Portfolio Sort Method	23
3	Sample Selection	26
3.1	Sample Period	26
3.2	Return Interval	27
3.3	Listing Period & Thin Trading	27
3.3.1	Index Sample Selection	28
3.3.2	Equity Sample Selection	29
4	Methodology and Results	32
4.1	Factor Selection	32
4.1.1	Determining the Appropriate Number of Factors	33
4.1.2	Selecting the Appropriate Index Proxies	33
4.1.3	Rolling Window of Eigenvalues	39
4.2	Beta Estimation	41
4.2.1	Number of Regressors	42
4.2.2	Significance of Beta Estimates	45
4.2.3	Magnitude of Beta Estimates	52
4.2.4	Stability of Beta Estimates	54
4.3	Empirical Evaluation	57
4.3.1	Resource Stocks	58
4.3.2	Non-Resource Stocks	60
5	Conclusion	63
A	Loading Plots for Alternative Factoring Methods	66

List of Tables

1	Breakdown of Sample Selection	28
2	Loadings on All-Share & ICB Industry Classified Indices	37
3	Correlation Coefficients Between Selected Indices and <i>ArcelorMittal SA</i>	44
4	Comparison of Beta Stability	56
5	Results for Portfolios Formed on Resource (RESI) Stock Betas	59
6	Results for Portfolios Formed on Non-Resource (FINI) Stock Betas	61

List of Figures

1	Illustrative Least Triangles Fitting Line	19
2	Distribution of Shares Based on Thin Trading	30
3	Scree Plot of Eigenvalues	34
4	Loadings Plot of Indices	35
5	Rolling Window of Eigenvalues	40
6	Illustrative OLS and RMA Regression Lines	43
7	Illustrative Index and Stock Volatility Chart	44
8	Loadings Plot of Equities	46
9	Distribution of Stocks Based on Median P-Value of Betas	49
10	Distribution of Stocks Based on Significance of Betas	50
11	Tradeoff Between Beta Median P-Value and Proportion of Significant Betas .	51
12	Magnitude of Beta Values on the JSE	53
13	Median Beta Values on the JSE Over Time	55
14	Loadings Plot of Indices Using Minimum Residual Factoring	67
15	Loadings Plot of Indices Using Generalised Weighted Least Squares (GLS) Factoring	68
16	Loadings Plot of Indices Using Principle Factor Analysis	69
17	Loadings Plot of Indices Using Maximum Likelihood Factor Analysis	70
18	Loadings Plot of Indices Using Principal Components Analysis and a Varimax Rotation.	71

1. Introduction

The beta (β) measure occupies a central place in empirical finance. It refers to the strength of the linear relationship between the rate of return of a particular investment and the rate of return of the market as a whole. From its position within the Capital Asset Pricing Model (CAPM) of Treynor (1961*a,b*), Sharpe (1963), Lintner (1965*a,b*), Mossin (1966) and Black et al. (1972), beta has enjoyed wide recognition and prominence. In later research, beta took on a broader role as part of the more flexible Arbitrage Pricing Theory (APT) of Ross (1976), where a number of betas can represent the possible ‘priced’ factors which are operational in a particular market.

The more flexible nature of the APT has led to useful application in the context of the Johannesburg Stock Exchange (JSE). The JSE is different from many exchanges in having a large proportion of its market capitalisation attributable to mining and resource firms. The most relevant research in this area is conducted by Van Rensburg and Slaney (1997) and Van Rensburg (2002), who show that there exists a dichotomy in the returns generating process between resource related and non-resource related stocks. This is referred to as the ‘market segmentation’ problem because it suggests the presence of two separate markets, each with their own Security Market Lines. On this basis the authors argue for the use of a two factor APT model utilising appropriate index proxies for resource related and non-resource related risk and return for application on the JSE.

More recent analysis performed by Kruger (2005) and Laird-Smith et al. (2016) raises the question as to whether the APT factors identified by Van Rensburg (2002) continue to be appropriate for application on the JSE. The most obvious change required is to adjust for the next major reclassification which accompanied the replacement of the JSE Actuaries indices with the new joint venture FTSE/JSE Africa Index Series in 2002. In addition, both Kruger (2005) and Laird-Smith et al. (2016) show the emergence of a dual exposure for the previously selected *Financial and Industrials* index, rendering it undesirable as a proxy. This leads Laird-Smith et al. (2016) to select the *SA Financials* index and the *SA Resources* index as the most appropriate proxies for the first and second major factors respectively. However, values for these indices only begin at March 2006, which is used as the starting date of the analysis. This shorter sample period indicates the research is incomplete. Further analysis is required to determine whether these findings are sample specific or too heavily influenced by periods of financial crisis such as took place during a 2007/2008. There is also a practical limitation to using index proxies only available

from March 2006; in order to calculate beta values before this date, an updated analysis is required to select proxies which are live for the longer period under consideration.

Despite high expectations, beta has produced poor results in empirical tests. This is especially true on the JSE where beta has not only performed poorly, but produced highly counterintuitive results. The most relevant of these empirical tests are those performed by Van Rensburg and Robertson (2003) and Strugnell et al. (2011), who conduct portfolio-based studies using betas estimated from OLS regression. Van Rensburg and Robertson (2003) show beta to have, if anything, a significant *inverse* relationship with stock returns. In a larger study Strugnell et al. (2011) also show beta to have a significantly inverse relationship with stock returns, although this finding loses its significance when utilising a Dimson (1979) correction to account for the presence of thin trading. Because the CAPM and APT would predict a significantly *positive* relationship between beta and stock returns, these results are significantly at odds with the established theory.

There are a number of possible explanations for beta's poor empirical performance. Unlike some other firm characteristics, beta is a highly dynamic measure and can change markedly depending on the decisions taken when estimating it. This paper will make the case for an alternative beta regression method. Instead of the usual Ordinary Least Squares (OLS) regression approach, this paper argues for the use of Reduced Major Axis (RMA) regression as proposed by Camp and Eubank (1981). In doing so, the various advantages of RMA regression are highlighted. Chiefly, it will be shown that OLS regression chronically underestimates the magnitude of beta. Correspondingly it will be argued that RMA regression produces a more accurate measure of the relationship between the returns of an investment and the returns of the market and is thus less likely to underestimate the magnitude of the beta parameter. Examples of this have been shown by Tofallis (2008) using a selection of stocks comprising the *Dow Jones Industrial Average* and by Laird-Smith et al. (2016) using stocks comprising the JSE *Top40* index. If beta values are being underestimated, there is reason to suspect that they will result in incorrect portfolio sorts when evaluated and therefore incorrect empirical results. Alternative beta estimation methods are not considered by either Van Rensburg and Robertson (2003) or Strugnell et al. (2011), except for the Dimson (1979) corrections applied to OLS betas to account for thin trading.

Another explanation of beta's poor empirical performance is to do with the significance of the actual beta values being estimated. It is not clear from the work of Van Rensburg and Robertson (2003) and Strugnell et al. (2011) whether the beta values they evaluate were

considered significant or not significant at the point at which they were estimated. The significance tests performed all take place at the *evaluation* stage of the analysis, where the expected returns-beta relationship is tested. This leaves open the possibility that beta *is* a significant predictor of stock returns, but only for a subset of stocks (those for which it can reliably be estimated). The only way of determining whether this is the case, is if the beta values of the various stocks are tested for significance at the point at which they are *estimated*. Thereafter, if necessary, the sample can be filtered to include only those stocks with significant beta values. While Van Rensburg and Robertson (2003) do make use of a thin trading filter, which limits the sample of stocks, it isn't clear from the analysis whether this would exclude all the stocks which did not produce significant beta values.

In summary, the topics examined in this paper fit into three broad categories. They are:

1. *Factor Selection*, which investigates the identity of beta by examining the common sources of market variation. In doing so, it is established what factors ought form the basis of the appropriate APT model.
2. *Beta Estimation*, which addresses questions of how beta values ought be calculated, specifically those regarding the appropriate regression method and the significance of beta estimates.
3. *Empirical Evaluation*, which uses statistical tests to establish whether beta has any power in predicting stock returns.

These topics will be used as the subsections of both the **literature review** (which will follow immediately below in **Section 2**) and the **methodology and results** (**Section 4**). The details and considerations of the **sample selection** are discussed in **Section 3**.

2. Literature Review

Asset pricing spans an extensive and often interrelated range of subject areas. As mentioned, in order to assess the relevant literature, this review will follow the progression as laid out in Section 1. The aspects relevant to the **factor selection** will be discussed in **Section 2.2**, then the aspects related to **beta estimation** will be discussed in **Section 2.3**. Lastly, the resultant **empirical evaluation** of the various beta measures will be discussed in **Section 2.4**. Before this however it is important and instructive to provide an overview of the asset pricing landscape and **background** to the various asset pricing models (**Section 2.1**).

As may be expected, the amount of research in this field does not allow for an exhaustive examination of all the aspects of the literature. Readers seeking a more detailed assessment, as well as for references consulted more broadly for the purpose of this review, are directed towards surveys from Campbell (2000), Cochrane (2011) and Goyal (2012).

2.1. Background

Arguably the central investigation of empirical finance is that of how to accurately price capital assets which are traded in the market. In order to do so, researchers and academics are tasked with uncovering why it is that some assets are able to generate different rates of return to others. The most famous example of this comes in the form of the equity premium; where common stocks listed on an exchange are able to garner a higher rate of return than safer asset classes such as treasury bills¹. While the theory underpinning this research has applications across investments types, the presence of the equity premium has led to the focus of the empirical research being on those higher return investments such as common stocks.

As is the case with all empirical fields, models are created in an attempt to explain the existence of stylised facts. These are findings in the data that are so widely observed and robustly tested within a given context that they are accepted as being true. The most basic of these facts, as illustrated by the equity premium, is the observation that risk taken on by investors needs to be compensated with higher returns. In order to be compelling, financial models must provide more than a statistical analysis of stylised facts, they need to provide the theoretic underpinnings which ground those facts in reality. A unique feature of finance is the ample amount of data available to practitioners. In addition to the regulatory and disclosure requirements that are often in place for financial institutions, the nature of

¹The first observation and coining of the ‘equity premium puzzle’ is by Mehra and Prescott (1985).

financial markets is such that it generates a large amount of data as part of its everyday functioning. This enables us to understand why rates of return differ between assets and asset classes (Goyal, 2012).

Another somewhat unique feature of finance is the degree to which uncertainty and risk play a part in not only the empirical evaluation of the models (as happens in many disciplines) but also in the theoretic foundations of those models. In order to understand the nature of the equity premium, one must begin with the observation that, while common stocks garner a higher rate of return in comparison to most well-secured debt, this kind of investment comes with a correspondingly larger degree of risk. The fact that all investors face risk in their investments means that any valuation model must take into account the effect of this risk on the resulting market price and, correspondingly, the return on the asset. However, this tradeoff between risk and expected return is not direct. The models that have proven most capable are those with a nuanced approach to the risk-return tradeoff and diversification (Goyal, 2012).

Academic work undertaken at the formation of the profession was geared towards understanding the effect of diversification on the pricing of assets. This work culminated in two of the most famous and celebrated financial models already mentioned: the **Capital Asset Pricing Model (CAPM)** and **Arbitrage Theory of Capital Asset Pricing (APT)**². The success and debate surrounding these and related diversification models has meant that literature in this area is now designated as ‘Portfolio Theory’, with the emphasis placed on risk and return of investments as part of a larger holding of assets, each with their own risk and return profile. The central tenant of these models is that the presence of diversification leads the joint behaviour of these assets to differ from the behaviour of their constituent parts.

2.1.1. The Capital Asset Pricing Model (CAPM)

The CAPM builds upon the work of Markowitz (1952), who pioneers the mean-variance framework, the elements of which refer to measures of return and risk respectively. In doing so, Treynor (1961*a,b*), Sharpe (1963), Lintner (1965*a,b*), Mossin (1966) and Black et al. (1972) variously postulate expected returns on any particular asset to be dependent on only a single factor: sensitivity to the the market portfolio. The logic that underlies this conclusion made the CAPM a powerful and intuitive model from which the finance profession was able to build. If investments can be bought and held in combination with

²Commonly referred to as ‘Arbitrage Pricing Theory’ from where the acronym is derived.

one another, some of their movements will work against or hedge one another, sacrificing some return, but also mitigating risk in the process. Portfolios can thus optimise return for various levels of risk based on historical data. This is referred to (under the mean-variance framework) as the efficient frontier. Given that the efficient frontier represents the highest level of return for a given level of risk, all investors are incentivised towards some point upon it, with differences emerging only depending on the investor's risk appetite. The addition of a risk-free asset creates a larger opportunity set, one that begins at the axis representing zero risk. If investors are able to trade the risk-free asset, they are incentivised hold it along with one of the riskier portfolios up to the point where they can garner the highest expected return at the lowest possible level risk. This manifests in investors combining the risk-free asset with the optimal risky portfolio to create a new efficient frontier. The portfolio which optimises in this regard is that which lies at the tangency between the zero-risk asset on the axis and the efficient frontier of portfolios. Given that no other portfolio meets this optimality condition, it is often referred to as the 'market portfolio': the combination of assets that ought always be utilised when crafting the risky portion of one's investments. Because it shows the optimal way that capital should be prioritised, the tangency line from the risk-free asset to the market portfolio is referred to as the 'Capital Allocation Line'. An important conclusion of this is, given the presence of a risk-free asset, the use of the market portfolio (as the portion of assets which carry risk) ought be uniform across investors. Unlike in the case of only risky portfolios (where the point at which an investor gravitated was dependent on their own individual risk appetite) the presence of a risk-free asset means that all but one of those portfolios (the market portfolio) are rendered mean-variance inefficient.

As a result of this framework, the authors of the CAPM are able to delineate between two types of risk that investors are capable of taking on: unsystematic risk, which is specific to a particular asset and can resultantly be diversified away, and systematic risk which is not tied to any particular asset and thus cannot be diversified away. The mechanism they identify follows logically from the statements above; with the ability to diversify one's portfolio and the availability of a risk-free asset, investors are always incentivised to diversify away unsystematic risk and, as a result, will only hold the market portfolio. Consequently, the value of any given asset is determined by its value as part of a well-diversified (market) portfolio. Given that investors are only interested in taking on risk for which they are compensated, the systematic risk of an asset (under the CAPM) ought be the singular determinant of its value. The fact that different assets will exhibit different sensitivities to the market means that the measure of that market sensitivity is correspond-

ingly the measure of the value of the asset by way of the return it is able to generate. This sensitivity parameter is called ‘beta’ and was first referred to this way by Sharpe (1963).

Beta is typically estimated as the result of a regression technique³ and represented with the greek letter β when used in mathematical expressions. These regressions ordinarily rely on a broad market index to proxy for the market portfolio. In the United States this would ordinarily be the *Standard & Poor’s (S&P) 500* index. For a time, this link between beta and the CAPM meant the two were considered synonymous. This notion is re-enforced with the beta measure used as a test for the CAPM. If higher beta stocks do not provide a higher level of return than lower beta stocks, it is indicative of beta (the market) not being the motivating force behind stock returns and thus evidence against the CAPM. The same is true in reverse; if higher beta stocks do provide a higher level of return than low beta stocks, it is evidence in favour of the CAPM. The result is that testing the betas produced by the CAPM is in fact testing of the model itself.

The CAPM is typically shown in one of the following two forms, stemming from Sharpe’s (1963) original usage. Either as:

$$R_i = \alpha + \beta (R_m) \tag{1}$$

or as:

$$R_i - r_f = \alpha + \beta (R_m - r_f) \tag{2}$$

In both cases (1 and 2) where,

R_i are the returns on a particular investment (i), usually in percentage form,

α is the intercept term of the fitted line,

β is the slope of the fitted line and

R_m are the returns on the market.

And in the second case (2) where,

r_f is the rate of return available on a risk-free asset.

Although the CAPM provided a powerful and intuitive framework for asset pricing,

³As previously indicated, the pertinent aspects of beta estimation are included in Section 2.3.

subsequent debate and poor results in empirical testing meant the CAPM was never able to rise to the expectations placed upon it. A major theoretical criticism levelled against the CAPM was proffered by Roll (1977), who argues that the mean-variance efficiency of the market portfolio is in fact the only testable assumption of the CAPM. Further, given that the composition of the market portfolio is not identifiable in reality, it follows that the model itself is fundamentally untestable. This applies equally to the standard practice of adopting a market index to proxy for the market portfolio; the index may be mean-variance efficient whereas the market portfolio may not be. In the South African case, where the market portfolio is usually proxied for by the JSE All Share Index (ALSI), the performance of the CAPM is equally poor. In addition to the theoretic shortcomings, Van Rensburg (2002) shows the ALSI to be mean-variance inefficient when taking into account the opportunity for offshore investments. This essentially invalidates the ALSI for its typical application as a proxy for the South African market.

The CAPM was initially expected to be a hugely influential model. However, these theoretical criticisms, together with the poor empirical results already mentioned, meant that it was never able to rise to the expectations placed upon it. It did however lay the groundwork for what became another influential school of thought in the form of the Ross (1976) APT.

2.1.2. Arbitrage Pricing Theory (APT)

While the CAPM has been shown to have shortcomings, the distinction between systematic and unsystematic risk has continued to inform asset pricing frameworks in other spheres. A notable instance of this is in the APT. It is generally agreed upon that investors are compensated for holding only systematic risk⁴. As a result, contemporary debate is centred mainly on what constitutes systematic risk. The APT provides a more general framework than the CAPM, acknowledging and accounting for the existence of a number of factors that underlie the returns generating process. This has a strong intuitive basis; a firm's profitability may be subject to a number of forces. For a mining company, for example, these may include factors such as the price of the commodity they are selling or the risks related to their labour relations. Drawing upon the distinction between systematic and unsystematic risk, the factors that cannot be diversified away will be 'priced' into the cost of the asset.

APT models take the form (Ross, 1976; Van Rensburg, 1997):

⁴See Goyal (2012) for a survey.

$$R_i = E(R_i) + \sum_{k=1}^k \beta_{ik} f_k \quad (3)$$

where,

R_i are the realised returns on a particular investment (i),
 $E(R_i)$ is the expected rate of return on a particular asset (i) at the beginning of the period,
 f_k is the k th risk factor that impacts on asset i 's returns and
 β_{ik} is the coefficient that measures the sensitivity of realised returns of investment (i) to movements in factor k .

Like the CAPM, APT models are estimated from various regression techniques. The returns on various stocks are regressed on the returns generated from each of the the systematic risk factors under investigation. As a result, following the same naming convention as the CAPM, an asset will have a number of 'betas' to represent its sensitivity to these priced factors rather than the single 'beta' envisioned by the CAPM.

Some forms of the APT will appear to be very similar to the CAPM. This is the result of a similar naming convention and emphasis on broad and non-diversifiable sources of risk and return. As is made clear by Sharpe (1984, 23), the CAPM is not only reconcilable with the APT, it can in fact be seen as a restricted version of the APT (albeit with additional assumptions). Importantly however, the APT is more flexible; it is not encumbered like the CAPM which relies on the existence of a market portfolio. The APT has less restrictive assumptions⁵ by instead postulating the existence of a number of systematic factors driving risk and return.

2.2. Factor Selection

After having established that a multi-factor approach is possible, the question naturally arises as to how to identify and select the appropriate factors for the APT. Unlike in the case of the CAPM, the identity and economic underpinning of the priced factors is not pre-defined. As should be evident, the process undertaken to identify and subsequently select these factors will have a material impact on all the subsequent analysis. A misidentified or unidentified factor will result in a bias and/or entirely omitted set of beta estimates. These betas will in turn be used in empirical evaluation. As such, the method of factor selection

⁵A full list of these assumptions and differences is given by Copeland and Weston (1983).

constitutes *a priori question* in the field of asset pricing, one that must be answered in advance of further questions and upon which the validity of eventual testing will depend.

2.2.1. Approaches

Broadly speaking, there exist two approaches to determining the appropriate factors for inclusion in a Ross (1976) APT Model. The first is the **factor analytic approach** which makes use of factoring methods such as factor or principle component analysis to analytically derive the priced factors. The second is the **pre-specified variable approach** which entails examination of already existing macroeconomic variables as candidate factors and then evaluating the sensitivity of security returns to those factors. A useful explanation of the differences between the approaches is that the factor analytic approach begins with an examination of the data and uses the results of that examination to infer the governing fundamentals whereas the pre-specified variables approach begins with variables that represent underlying fundamentals and tests those variables for statistical viability.

The advantage offered by the factor analytic approach, as given by Ross (1976) who first suggests it, lies in the usefulness of the systems equations methodology which are part of the factor analysis procedure. This approach is tested by Roll and Ross (1980), Chen (1983) and more extensively by Lehmann and Modest (1988).

The main weakness of the factor analytic approach (and where there has been the most extensive criticism) is regarding the analytically derived factors. It is argued that these factors may not be identifiable or interpretable and, as a result, may not have a substantive link to any macroeconomic phenomenon (McElroy and Burmeister, 1988). The related problem is that such an approach allows researchers to ‘mine’ or ‘trawl’ for factors on the basis of only spurious correlation. When expressed in this way, the corresponding benefit of the pre-specified variable approach becomes clear; the use of already existing macroeconomic variables will better ensure economic substance. A further weakness of the factor analytic procedure is that the analytically derived factor scores will, by their nature, be standardised to a mean of zero and a unitary variance; this makes any inferences that rely on the magnitude of these changes dependent on essentially arbitrary transformations of these factor scores (Francis, 1986).

The work of Chen et al. (1986) allows for a successful reconciliation of these techniques. By utilising the systems equations methodology in the factor analytic approach at the out-

set of the process and subsequently seeking out proxies for macroeconomic variables, the best of both methods can be achieved. The correct utilisation of identified proxies remains the result of the systems equation methodology, meaning its usefulness is retained at the exploratory stage of the analysis. In addition, the understandable concerns regarding the spurious or uninterpretable economic underpinning of the factors is alleviated. Finally, the proxies themselves (unlike the analytically generated factors) allow for parameter estimates which are not standardised to a mean of zero and a unitary variance.

2.2.2. *Application on the Johannesburg Stock Exchange (JSE)*

As mentioned, the more flexible nature of the APT has resulted in useful applications on the JSE. Under the CAPM, the most obvious candidate for the market proxy would be the ALSI, which represents 99% of the JSE's market capitalisation. However, as mentioned, the relevant literature has established that there exists at least one additional source of risk and return in the form of resource-related risk. Historically, mining and resource stocks have constituted a sizeable portion of the JSE's market capitalisation. In addition, the literature has shown the movements of resource related stocks to be suitably distinct from that of the non-resource related stocks.

Research into the changing behaviour of mining stocks on the JSE was first undertaken by Campbell (1979), who observed that, despite a fall in the beta calculated for the *JSE Industrial Index* during the period from the mid-1960s to the mid-1970s, the beta calculated on the *JSE Gold Index* in fact rose over the same period⁶. This behaviour was the same for the constituent shares of those indices. The betas for these shares, however, when regressed against their respective market indices, did not demonstrate the same kind of change. These observations lead Campbell (1979) to conclude that the shares from the different sectors had their own risk profile and corresponding Security Market Line (SML).

Various research findings followed from this observation. Gilbertson and Goldberg (1981), in a somewhat limited study involving only three shares, utilised a two index linear factor model with the *JSE Mining* and *Industrial* indices used as regressors. This comparison showed the two factor model to have superior explanatory power when evaluated using the R^2 statistic. In the case of one of the shares, the beta estimates from both the regressors were shown to be statistically significant, indicating the possibility of stocks with dual exposures that would not be captured by a single factor model.

⁶Both calculations regressing on the ALSI as the market proxy.

While this analysis from Gilbertson and Goldberg (1981) is revealing, it does not fully address the question of factor selection. Although the candidate variables are shown to be statistically significant, this significance is not compared against that of other candidate variables. The question is left open as to whether any one proxy of, for example, resources risk (such as the mining index) was superior to any other proxy (such as a commodity price or index). Further research undertaken by Page (1986) works to alleviate this concern with the application of principle component analysis (PCA) to a range of shares and portfolios of JSE returns for the period 1973-1982. In doing so, Page (1986) shows two of the three⁷ analytically generated risk factors to be associated with a statistically significant risk premium. Furthermore, it is again shown that these remaining risk factors have a higher explanatory power than their CAPM equivalents. A varimax rotation is then used to uncover the identity of the underlying factors. The varimax rotation is part of a family of possible rotations which are ‘orthogonal’ in nature. Orthogonal rotations maintain the uncorrelated nature of the underlying factors (Kaiser, 1958). Even under this relatively strict criteria, Page (1986) shows there to be two major factors: one comprising of mainly mining shares and the other comprising of mainly industrial shares.

The work done by Page (1986) is however weakened by its reliance on the analytically generated factors, the shortcomings of which are highlighted in Section 2.2.1. This has lead Van Rensburg and Slaney (1997) and Van Rensburg (2002) to utilise the approach of, amongst others, Chen et al. (1986) to reconcile the factor analytic approach with the pre-specified variables approach in the manner described in Section 2.2.1. In both of these papers, observable macroeconomic proxies for the analytically generated factors are explicitly sought. Van Rensburg and Slaney (1997) utilise a number of permutations of the factor analysis procedure, performing principle components, principle factor analysis and maximum likelihood analysis on both the correlation and covariance matrix of JSE stock returns. Both promax⁸ and varimax techniques were then employed for the purpose of rotation. The findings of the factor selection procedure in this case showed the *All Gold Index* and the *Industrial Index* to be the most appropriate factors for inclusion in the APT at the time.

Another important element of Van Rensburg and Slaney (1997)⁹ is the testing of the single and two factor models by means of a primary and secondary regressions. This procedure

⁷As determined by a scree test.

⁸The promax rotation, unlike in the case of Page (1986), does not force the factors to remain uncorrelated. For this reason, it is sometimes referred to as an ‘oblique’ rotation.

⁹Later updated in Van Rensburg (2002).

regresses the returns of the various sectoral indices under consideration in the model in order to find their average explanatory power as measured by R^2 . These results indicate the two factor model to have superior explanatory power. Next, the residuals of the primary regressions are regressed on the alternative model in order to determine whether a material amount of variation remaining after the first regression can be explained by the competing model. The results of these secondary regressions show that, whereas the single factor model is capable of explaining only a negligible amount of the residuals of the two-factor model, the two factor model explained almost 17%¹⁰ of the residuals of the single factor model. This acts as strong evidence for the fact that the models' residual errors are cross-correlated with one another and that the selected factors in the two factor model are capable of explaining this variation whereas the single factor model cannot. The presence of cross-correlated residuals is a serious weakness of the single factor model as it means the model violates the diagonally assumption (Sharpe, 1963) and results in downwardly bias beta values.

Van Rensburg (2002) updates and recasts the procedures of Van Rensburg and Slaney (1997). A number of events give impetus to this re-examination. Firstly, the reclassification of the JSE Indices as of March 2000 created a sometimes dramatic change to the composition of the various indices involved in the factor selection methodology. These changes naturally raise the question as to the appropriate proxy selection thereafter. Secondly, the fall of the gold sector as a proportion of the overall JSE capitalisation raises suspicion as to whether the variant of the previous *Gold Index* would remain a suitable proxy for resource risk. Lastly, the changing make-up of the South African economy and financial landscape (such as the large growth in the financial sector) creates the need to re-evaluate the nature of the market proxies more generally. The findings show these considerations to be well-founded. With the factor selection methodology remaining largely the same, the JSE *Financial and Industrial Index* (FINDI) and *Resources Index* (RESI) are identified as the most appropriate candidate factors after a promax rotation.

A number of other items of research are included in both Van Rensburg and Slaney (1997) and Van Rensburg (2002). The first is a series of Wald tests, which are performed in the case of the two factor model to indicate whether there exists a significant difference between the two estimated coefficients. If it can be shown that there is a material difference between these coefficients, it is a further demonstration of the dichotomy in the generating process between the two effects. The findings show this to be the case for 70 of the 84 shares

¹⁰When this procedure was update in Van Rensburg (2002) the results were similarly significant.

sampled¹¹. Another notable addition of Van Rensburg (2002) interrogates the mean-variance efficiency of the *JSE All Share Index* (ALSI). The impetus for this examination was, at the time of writing, an increasing demand for investors in South Africa to gain offshore exposure and to hedge away currency risk. By constructing the mean-variance frontier with an allowance for offshore holdings (as proxied for by the *Dow Jones Industrial Index*) along with the relaxation of exchange controls, it was demonstrated that the ALSI was not an appropriate proxy for inclusion in a single index model (CAPM) on account of the fact that it did not lie on the efficient frontier.

Significant changes to both the South African economy and to the JSE have taken place since Van Rensburg (2002). The most obvious of these is the replacement of the *JSE Actuaries* indices with the new joint venture *FTSE/JSE Africa Index Series* which took place in 2002. This again creates a fundamental need for re-specification. Some of this work has been started by Kruger (2005) in the process of creating a framework for understanding the benchmark risk in South African equity portfolios. At the time, many of the indices were very new and so a blend of data was used drawing on the new FTSE Global Index Methodology and the historical Financial-Industrial and Resources sectors from *I-Net Bridge* going back to 1966. Factor analysis was then applied in a manner similar to Van Rensburg and Slaney (1997) and Van Rensburg (2002). In this case a varimax rotation was used. Due to the orthogonal nature of the varimax rotation, the results are slightly more difficult to interpret than if an oblique rotation had been used. Again resources dominate the second major factor. The loadings plot shows that a number of indices appear to be more appropriate proxies for the first factor than the previously selected FINDI index. The FINDI index appears to have slightly more of a dual exposure than it did previously, however this difference is not found to be statistically significant. On these grounds and to retain consistency, the FINDI is again selected to proxy for the first factor.

More recent analysis is conducted by Laird-Smith et al. (2016) using the same principle factor analysis as used in Van Rensburg and Slaney (1997), Van Rensburg (2002) and Kruger (2005). In this case a broader range of indices are used from *Thomson Reuters Datastream*. These indices include the more recently constructed SA Resources Index for which the earliest available data points are March 2006. The findings for resource shares remain the same. As was the case with Kruger (2005), the results show a dual loading on the part of the FINDI, which in this case lies somewhere in the middle of the axis representing the first factor. This finding may be partly due to the comparatively small sample size, which contained only 8

¹¹Passing the test at the 99% significance level. A number of additional shares passed at the 95% level.

years of monthly returns. This time period was chosen to maximise the number of indices which could be included in the sample. Because of this, the SA Resources Index and the SA Financials index, which have their earliest starting points March 2006, could be selected as appropriate index proxies. While this may be appropriate for studies using data from only 2006 onwards, the shorter listing periods of these indices will not allow for analysis going back any further. Any beta values that are to be calculated for any period before 2006 will require index proxies going back further, preferably for the entire period under consideration. In addition, the shorter time period of the study may indicate the results are only capturing a fairly recent phenomenon. A study of a larger sample period is needed to establish whether the apparent dual loading of the FINDI is material or not.

2.3. Beta Estimation

After having established the appropriate factors for inclusion, the next question that arises is how to appropriately go about estimating the beta coefficients. There are many aspects to this, ranging from practical considerations to broader questions over the nature of error in the market proxy. While most of the research in this area makes allowances for some of these considerations, they are seldom brought together and dealt with concurrently. As mentioned previously, these areas are often interrelated.

One of the main sources of bias results from a misspecified market proxy and is dealt with in the previous section. The remaining aspects with bearing on beta estimation are discussed as follows: practical details of beta measurement are discussed immediately below in **Section 2.3.1**. Techniques with the aim of correcting for bias in beta estimates are contained in **Section 2.3.2**. This includes those dedicated to dealing with the problem of thin trading. Finally the broader topic of what kind of estimation method is preferable is addressed in **Section 2.3.3**.

2.3.1. Measurement Considerations

Share prices are a time series of data points. In order to estimate relationships between them they need to be broken up into various time periods. The first and most obvious instance of this occurs in the length of the return interval. Share price and index returns can be taken over daily, weekly or monthly periods. There is good reason to suspect that intervals of any longer or shorter time periods than this will either be too long (and so aggregate away meaningful relationships) or too short (and as a result mask changes which occur over a longer time). Returns are typically percentage changes between the closing

value measured at the end of the period and the opening value of that period (Bradfield, 2003).

After having established the return interval for which returns are to be measured, the question that naturally follows is how long the estimation period should be. Again, as with the return interval, there is a balancing act required when trying to develop an accurate understanding of share price returns. The estimation period ought be long enough so as not to fall victim to short term phenomena or spurious correlation. Correspondingly however, the estimation period cannot be so long so as to aggregate away substantive economic realities. Many aspects of both the economy and the financial system will change over very long periods of time. Periods that are too long may thus not capture critical details. The obvious additional advantage of the longer period is having a larger number of data points from which to make inferences; this is especially true in the case with longer return intervals, where data points are fewer.

Bradfield (2003) gives a well-rounded account of these dynamics and how they may be applied in the South African market. Blume (1975), Eubank and Zumwalt (1979), Corhay (1992) and Bradfield (2003) all highlight the apparent consensus that has developed over the use of monthly intervals when examining excess returns. This trend has carried over into the contemporary literature as a means to balance the competing concerns of having a sufficiently long period for which returns are able to accumulate, but no so long so as to lose details of share price movements. In a similar vein, Bradfield (2003), along with Gonedes (1973) and Kim (1993), highlights the related consensus that five years ought be the preferred estimation period. Given the fact that monthly returns will yield only 12 data points for a year's worth of returns, a longer estimation period may be necessary to estimate beta correctly while at the same time noting that periods longer than this may will fail to capture the reality of the market in the present, a large number of factors having changed over, say, a ten year period. This trend has also continued in the contemporary literature, notably in the work of Strugnell (2010) and Strugnell et al. (2011).

2.3.2. Correcting for Bias Estimates & Thin Trading

In the international literature, Scholes and Williams (1977) highlight the problem of non-synchronous trading, where stocks that are traded less frequently will exhibit a kind of measurement error¹². These stocks will appear to be serially correlated, not on account of actual trading patterns, but on account of not being traded in the first place. The

¹²Details of other kinds of measurement error will be given in Section 2.3.3.

problem of shares being traded less frequently is commonly referred to as ‘thin trading’. Unsurprisingly, Scholes and Williams (1977) show the effect of this to be most egregious when estimating betas from daily data.

A possible solution to this problem, used in some of the literature, is to omit those securities that were not sufficiently traded. While this approach is tempting, it introduces the possibility that conclusions drawn from the analysis may only apply to shares that are sufficiently well traded. This is also a problem if one is going to investigate the existence of a liquidity premium or size effect alongside beta, which will necessitate examination of stocks that are traded with less frequency (Strugnell et al., 2011).

The problem of thin trading is amplified in the case of the JSE, where the amount of trading is often far less than that of larger exchanges such as the London Stock Exchange¹³. In order to overcome this problem, Strugnell et al. (2011) utilise two related corrections. The first is suggested by Scholes and Williams (1977), who submit that an unbiased and consistent beta estimate can be achieved by introducing a lag effect and thereby smoothing beta over the preceding, contemporaneous and succeeding period’s excess returns. This is further expanded by Dimson (1979) who argues that a further lag should be introduced and that the extent of the lag ought be proportional to the extent of thin trading of the particular security. In the case of the JSE as analysed by Strugnell et al. (2011), this can amount to a lag of up to five months.

2.3.3. Estimation Method

Ordinary Least Squares (OLS) regression is the most commonly used method for beta estimation. This is also true in the South African literature; although in some cases beta is subjected to a correction to account for thin trading, the beta estimates being corrected originate from OLS regressions. This was the case in Strugnell et al. (2011) as mentioned above. However, some researchers question whether this use of OLS regression is appropriate. In particular, Tofallis (2008), in combination with the earlier work of Camp and Eubank (1981), makes the case for a different estimation method. Their contention is that ‘Reduced Major Axis’ (RMA¹⁴) regression ought replace OLS as the preferred

¹³This difference in exchange liquidity was illustrated during the process of incorporating JSE stocks into the FTSE Global Index Classification. In doing so, the FTSE standard index liquidity guidelines had to be relaxed somewhat so that they may include JSE shares with lower market capitalisations (Kruger, 2005).

¹⁴The acronym can generate confusion as there also exists ‘Ranged Major Axis’ regression, which is different. To avoid this, some researchers employing Reduced Major Axis also sometimes use its other name, which is ‘Standardised Major Axis’ (SMA).

method of beta estimation. RMA regression uses the *geometric mean functional relationship* between the variables of a regression to find the fitting line. There exist a number of both historical and contemporary naming conventions for this estimation method. Currently the most widespread naming convention is ‘Total Beta’ (Damodaran, 2002; Laird-Smith et al., 2016) when referring to the beta values themselves. However, this paper will refer to these beta values as ‘RMA betas’. This is to make clear the differences between them and ‘OLS betas’ with which they are contrasted.

There are various characteristics of RMA regression which make its application appealing for the purpose of beta estimation. RMA regression belongs to a family of estimation methods broadly known as *error-in-variables* models or *measurement error* models. The appeal of error-in-variables models is that they allow for the possibility of error in both the dependent and so-called independent variables of the regression. When applied to market models in finance, this amounts to allowing for the possibility of error in the returns generated from the market portfolio (or factor) as well as those generated from the investment or stock under consideration. In this respect, they harken back to one of the most serious and originating critiques of the CAPM: that being Roll’s Critique (1977). As noted previously, Roll’s Critique (1977) suggests that: because the market portfolio is specified to be the universe of all investable assets¹⁵ and that this universe of assets is necessarily unobservable, the CAPM will remain fundamentally untestable as a result.

This kind of error-in-variables problem can be addressed in a number of ways. Naturally, all error-in-variables models, by their nature, will account for error in the so-called independent variable in some way or another. However, not all error-in-variables models will accomplish the task satisfactorily; many of these models, for example, require knowledge of the underlying ratio between the error present in the independent variable and the error present in the dependent variable. The problem with this is that this ratio is often just as illusive as knowledge of the market portfolio itself, resulting in the same problems as identified in Roll’s Critique (1977). Contrastingly, however, RMA regression does not require knowledge of this kind. RMA will minimise error in both the variables of a regression. This is accomplished by minimising the sum of the areas of the triangles formed between each of the observations and the fitting line (Tofallis, 2008). Because of this fact, the line is often known as the ‘least triangles’ fitting line (Tofallis, 2008). This is illustrated by Figure 1.

¹⁵It has been noted that in order to even begin to encompass the universe of all investable assets, one would have to catalogue and measure more abstract investments such as those made in education and other human capital (Mayers, 1972).

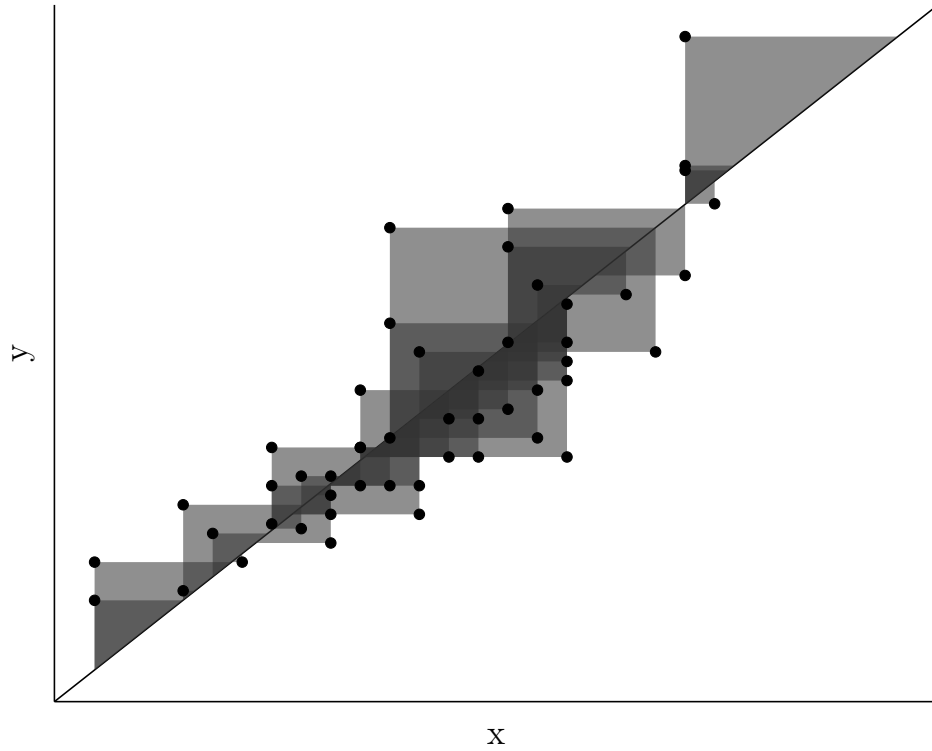


Fig. 1. Illustrative Least Triangles Fitting Line

Tofallis (2008) provides a number of additional reasons for the use of RMA regression. One of these is that the equation used to estimate it more closely aligns to the traditional understanding of beta among market participants. It is commonly asserted that stocks with beta values smaller than 1 are defensive in nature. It is correspondingly asserted that stocks with betas greater than 1 are more risky in their nature. If this is the case, it should be that beta is the *relative volatility* between investment returns and market returns. This identity is, in fact, precisely summarised in the RMA beta equation (Tofallis, 2008):

$$\beta^* = (\text{sign of } r) \frac{\sigma_i}{\sigma_m}, \quad (4)$$

where:

β^* is a parameter estimated from the Geometric Mean Functional Relationship,
 r is the correlation between the investment's returns and the market's returns,
 σ_i is the standard deviation of the investment's returns and
 σ_m is the standard deviation of the market's returns.

There is a further intuitive basis provided by Tofallis (2008) for the use of RMA regression

over OLS. This is to do with the treatment of the so-called ‘independent variable’. Under OLS regression the returns of the investment are acknowledged as containing error, however it is assumed that the returns of the market are measured without error. This is a very strict assumption to make and it is one that almost never reflects the reality of typical returns. Most returns used in asset pricing models are widely acknowledged as being accompanied by frequent excess volatility (Shiller, 1992). This is true of both the returns on the ‘market’ and those of the investment. If OLS regression is used under these conditions it will result in downwardly bias beta estimates. To show this, consider that two possible least squares regression lines can be found depending on which of the variables is treated as containing error. The usual OLS y on x regression line (which corresponds to the typical beta value) and the reverse OLS or x on y regression line which minimises the error in the other direction. Error-in-variables models such as RMA will produce fitting lines somewhere between these two extremes. It follows that, for positive beta values, typical OLS regression will underestimate the magnitude of beta. Tofallis (2008) demonstrates that this can be shown analytically; the formula for OLS beta (β) is:

$$\beta = r \frac{\sigma_i}{\sigma_m}, \quad (5)$$

and the formula for RMA beta is given in (4). It follows that β^* in terms of β is:

$$\beta^* = \frac{\beta}{r}. \quad (6)$$

Since r is always bound between -1 and 1 it follows that $|\beta^*|$ will always be larger than $|\beta|$. It also follows that the extent of this difference will depend on the strength of the correlation between the variables. Tofallis (2008) examines a selection of stocks which form part of the Dow Jones Industrial Average and shows the difference between the competing betas to be significant. In the South African market, Laird-Smith et al. (2016) perform a similar analysis using stocks that form part of the JSE Top40 index; again the differences are found to be significant.

Another advantage of RMA is scale invariance: if either variable is scaled either up or down by a constant term, the RMA regression line will scale exactly in proportion with that constant. The same is not true for some other regression lines. This means that the scale on which the variable is measured can have an impact on the actual regression estimates. Given that the market proxy is often an index or macroeconomic variable, whose returns are

determined on a different scaling system to that of an individual stock, this would appear to be a very important consideration. RMA regression is one of the regression techniques that is impervious to this kind of scaling (Tofallis, 2008).

While there have been other beta estimation methods used over the many studies in this field, very few methods have the advantages of RMA regression as have been highlighted above. Overall, this catalogue of advantages makes RMA an enormously compelling technique for beta estimation. At the same time, showcasing these advantages makes plain the shortcomings of OLS. Consequently, it also casts doubt on the tests of asset pricing models that have been conducted up to this point. If OLS beta estimates are downwardly biased on a systematic basis, it naturally raises the question as to whether rejection of beta by, for example Van Rensburg and Robertson (2003) and Strugnell et al. (2011), would have occurred if a more accurate estimator such as RMA had been used. To determine this, it is necessary to establish rolling windows of both OLS and RMA beta estimates and subject them to the same empirical tests.

2.4. Empirical Evaluation

While the above sections covering factor selection and beta estimation are necessary components of evaluating the CAPM/APT (beta), they are of course only preliminaries to the empirical testing. The CAPM/APT can only be valid if the expected return-beta relationship holds. Tests of this relationship generally fall into one of two categories: **regression-based methods (discussed in Section 2.4.1)** and those that use **portfolio sorts (discussed in Section 2.4.2)**.

2.4.1. Regression-Based Methods

The earliest and best understood empirical tests of beta are regressions performed on the the Security Market Line (SML), a technique which finds its root in the original CAPM literature. After having established the Security Characteristic Line (SCL) and estimated the relevant betas, the natural progression is to perform a second-pass regression of security returns using those betas as regressors. In the case of beta, this is akin to estimating and evaluating the SML. Early tests of this nature were performed by Lintner (1965*a*) and later by Miller and Scholes (1972). These tests do detect the expected returns-beta relationship, although estimate it as being far weaker than would be the case if the CAPM were true. Additionally, these tests reveal a substantial constant term, significantly at odds with the predications of the CAPM, indicating there is a substantial portion of returns unexplained

by the sensitivity to the market (beta).

At this point in the international literature, two significant concerns were raised regarding the CAPM which may provide an explanation for the model's poor empirical performance. The first was Roll's Critique (1977), which has been discussed in Section 2.3.3. The second is what as come to be known as the *error-in-variables* problem. Two related observations combine to make the error-in-variables problem a serious stumbling block for practitioners looking to utilise the CAPM. The beta measure is, in itself, an estimated quantity, meaning that it requires a specialised treatment when being incorporated as an explanatory variable in a regression. The framework of OLS regression is that, while the dependent variables are defined as having variation, the practitioner knows the level of the so-called independent variable. This is not the case with beta; because it is, in itself, estimated, it will have a distribution of its own that needs to be accounted for. This, in combination with the second observation (that stock returns are highly volatile in the first place, a theme described in the previous section), means that traditional applications of OLS regression are invalid for the purpose of evaluating the SML.

Some of the corrections proposed to the second-pass regression are discussed in the next section. Before doing so however, it is important to highlight that the problem of measurement error in beta (often referred to as the error-in-variables problem) is different from the concerns and discussions contained in Roll's Critique (1977) and in Section 2.3.3. The substance of Roll's Critique (1977) is that the market portfolio is fundamentally unobservable and subject to error. This problem is thus localised to the *estimation* stage of the process, whereas the error-in-variables problem occurs at the subsequent *evaluation* stage. The fact that these problems are so similar can be a source of confusion. The problems are so similar in fact that they may well have related solutions. In the previous section, the use of RMA regression is argued for as a more realistic and revealing tool for beta estimation on account of - among other advantages - being designed to estimate functional relationships between variables which are both measured with error. This advantage at the *estimation* stage of the analysis can be applied in a very similar way at the *evaluation* stage. This similarity between the problems also carries through the same consequences of using OLS regression. The nature of the error-in-variables problem is that OLS regression will tend to underestimate the strength of the linear relationship between the variables being considered.

2.4.2. Portfolio Sort Method

Even if an alternative regression method isn't used at the evaluation stage as suggested in Section 2.4.1, there is good reason to suspect that authoritative conclusions can still be drawn using traditional tests if other corrections are made. One of the most common ways of establishing whether beta is a significant driver of stock returns is to create a set of portfolios created from stocks, where the stocks have been sorted on the basis of their respective betas. Portfolios are formed by dividing the sorted stocks into groups, sometimes referred to as fractals. Within the portfolios, stocks can be equally weighted or weighted by their market capitalisation. The returns on these portfolios are then contrasted with one another by means of statistical tests. This technique was first utilised by Black et al. (1972) in response to the concerns of measurement error in the beta estimates. This concern is well established by Miller and Scholes (1972), who make use of simulated data (constructed to follow the SML relationship) and show how it fails to reveal the expected return-beta relationship even when designed to do so. These findings have meant that the portfolio sort method has become prominent in the contemporary literature, both in South Africa and internationally.

The strength of the portfolio sort method is also a shortcoming. By grouping stocks into portfolios, there will necessarily be fewer data points upon which to produce regression estimates. The advantageous aspect of the technique, it is contended, is that it will diversify away the firm-specific part of returns, thereby obtaining more accurate beta estimates. Statistically, this has the advantage of reducing variation; because stock returns are well understood to be highly variable, whereas portfolios of stocks have far less variance. In order to ensure the best of this trade-off, the portfolios used should have the widest possible distribution of betas; this is best accomplished by sorting the beta estimates by their magnitude, dividing them into fractals and forming the portfolios based on those groupings. Early tests of the CAPM internationally, such as those performed by Fama and MacBeth (1973) provided mixed results regarding the expected return-beta relationship, concluding that, although the tests failed to reject the CAPM hypothesis, they produced a SML that was too flat. These mixed results in respect of the CAPM/APT have persisted into the contemporary international literature¹⁶.

As mentioned at the outset of this paper, beta has also shown very poor empirical performance for stocks listed on the JSE. In particular, Van Rensburg and Robertson (2003)

¹⁶See Goyal (2012) for a survey.

show that there is, if anything, an inverse relationship between beta and stock returns. This surprising and counterintuitive finding raises further questions as to exactly what would give rise to a pattern significantly at odds with the APT. This work is given added importance on account of the fact that the two factor APT model of Van Rensburg and Slaney (1997) is utilised instead of the flawed single-index (CAPM) model. Although in this case resource shares were too few to provide meaningful sample, the remaining stock betas were regressed on the FINDI, which makes the findings robust to the market segmentation problems discussed in Section 2.2. Further and updated analysis is performed by Strugnell et al. (2011), in which a larger sample is taken. This is made possible by changes in the market and availability of information. Further, the Dimson (1979) Aggregated Coefficients method is applied at the estimation stage to correct for the problem of thin trading. In this case, the negative expected return-beta relationship persists, but is robbed of its statistical significance when the Dimson Aggregated Coefficients method is used for beta estimation. Nevertheless, even without this negative relationship, there still exists nothing by way of evidence in favour of beta being a significant (let alone singular) driver of stock returns on the JSE.

There exist a number of possible explanations for this poor empirical performance on the JSE. The first and most obvious is that the CAPM and two factor APT are flawed models in that they do not explain the underlying reality of the South African financial market. Another possible explanation of beta's poor performance is that the analysis undertaken in this area has not taken into account the market segmentation problem outlined in Van Rensburg and Slaney (1997), Van Rensburg (2002) and motivated for in Section 2.2.2. As indicated, if an incorrectly specified market proxy is utilised for the purpose of beta estimation, the resulting estimates will be bias downwards and the subsequent portfolio sort will be incorrect as a result. This may explain the findings of Strugnell et al. (2011), who uses the ALSI as a market proxy for all shares in the sample.

Another possible explanation for this poor performance is the use of the incorrect beta estimation method. As argued for by Tofallis (2008) and in Section 2.3.3, if beta estimation is performed without the use of some kind of error-in-variables technique, it will also result in downwardly bias beta estimates and an incorrect portfolio sort. While the South African literature has made some correction for beta bias to account for thin trading, it would still benefit from a explicit error-in-variables estimation method. When combined, these elements provide a powerful impetus for an update and re-examination of the various aspects of beta on the JSE. Whether those beta estimates are eventually tested using a

regression based method or with the traditional portfolio sort method, the results will only be compelling in the instance where the factors have been properly selected and betas have been correctly estimated.

3. Sample Selection

Two main datasets are used in the ensuing analysis. The first of these is the time series of closing prices generated by each of the indices that form part of the *FTSE/JSE Africa Index Series*. The second of these is the time series of closing prices generated by the shares comprising those indices. Cumulatively, this amount to all of the JSE indices and all of the shares which constitute them respectively. In the interest of drawing the most broadly representative sample of securities, all indices and their constituent shares were initially considered for inclusion in the sample set. Also in the interest of being as broad as possible at the outset, the time series of closing prices at the end of every trading day is extracted for each of the various equities and indices. The required data was obtained from the Bloomberg (2016) application program interface (API) and is extract from the terminal at the University of Cape Town main library.

3.1. Sample Period

Part of the objective of drawing a broadly representative sample is to draw the data from as long of a time period as possible. Having a smaller sample period will naturally limit the applicability of the analysis performed on that particular dataset. In doing this, however, practitioners should be cautious of the possibility that their findings may be time-varying. This has an acute importance in the case of the JSE where, historically, resource stocks have dominated the exchange but, as the South African market has developed, the non-resource stocks have grown to occupy a much larger position than they did previously (Van Rensburg, 2002). This leads to the possibility that conclusions arrived at using longer time periods may be describing realities present at only at a certain point in the sample period, rather than consistent phenomena.

Bloomberg's JSE index data begins at July 1995. While some of the equities have price histories on Bloomberg going back further than this, given that the majority of the analysis utilises the relationship between stocks and indices, July 1995 is also selected as the start date of the share data, establishing a synchronous time period across the datasets. This provides approximately 20 years of data spanning the two datasets. The length of this sample period is similar to those used in related studies of the JSE which have already been mentioned. There is also little danger of the sample being too long, however the conclusions arrived at will be tested to ensure they are not time-varying.

Rebasing of the indices is required to account for the changes in index structure that have oc-

curred in the past 20 years. As mentioned, the most consequential of these changes occurred in 2000 and 2002, however smaller changes have also occurred since then. Because Bloomberg provides the index data in rebased form, no further correction for this is necessary.

3.2. Return Interval

As mentioned in Section 3, the full set of daily closing prices for both indices and equities was extracted and available for analysis. Share and index returns can be calculated over any time interval as long as the data is available for that portion of the time series. Realistically however, there is a balancing act required when making such a determination as described in Section 2.3.1. Daily, weekly and monthly periods are the usual candidates for the computing of return intervals in the majority of the relevant literature. From the perspective of the JSE, a daily return interval is untenable on account of thin trading. The same is true, although to a lesser extent, of weekly return data (Strugnell, 2010). By contrast, return intervals of longer than one month will inevitably aggregate away meaningful relationships that are present. Even if this were not the case, the longer time interval also unnecessarily limits the number of data points available for the analysis.

As a result of this, a monthly return interval is chosen for the ensuing analysis. This decision is taken to balance out the competing concerns discussed immediately above and because monthly returns are (for the same reasons as outlined) the most widely utilised return interval in the relevant South African literature. This improves the chances that the findings will be more easily and reliably compared between related studies. The return generated by a share or index in a given month is calculated by taking the percentage change between each month's closing price and the closing price in the month immediately preceding it.

3.3. Listing Period & Thin Trading

The nature of stock market listing and index construction means that not all the time series of returns are available for the entire sample period. Especially with the selected longer time interval, only some indices and individual equities are likely to have a continuous listing for the approximately 20 years under consideration. Practically, this means that there are missing values in the dataset. Missing values present a problem for some of the statistical techniques applied in this paper; notably principle component analysis cannot be performed on a dataset with missing values. However, as will be made clear in the applicable sections, some of the analysis can be performed even with missing data. For this reason it is necessary to specify two layers of the sample selection in each case: one layer

allowing for the presence of missing values and one layer that does not. A summary of the sample selection is provided in Table 1, with explanations for indices and equities following in Sections 3.3.1 and 3.3.2 respectively.

Table 1: Breakdown of Sample Selection

	Indices	Shares
Sample initially considered	85	580
No value in period under consideration	(4)	(9)
Insufficient number of observations	-	(110)
Excluded on the basis of thin trading	-	-
Excluded as duplicate	(2)	-
Broad sample	79	461
Not live/listed for the entire period under consideration	(39)	(369)
Restricted sample	40	92

3.3.1. Index Sample Selection

By their nature, indices are broad market indicators. They serve this function by aggregating together securities in the market based on some criteria. This has a number of useful properties. Firstly, while indices do sacrifice some depth of data, this aggregation means they are less likely to have missing values. Secondly, as has been mentioned, it is useful as a tool for exploratory research where similar market sectors are often the criteria used to determine index membership. Last and relatedly, these indices, because of their nature, are often useful candidate factors for inclusion in the kinds of asset pricing models already mentioned in Section 2. When used in this way they ordinarily stand-in or proxy for some underlying market factor.

All JSE indices available on Bloomberg are initially considered for inclusion in the dataset. As mentioned, both small and large changes in the JSE index structure mean that not all indices are live for the entire period under review. Of the full set of 85 indices initially considered for inclusion, 4 are excluded on account of not being live for any portion of the period under review. 2 indices are excluded as being duplicates. None of the indices exhibit thin trading to the extent that they may be worthy of exclusion from the sample. Of those 79 remaining indices, 40 are listed for the entire sample period. Consequently, those 40 indices

will constitute the limited sample layer for indices, whereas the set of 79 will constitute the broader layer.

3.3.2. *Equity Sample Selection*

The full set of JSE equities available on Bloomberg is initially considered for inclusion in the sample set. This amounts to 580 stocks variously listed over the existence of the JSE. Shares are considered from both the *Main Board* and the *AltX*. 9 of these were not listed for any amount of time during the period under consideration and were thus excluded. This leaves 571 equities remaining in the sample. Additionally, not all of the remaining equities are useful for the analysis; some have not or were not listed for a sufficient period to produce a five year beta statistic. For this reason, equities with fewer than 60 observations are excluded from the dataset. This amounts to 110 equities and leaves 461 remaining in the sample.

Thin trading poses a much larger problem in the case of equities than it does in the case of indices. The time series of equity returns was inspected for the presence of zero (0) values. If zero values make up a sizeable proportion of the time series of an equity's returns, it is likely to lead to a distortionary effect on the analysis. One consequence of this is that zero values will naturally be correlated with one another, but not on the basis of any underlying market factor.

The extent of thin trading which takes place on a particular equity is best understood by means of a calculation. During time periods when an equity is not listed, the time series of returns contains an *NA* value. During times when the equity is listed, but not traded, the time series of returns will contain a zero (0) value. At all other times, the time series will contain the share's return. Thus, the extent of thin trading on an equity can be approximated by the number of zero values in the time series of returns divided by the number of non-NA values in the same time series of returns. In the case of monthly returns as used in this paper, this represents the proportion of months for which the stock was listed where it did not trade. This calculation is performed on the 461 shares remaining in the sample. The distribution of these proportions is shown in Figure 2.

For a stock to not trade in a given month is a strong indicator of thin trading. While it may be common for some of the smaller capitalisation stocks to not trade on a given day, it is far less likely that this will occur during a month and even less so over a number of months. As a result, when examining Figure 2, it makes sense that most stocks which

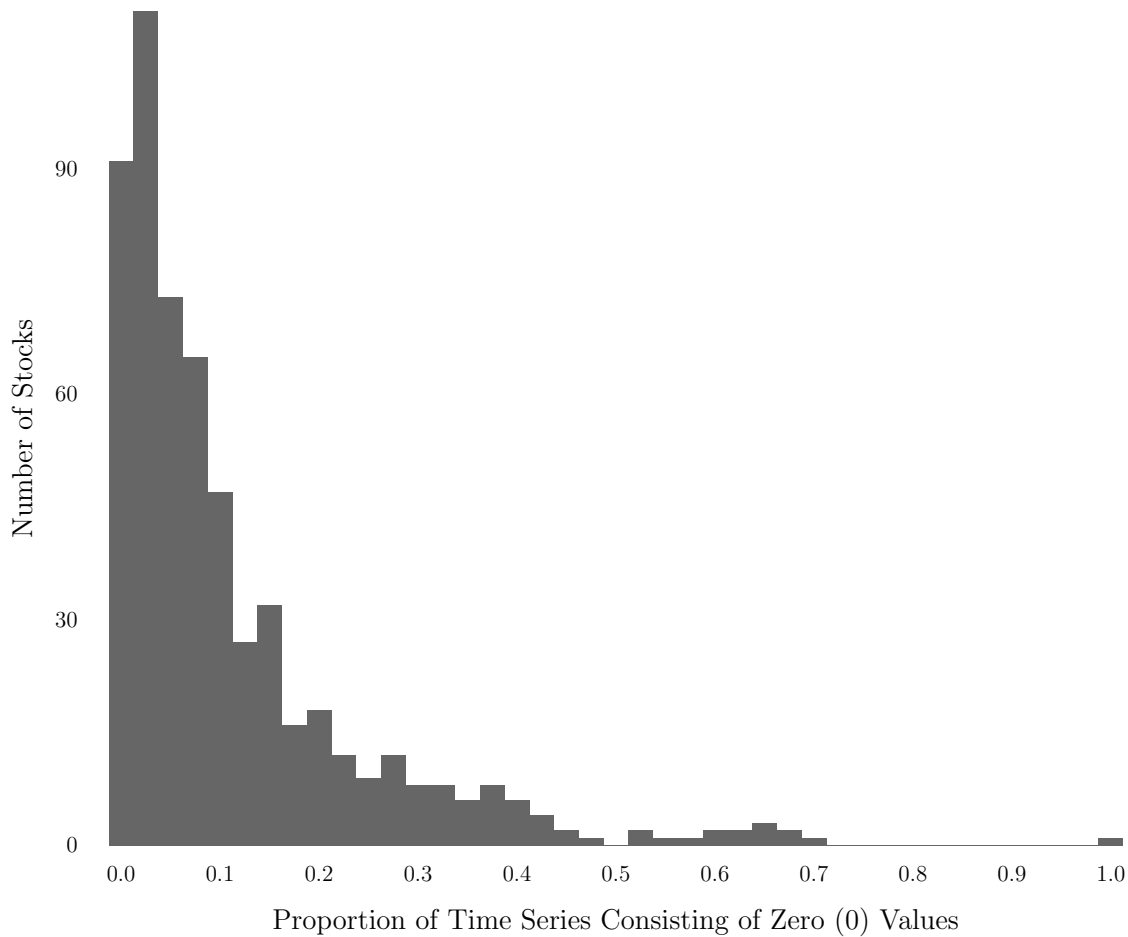


Fig. 2. Distribution of Shares Based on Thin Trading

remain in the sample have a low proportion of their time series of returns occupied by zero values. As can be seen, the majority of stocks are shown to have their proportion at a level less than 10%. Naturally it does not make sense to exclude such stocks on account of such a small amount of thin trading. To do so would severely limit the number of stocks available for analysis. Another reason for this is that, at such small values, the quality of the data can sometimes play a part, for example, showing a zero value when the stock was in fact delisted (Strugnell, 2010).

Regarding the other stocks in the sample however, it can be seen that thin trading does play a role. A sizeable number of stocks do show evidence of thin trading, with more than 10% of their values being zero. This raises the question of what to do with these stocks. Some of the prior literature, for example Van Rensburg and Robertson (2003), chooses to exclude these stocks. If they are to be excluded one would need to set a level at which to exclude some while leaving others. The level chosen for this risks being

arbitrary. In addition, this decision will limit the applicability of the analysis. Following the example of Strugnell et al. (2011), this paper does not make use of a thin trading filter. Instead, all the shares that remain in the sample are retained and the effect of thin trading will be examined and controlled for in the later analysis. Of these 461 equities remaining in the sample, 92 were listed for the entire sample period and thus will constitute the limited sample layer for equities. The set of 461 will constitute the broad layer.

4. Methodology and Results

The methodology and results will follow the same three broad areas as outlined in the literature review and at the outset of the paper. For consistency and ease of reading, they are sub-divided the same way below rather than having separate sections. The methodology and results of the **factor selection** will follow immediately below in **Section 4.1**. **Beta estimation** will follow thereafter in **Section 4.2**. Finally, the **empirical evaluation** is performed in **Section 4.3**.

Two programming languages are used in arriving at these results: *Python* and *R* (R Development Core Team, 2016). These are both mature platforms of data analysis and come with useful programming interfaces. For Python, the programming is performed in a series of *Jupyter* (formally *IPython*) notebooks. For R, the programming is performed in a series *R Notebooks* made available in the RStudio (2016) v1.0.136 Preview. Both of these tools provide a useful format for programming, especially for the kind of multi-step data analysis featured in this paper. Blocks of programming code can be executed separately and their output can be inspected within the notebook application.

Where possible, the analysis that is performed in one programming environment is repeated in the other for the purpose of validation. However there are cases where the breadth of statistical packages available in R are not available in Python. In these cases the functionality of the *rPy2* project allows for the importing of R functions and manipulating of R objects within the Python framework so long as there is a current version of R running alongside it. All of the charts and graphics displayed as part of the results are created in R using the *ggplot2* package (Wickham, 2009). Unless otherwise stated, results are rounded to two decimal places.

4.1. Factor Selection

As mentioned in Section 2.2.2, the research conducted by Van Rensburg and Slaney (1997), Van Rensburg (2002), Kruger (2005) and Laird-Smith et al. (2016) has formalised the approach to selecting factors for the JSE. In this case, the use of the restricted sample ensures that all candidate proxies are live for the entire period under consideration. It also removes the difficulties created by the presence of missing values, which are a problem for the factoring methods being employed.

In the interest of being comprehensive, a number of factoring methods were used for the

purpose of factor selection. These were: a) principle component analysis (PCA) b) minimum residual factoring c) generalised weighted least squares (GLS) factoring d) principal factor analysis and e) maximum likelihood factor analysis. The results arrived at are robust to whichever method was used and consequently only the results of the principal component analysis are shown here¹⁷. The loadings plots for the other factoring methods are included in Appendix A.

4.1.1. Determining the Appropriate Number of Factors

The decision of how many factors to include in the model is a balancing act between the parsimony offered by including fewer factors and the additional explanatory power offered by including a larger number. In order to determine this, the extracted factors are automatically ordered as part of the factoring methods on the basis of their explanatory power. The explanatory power of a particular factor is measured by its eigenvalue. The first 15 of these components are displayed as a scree plot in Figure 3.

As was the case in Van Rensburg and Slaney (1997), Van Rensburg (2002), Kruger (2005) and Laird-Smith et al. (2016), the explanatory power offered by the first two factors is larger than those that occur thereafter. General practice with this technique, beginning with Cattell and Jaspars (1967), is to extract factors at the point at which the scree plot flattens out. In this case, the third factor does not appear to offer any greater explanatory power than the fourth or subsequent factors. As such, the first two factors are selected to form part of the APT model. The continued prominence of the second factor continues to re-enforce the findings arrived at in the prior literature: that the CAPM or single index model is inappropriate for application to the JSE. The evidence continues to show that there exist, at minimum, two factors necessary for inclusion in the appropriate APT model.

4.1.2. Selecting the Appropriate Index Proxies

As mentioned, one of the salient criticisms of these kinds of factoring methods is that the factors themselves may not correspond to any underlying influence or reality. The danger therefore exists that the factors identified may simply be artefacts of statistical arrangement rather than any substantive economic reality. If this is the case it is unlikely that these factors will explain or predict returns into the future. It is for this reason, rather than employing the constructed factor scores, that an explicit index proxy is sought (Chen et al.,

¹⁷The term ‘factor’ and ‘component’ are used interchangeably throughout the explanations even though the reported results are for principle components.

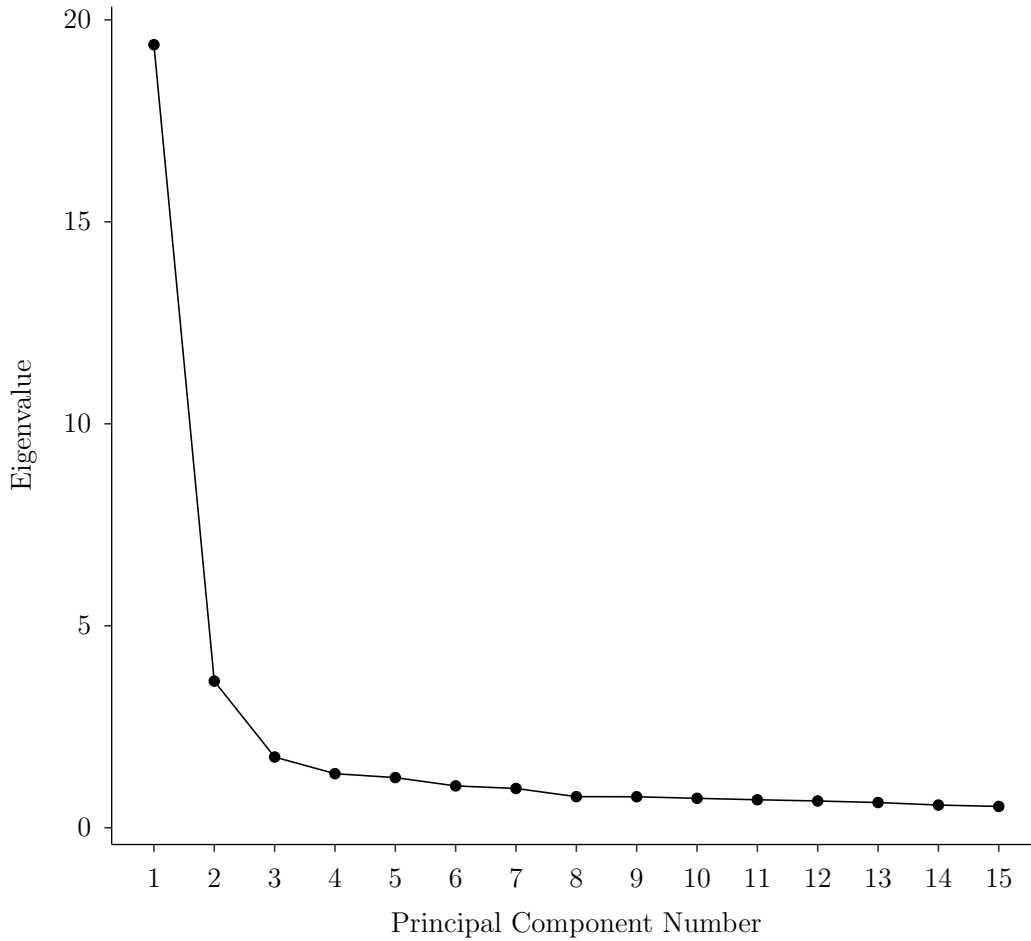


Fig. 3. Scree Plot of Eigenvalues

1986; Van Rensburg and Slaney, 1997; Van Rensburg, 2002).

As a separate consideration, it is widely observed by practitioners of this technique that the raw component loadings are difficult to interpret when trying to identify the relevant index proxies. This is as a result of the factoring methods themselves: factors are automatically calculated and ordered such that the first factor contains the maximum amount of variation. The most widely accepted approach to dealing with this is to employ a rotation of the matrix of factor loadings. As discussed previously, two rotation methods are available: orthogonal and non-orthogonal (or oblique) rotations. A varimax and promax rotation are performed on the loadings of the first two principal components.

The effect of the rotation is the similar in both cases, however the promax rotation is shown for illustration in Figure 4 as it allows for greater flexibility. This is because oblique

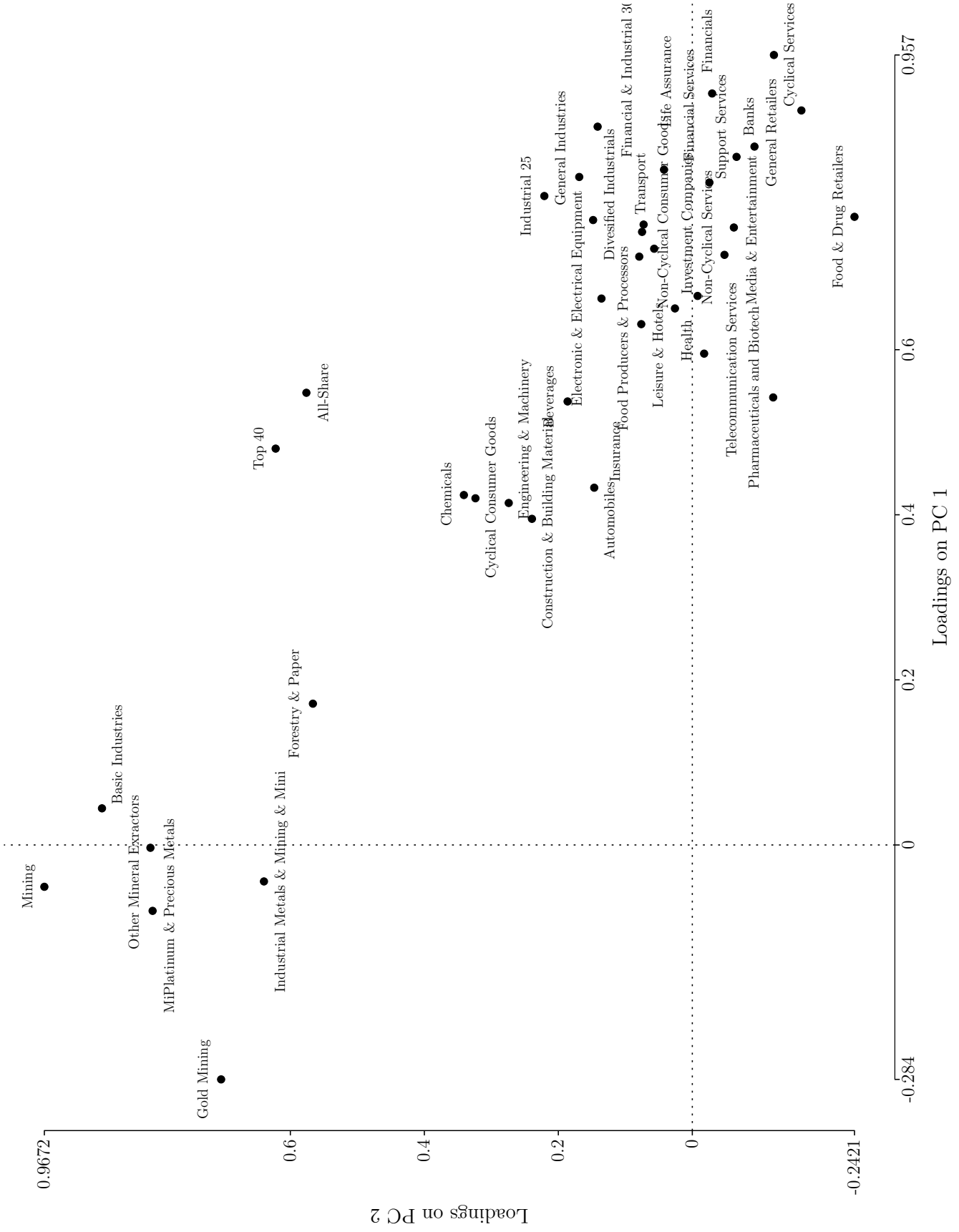


Fig. 4. Loadings Plot of Indices

rotations are not constrained into forcing the factors to remain uncorrelated (Hendrickson and White, 1964). A certain amount of correlation between factors will sometimes occur in the performance of the various indices (as will be shown in the next section). By allowing for the possibility that the factors are somewhat correlated, oblique rotations allow for a better decomposition of their returns. The varimax rotation is shown for comparison in Figure 18, included in Appendix A.

The loading plot shown in Figure 4 broadly resembles that of Van Rensburg and Slaney (1997), Van Rensburg (2002), Kruger (2005) and Laird-Smith et al. (2016). Again it is shown that there is a dichotomy in the returns generating process for shares listed on the JSE. Mining and resources related indices load highly on the second factor, while most other indices load highly onto the first factor. The All-Share and Top 40 indices, because they comprise of stocks in both sectors, have moderate loadings on both the first and second factors.

The most appropriate index proxy for these major factors is found using two criteria (Van Rensburg and Slaney, 1997; Van Rensburg, 2002). The first is that the index should have as close of a loading to 1 with the factor for which it is aiming to proxy. The second criteria is that the index should have a loading as close to 0 with the other major factor. These criteria aim to ensure that the chosen index is most likely to reflect the returns of the factor it is aiming to proxy for and is least likely to be influenced by returns in the other major factor. The index proxy that best meets these criteria can be inferred by the loading plot in Figure 4, but for a closer inspection of some elements the loadings of the All-Share and ICB industry-level indices are shown along with the number of their index members in Table 2.

It is useful to begin with the second largest of the two major factors because its identity is less ambiguous than the first. It is clear that the second major source of returns generation for the JSE is linked to the performance of resource related stocks. This finding is consistent with the prior research. In Van Rensburg and Slaney (1997) the second factor was proxied for by the *All Gold* index. When the analysis was repeated in Van Rensburg (2002) under a new index regime, the *Resources Index* (at the time JSE Index Code CI11) was found to be the most appropriate proxy. In the work of Kruger (2005), performed under the current index structure, the Resources index (changed to JSE code J000) was found to be the most appropriate. In the work of Laird-Smith et al. (2016) the *SA Resources* index¹⁸ was found most appropriate; however, as already mentioned, it cannot be used for longer term analysis.

¹⁸This combines the *Oil & Gas* as well as the *Basic Materials* indices.

Table 2: Loadings on All-Share & ICB Industry Classified Indices

Index Name	Index Members [†]	Loadings	
		Component 1	Component 2
All-Share	169	0.53	0.57
Basic Industries	27	0.00	0.90
Cyclical Consumer Goods	17	0.41	0.32
Cyclical Services	21	0.98	-0.15
Financials	62	0.93	-0.06
General Industries	27	0.81	0.15
Non-Cyclical Consumer Goods	7	0.76	0.05
Non-Cyclical Services	4	0.73	-0.07

[†] As at 17th March 2015.

In this updated analysis two index proxies appear most promising: the ICB industry-level *Basic Materials*¹⁹ index (JSE index code J510) and the more granular sector-level *Mining* index²⁰ (JSE index code J177). It is worthwhile to note that many of the candidate proxies are in close proximity to one another on the loadings plot. When this is the case the factoring method or rotation used in the analysis can impact which proxy falls closest to the loadings of 1 and 0 on the respective axis. While it appears as if the Basic Materials index better meets this criteria, a different factoring method and/or rotation would sometimes result in the Mining index being the appropriate proxy according to this criteria. The effect of using different factoring methods and rotations can be seen by examining their loading plots included in Appendix A. In these cases, the index proxy that offers the greatest economic interpretability should be selected. All of the constituent members of the Mining index are also included in the Basic Materials index on account of the fact that one is a sub-index of the other. Having an index that aggregates a greater number of equities together will serve as a better proxy. For this reason the Basic Materials index is selected as the appropriate proxy for the second factor²¹.

Before considering the identity of the first factor, it is useful to consider a further implication of the proximity of the various candidate proxies. As discussed immediately above, the fact that many of the acceptable proxies lie very close to one other in the loadings plot means

¹⁹Referred to as the *Basic Industries* index on Bloomberg, but *Basic Materials* in the JSE Documentation. It is represented by ICB industry code 1000.

²⁰The mining index is a ICB sector-level classified index with a code of 1770.

²¹The reasons for the differences between this selection and that of Kruger (2005) is to do with changes in index structure. The Mining index now falls under the Basic Materials index.

there may be a variety of acceptable APT models that could be employed by practitioners, depending on their needs. They may have some reason to make use of, for example, the Mining index instead of the Basic Materials index. For this reason it is more useful to state the APT model in slightly more general terms: instead of stating a definitive index proxy, the APT ought be stated descriptively by using a name for the factor rather than the name of the index. This also removes the need for a continual change in abbreviations. In the case of the resource-based factor, it can be specified in the same way as it is done in Van Rensburg (2002) and Kruger (2005) as ‘RESI’ where it represents one of the acceptably close index proxies. As mentioned, this paper will make use of the Basic Materials index as the RESI proxy.

The first factor is the more general *market* factor, the identity of which follows more closely with the non-mining elements of the South African economy. In the previous research conducted by Van Rensburg (2002), it has been proxied for by the *Financial and Industrial* index (JSE index code J213) or ‘FINDI’. However, the current pattern of loadings shows the FINDI to be slightly removed from the axis of the first factor. Unsurprisingly, the FINDI appears directly between the financials (which lie more directly on the axis) and the industrials, which do not. This shows an emergence of what may be a dual exposure for industrial shares. This possibility has already been shown in the work of Kruger (2005) and Laird-Smith et al. (2016) and renders the FINDI an undesirable proxy for the reasons already indicated. A working hypothesis for this change may be that the growing internationalisation of industrial companies on the JSE means their returns are more closely aligned with dollar-denominated resource prices, which are obviously closely bound up with the second factor. Although this naturally merits separate investigation beyond the scope of this paper.

The most likely candidates for an index proxy of the first factor when examining the ICB industry-level indices are the *Financials* index and the *Cyclical Services* index. The Financials index in this case would be preferred as it has a loading closer to 0 on the second factor. One could of course also choose from the indices that are not ICB industry level: the most likely candidates from the loadings plot being Life Insurance, Financial Services or General Industries. As already stated, at the point where the indices are this close together, the decisions over factoring method and rotation can lead to different choices of proxy. It is again pertinent to consider substantive economic reality as a criteria as opposed to simple statistical construction. As indicated in Table 2, the Financials index has an approximately six times larger number of constituent members than the Cyclical Services index. This means

that the factor loadings on this index are more likely to represent genuine economic realities, rather than dominance by a few key stocks, as can occur on smaller exchanges such as the JSE (Kruger, 2005). Further to this, the Financials index is in fact the broader industry-level parent index of the other possibilities such as the Financial Services or Life Insurance indices. The combination of these arguments point to the Financials index (FINI) as the appropriate proxy for the first factor. Together with the resource related (RESI) component, the two factor APT model for the JSE is thus:

$$R_i = \alpha_i + \beta_{F,i}R_{FINI} + \beta_{R,i}R_{RESI} + u_i, \quad (7)$$

where

R_i are the returns for the investment i ,

α_i represents the excess returns attributable to investment i ,

$\beta_{F,i}$ represents the beta estimate for the regression of returns on investment i on the FINI,

$\beta_{R,i}$ represents the beta estimate for the regression of returns on investment i on the RESI,

R_{FINI} are the returns on the Financials index (FINI),

R_{RESI} are the returns on the Basic Materials index (RESI) and

u_i represents the error term for the regression.

4.1.3. Rolling Window of Eigenvalues

The analysis in Section 4.1.2 shows how the South African financial market is beginning to change. The necessary changes in the appropriate index proxies for the JSE will naturally raise the question as to how the primary factors may have changed over time. In order to investigate this, it is possible to employ the same principal component analysis as performed in Section 4.1.2, but with some differences. Instead of performing the analysis over the entire sample period, one can establish a rolling window function to perform the analysis at every timestamp²². This can be done on the broader set of equities²³ in the sample as items with missing values can be dropped when they have missing values in that particular time window, but retained in the case where they do not. In doing so, a rolling window of eigenvalues is established which can be used to calculate the proportion of the variation explained by the various components over time, with each timestamp showing the percentage of variance explained in the time window leading up to that timestamp. In this case a two-year rolling window period is used and the results of this are presented in Figure 5.

²²This necessitates specifying a minimum number of observations to ensure valid results. In all these cases, the minimum value for the analysis is the same as the window size.

²³The reason to make use of equities in this case instead of indices is to aid in the interpretation of the proportion of variation figure.

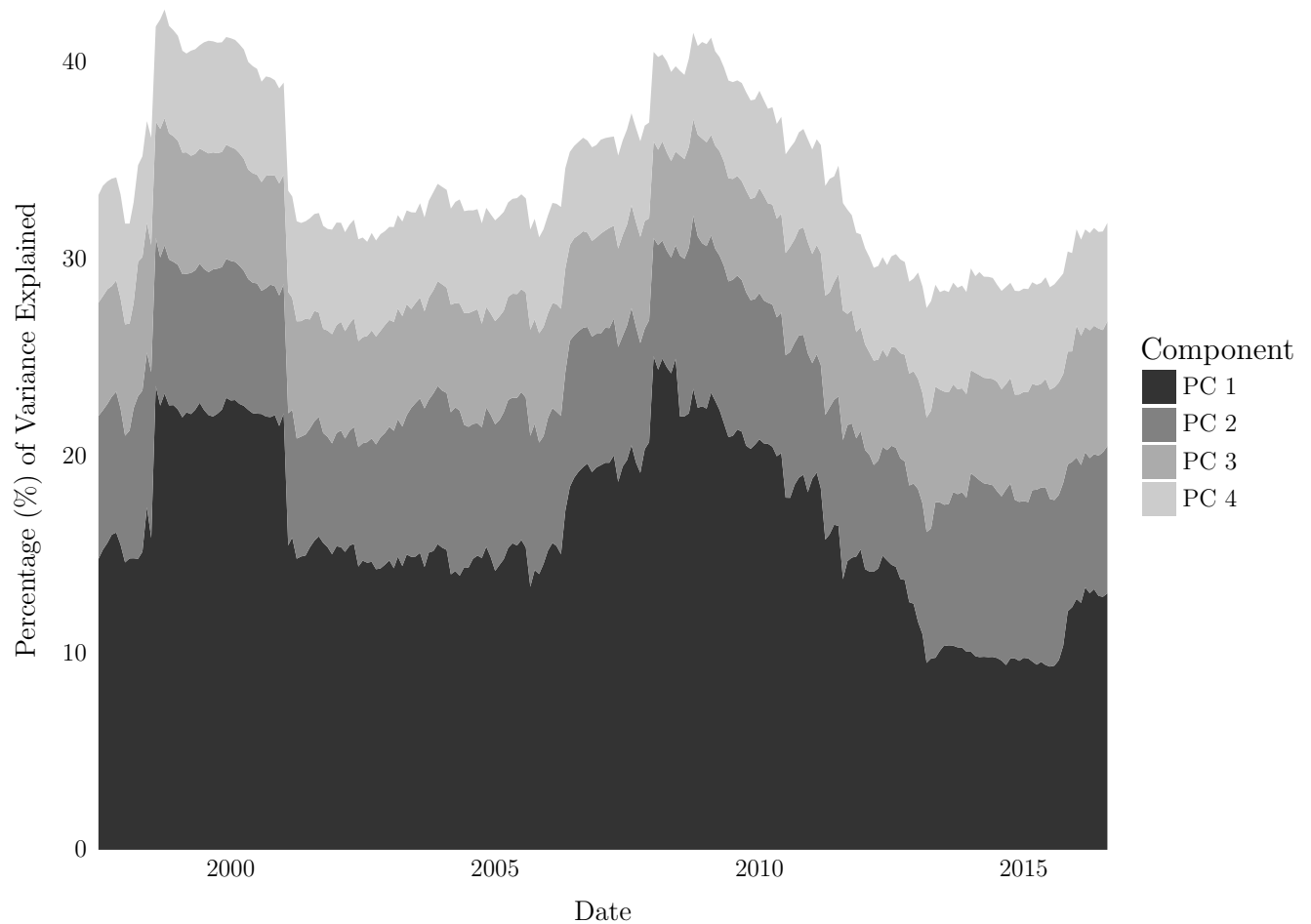


Fig. 5. Rolling Window of Eigenvalues

It has already been established from the scree plot in Figure 3 that the first factor accounts for the largest portion of the variation in the sample. From Figure 5 it is clear to see that the remaining factors, of which factors 2-4 are shown, remain relatively consistent in the amount of variation they explain over time. The first factor is far more volatile and is the major driver of overall explanatory power. This is a promising result for the two-index model: if the remaining factors (2-3) showed large differences in explanatory power over time, it would indicate the need for constant reassessment of the model’s validity. If the factors can be shown to be reasonably constant, the model can be relied upon more readily by practitioners.

Another useful aspect of these results is to evaluate them in terms of systematic risk. The way to do this is by examining the first factor, which is best measure to show how synchronous the returns on the JSE are: the larger the proportion of variation explained by the first

factor, the more the stocks on the JSE are moving together at any particular point in time. The chart shows the first eigenvalue undergoing a sharp increase in the period leading up to the financial crisis of 2007/2008 but decreasing thereafter. Similar analysis on these kinds of cross-correlations has been performed elsewhere, notably by Zheng et al. (2012) who perform PCA on the Dow Jones economic sector indices. Their analysis shows a similar pattern to that of Figure 5, where there is a spike in the first eigenvalue leading up to the financial crisis.

4.2. Beta Estimation

Section 4.1.2 establishes the appropriate form of the APT. This section will turn attention to the beta values themselves and the aspects involved in their estimation. The first question to consider is which regression technique is best suited to obtain the most accurate beta estimates. Section 2.3.3 has already provided the details of this, but some are worth emphasising. There is a strong analytical and empirical basis for believing that OLS is an inappropriate regression method for the estimation of beta values. This is due to its underlying assumptions; in particular, the assumption that suggests that one of the variables of the regression is subject to error while the other is not. This follows from the designation of one variable as being the ‘independent variable’ and the other as the ‘dependent variable’ (Tofallis, 2008). For the purpose of conventional OLS beta estimation in finance, the returns on the market portfolio are taken as being the independent variable whereas returns on the equity are taken as being the dependent variable. As argued for previously, there is little justification for this approach. It is very difficult to find a clear distinctions between the two variables in the regression and even more difficult to do so in such a way so as to determine that one is completely without error while the other is not. This point underscores one of the primary advantages of the RMA regression method when compared to OLS. RMA regression does not make any assumptions as to the degree of error in either of the variables of the regression. As a result, the variables are treated similarly to one other. The RMA regression line is also symmetric, so it is of no consequence which variable is plotted on the x or y axis.

The next important consequence of this is that regressions which make use of OLS will often significantly underestimate the magnitude of beta (Tofallis, 2008). To illustrate this, consider an example of beta estimation using the monthly returns of *ArcelorMittal South Africa Limited* (ACL) for the period December 2004 to December 2009²⁴. The scatterplot

²⁴A five year estimation period consistent with the beta estimates evaluated in this paper.

of the returns against those of the RESI are shown in Figure 6 along with the estimated OLS and RMA regression lines²⁵. The pattern of the points makes clear that there is a strong relationship between the variables. The question is which estimator better captures the magnitude of that relationship. As can be seen, the RMA regression line is significantly steeper than that of the OLS. The estimated OLS beta value is approximately 0.85 whereas the RMA beta value is approximately 1.37. As Tofallis (2008) points out, it is often cited that stocks with beta values less than one are considered defensive on account of the fact that they offer less volatility than the market. If one were to analyse ACL using the OLS regression, one would conclude that the stock is defensive. However, this would be a mistake. To show this, consider the same set of returns, but plotted linearly in Figure 7. The solid line represents the returns on the RESI whereas the dashed line represents the returns of ACL. An inspection of the plot shows that the relationship between the returns is not less than one. While it is clear there is a strong relationship between them, the returns of ACL vary considerably more than the returns of the RESI. While there are some months where the returns are in line with one another, there are many more where the returns on ACL exceed the returns on the RESI by some distance. For example, during the simultaneous spike in returns that occurred in February 2008, the ACL stock rose by approximately 29% whereas the RESI rose by only 18%. This relationship is far better captured by RMA regression which (using the ‘least triangles’ line) maps the functional relationship between the data. In doing so it gives a far more accurate measure of the strength between share and market returns (Tofallis, 2008).

4.2.1. Number of Regressors

The appropriate form of the APT is shown in Section 4.1 and takes the form of a two-index model using the RESI and the FINI as the applicable index proxies. This can be considered the parsimonious ‘base-case’. In order to evaluate beta, it is necessary to consider whether one or both of these factors have a significant relationship with returns in the case of any single stock. Van Rensburg (2002), in unreported results using indices, concluded that the majority of JSE listed equities had a strong relationship with one of the two factors, but not with both. This assessment was made using OLS regression of the two factors and examining the betas for significance using a t-test.

However, caution should be exercised when deciding what kind of significance test to employ. This is especially true when the two regressors are correlated. To illustrate how this may occur, consider the correlations between the returns of the ALSI, RESI, FINI and ACL

²⁵The method of determining the RESI as being the appropriate index proxy is given in Section 4.2.2.

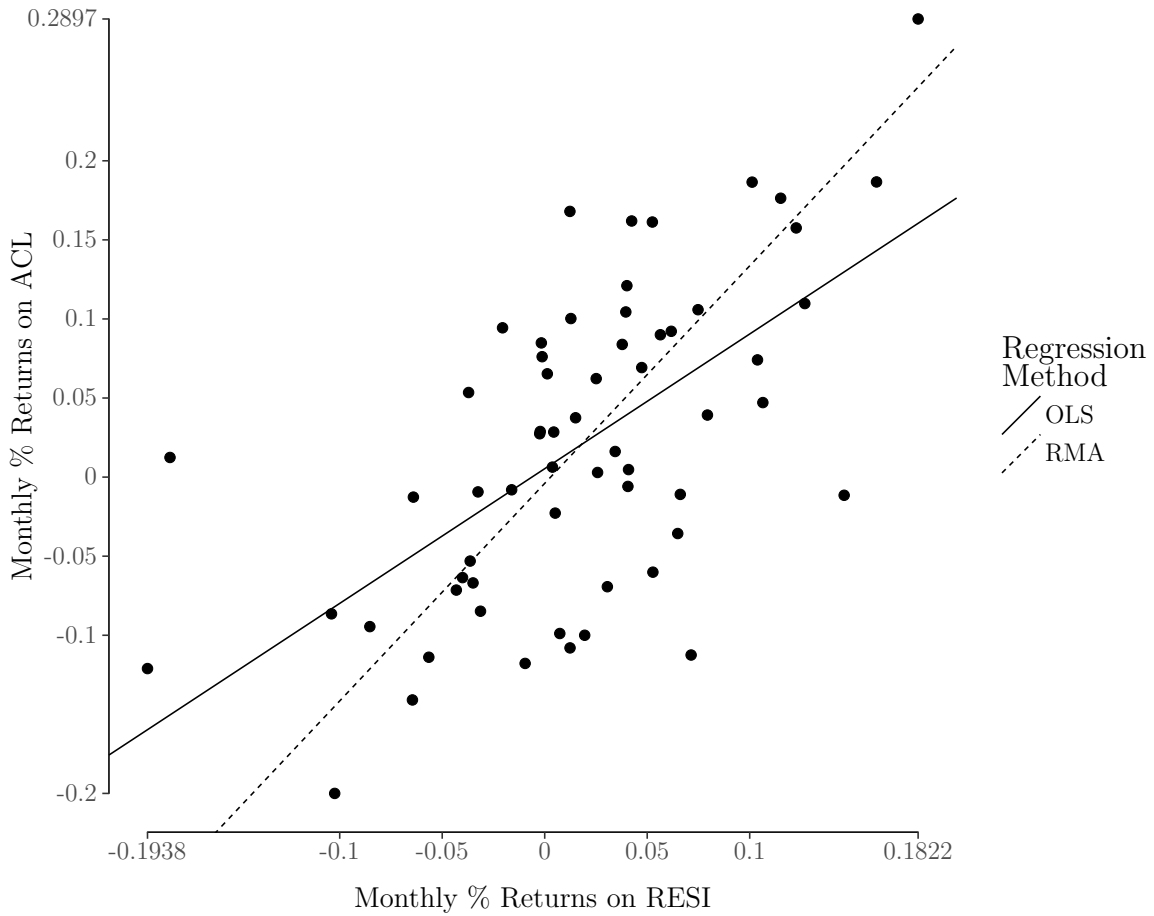


Fig. 6. Illustrative OLS and RMA Regression Lines

shown in Table 3. In line with the above example, there is a strong correlation between the RESI and ACL, which is approximately 0.6. There is also, however, a correlation between the FINI and ACL, approximately 0.27. From this one may be tempted conclude that both factors contribute to the share's returns. However, the correlation between the FINI and the RESI is 0.37, greater even than the correlation between the FINI and the ACL. From this one should conclude that the correlation between FINI and ACL is in fact spurious and only exists as a result of the correlation between the FINI and the RESI.

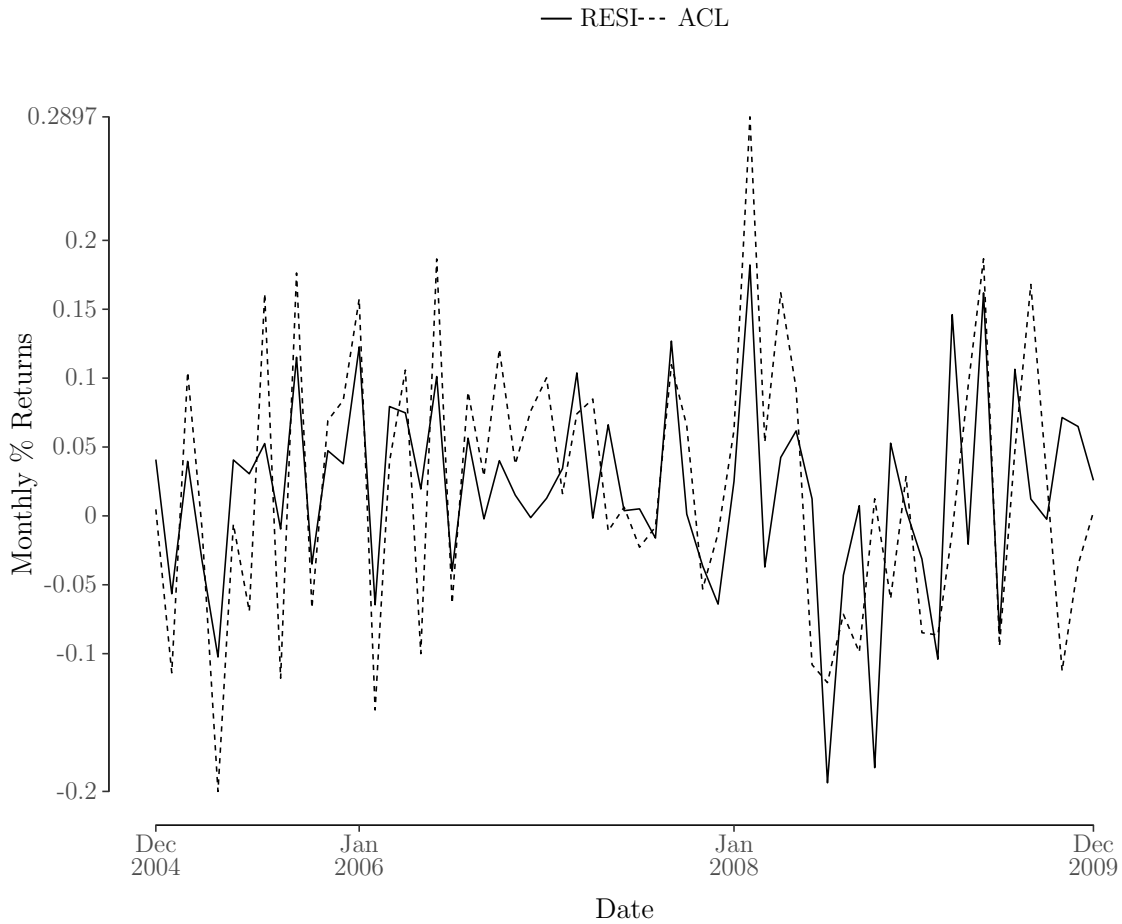


Fig. 7. Illustrative Index and Stock Volatility Chart

Table 3: Correlation Coefficients Between Selected Indices and *ArcelorMittal SA*

	All-Share	Financials	Basic Materials	<i>ArcelorMittal SA</i>
All-Share	1			
Financials	0.75	1		
Basic Materials	0.74	0.37	1	
<i>ArcelorMittal SA</i>	0.49	0.27	0.60	1

This thinking can be applied to the whole sample in a manner similar to how it was applied to indices in Section 4.1. By applying PCA to the correlation matrix of share returns, common sources of variation (factors) are grouped together and ordered by their importance. On account of the fact that indices simply aggregate the performance of their constituent stocks, there is strong reason to suspect that the results of the share PCA will be similar to the index PCA already performed. This proves to be correct,

as the loading plot shown in Figure 8 displays the same pattern of loadings as was shown in Figure 4, with resource stocks loading on the second factor and the other non-resource related stocks loading on the first factor. In this case, the promax rotation was again employed. The advantage of doing so is to decompose the possible correlation between the explanatory variables, in this case the correlation between the FINI and the RESI.

To make the results more readable, the loading plot is overlaid with the ICB industry classifications and the size of the data points are scaled according to the equity's market capitalisation²⁶. Thus the larger the stock is in terms of market capitalisation, the larger it will appear on the loadings plot. The shape of the marker is determined by its ICB classification. The same pattern is evident as was the case with the indices. The circular markers indicating the resource related stocks are centred around the second factor and the remaining stocks are centred around the first factor. In confirmation of the analysis in Section 4.1.2, very few stocks load moderately on both of the primary factors; most stocks lie between -0.2 and +0.2 on one of the axes. This shows that only a single regressor is necessary in the case of most stocks listed on the JSE. For simplicity, this paper will use only a single regressor for any given stock, the one which corresponds to its most significant index proxy, either resource related (RESI) or non-resource related (FINI). How that is determined will be discussed in Section 4.2.2.

4.2.2. *Significance of Beta Estimates*

Beta is usually assessed by means of an empirical evaluation of the expected returns-beta relationship. On the JSE, Van Rensburg and Robertson (2003) and Strugnell et al. (2011) do this by means of a portfolio-based assessment. A similar assessment to these will also take place in the next section. However beta can also be considered significant or non-significant at the point it is *estimated*, before any empirical evaluation has taken place.

This emphasis placed on only the final test may be justified in some instances. It stands to reason that betas which are not significant at the estimation stage will not be useful in the eventual empirical evaluation. However, it is not clear whether this is the case for the JSE under tests conducted by Van Rensburg and Robertson (2003) and Strugnell et al. (2011). In these cases there is a possibility that a portion of the beta values used in the formation of portfolios were statistically significant while others were not. This would imply that the portfolio sorts, as conducted in this case, included stocks with both significant and non-

²⁶As at their last traded date in the sample period.

significant betas. This leaves open the possibility that beta *is* a significant predictor of stock returns, but only for a subset of stocks (those for which it can be reliably estimated). Importantly, the only way we can know if this is the case is if the significance tests are conducted at the estimation stage and the sample is filtered to only include stocks with significant betas.

Some mechanisms already account for this implicitly. The use of a thin trading filter in Van Rensburg and Robertson (2003) will have the effect of removing some stocks with non-significant betas. Naturally stocks with a large number of non-trading periods will be less likely to produce valid beta statistics. This may still be imprecise however. It remains a better approach to conduct the significance tests directly rather than rely on a thin trading filter. It is for this reason that a thin trading filter is not applied in the sample selection (Section 3.3.2).

Before establishing which beta values are significant however, it is necessary to discuss what ought dictate whether a beta value is considered significant or not. Helpfully, tests of significance form part of the relevant regressions and often accompany the output when performed with any number of statistical packages. The most commonly employed is the t-test of significance of the coefficients in an OLS regression. One may question whether this is appropriate given the concerns raised over OLS regression in Section 2.3.3. However, the same concerns do not in fact apply. This is because the significance of a regression line is a separate matter from its position. As Jolicoeur (1975) put it: only once it has been shown that there is a significant relationship between y and x does the question then turn to finding the best line to capture that relationship. If one were to search for an alternative test, one may consider the well-established test performed on the correlation coefficient (r)²⁷. However, as pointed out by Ricker (1984), the OLS t-test performed on beta is in fact simply an algebraic rearrangement of the same test performed on r . For this reason and for the sake of consistency, the OLS p-value will be used as a test for the significance of beta values.

To do this, another two rolling regressions are established. The returns of the stocks remaining in the sample are regressed against both the returns of the FINI and the returns of the RESI in separate regressions. As discussed earlier, only five year beta values are computed for the purpose of this paper, which means a window size of 60 months is employed. Correspondingly, within the rolling window function, it is specified that any given regression must contain at least 60 data points. This ensures that no partial beta

²⁷Also tested by means of a t-test.

values are calculated. One advantage of using a rolling window is that it does not demand that shares are listed for the entire period under consideration. Shares will generate appropriate estimators for the period they are listed, so long as that period is sufficiently long to generate a viable beta statistic. This amounts to performing the relevant OLS regression on every data point within each time series of share returns after it exceeds 60 observations. The p-values of each regression is calculated and tabulated, creating a time series of p-values. It follows that the number of regressions performed on a particular share will be $n - 60$ where n stands for the number of observations in that share's time series of returns. As a result, a new time series of $n - 60$ p-values is found for 379 of the 461 stocks remaining in the sample. While the other 82 stocks did have greater than 60 observations, these were not continuous and thus did not have sufficient values to calculate even a single valid beta statistic. On this basis, these 82 stocks were excluded from the sample.

To make the assessment of significance, each stock's time series of p-values is summarised by its median. This is done for both the RESI and FINI p-values and can be used to determine into which category each stock belongs. If the median RESI p-value is lower, it shows the stock to have a risk-return profile which is resource related and if the FINI p-value is lower, it shows the stock to have a risk-return profile which is non-resource related. The distribution of these median p-values is shown in Figure 9. As can be seen, the majority of shares have median p-values which would indicate their betas to be statistically significant a large proportion of time. This is not overwhelming however. Of the 379 shares remaining in the sample, 212 had median p-values less than 0.05. The next most significant grouping (30 shares) has a median p-value between 0.05 and 0.1. The remaining 137 shares have median p-values spread out seemingly evenly thereafter.

These results provide a mixed account of beta. The fact that some stocks exhibit betas which are significant, but some do not, underscores the point made directly above. Some beta values being used in previous portfolio sorts may not have been significant. There is therefore a risk that these betas may have distorted the results of the significant betas. Whether this would have been corrected by a thin trading filter or another kind of beta correction remains unclear.

The stocks with median p-values between 5% and 10% are of particular interest for being

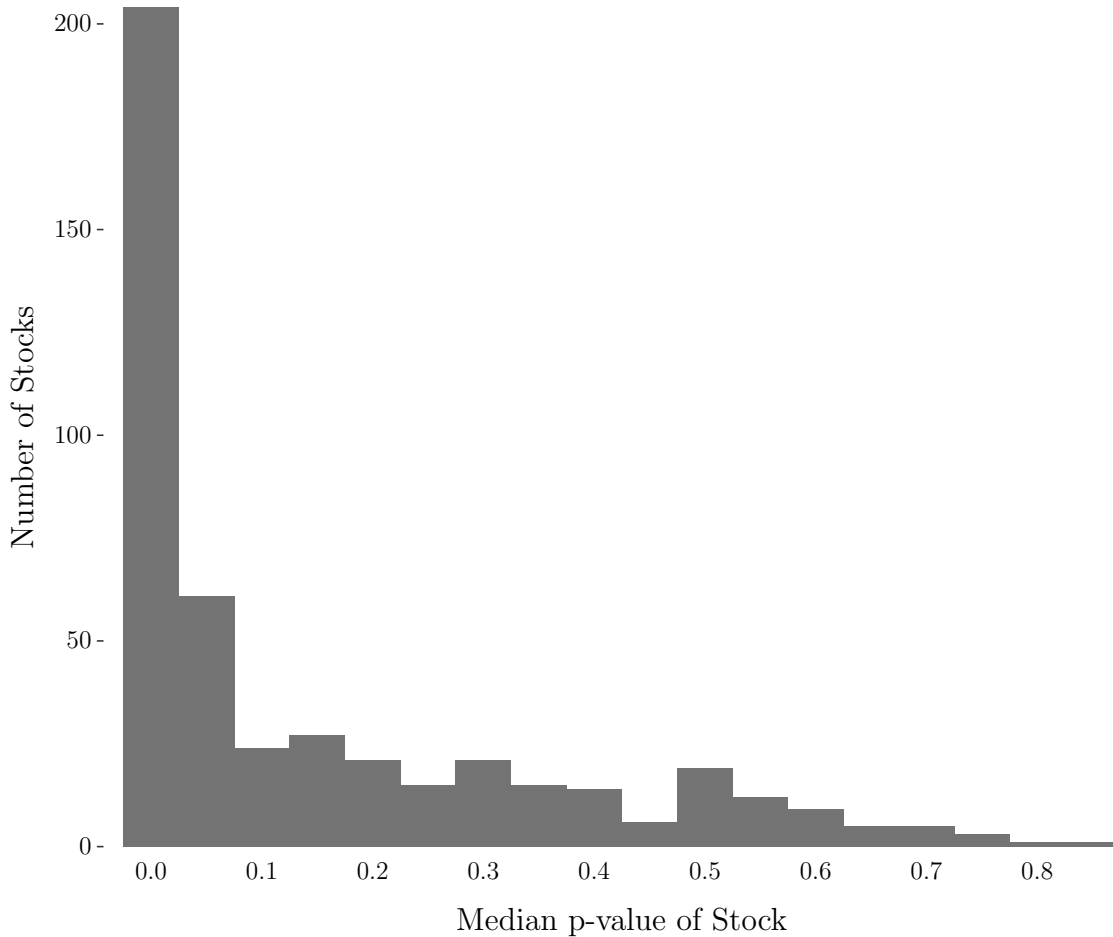


Fig. 9. Distribution of Stocks Based on Median P-Value of Betas

on the margin between what is typically considered significant or non-significant²⁸. Another kind of calculation is needed to explore this. The 5% level is the most commonly used and understood benchmark for significance and for that reason it will be used as a baseline for the exploration. Within the time series of a stock's p-values there will be those which fall below 5% and those which do not. If each p-value is evaluated this way it can be established what proportion of betas in a stock's time series are significant. Within each stock's time series this is performed by dividing the number of p-values below 5% by the total number of p-values in the series. If the proportion is 1, it implies that all betas calculated for a particular stock were significant. If it is 0 it implies that none of the betas calculated for a particular stock were significant. The distribution of these proportions is shown in Figure 10 where again the dichotomy is made clear. For the majority of stocks, beta is either

²⁸Given a fixed sample size, there is generally a tradeoff involved in setting the significance level of a statistical test. A stricter (lower) significance level will reduce the type I error rate, but will increase the type II error rate.

significant for the entire listing period or for no part of the listing period. This is shown by the peaks at either end of the distribution.

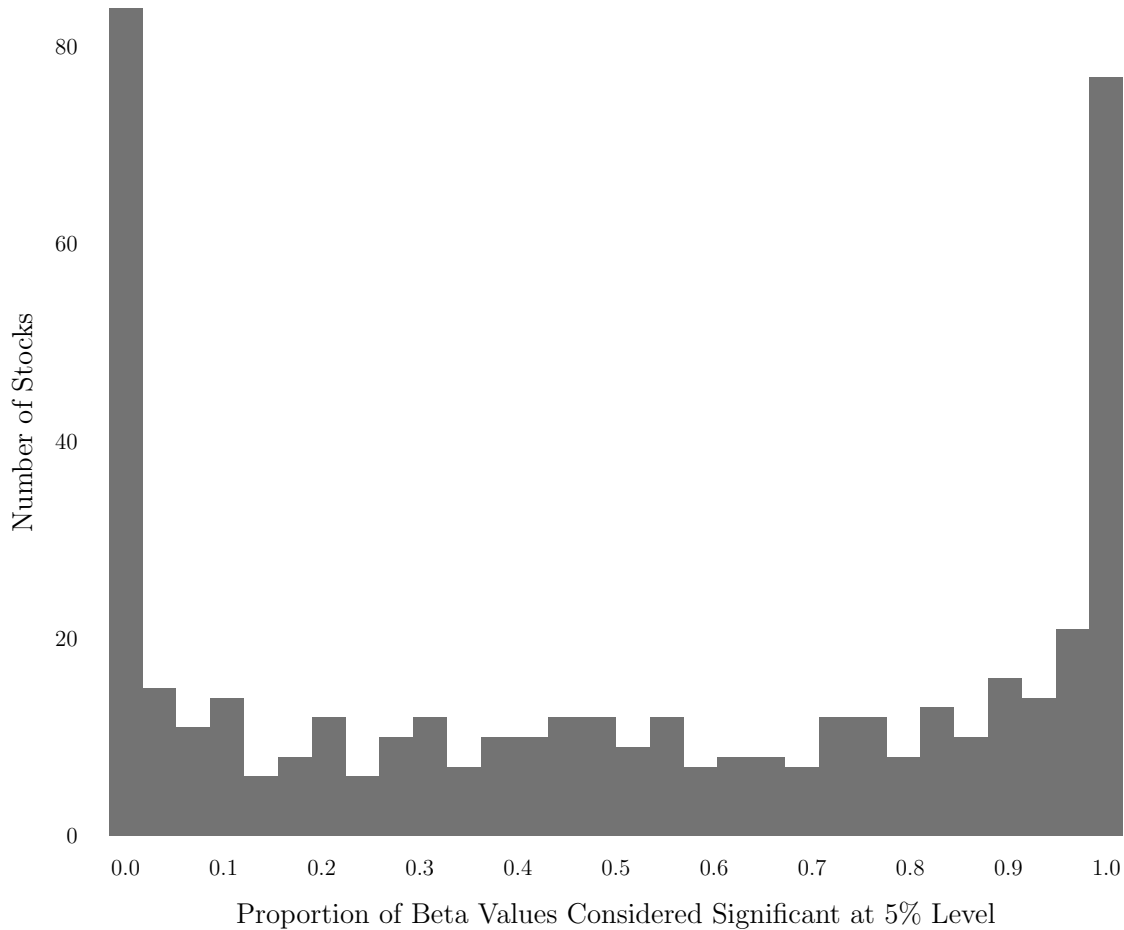


Fig. 10. Distribution of Stocks Based on Significance of Betas

Figure 10 however also shows a middle ground indicating that, for some stocks, beta drifts into and out of significance at some point or another. A balancing act is therefore required to establish what level of significance is appropriate to include some stocks and exclude others. If a stock's beta is only significant a small proportion of the time, it will bias the results of an empirical evaluation. However if a high level of significance is demanded, it will unnecessarily limit the sample of securities available for the portfolio sort, reducing the applicability of the conclusions arrived at. To show this tradeoff more clearly, these two distributions are plotted against one another in Figure 11. This makes the relationship clear: the smaller the median p-value, the greater proportion of the betas which are considered significant. This much may appear obvious by the way the measures are constructed, how-

ever this tradeoff is not direct; even at relatively low median p-values, many stocks still drift into and out of significance at various stages. There may be a number of reasons for this. For some stocks there may well be periods of firm-specific distress, where the stock does not follow the pattern of the market. This is not a sufficiently good reason to exclude the stock however. Conversely, this isn't a reason to be too lenient with p-values either. Figure 11 shows that, beyond the 5% median p-value (as shown by the dotted line), the proportion of significant beta values drops off sharply. Consider if a 10% filter were to be used, this would introduce stocks into the portfolio sort which were not significant at the 5% level for any part of their listing period; this is indicated by the points in the bottom left corner of the plot.

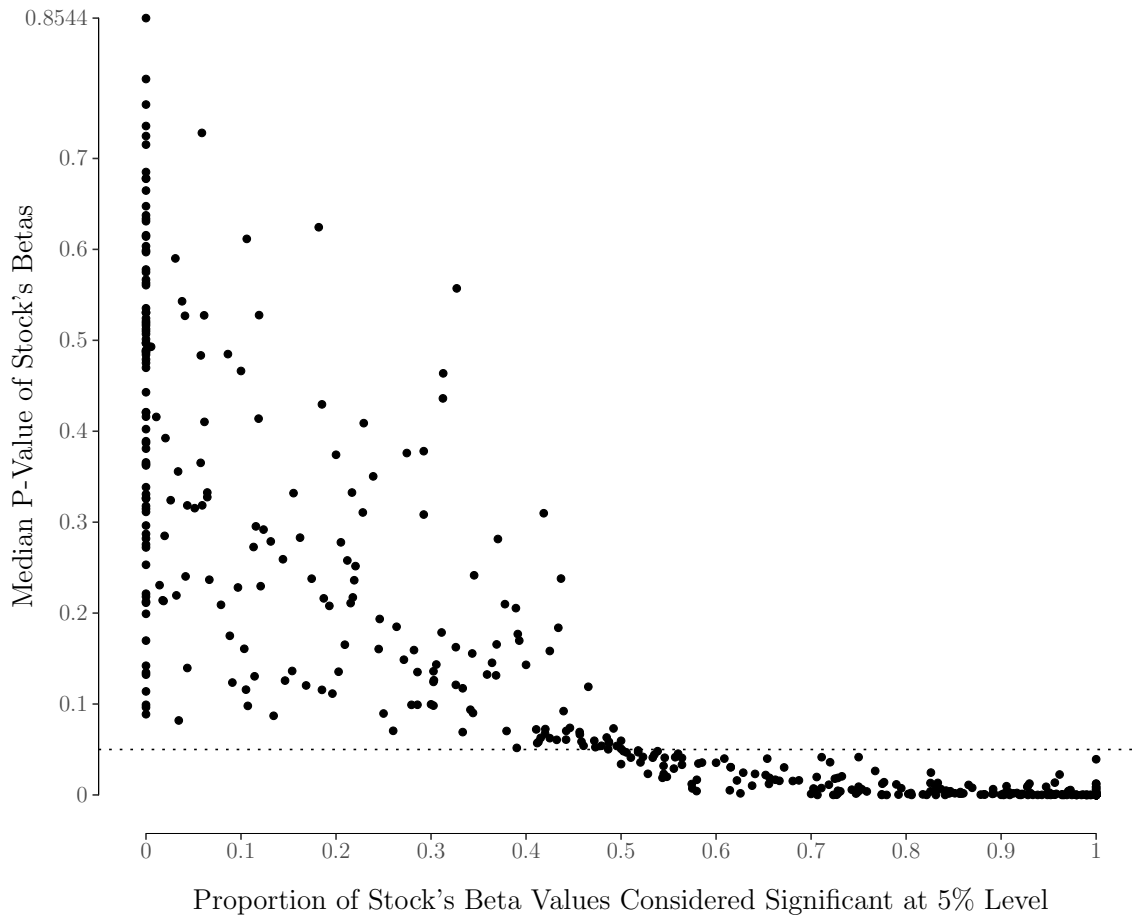


Fig. 11. Tradeoff Between Beta Median P-Value and Proportion of Significant Betas

These results show a 5% median p-value to be a well justified significance filter for the JSE. It ensures that all beta values brought forward into the portfolio sort are significant for the majority of their sample period while still being broad enough to allow for the possibility

of betas not being significant for a portion of the stock's listing. This ensures that stocks aren't aggressively excluded. The result is that, of the 379 stocks remaining in the sample, 235 are retained for the ensuing analysis. 169 of these have a lower median p-value when regressed on the FINI than with the RESI and are thus categorised and grouped as 'non-resource' stocks going forward. Correspondingly 66 of these have a lower median p-value when regressed on the RESI than with the FINI and are thus categorised as 'resource stocks'.

Overall these results provide a mixed account of beta on the JSE. While some beta values are found to be consistently significant, others are found to almost never be significant. In between there are a set of stocks whose beta value is significant most of the time, but which encounter periods where they are not significant. In part this is a negative result for beta, which cannot claim any kind of universal applicability among stocks listed on the JSE. The results which follow do not purport to extend beyond these securities identified as having significant betas. In another sense, these results appear to validate the suspicion that empirical tests of beta on the JSE should employ a significance filter of some kind. Only by testing significant beta values as opposed to non-significant ones can we be assured that our conclusions regarding those betas are authoritative.

4.2.3. Magnitude of Beta Estimates

The example of ACL at the beginning of this section illustrates how OLS regression can underestimate the magnitude of beta. It also shows how an error-in-variables method such as RMA can provide a more accurate measure of a share's sensitivity to its appropriate index proxy. Analytically, it has already been shown in Section 2.3.3 that OLS regression will always result in a lower magnitude beta estimate than RMA regression in the case of a single regressor. However, it remains important to establish the extent of these differences; a small difference between these two beta measures may have only a negligible effect, but large differences will have a material impact on the conclusions arrived at.

Some introductory analysis of this has already been performed by Tofallis (2008) and Laird-Smith et al. (2016). In each case a selection of RMA and OLS betas for major shares are examined at a fixed point in time. The analysis of this paper aims to go further by incorporating all significant share betas and examining them over a rolling window period.

To assess the extent of these differences between the beta measures, another rolling regression is conducted in a similar manner to the one in Section 4.2.2. Again a single index regression is performed however on this occasion using only the appropriate index proxy: the RESI in

the case of resource related stocks and the FINI in the case of non-resource related stocks. This analysis is performed twice: in the first case using OLS regression and in the second case using RMA regression. In each case a beta value is calculated. Similar to the rolling regression already performed, this results in the calculation $n - 60$ beta values in the time series of every share where n is the number of observations in that time series. Similarly, a minimum of 60 observations is set to ensure no partial beta values are calculated. In order to display these results, each share's time series of betas is summarised by its median value. The distribution of those median values is displayed in Figure 12.

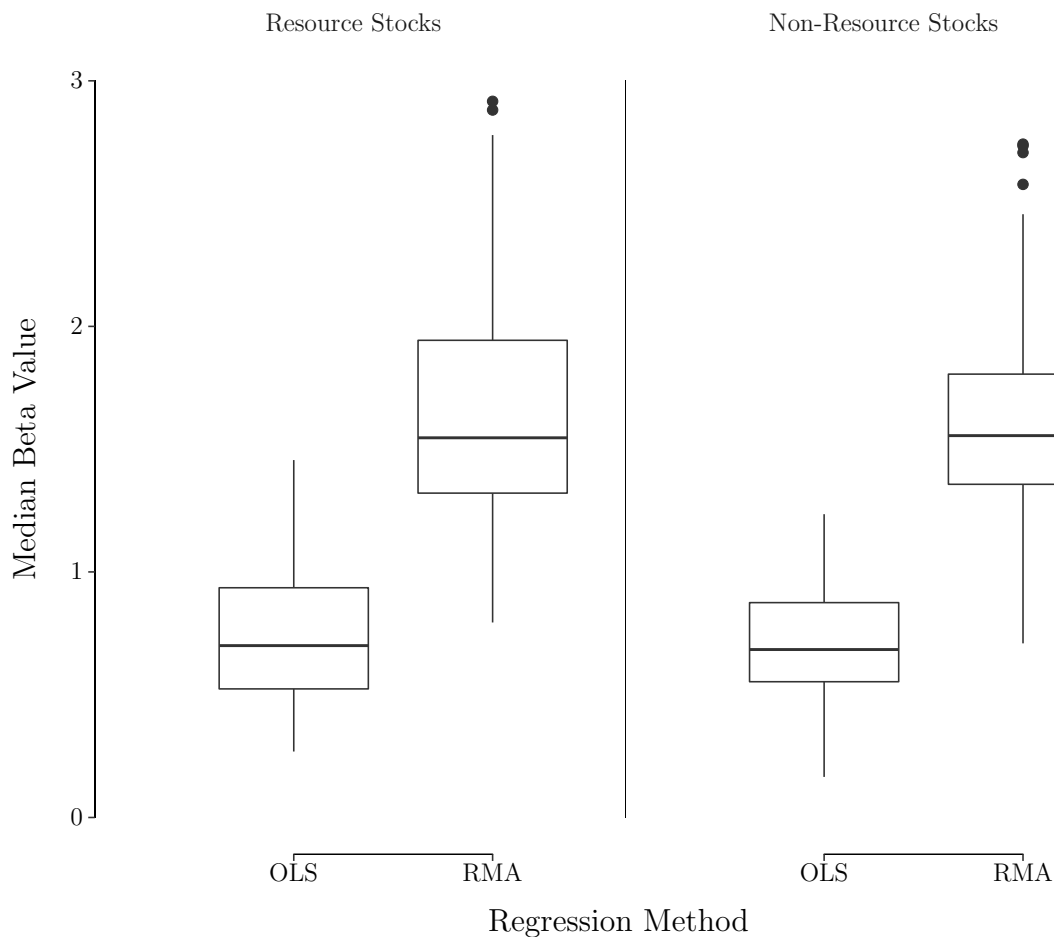


Fig. 12. Magnitude of Beta Values on the JSE

It is clear from the results that the choice of regression method has a material impact on the magnitude of the beta values which are calculated. Taking the case of shares which have the FINI as their appropriate index proxy: the OLS beta values range from 0.17 to 1.24 whereas the RMA beta values range from 0.71 to 2.74. The same pattern can be seen

in the case of shares which have the RESI as their appropriate index proxy.

It is worth emphasising the point made previously by Tofallis (2008) and in Section 2.3.3. One of the main interpretations and uses of beta is to describe the strength of the relationship between the returns of the investment and the returns of the market as a whole. Consequently betas greater than 1 are taken to imply that an investment is riskier than the market and betas smaller than 1 are defensive or less risky than the market. In the case of these stocks, the t-tests prove the market is the underlying driver of returns. The OLS beta measure however, fails to capture this relationship. Instead OLS beta would conclude that the vast majority of these stocks were defensive. The upper quartile of the OLS betas are 0.94 in the case of resource stocks and 0.87 in the case of non-resource stocks. This would imply that more than three quarters of stocks in the sample were defensive in nature. This simply isn't the case however. Most of the stocks in the sample have volatilities greater than that of the market. Given that the market has been shown to be the driver of returns for these stocks, one should then expect beta values larger than one. RMA regression provides the better estimator precisely because it uses this ratio of volatilities for calculating beta as shown in Equation (4). The results are clear in examining Figure 12; RMA betas occupy a more realistic range of values than their OLS counterparts. Some will be defensive; for example the lower bound of RMA betas for non-resource stocks is 0.71. Others however will be naturally be larger on the basis of their larger risk as measured by their standard deviation.

Another advantage of performing the various regressions on a rolling window basis is the ability to detect whether any of the findings are time-varying. This can be done with only a slight change to the procedure immediately above. Instead of completing the calculation by summarising each share by the median beta value in its time series, one can take the median beta value of all the shares at each timestamp. The results of this are shown in Figure 13. Again the results show a substantial difference between the beta values calculated from the two regression methods. The median RMA values shown by the dashed line are consistently above their OLS equivalents shown by the solid line. This demonstrates that the differences in magnitude are not time-specific.

4.2.4. Stability of Beta Estimates

The stability of beta values ought be a serious consideration when making an evaluation of regression methods. If beta estimates are found to be unstable they cannot be relied upon to produce a reasonable measure of the strength between the returns of the market and that of the investment they are examining (Tofallis, 2008). If beta is constantly changing, it

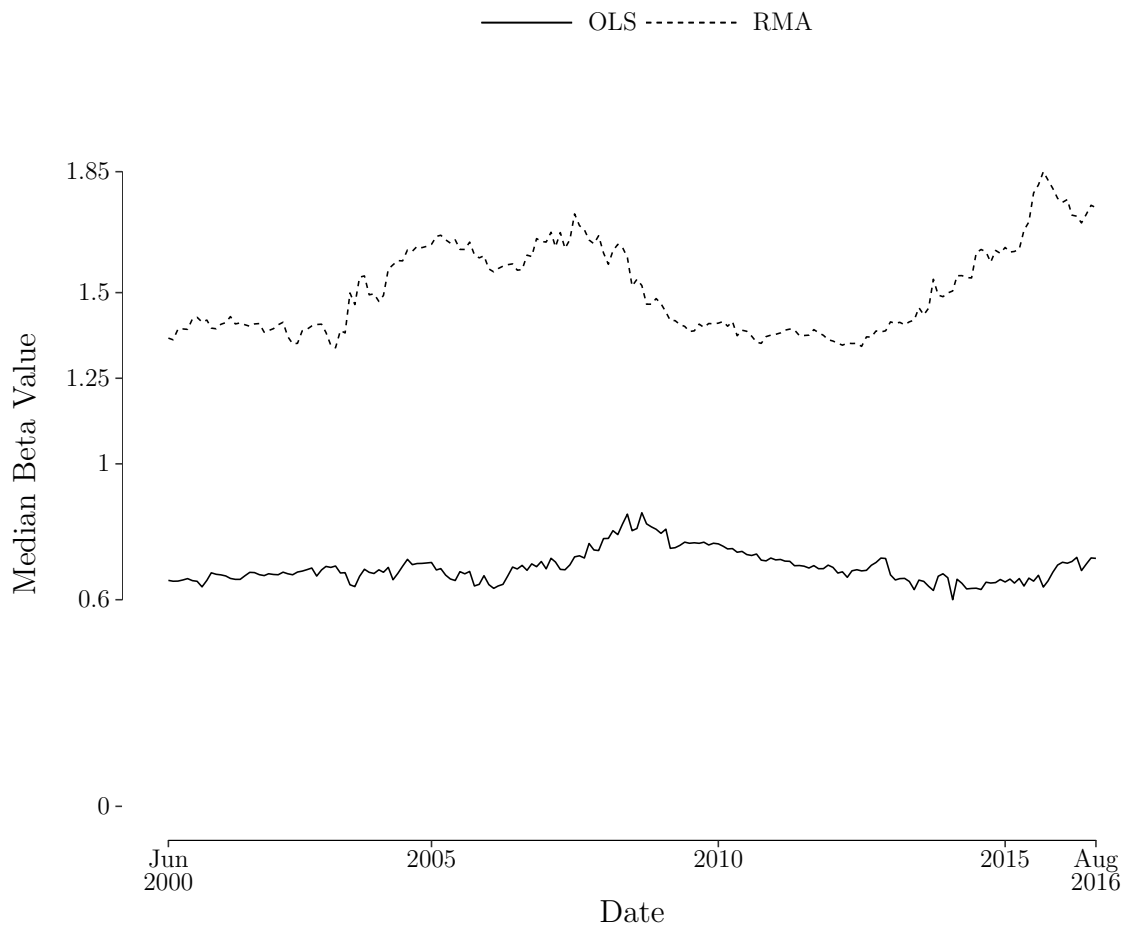


Fig. 13. Median Beta Values on the JSE Over Time

implies the model is in constant need of recalibration. In addition, such models are unlikely to make accurate predictions, even if subject to constant re-evaluation.

In the case of Laird-Smith et al. (2016), RMA beta is shown to be more stable when assessed by percentage changes over three discrete points in time. These points are anchored by the the financial crisis of 2007/2008 with one period being taken before the financial crisis, the next period during the crisis and the final period after the crisis. In the case of a rolling window of betas however, there is no need to define such periods, which run the risk of being arbitrarily defined. It also allows for a great deal more data points as illustrated by Figure 13. The most accepted measure of stability is standard deviation. However having only a few discrete periods available in that case did not allow for standard deviation to be computed. Again the advantage of using the rolling window period is made clear: by calculating a time series of beta values, the standard deviation of the various betas can be computed.

This computation is performed in a similar way to those done in the previous section. In the case of each stock, a time series of beta values is already available. This is from the rolling regression performed in Section 4.2.3. The standard deviation (σ) of each stock's beta is then calculated. This results in a set of beta standard deviations: two for each stock remaining in the dataset (one OLS and one RMA for each stock). This set of standard deviations is then summarised by its median value, grouping by the regression type (OLS or RMA) and appropriate index proxy (either resource stocks or non-resource stocks). When making this assessment, it is also important to consider the context from which these values emerge. The standard deviation of a particular variable is dependant on the magnitude of that variable. Values which are measured in larger units will have a larger standard deviation. For this reason it is necessary to compensate for the magnitude of the measure being studied such that it may be comparable with other measures. The most obvious way of doing this is by dividing the standard deviation of a set of values by the mean of those values (σ/\bar{x}). This will compensate for the magnitude of the variable being evaluated. This is done for every standard deviation value already calculated resulting in two compensated standard deviation values for every stock. As was the case with the plain standard deviation values, these are also summarised by their median. Both sets of results are displayed in Table 4.

Table 4: Comparison of Beta Stability.

Share Category	σ		σ/\bar{x}	
	OLS	RMA	OLS	RMA
Resource Stocks	0.19	0.24	0.27	0.15
Non-Resource Stocks	0.16	0.22	0.24	0.14

As shown in the table, the standard deviation results behave as expected. The larger range in which RMA betas exist result in a higher standard deviation figure when compared with OLS. Of greater interest are the measure values which compensate for the greater magnitude of RMA betas. By this standard, the RMA estimator shows itself to be more stable when adjusted for its magnitude. The difference between OLS and RMA is large in both cases. For resource stocks the value is 0.15 for RMA beta whereas the OLS value is 0.27.

It can be concluded that OLS beta is more stable in absolute terms when compared to RMA beta. However these results should be evaluated along with those of the previous sections. The reason OLS appears to be more stable is due to the beta estimates which result from it being downwardly bias. This much can be established by using a stability measure which compensates for magnitude and uses that as the basis for comparison. When that is

performed in this case it shows RMA beta to be more stable than OLS. It can therefore be concluded that not only does the RMA beta measure provide a far more accurate estimator for the strength of the beta relationship, it also does this while remaining more stable relative to its magnitude. Overall, this provides a compelling rationale for the use of RMA regression over that of OLS.

4.3. Empirical Evaluation

The APT can only be considered valid if there is shown to be a relationship between beta and expected returns. To establish this, an empirical evaluation is required to determine whether beta has the ability to forecast stock returns. The above sections provide the elements necessary for this empirical evaluation of beta to take place. Section 4.1.2 provides the appropriate format of the APT and the appropriate index proxies for estimation. This corrects for the market segmentation problem and enables the separation of results based on the appropriate index proxies. Furthermore, the assessment made in Section 4.2.2 regarding the significance of beta estimates on the JSE allows for the use of a significance filter where only stocks with a median beta p-value smaller than 5% are included as part of the empirical results. While this decision does limit the applicability of the findings, it ensures that credible empirical tests can take place; if betas are not found to be significant at the estimation stage, they will not produce a viable result of any kind when subjected to empirical evaluation. It is better in such circumstance to examine the subset of stocks with significant beta values to see if those betas have power as predictors of returns for those particular stocks.

Before moving on to these empirical results, it is necessary to mention some of the corrections made to beta in previous studies. Notably, the work of Strugnell et al. (2011) performs two beta corrections to correct for the presence of thin trading. These corrections are those proposed by Scholes and Williams (1977) and Dimson (1979). The Dimson correction is an extension of Scholes and Williams correction that make the extent of beta smoothing proportional to the extent of thin trading. While these do appear to be justified, it is decided at this point not to make any further corrections to beta for the purpose of this evaluation. The reason for this is that it is unclear how the possible beta corrections will interact with the RMA beta measure proposed. Scholes and Williams (1977) for example show their estimator to be consistent and unbiased if the periods of thin trading are shown to be independently and identically distributed. It isn't clear whether the same would be true in the presence of RMA regression. For this reason it is suggested as a topic of further

study rather than being incorporated herein.

The empirical results arrived at follow the general approach used by Van Rensburg and Robertson (2003) and Strugnell et al. (2011), who make use of the portfolio sort methodology. This has the advantage of making the results more comparable with those arrived at by these studies. Because beta is the only attribute examined in this case, only a one-way portfolio sort is necessary. As was the case with the rest of the analysis, delisted or otherwise inactive stocks are not removed from the sample, which mitigates against the effects of survivorship bias.

The betas required for this portfolio sort have already been calculated as part of the analysis conducted in Section 4.2.3. Following the methodology of Van Rensburg and Robertson (2003) and Strugnell et al. (2011), these betas are used as the basis for the formation of equally weighted quintile portfolios at each month-end. Equally weighted portfolios are selected to avoid the well-established concentration problem on the JSE (Kruger, 2005). Following Van Rensburg and Robertson (2003), the simulated portfolios of this paper are reconstituted monthly to prevent information loss. The returns generated on each of these portfolios are calculated over ensuing 1, 3, 6 and 12 month holding periods. The mean value of these returns is then calculated for each quintile across the sample period. The results of resource (RESI) stocks are shown in Table 5, with discussion following in Section 4.3.1. The results of non-resource (FINI) stocks are shown in Table 6 with discussions in Section 4.3.2. In each case, the differences in mean returns between selected quintile portfolios are calculated and assessed for significance using a t-test. Results of this are reported at various levels of significance as indicated.

4.3.1. Resource Stocks

The resource related stocks present fewer statistically significant findings than those of non-resource stocks. This is likely due to the smaller sample of resource shares available to form part of the portfolio sort. Van Rensburg and Robertson (2003) do not display the results for resource stocks on this basis. In the case of Strugnell et al. (2011), resource and non-resource stocks are analysed together with possible negative implications as have been discussed previously. In the case of portfolios sorted on OLS betas, none of the pairs of mean returns are found to be significantly different from one another. In the case of RMA regression however, both the 1 month and 3 month returns between portfolios 5 and 2 are shown to be statistically different from one another at the 5% level. These results are in line with what the APT would predict: the portfolios formed on larger beta stocks are shown

Table 5: Results for Portfolios Formed on Resource (RESI) Stock Betas

Holding Period	Regression Type	Portfolio					Portfolio Differences		
		1	2	3	4	5	5-1	2-1	5-2
1 Month	OLS	1.45%	1.74%	0.78%	1.09%	1.39%	-0.05%	0.29%	-0.34%
	RMA	1.27%	0.68%	1.53%	1.29%	1.36%	0.09%	-0.59%	0.68%*
3 Month	OLS	4.69%	4.84%	2.46%	3.87%	4.22%	-0.46%	0.15%	-0.61%
	RMA	3.63%	2.94%	5.27%	3.93%	4.75%	1.12%	-0.69%	1.81%*
6 Month	OLS	9.15%	8.84%	6.94%	10.66%	7.75%	-1.40%	-0.31%	-1.09%
	RMA	6.51%	8.59%	12.07%	7.20%	9.73%	3.23%	2.09%	1.14%
12 Month	OLS	22.27%	15.79%	17.12%	28.42%	20.14%	-2.13%	-6.48%	4.35%
	RMA	16.91%	14.69%	27.45%	19.94%	20.15%	3.24%	-2.22%	5.46%

* $p < 0.05$

** $p < 0.025$

*** $p < 0.01$

to have higher returns.

What is of greater interest however is the overall pattern of stock returns. For portfolios formed by OLS betas, all holding period returns between portfolios 5 and 1 have negative values. This is the opposite of what the APT would predict with larger beta stocks having lower returns. When portfolios are formed on RMA betas however those same differences are all positive. This amounts to a substantial reversal of the pattern of portfolio returns based solely on the regression method used to estimate the beta values. These results should be viewed sceptically however. In the first place, the sample of resource stocks is smaller and only a few statistically significant results are found. In addition, these results are not consistent over the various holding periods. Overall, it appears questionable as to whether the portfolio sort methodology is properly suited to the evaluation of RESI betas.

4.3.2. Non-Resource Stocks

The results for non-resource stocks are not subject to the problems of having a small sample as was the case for resource stocks. In this case a greater number of statistically significant conclusions are arrived at. These all occur when testing differences between the portfolios 1 and 5. In the case of portfolios formed on OLS betas, these differences are found to be significant for all holding periods. In the case of the one month holding period, this difference was found to be significant at the 1% level. All of these differences are negative, resulting in the same surprising and counterintuitive finding arrived at in Van Rensburg and Robertson (2003) and Strugnell et al. (2011). OLS beta appears to have, if anything, an inverse relationship with stock returns.

However, when portfolios are formed on RMA beta, a completely different conclusion is arrived at. When examining the differences in returns between the RMA sorted portfolios, they are almost always found to be positive in the cases where OLS beta portfolios were significantly negative. This is the same reversal pattern as was shown in the case of resource stocks (Section 4.3.1). In most cases, this removes whatever statistical significance was present under OLS. The only exception is the case of a 3 month holding period, where returns between RMA portfolios 1 and 5 are found to be significantly different at the 5% level. In addition, while not significant, the differences between RMA portfolios 1 and 5 in the case of 1 and 12 month holding periods result in p-values of 9.7% and 9.1% respectively, placing them on the margin of significance at the 10% level. These results naturally raise the suspicion as to what would occur in the presence of a thin trading filter or beta correction.

Table 6: Results for Portfolios Formed on Non-Resource (FINI) Stock Betas

Holding Period	Regression Type	Portfolio					Portfolio Differences		
		1	2	3	4	5	5-1	2-1	5-2
1 Month	OLS	1.44%	1.59%	0.99%	1.25%	1.25%	-0.19%***	0.15%	-0.34%
	RMA	1.11%	1.31%	1.26%	1.37%	1.60%	0.49%	0.20%	0.29%
3 Months	OLS	4.26%	4.65%	3.35%	3.93%	3.75%	-0.51%***	0.40%	-0.90%
	RMA	3.54%	3.89%	3.71%	4.39%	5.31%	1.77%*	0.35%	1.42%
6 Months	OLS	7.89%	10.23%	6.77%	6.49%	7.34%	-0.55%*	2.35%	-2.90%
	RMA	7.40%	8.31%	8.10%	8.40%	8.41%	1.01%	0.90%	0.11%
12 Months	OLS	18.23%	20.64%	17.29%	15.17%	15.90%	-2.33%***	2.42%	-4.75%
	RMA	14.61%	16.89%	18.18%	18.79%	18.12%	3.50%	2.27%	1.23%

* $p < 0.05$

** $p < 0.025$

*** $p < 0.01$

The broader trend of these results is clear. As was the case for resource stocks, almost all negative differences computed are reversed into positive differences when examining portfolios sorted on RMA betas as opposed to OLS. In the case of a three month holding period, returns between RMA portfolios 1 and 5 are found to be significantly different at the 5% level. This implies that a set of portfolio differences went from being *significantly negative* under the OLS methodology to *significantly positive* under RMA. Consequently the conclusion arrived at changes from being surprising and counterintuitive to being one which is in support of the APT.

In general, it is shown that completely different portfolio sorts can result depending on whether OLS or some other kind of regression method (such as RMA) is used as the basis for beta estimation. As has been argued for in Section 4.2.3, RMA provides a more accurate measure of the strength of the relationship between stock and market returns. It therefore follows that, in addition to the downward biasing of beta estimates, the OLS regression method is also distorting the results of portfolio sorts, leading to the erroneous conclusion that there is an inverse relationship between beta and stock returns. When the analysis is repeated with the rival RMA beta measure, the results are closer to what would be predicted by the APT, significantly so in some cases.

As mentioned, an item of particular interest would be to examine how possible corrections to RMA betas would alter the results of these tests. As discussed, these beta values, both RMA and OLS, will still have been affected by any thin trading remaining in the sample. If this correction meant that a greater number of RMA betas were significant in their ability to predict stock returns, it would serve as strong evidence in favour of the APT. This would be especially powerful considering most of the evidence presented thus far has been to the contrary. Furthermore, such a result would raise questions as to how other anomalies that have been identified would behave in the presence of the APT; in particular, whether they would still be considered significant after compensating for market related (beta) risk.

5. Conclusion

This paper has carried out an examination of the nature and significance of market betas on the Johannesburg Stock Exchange (JSE). This has required not only an empirical evaluation of returns-beta relationship, but also an investigation into various aspects that play a role in beta's estimation. In doing this, it has been demonstrated how interconnected these elements are with one another; decisions taken at the earlier stages of analysis are shown to have consequences for the beta estimates which are used in the empirical evaluation. These conclusions follow the same three broad categories as have been used throughout this paper: 1) Factor Selection 2) Beta Estimation and 3) Empirical Evaluation.

In the context of the APT, the identity of market betas is not pre-defined and thus raises the question of **factor selection**. This paper has updated and broadened the factor analytic work carried out by Van Rensburg and Slaney (1997), Van Rensburg (2002), Kruger (2005) and Laird-Smith et al. (2016). It is demonstrated that there continues to be a dichotomy in the returns generating process between resource and non-resource stocks on the JSE. Consequently it is again argued that a two factor Ross (1976) APT model ought be considered the parsimonious 'base case' for application to shares listed on the JSE. Furthermore it is argued that the ICB industry-level Basic Resources (RESI) index (JSE code J510) ought used as the updated proxy for resource related risk and return. This is consistent in its nature with earlier proxies however it is updated to reflected the rebased and current index structure. Unlike that specified in Laird-Smith et al. (2016), this proxy is also continuously listed for the entire sample period, which means it can be relied upon for long-term and historical analysis. For non-resource stocks, it is argued that the ICB industry-level Financials (FINI) index (JSE code J580) ought used as the appropriate index proxy. Again it is confirmed that the previously selected Financial and Industrial index (FINDI) is undesirable on account of the dual exposure exhibited by industrial shares. The exact nature and dynamics of this change is suggested as a topic of further study. This paper has also demonstrated that use of either the FINI or the RESI is appropriate for the majority of stocks listed on the JSE; it is seldom that stocks exhibit exposure to both factors.

Regarding **beta estimation**, this paper has made the case for an alternative method to the traditional OLS regression approach. Following the arguments made by Camp and Eubank (1981), Tofallis (2008) and Laird-Smith et al. (2016), it is argued that Reduced Major Axis (RMA) regression ought be employed for the purpose of beta estimation. A number of arguments are detailed for this. Chiefly, it is shown that RMA regression provides a more

accurate measure for the strength of the relationship between the returns of an investment and the returns of the market. Correspondingly, it is shown how OLS regression chronically underestimates the strength of this relationship. In addition, by making use of a number of rolling window regressions, it is shown that RMA beta values are more stable than OLS beta values when compensated for their magnitude. These characteristics make RMA regression a more appropriate estimation method than OLS for the purpose of beta estimation.

Furthermore, the use of a rolling window function to estimate betas allows for an assessment of their statistical significance over time. What is shown is a divergence in the significance of betas on the JSE. These fit into three main groupings. The largest grouping of stocks have betas which are significant for a large or entire portion of their listing period. The second largest grouping have betas which are almost never significant for any period of their listing. Lastly, there is a grouping of stocks whose betas are intermittently significant. Stocks with beta values that are not significant present a problem for empirical tests of the expected returns-beta relationship. To account for this, it is suggested that a ‘significance filter’ ought to be employed to limit empirical analysis of the beta-returns relationship to only those stocks for which the beta values are shown to be significant the majority of the time. It is shown that, on balance, a median p-value of 5% should be used when making such a determination. At this level, it is shown that betas are still significant for the majority of their listing period.

After making use of this significance filter, the **empirical evaluation** of both OLS and RMA betas is carried out using the same broad methodology of Van Rensburg and Robertson (2003) and Strugnell et al. (2011). The results for RESI stocks do not show consistently significant results for either beta measure, likely owing to the smaller sample of resource shares available in the sample. This is consistent with the findings of Van Rensburg and Robertson (2003). For non-resource (FINI) shares however, the results show a deep contrast between the OLS and RMA regression methods. In the case of portfolios constructed using OLS regression, the results are similar to those of Van Rensburg and Robertson (2003) and Strugnell et al. (2011), showing a significantly inverse relationship between OLS beta and stock returns. However when this analysis is repeated using portfolios sorted on RMA betas, it results in a reversal of this pattern. Many of the differences that were negative when portfolios were sorted on OLS betas are shown to be positive when the portfolios are sorted on RMA betas. While not all of these pairings result in statistically significant positive values, many of the p-values are small and would be considered significant at the 10% level. In one case this actually resulted in a set of portfolio differences being significantly negative under OLS sorted portfolios to being significantly positive under RMA sorted portfolios.

This amounts to a dramatic reversal in the assessment and outlook of beta. Under an OLS regressions methodology, beta shows a negative relationship with returns, a result which is both surprising and significantly at odds with the established theory. However under the RMA methodology beta is shown in some cases to be a significant positive predictor of stock returns, in line with established theory.

Neither the OLS nor RMA beta values examined in this paper were compensated for the effects of thin trading. This raises the question as to whether RMA portfolio differences close to significance currently (under 10%) would have been considered significant if a beta correction or thin trading filter had been applied. For this reason it is suggested as an avenue of further research. In the case of a Scholes and Williams (1977) or Dimson (1979) beta correction, it would need to be examined how the beta correction interacted with the proposed RMA beta measure.

Overall, this paper has made clear the advantages of using RMA regression over that of OLS. Not only does RMA beta provide a more accurate measure of the strength of the relationship between investment and market returns, it is also shown that the chronic underestimation of beta is also distorting the results of portfolio sorts, leading to the erroneous conclusion that there is an inverse relationship between beta and stock returns. For these reasons it is suggested that RMA regression take the place of OLS as the preferred method of beta estimation for stocks listed on the JSE.

Appendix A. Loading Plots for Alternative Factoring Methods

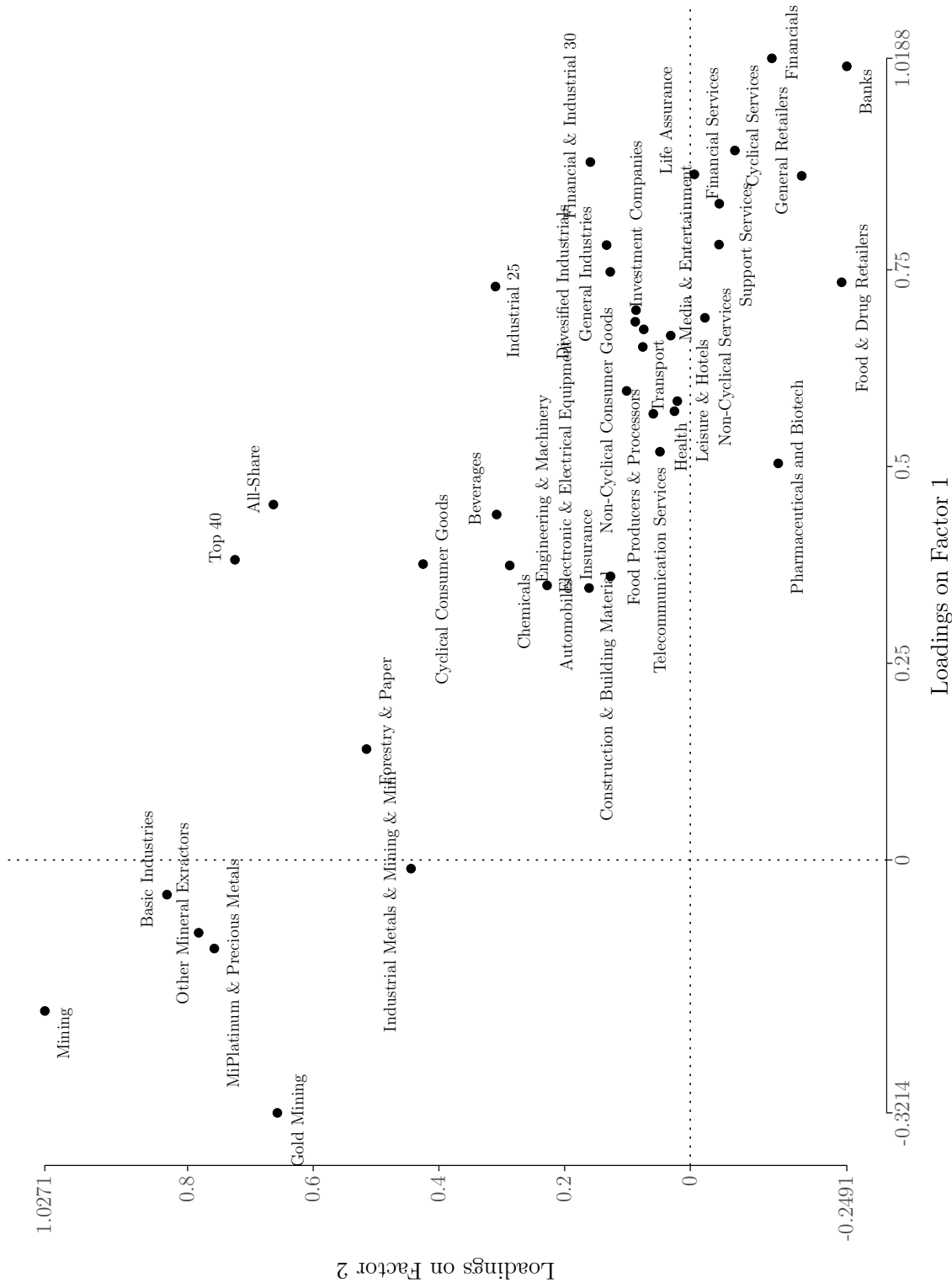


Fig. 14. Loadings Plot of Indices Using Minimum Residual Factoring

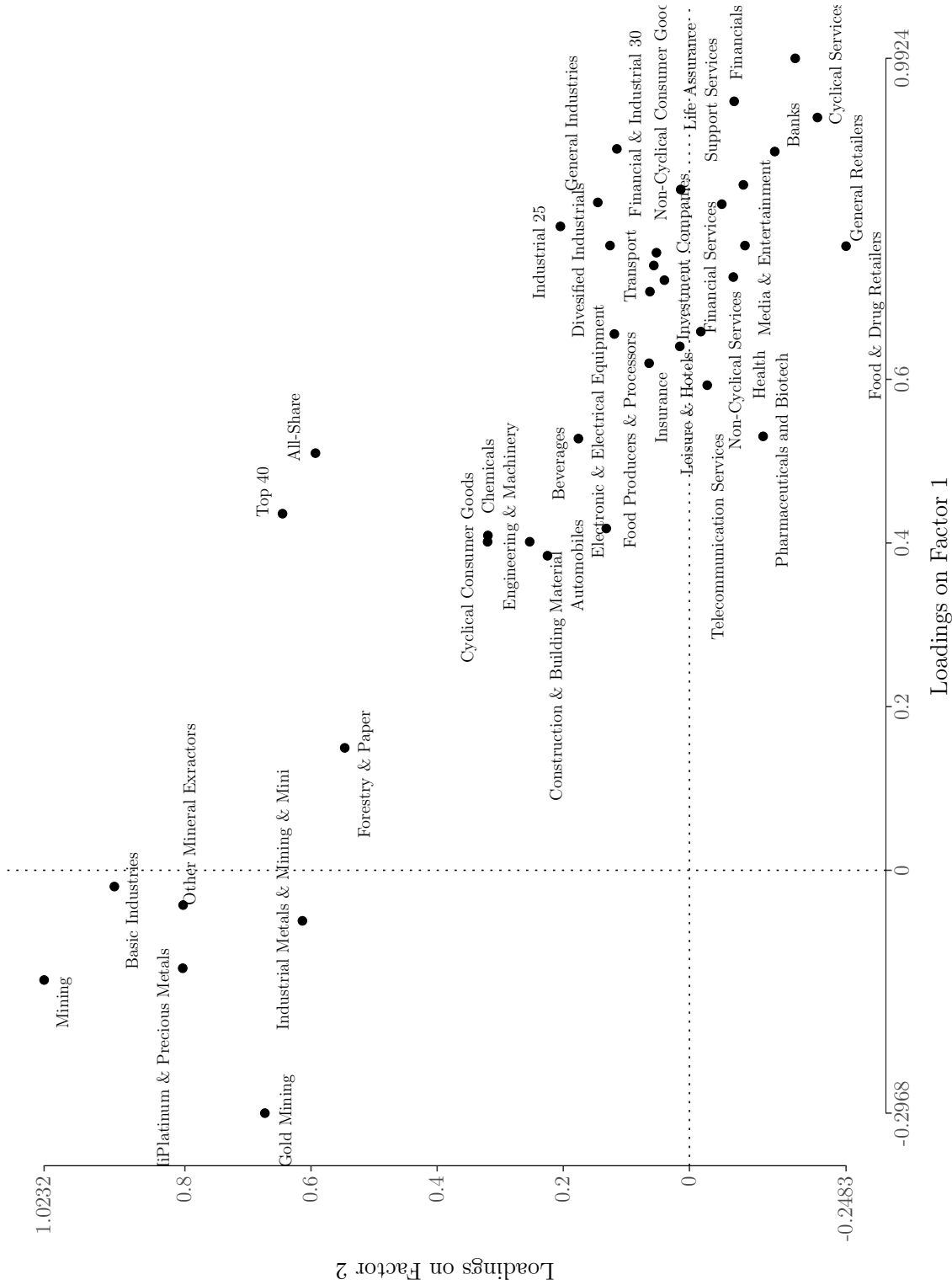


Fig. 15. Loadings Plot of Indices Using Generalised Weighted Least Squares (GLS) Factoring

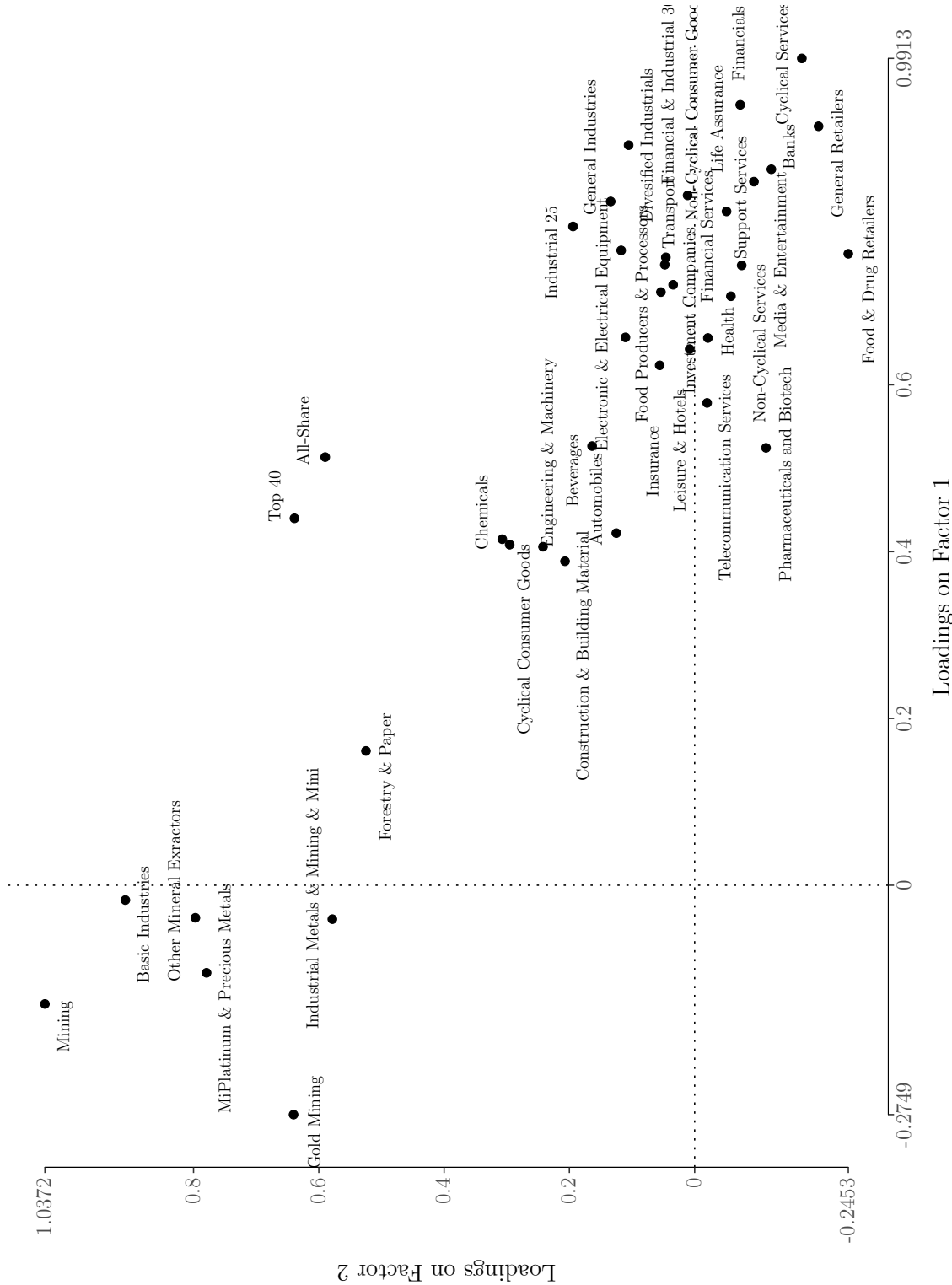


Fig. 16. Loadings Plot of Indices Using Principle Factor Analysis

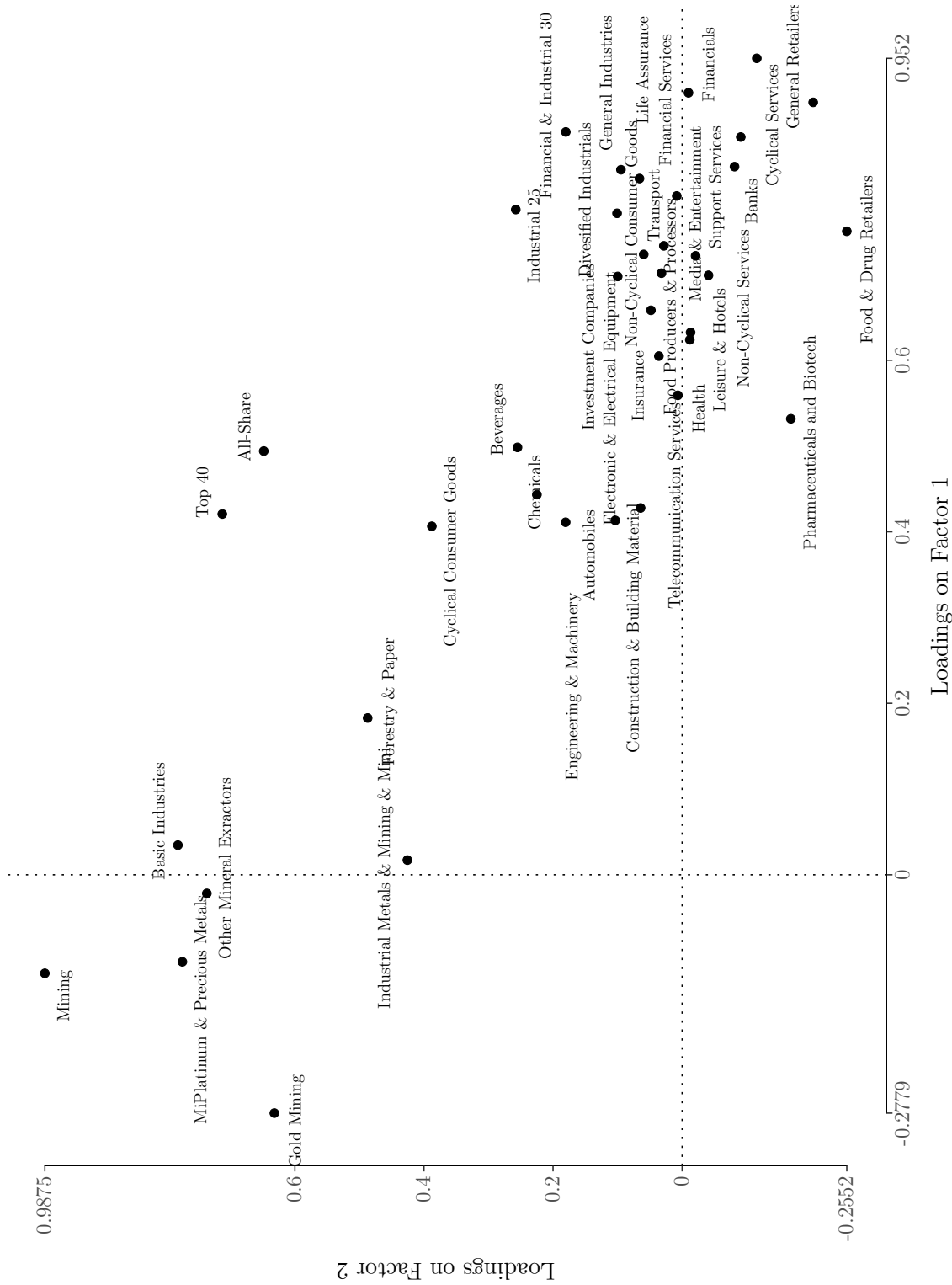


Fig. 17. Loadings Plot of Indices Using Maximum Likelihood Factor Analysis

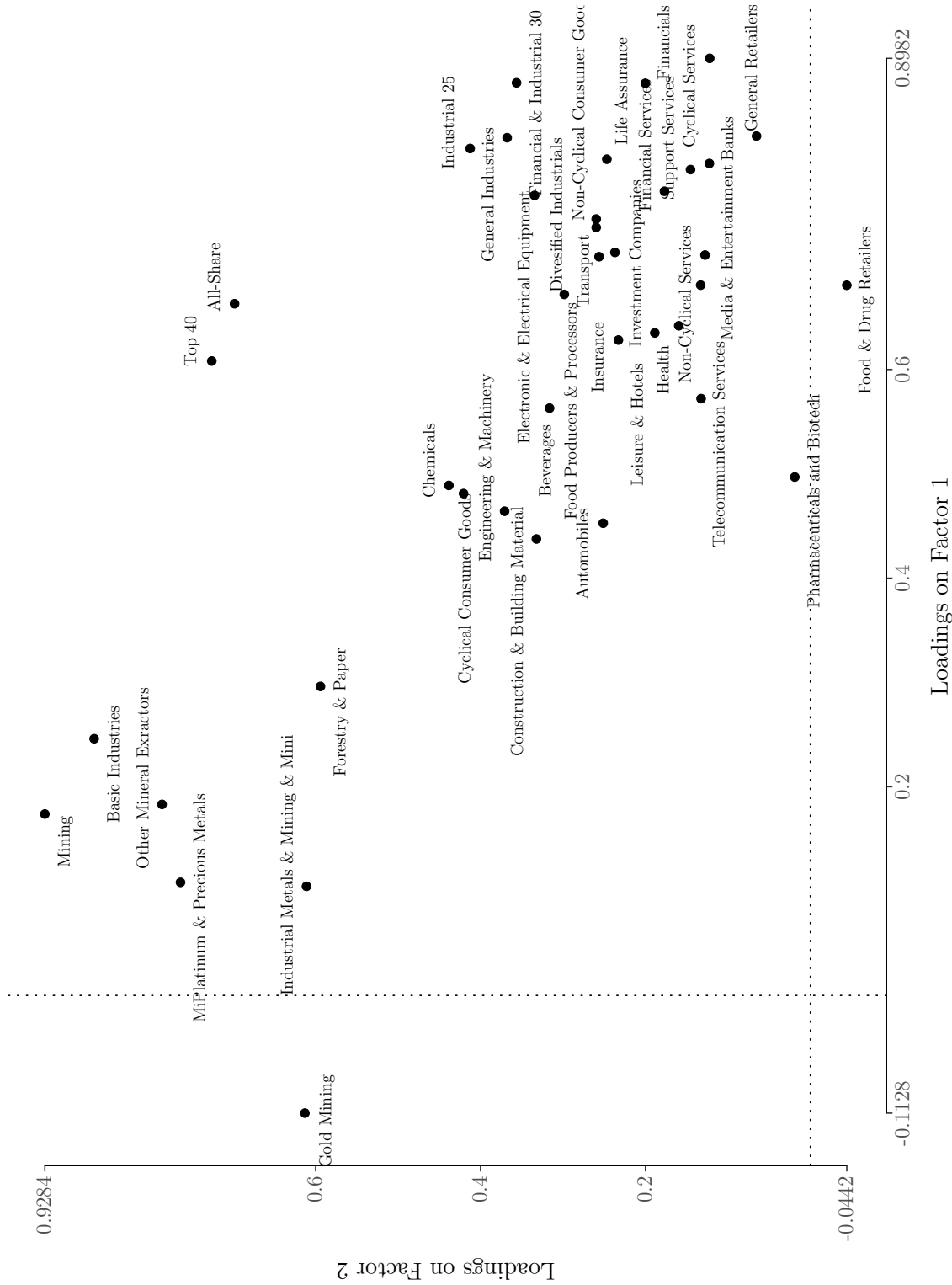


Fig. 18. Loadings Plot of Indices Using Principal Components Analysis and a Varimax Rotation.

References

- Black, F., Jensen, M. C. and Scholes, M. 1972. The Capital Asset Pricing Model: Some Empirical Tests, *Studies in the Theory of Capital Markets* 81:79–121.
- Bloomberg 2016. ‘Bloomberg Professional’.
- Blume, M. E. 1975. Betas and Their Regression Tendencies, *The Journal of Finance* 30(3):785–795.
- Bradfield, D. 2003. Investment basics XLVI: on estimating the beta coefficient, *Investment Analysts Journal* 32(57):47–53.
- Camp, R. C. and Eubank, A. A. 1981. The Beta Quotient: A new measure of portfolio risk, *The Journal of Portfolio Management* 7(4):53–58.
- Campbell, G. 1979. Risk and Return on the Johannesburg Stock Exchange, Unpublished mba thesis, University of the Witwatersrand.
- Campbell, J. Y. 2000. Asset Pricing at the Millennium, *The Journal of Finance* 55(4):1515–1567.
- Cattell, R. B. and Jaspars, J. M. F. 1967. A General Plasmode for Factor Analytic Exercises and Research, *Multivariate Behavioral Research Monographs* 3:1–212.
- Chen, N. 1983. Some Empirical Tests of the Theory of Arbitrage Pricing, *The Journal of Finance* 38(5):1393–1414.
- Chen, N., Roll, R. and Ross, S. 1986. Economic Forces and the Stock Market, *The Journal of Business* 59(3):383–403.
- Cochrane, J. H. 2011. Presidential Address: Discount Rates, *The Journal of Finance* 66(4):1047–1108.
- Copeland, T. E. and Weston, J. F. 1983. *Financial Theory and Corporate Policy*, Vol. 5, Addison-Wesley Publishing Company, Inc.
- Corhay, A. 1992. The intervalling effect bias in beta: A note, *Journal of Banking and Finance* 16(1):61–73.
- Damodaran, A. 2002. *Investment valuation*, 2 edn, Wiley, Hoboken.

- Dimson, E. 1979. Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7(2):197–226.
- Eubank, A. A. and Zumwalt, J. K. 1979. An Analysis of the Forecast Error Impact of Alternative Beta Adjustment Techniques and Risk Classes, *The Journal of Finance* 34(3):761–776.
- Fama, E. and MacBeth, J. 1973. Risk, return, and equilibrium: Empirical tests, *Journal of political economy* 81(3):607–636.
- Francis, J. C. 1986. *Investments: Analysis and Management*, McGraw-Hill series in finance, McGraw-Hill.
- Gilbertson, B. and Goldberg, M. 1981. The Market Model and the Johannesburg Stock Exchange, *Investment Analysts Journal* 10(17):40–42.
- Gonedes, N. J. 1973. Evidence on the Information Content of Accounting Numbers: Accounting-Based and Market-Based Estimates of Systematic Risk, *The Journal of Financial and Quantitative Analysis* 8(3):407–443.
- Goyal, A. 2012. ‘Empirical cross-sectional asset pricing: A survey’.
- Hendrickson, A. E. and White, P. O. 1964. Promax: A Quick Method For Rotation to Oblique Simple Structure, *British Journal of Statistical Psychology* 17(1):65–70.
- Jolicoeur, P. 1975. Linear Regressions in Fishery Research: Some Comments, *Journal of the Fisheries Research Board of Canada* 32(8):1491–1494.
- Kaiser, H. F. 1958. The varimax criterion for analytic rotation in factor analysis, *Psychometrika* 23(3):187–200.
- Kim, D. 1993. The Extent of Nonstationarity of Beta, *Review of Quantitative Finance and Accounting* 3(2):241–254.
- Kruger, R. 2005. A framework for evaluating the benchmark risk of South African equity portfolios, Master of business science, University of Cape Town.
- Laird-Smith, J., Meyer, K. and Rajaratnam, K. 2016. A study of total beta specification through symmetric regression: the case of the Johannesburg Stock Exchange, *Environment Systems and Decisions* 36(1):114–125.
- Lehmann, B. N. and Modest, D. M. 1988. The Empirical Foundations of the Arbitrage Pricing Theory, *Journal of Financial Economics* 21:213–254.

- Lintner, J. 1965a. Security Prices, Risk, and Maximal Gains From Diversification, *The Journal of Finance* 20(4):587–615.
- Lintner, J. 1965b. 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.
- Markowitz, H. 1952. Portfolio Selection, *The Journal of Finance* 7(1):77–91.
- Mayers, D. 1972. Non-marketable Assets and Capital Market Equilibrium Under Uncertainty, *Studies in the Theory of Capital Markets* pp. 223–248.
- McElroy, M. B. and Burmeister, E. 1988. Arbitrage Pricing Theory as a Restricted Nonlinear Multivariate Regression Model: Iterated Nonlinear Seemingly Unrelated Regression Estimates, *Journal of Business & Economic Statistics* 6(1):29–42.
- Mehra, R. and Prescott, E. C. 1985. The equity premium: A puzzle, *Journal of Monetary Economics* 15(2):145–161.
- Miller, M. H. and Scholes, M. 1972. Rates of return in relation to risk: A reexamination of some recent findings., *Studies in the theory of capital markets* 23.
- Mossin, J. 1966. Equilibrium in a capital asset market, *Econometrica* 34(4):768–783.
- Page, M. J. 1986. Empirical testing of the arbitrage pricing theory using data from the Johannesburg Stock Exchange, *South African Journal of Business Management* 17(1):38–42.
- R Development Core Team 2016. ‘R: A Language and Environment for Statistical Computing’.
URL: <http://www.r-project.org>
- Ricker, W. 1984. Computation and uses of central trend lines, *Canadian Journal of Zoology* 62(10):1897–1905.
- Roll, R. 1977. A critique of the asset pricing theory’s tests Part I: On past and potential testability of the theory, *Journal of Financial Economics* 4(2):129–176.
- Roll, R. and Ross, S. A. 1980. An Empirical Investigation of the Arbitrage Pricing Theory, *Journal of Finance* 35(5):1073.
- Ross, S. A. 1976. The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory* 13:341–360.

RStudio 2016. 'R Notebooks'.

URL: <https://blog.rstudio.org/2016/10/05/r-notebooks/>

Scholes, M. and Williams, J. 1977. Estimating betas from nonsynchronous data, *Journal of Financial Economics* 5(3):309–327.

Sharpe, W. F. 1963. A Simplified Model for Portfolio Analysis, *Management Science* 9(2):277–293.

Sharpe, W. F. 1984. Factor models, CAPMs, and the ABT (sic)., *Journal of Portfolio Management* 11(1):21–25.

Shiller, R. J. 1992. *Market Volatility*, MIT Press.

Strugnell, D. 2010. Beta, size and value effects on the JSE, 1994–2007, PhD thesis, University of Cape Town.

Strugnell, D., Gilbert, E. and Kruger, R. 2011. Beta, size and value effects on the JSE, 1994–2007, *Investment Analysts Journal* 40(74):1–17.

Tofallis, C. 2008. Investment volatility: A critique of standard beta estimation and a simple way forward, *European Journal of Operational Research* 187(3):1358–1367.

Treynor, J. 1961a. Market Value, Time, and Risk, Unpublished manuscript.

Treynor, J. 1961b. Toward a theory of market value of risky assets, Unpublished manuscript.

Van Rensburg, P. 1997. Investment Basics: XXXIV. The arbitrage pricing theory, *Investment Analysts Journal* 26(46):60–64.

Van Rensburg, P. 2002. Market Segmentation on the Johannesburg Stock Exchange II, *Studies in Economics and Econometrics* 26(1):83–99.

Van Rensburg, P. and Robertson, M. 2003. Size, Price-to-Earnings and Beta on the JSE Securities Exchange, *Investment Analysts Journal* 58:7–14.

Van Rensburg, P. and Slaney, K. 1997. Market Segmentation on the Johannesburg Stock Exchange, *Journal for Studies in Economics and Econometrics* 21(3):1–23.

Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.
URL: <http://ggplot2.org>

Zheng, Z., Podobnik, B., Feng, L. and Li, B. 2012. Changes in cross-correlations as an indicator for systemic risk., *Scientific reports* 2:888.