Are macroeconomic factors adequate proxies for systematic influences in stock returns? A South African perspective

(Paper RIAJ 2019 - 0122)

1. Jan J. Szczygielski

   University of Pretoria; Private Bag x20; Hatfield; Pretoria; South Africa; 0028

   kuba.szczygielski@up.ac.za
2. Leon M. Brümmer

   University of Pretoria; Private Bag x20; Hatfield; Pretoria; South Africa; 0028

   leon.brummer@up.ac.za
3. Hendrik P. Wolmarans (corresponding author)

   University of Pretoria; Private Bag x20; Hatfield; Pretoria; South Africa; 0028

   Email: hendrik.wolmarans@up.ac.za
4. Adam Zaremba

   University of Dubai, PO Box 14143; Academic City; Dubai; United Arab Emirates

   Poznań University of Economics and Business, Poznań; Poland

   Email: azaremba@ud.ac.ae

Declarations of interest: none.

ABSTRACT

We investigate whether macroeconomic factors adequately proxy for systematic influences in stock returns within the South African context. We also investigate whether a commonly used solution to factor omission in macroeconomic factor models, the residual market factor, adequately reflects systematic influences not reflected by a set of macroeconomic factors. Our contribution lies in precisely quantifying the ability of macroeconomic and residual market factors to proxy for systematic drivers of returns. Systematic influences are represented by statistically derived factor scores which are then related to a set of carefully selected macroeconomic factors. We find that the identification of macroeconomic factors that proxy for systematic influences is a challenge in itself. Once identified, macroeconomic factors are poor and unstable proxies for systematic influences and the use of a residual market factor does not significantly improve the approximation of factor scores. Our conclusion is that macroeconomic linear factor models are likely to be underspecified, even if a residual market factor is included. This has implications for researchers, investors, econometricians and economists that rely on macroeconomic factor models to study financial markets.

Keywords: macroeconomic factors, factor scores, linear factor model, systematic influences, residual market factor, underspecification

JEL classification: C01, C22, C58, G12

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

1. Introduction

The foundation of multifactor time-series models that relate stock returns to factors representative of systematic economy-wide influences is the linear factor model that underpins the Arbitrage Pricing Theory (APT) (Szczygielski and Chipeta, 2015; Elton & Gruber, 2018). However, studies that seek to model the return generating process using factor sets suggest that macroeconomic factors are poor proxies for systematic influences in stock returns (Chang,1991; Chen & Jordan,1993; Connor, 1995; Van Rensburg, 1997, 2000; Middleton & Satchell, 2001; Spyridis, Sevic & Theriou, 2012). Nevertheless, macroeconomic factors have an advantage over characteristic-based factors such as size, value and momentum and statistical factors in that they can readily be linked through economic reasoning to risks that investors face (French, 2017). Furthermore, equity factors constructed for the purposes of factor investing are dependent upon the macroeconomic environment (Amenc, Esakia, Goltz & Luyten, 2019). Macroeconomic factors in models therefore matter from an inferential and investor perspective and cannot be readily discarded. To address the poor ability of macroeconomic factors in explaining returns, a residual market factor can be used to control for influences not reflected in a factor set. The residual market factor constitutes the residuals of an auxiliary regression of a factor set onto a broad market index and fulfils the role of a catch-all proxy for omitted factors and is widely viewed as addressing the problem of factor omission (Deetz, Poddig, Sidorovitch & Varmaz, 2009; Czaja, Scholz & Wilkens, 2010). Furthermore, a general gap and oversight in the literature is that macroeconomic linear factor models do not explicitly acknowledge that total return variation is explained by systematic *and* idiosyncratic components (Greene, 2012). Macroeconomic models rely upon macroeconomic factors to proxy systematic influences that explain systematic variation but report upon total variation explained. If systematic components have a low level of explanatory power in the first instance, that is, most of the variation is attributable to idiosyncratic components, then the $\bar{R}^2$ will be a poor indicator of the ability of macroeconomic factors to proxy for systematic influences. This would be the case if adequate diversification is not fully achievable and idiosyncratic components are not completely diversified away (Wei, 1988).

We offer a comprehensive investigation into the ability of macroeconomic factors and the residual market factor to represent systematic influences that are believed to drive stock prices. Additionally, we consider the most extensive set of factors to date in our search for proxies relative to other similar South African studies and most international studies (Van Rensburg, 1996, 2000; Birz & Lott, 2011). Moreover, our contribution lies in that we precisely quantify the ability of macroeconomic factors and a residual market factor (or factors) to proxy for systematic influences. Relatedly, another contribution is that we separate and measure systematic components directly by using factor analysis. We therefore address the shortcomings of previous studies that draw inferences upon measures of total variation. Finally, and as part of robust checks, we extend our unrestricted specification to include South African versions of characteristic-based Fama and French (1993) size and value factors, a Carhart (1997) momentum factor and two additional factors, namely the Fama and

French (2015) profitability and investment factors. Our study focuses on the South African stock market, a large emerging stock market that has not hitherto been explored to the same extent as its European, South American and Asian peers.

In our analysis, we use factor analysis to derive factor scores that represent systematic components in returns. Macroeconomic factors are then tested to determine whether and how well these factors proxy for systematic components. We then consider the efficacy of the residual market factor in addressing underspecification. We also attempt to gain insight into one of the potential reasons for the poor ability of macroeconomic factors to proxy for systematic influences. To do so, we estimate breakpoint and rolling regressions to investigate the stability of the relationships between factor sets and representations of systematic influences.

Key findings are that the identification and selection of macroeconomic factors that qualify as proxies for systematic influences is a challenge. Once identified, macroeconomic factors are shown to be poor proxies. Although this is recognized in the literature, it is also shown to be the case in the South African context. The use of orthogonal residual market factors does not significantly improve the approximation of derived factor scores. This suggests that this approach is unlikely to be effective in addressing factor omission in macroeconomic factor models. A potential reason for the poor ability of macroeconomic factors to proxy for systematic influences is the instability of systematic-macroeconomic factor relationships which appears to be related to the economic cycle. Our results are robust to the inclusion of characteristic-based factors which contribute to the approximations of systematic influences. We go onto outline the implications of our findings and recommend that a factor analytic augmentation should be considered together with macroeconomic factors in macroeconomic model specifications to account for omitted and unspecified factors.

The study should be of interest to researchers, investors, econometricians and economists that apply macroeconomic factor models. A failure to adequately reflect influences that impact returns will have an adverse impact on the econometric modelling of return-macroeconomic factor relationships and the asset pricing applications of such models. It will also impact inference making and result in misleading insights. In summary, macroeconomic linear factor models are likely to suffer these consequences and the use of a residual market factor is unlikely to alleviate factor omission bias. This should not be ignored.

2. Literature Review

Any representation of the return generating process requires that the number of factors is stipulated. By deriving factor scores by applying factor or principal component analysis techniques, composite representations of systematic influences can be obtained. Roll and Ross (1980) derive a five-factor structure from returns on stocks listed on the New York Stock Exchange (NYSE) and the American Stock Exchange (ASE). Elton and Gruber (1988) derive a four-factor structure to describe the return generating process of the Japanese stock market. Yli-Olli and Virtanen (1993) identify three to four factors in Finnish stock returns.

For the South African stock market, Page (1989) suggests a two-factor structure for returns on 30 randomly selected JSE stocks. Biger and Page (1993) propose a three-factor structure for unit trust returns. Van Rensburg and Slaney (1997) suggest that there are at least two but no more than three factors. Chimanga and Kotze (2009) apply principal component analysis JSE and propose a two-factor model. As with developed markets, the South African stock market is characterised by multiple systematic factors.

The limitations of statistical techniques in factor extraction and identification and the interpretability of statistical factors are what spurs the use of pre-specified macroeconomic factors (Priestley, 1996; Chimanga & Kotze, 2009: 82). In such studies, macroeconomic factors are assumed to proxy for systematic influences in stock returns, represented by factor scores (Chen & Jordan, 1993; Connor, 1995; Panetta, 2002). The seminal works of Chan, Chen and Hsieh (1985), Chen *et al.* (1986) and Hamao (1988) are the first to formally establish a link between returns and macroeconomic factors. All three works refer to the dividend discount model to justify the consideration of macroeconomic factors which postulates asset prices as a function of expected cash-flows and a discount rate. They argue that macroeconomic aggregates impact returns through their impact on expected cash flows and/or a discount rate, which in turn, are reflected in asset prices (see also Birz & Lott, 2011: 2793; Al-Tamimi, Alwan & Rahman, 2011: 3). Chan *et al.* (1985) demonstrate that the monthly growth rate in industrial production, the unanticipated inflation rate, changes in expected inflation, changes in the term structure of interest rates, the default spread and changes in the business cycle impact expected returns. Chen *et al.* (1986) find that the monthly industrial production growth rate, the unexpected inflation rate, the default spread and the term structure of interest rates explain returns. Hamao (1988) find that these same factors explain returns for the Japanese stock market.

Although appealing, macroeconomic factors appear to be poor proxies for systematic influences. Middleton and Satchell (2001) argue that the problem of underspecification can only be avoided if factors are derived statistically and a sufficiently significant number of factors is arrived at. Van Rensburg (2000) states that the assumption of uncorrelated residuals across assets, $E(\varepsilon_{it}, \varepsilon_{jt}) = 0$, is likely to be violated for specifications that employ pre-specified macroeconomic factors as explanatory factors. A violation of this assumption suggests that residual co-movement is attributable to the inability of a factor set to adequately proxy for factors that drive returns (Elton, Gruber, Brown & Goetzmann, 2014). Empirical evidence that macroeconomic factors by themselves are poor proxies for systematic influences is readily found in studies that estimate macroeconomic factor models. Burmeister and Wall (1986) and Connor (1995) show that macroeconomic factors by themselves explain less than 30% of the total variation in US returns. Van Rensburg (1997) extracts two factors that explain an additional 43.9% of variation from the residuals of a macroeconomic factor specification explaining returns on the JSE. This implies that there are common factors that are relegated to the residuals that are not reflected by the macroeconomic factor set. Szczygielski and Chipeta (2015) propose an eight-factor model to describe returns on the JSE All Share Index. Macroeconomic factors by themselves explain around a fifth of the variation in returns. Chen and Jordan (1993) extract five statistical factors from returns for firms in the CRSP database and find that the macroeconomic factors approximate between 2.2% and 49.5% of the factor scores. The low level of explanatory power suggests that the macroeconomic factor set fails to adequately approximate systematic influences.

The solution to factor omission in macroeconomic factor models is to include a residual market factor, which is assumed to act as a catch-all proxy for omitted influences (Deetz *et al,* 2009; Czaja *et al.,* 2010). Burmeister and Wall (1986) report substantial increases in the $\bar{R}^2$s when a residual market factor derived from returns on the S&P500 is included in their respective specifications. Chang (1991) compares the mean residuals of a (purely) macroeconomic factor model and a model that comprises the same set of macroeconomic factors and a residual market factor. For two of the three subperiods, mean residuals from the macroeconomic factor model are significantly different from zero. The inclusion of the residual market factor results in a substantial reduction in mean errors and insignificance, suggesting that the residual market factor effectively reflects omitted information. Van Rensburg (2000) reports that the $\bar{R}^2$ for a model relating returns on the JSE All Share Index to five factors increases from 0.29 to 0.91 when two residual market factors derived from two industrial sector indices are incorporated into the model. A limitation of Van Rensburg's (2000) study is that it considers the efficacy of residual market factors at an aggregate market level, for a single series. The inclusion of residual market factors has a positive impact on the explanatory power of models utilizing macroeconomic factors. Nevertheless, as only the true market portfolio will reflect all return generating factors, and such a portfolio is impossible to identify or construct, a broad market index is unlikely to produce a residual market factor that fully reflects omitted systematic influences (Born & Moser, 1988; Gadzinski, Schuller & Vacchino, 2018).

The preceding studies suggest that macroeconomic factors are poor proxies for systematic influences in returns. However, they do not directly relate macroeconomic factors to purely systematic components and establish their adequacy as proxies. Also, the ability of a residual market factor to proxy for purely systematic components is not considered. We address these gaps in this study.

3. Data and Methodology

3.1. Data and Sample Period

The exploration of the factor structure of returns begins with factor analysis of the return series comprising the sample. Monthly index data is obtained from the IRESS Expert database, comprising industrial sectors constituting the South African stock market, the JSE, and spans the period January 2001 to December 2016. Only industrial sectors with a full data history are included in the sample, constituting data for 26 industrial sectors (out of a total of 33 sectors at the time of writing). Month-end data is used and the risk-free rate used to derive (continuously compounded total monthly) excess returns is the closing yield on the R186 government bond (Nel, 2011, PWC, 2015). [i]

The sample period coincides with numerous notable economic events that have a potential impact on the South African stock market, given its integration with global financial systems. These include the terrorist attack on the World Trade Centre in 2001, the war on terror in Iraq (2003), the ongoing emergence of China and India as economic powers, the Russian financial crisis (2008-2009, 2014-present), the sub-prime mortgage crisis of 2007 and 2008 and the Great Recession that follows, the European sovereign debt crisis (2009 onwards) and the early stages of the British withdrawal from the European Union (2016 onwards). South Africa has also experienced changing socio-economic conditions during this period. These are a commodity driven economic boom prior to 2008, increasing trade with India and China, a recession for the first time since 1994, the hosting of the FIFA World Cup in 2010 and a shifting political landscape (2016 municipal elections).

3.2. Factor structure analysis

A scree test is first conducted on the return correlation matrix. The scree plot and the resultant flexion point are indicative of the number of true systematic factors in returns (Merville, Hayes-Yelken & Xu, 2001). The second test applied to identify the number of factors is the minimum average partial (MAP) test. This approach is congruent with the assumption of uncorrelated residuals, $E(\varepsilon_{it}, \varepsilon_{jt}) = 0$, which underlies all linear factor models (Zwick & Velicer, 1986). Factor scores are derived using the Bartlett (1937) method, which produces scores that are most likely to represent the true factor scores (Bartlett, 1937; DiStefano, Zhu & Mîndrilă, 2009). To simplify the factor structure and to ensure that the identified macroeconomic factors are related to as few factor score series as possible as opposed to multiple factor score series – thereby aiding interpretability - an orthogonal varimax rotation is undertaken. Therefore, varimax rotation minimizes the number of specific macroeconomic factors that load onto each extracted factor (Yong & Pearce, 2013).

3.3. Identification of macroeconomic factor proxies for systematic influences

Next, we identify a broad set of macroeconomic factors that are potential proxies for systematic influences. The Quantec EasyData database is used to obtain macroeconomic data to construct macroeconomic factors. Factors are categorized, with categories representative of real activity, prices, cyclical indicators, exchange rates, monetary factors, commodities, interest rates and trade (see Table 1B in Appendix 1 for a list of factors considered). In the spirit of macroeconomic linear factor model literature, the dividend discount model is used in the preliminary screening of factors. The model assumes that asset prices are a function of expected cash-flows and/or the discount rate and therefore any macroeconomic factor innovation that impacts either of these or both can be considered (Birz & Lott, 2011; Al-Tamimi *et al.,* 2011):

$$S_{it} = \sum_{t=1}^{\infty} \frac{E(D_{it})}{(1+z)^t} \qquad (1)$$

where S_{it} is the price level for index i at time t, E is an expectation operator relating to expectations about future cash flows as denoted by D_{it}, and the discount rate prevailing between t and $t + \tau$ is denoted by z.

Macroeconomic factors should enter the linear factor model as innovations/unanticipated changes (Birz & Lott, 2011; Bessler & Kurmann, 2014). To derive innovations, correlograms are inspected and if significant correlation is detected at a specific order(s), up to 12 orders, an autoregressive time series model is estimated to remove components that permit predictability. The residuals of this specification are taken as a representation of unexpected components. The broad set of qualifying factors, in innovations, is listed in Table 1B of Appendix 1. For a factor to be considered as having a systematic impact, it must fulfil two conditions, namely 1) it must be correlated with at least half of the return series in the sample and 2) it must be correlated with returns on the market aggregate, the JSE All Share Index. As macroeconomic data is often subject to revisions and/or lags in announcements, each factor enters the factor-return correlation matrix contemporaneously and with up to three lags (Bilson, Brailsford & Hooper, 2001; Panetta, 2002). Correlations between returns on the JSE All Share Index and macroeconomic factors are tested using Pearson's (ordinary) correlation and confirmed using non-parametric Spearman's (rank) correlation coefficients (Bishara & Hittner, 2012). Factors that meet the two conditions are taken forward in the analysis.

3.4. Factor score regressions

To establish whether the macroeconomic factors retained for further analysis are proxies for the systematic influences in returns, the approach of Chen and Jordan (1993), Panetta (2002) and Spyridis *et al.,* (2012) is adopted. First, the retained macroeconomic factors are regressed onto the (rotated) factor scores for each statistical factor, F_{nt}:

$$F_{nt} = \alpha + \sum_{k=1}^{K} b_{nk} f_{kt} + \varepsilon_{nt} \tag{2}$$

where b_{nk} is the sensitivity of statistical factor *n* to macroeconomic factor *k*. It follows that if macroeconomic factors are significantly related to at least some of the statistical factors in this multifactor analysis, then these factors are proxies for systematic influences. Equation (2) is then augmented with a residual market factor derived by regressing returns on the JSE All Share Index onto the set of identified macroeconomic factors and using the residuals as a factor in a subsequent regression:

$$F_{nt} = \alpha + \sum_{k=1}^{K} b_{nk} f_{kt} + b_{nM\varepsilon} M\varepsilon_t + \varepsilon_{nt} \tag{3}$$

where all parameters are as in equation (2), but $M\varepsilon_t$ is the residual market factor and $b_{nM\varepsilon}$ is the coefficient on $M\varepsilon_t$. As the residual market factor is uncorrelated by construction with the remaining factors, it has no impact on coefficients associated with the remaining factors (Deetz *et al.,* 2009; Czaja *et al.,* 2010). A residual market factor can also act as a test factor; if identified macroeconomic factors are adequate proxies for underlying systematic factors, a second residual market factor will be redundant (Chang, 1991; Kryzanowski, Lalancette & To, 1994). Therefore, a second residual market factor, derived by regressing returns on the MSCI

World Market Index onto the macroeconomic factor set and the residual market factor derived from returns on the JSE All Share Index, is incorporated in addition to the existing factor set in equation (3):

$$F_{nt} = \alpha + \sum_{k=1}^{K} b_{nk} f_{kt} + b_{nM\varepsilon} M\varepsilon_t + b_{nIM\varepsilon} IM\varepsilon_t + \varepsilon_{nt} \tag{4}$$

where all parameters are as in equation (2) and (3) and $b_{nIM\varepsilon}$ is the sensitivity to the second residual market factor, $IM\varepsilon_t$, derived from the MSCI World Market Index.

As an extension of the analysis and a robustness test, we control for characteristic-based factors in the unrestricted model. We incorporate the Fama and French (1993) size (SMB_t) and value (HML_t) factors, the Carhart (1997) momentum (UMD_t) factor and two additional factors, the Fama and French' (2015) profitability (RMW_t) and investment (CMA_t) factor into equation (5) (we relegate these results to Table 1C in the supplementary appendix). Prior research has shown that certain characteristic-based factors are linked to the macroeconomic state and therefore, by extension, these factors should also proxy for pervasive influences in returns (Aretz, Bartram & Pope, 2010):

$$F_{nt} = \alpha + \sum_{k=1}^{K} b_{nk} f_{kt} + b_{nM\varepsilon} M\varepsilon_t + b_{nIM\varepsilon} IM\varepsilon_t + b_{nSMB} SMB_t + b_{nHML} HML_t + b_{nUMD} UMD_t + \\ + b_{nRMW} RMW_t + b_{nCMA} CMA_t + \varepsilon_{nt} \tag{5}$$

Equations (2), (3), (4) and (5) are estimated using least squares regression with Newey and West (1987) heteroscedasticity and autocorrelation consistent (HAC) standard errors. In this step, the intention is not only to confirm that the selected macroeconomic factors are proxies for systematic influences but also to determine how well these factors proxy for systematic influences. We treat the $\bar{R}^2$ as an indicator of the proportion of each statistical factor explained by the macroeconomic factor set, the macroeconomic factor set and $M\varepsilon_t$, the full factor set that includes both residual market factors and finally the full specification in equation (5) (Aretz *et al.,* 2010; Bessler, Kurmann & Nohel, 2015). The F-test is applied to test the joint significance of the coefficients and to confirm the overall significance of the approximation. This is also a test of the null hypothesis that the $\bar{R}^2$ is zero (Sadorsky, 2008). Consideration is also given to the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values (Spiegelhalter, Best, Carlin & Van der Linde, 2014).

3.5. Model stability and rolling regressions

We investigate the stability of macroeconomic factors in proxying for systematic influences by re-estimating equations (2), (3) and (4) as breakpoint least squares regressions, with breakpoints identified using the Bai-Perron test (Bai & Peron, 1998; Hansen, 2012). This test also identifies the timing of breaks and estimates the parameters associated with each segment (Carlson, Craig & Schwarz, 2000). For conciseness, the investigation of the stability of the representation of the factor scores is restricted to equation (4) (abridged results for equations (2) & (3) are reported in Table 1D of the supplementary appendix). Next and following Bessler & Kurumann (2014), we also re-estimate equations (2), (3), (4) and (5) for each factor as rolling regressions to determine whether the ability of the factor sets considered differs in proxying for systematic influences over time and during periods of economic upturns and downturns. If macroeconomic factors are shown to be poor proxies for systematic influences in returns, then this approach provides a possible explanation for such a finding. Macroeconomic factors may be good proxies for systematic influences only during short intervals, with overall significance attributable to short-term intermittent relationships that may be related to the economic state (McQueen & Roley, 1993; Panetta, 2002; Spyridis *et al.*, 2012).

4. Results and Discussion

4.1. The factor structure of the South African stock market

Figure 1 reports the results of the scree test and Table 1 reports the proportion of total variance explained by each factor and the results of the MAP test (Section 3.2).[ii]

Figure 1. Scree Plot of factor structure of the South African stock market

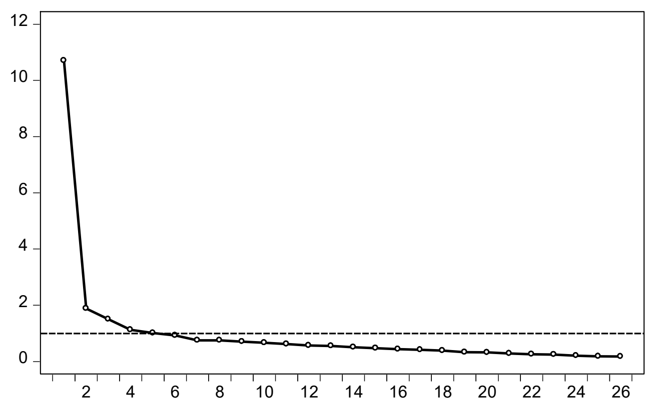


Table 1. Proportion of systematic variance explained and MAP test results

The

The

Panel A: Proportion of Variance Explained by Each Factor		
Factor Number	**Proportion of Variance Explained**	
1	0.412	
2	0.072	
3	0.058	
4	0.043	
5	0.039	
6	0.036	
7	0.029	
8	0.029	
9	0.027	
10	0.026	
Total	**0.770**	
Panel B: MAP Test Results		
Factors extracted	**Mean Communality**	**Mean Uniqueness**
3	**0.483**	**0.516**

Notes: Mean Communality is the mean proportion of common variance explained across return series by the statistical factors extracted on the basis of the MAP test. Mean Uniqueness is the mean proportion of variance across return series attributable to the return series themselves and not to systematic factors.

The distinct flexion point in Figure 1 indicates that two common factors characterize the structure of the South African stock market. Derived eigenvalues indicate that together, the first three factors explain a total of 54.2% of common variation in returns. This proportion is similar to that in Van Rensburg and Slaney (1997) who find that two extracted factors explain 50% of shared variation in returns whereas a third factor explains 4% of shared variation. The results of the MAP test reported in Panel B of Table 2 indicate a three-factor structure. The mean communality for a three-factor structure is 48.4%, suggesting that almost half of the common variation in South African stock returns is accounted for by these three common factors. These findings are in line with earlier work on the JSE (Biger & Page, 1993; Van Rensburg & Slaney, 1997). As the MAP test identifies the most important factors and the number of factors that result in the closest approximation of the assumption of uncorrelated residuals, factor scores are derived for a three-factor solution (Elton & Gruber, 1988; Elton *et al.*, 2014).

4.2. Proxy factor selection and factor score regressions

We identify only a handful of factors that qualify for further consideration (see Section 3.3). These are (the innovations) in the number of building plans passed, BP_{t-1}, the domestic composite cyclical leading indicator, $LEAD_{t-1}$, business activity, BUS_t, fluctuations in the Rand-Dollar exchange rate, USD_t, world metal prices, MET_t, long-term government bond yields, LTY_t and the leading indicator for South Africa's trading partners, TLI_t. Although all factors and categories are considered, for some categories, no factors are chosen. While certain factors appear to be potential candidate factors, they are disqualified following further consideration. Broad reasons for the disqualification of other seemingly qualifying factors are the lack of a systematic impact, lack of a correlation with the JSE All Share Index and strong correlation with other qualifying factors. The

Table 2. Factor score regressions

	A. Macroeconomic Factors			B. Macroeconomic Factors & $M\varepsilon_t$			C. Macroeconomic Factors, $M\varepsilon_t$ & $IM\varepsilon_t$		
Factor	F_{1t}	F_{2t}	F_{3t}	F_{1t}	F_{2t}	F_{3t}	F_{1t}	F_{2t}	F_{3t}
α	0.007	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
BP_{t-1}	1.339**	-0.364	0.401	1.339**	-0.364	0.401	1.339**	-0.364	0.401
$LEAD_{t-1}$	15.467*	7.756	15.460*	15.467*	7.756	15.460**	15.467*	7.756	15.460**
BUS_t	1.661**	0.766	0.565	1.661**	0.766	0.565	1.661**	0.766	0.565
$USD\varepsilon_t$	-7.763***	-3.621	6.531***	-7.763***	-3.621	6.531***	-7.763***	-3.621	6.531***
MET_t	-1.479	8.713***	1.372	-1.479	8.713***	1.372	-1.479	8.713***	1.372
LTY_t	-146.237***	-28.506	40.792	-146.237***	-28.506	40.792*	-146.237***	-28.506	40.792*
$TLI\varepsilon_t$	10.485	48.334*	82.364***	10.485	48.334**	82.364***	10.485	48.334**	82.364***
$M\varepsilon_t$				5.430***	9.081***	15.897***	5.430***	9.081***	15.897***
$IM\varepsilon_t$							4.317	1.174	3.918*
$\bar{R}^2$	0.246	0.161	0.131	0.285	0.259	0.431	0.294	0.255	0.437
F-statistic	11.962***	2.641**	6.555***	11.690***	4.695***	21.852***	12.739***	4.570***	18.458***
AIC	2.727	2.997	3.057	2.679	2.879	2.637	2.671	2.888	2.632
BIC	2.863	3.133	3.193	2.832	3.032	2.79	2.840	3.058	2.801

Notes: The asterisks, ***, ** and *, indicate statistical significance at the respective 1%, 5% and 10% levels of significance. Least squares with Newey-West heteroscedasticity and autocorrelation consistent (HAC) standard errors used for estimation purposes. F_{1t}, F_{2t} and F_{3t} are the statistical factor scores of the three respective factors derived using factor analysis in Section 4.1. All factors are in innovations, defined as follows: BP_{t-1}- number of building plans passed, $LEAD_{t-1}$- domestic composite cyclical leading indicator, BUS_t - business activity, $USD\varepsilon_t$- (orthogonalized) fluctuations in the Rand-Dollar exchange rate, MET_t- world metal prices, LTY_t- long-term government bond yields, $TLI\varepsilon_t$- (orthogonalized) leading indicator for South Africa's trading partners, $M\varepsilon_t$- the residual market factor orthogonal to the macroeconomic factor set, derived from returns on the JSE All Share Index, $IM\varepsilon_t$- a second residual market factor orthogonal to the macroeconomic factor set and $M\varepsilon_t$.

selection of only seven macroeconomic factors out of a broader set demonstrates the challenges associated with identifying proxies for systematic influences in returns.

Having identified macroeconomic factors that are candidate proxies for systematic influences, the following specifications incorporating these factors and the residual market factors, are estimated (see Section 3.4):

$$F_{nt} = \alpha + b_{nBP}BP_{t-1} + b_{nLEAD}LEAD_{t-1} + b_{nBUS}BUS_t + b_{nUSD\varepsilon}USD\varepsilon_t + b_{nMET}MET_t + b_{nLTY}LTY_t + b_{nTLI}TLI\varepsilon_t + \varepsilon_{nt} \quad (6)$$

$$F_{nt} = \alpha + b_{nBP}BP_{t-1} + b_{nLEAD}LEAD_{t-1} + b_{nBUS}BUS_t + b_{nUSD\varepsilon}USD\varepsilon_t + b_{nMET}MET_t + b_{nLTY}LTY_t + b_{nTLI}TLI\varepsilon_t + b_{nM\varepsilon}M\varepsilon_t + \varepsilon_{nt} \quad (7)$$

$$F_{nt} = \alpha + b_{nBP}BP_{t-1} + b_{nLEAD}LEAD_{t-1} + b_{nBUS}BUS_t + b_{nUSD\varepsilon}USD\varepsilon_t + b_{nMET}MET_t + b_{nLTY}LTY_t + b_{nTLI}TLI\varepsilon_t + b_{nM\varepsilon}M\varepsilon_t + b_{IM\varepsilon}IM\varepsilon_t + \varepsilon_{nt} \quad (8)$$

where in equation (6), (7) and (8), F_{nt} is the factor score for factor *n* at time *t*. The respective betas, *b*s, represent the sensitivities of the factor scores to innovations in macroeconomic factor *k* in equation (6), and the residual market factor derived from returns on the JSE All Share Index, $M\varepsilon_t$, in equation (7) and also the second residual market factor derived from returns on the US Dollar denominated MSCI World Market Index, $IM\varepsilon_t$, in equation (8).[iii] Our final specification, testing robustness and the omission of specific influences, incorporates characteristic-based factors, SMB_t and HML_t, UMD_t, RMW_t and CMA_t, in addition to the factors in equation (8) above:

$$F_{nt} = \alpha + b_{nBP}BP_{t-1} + b_{nLEAD}LEAD_{t-1} + b_{nBUS}BUS_t + b_{nUSD\varepsilon}USD\varepsilon_t + b_{nMET}MET_t + b_{nLTY}LTY_t + b_{nTLI}TLI\varepsilon_t + b_{nM\varepsilon}M\varepsilon_t + b_{IM\varepsilon}IM\varepsilon_t + b_{nSMB}SMB_t + b_{nHML}HML_t + b_{nUMB}UMD_t + b_{nRMW}RMW_t + b_{nCMA}CMA_t + \varepsilon_{nt} \quad (9)$$

where all parameters are as in equation (6), (7) and (8).

The results in Table 2 indicate that all macroeconomic factors and the residual market factors are statistically significant across factor regressions. F_{1t} is approximated by a combination of factors across panels, namely BP_{t-1}, $LEAD_{t-1}$, BUS_t, $USD\varepsilon_t$ and LTY_t, mostly local in character, with this being especially the case for LTY_t which can be viewed as a proxy for local fiscal and political risks (Baldacci, Gupta & Mati, 2011). Interestingly, F_{2t} is strongly related to world metal prices, MET_t, suggesting that this factor mostly represents the impact of commodity prices on the South African stock market. It is also weakly related to the economic conditions experienced by South Africa's trading partners, $TLI\varepsilon_t$, providing support for the international orientation associated with commodity prices. Finally, F_{3t} is strongly related to the economic conditions experienced by South Africa's trading partners, $TLI\varepsilon_t$, and weakly to $IM\varepsilon_t$ (in Panel C of Table 2). Also, it is positively related to $USD\varepsilon_t$, as opposed to negatively as for F_{1t}, suggesting that it captures a different aspect of the information reflected in this factor. It is also related to the domestic composite cyclical leading indicator, $LEAD_{t-1}$, suggesting that F_{3t} mostly reflects global economic conditions and components of the domestic business cycle that are associated with global economic conditions. The F-statistic is statistically significant across all specifications, confirming the overall significance of the approximation of the factor scores.

The unrestricted model in Panel C of Table 2 produces the optimal approximation for F_{1t} and F_{3t}. The $\bar{R}^2$ for F_{1t} is 0.294 and for F_{3t}, it is 0.437 although the increase in the $\bar{R}^2$ relative to that in Panel B for both factor specifications is marginal. This suggests that most of the information that would be reflected by unorthogonalised returns on the MSCI World Market Index, is already reflected by the macroeconomic factors and $M\varepsilon_t$. For F_{2t}, the $\bar{R}^2$ is 0.255. Importantly, the $\bar{R}^2$ values for each factor regression are far below one. This is especially true for the factor regressions with only the macroeconomic factors as the proxy factors in Panel A; the respective $\bar{R}^2$s are 0.246, 0.161 and 0.131 for F_{1t}, F_{2t} and F_{3t}. Concerningly, in Panel B, the coefficient on $M\varepsilon_t$is highly statistically significant across factor regressions and in Panel C, the coefficient on $IM\varepsilon_t$ is statistically significant for F_{3t}. A highly significant sensitivity to the residual market factor in Panel B implies that the macroeconomic factor set fails to proxy for all systematic influences in returns and that the residual market factor proxies for systematic influences, in addition to the macroeconomic factors (Van Rensburg, 1995). The respective $\bar{R}^2$ s in Panel B remain low, namely 0.285 for F_{1t}, 0.259 for F_{2t} and 0.431 for F_{3t}. Nevertheless, the relatively lower AIC and BIC values suggest that the inclusion of $M\varepsilon_t$ translates into better factor score approximations and a more adequate representation of the process that describes these underlying influences (Spiegelhalter *et al.*, 2014). The significance of $IM\varepsilon_t$ for F_{3t} in Panel C implies that there are other factors that are not reflected in the macroeconomic factor set and $M\varepsilon_t$. As argued by Van Rensburg (2000) and Middleton and Satchell (2001), these results also suggest that linear factor models that employ macroeconomic factors will be underspecified and that residual market factors will not alleviate or resolve underspecification.

There are a number of reasons why macroeconomic factors may be poor proxies for systematic influences. That macroeconomic factors by themselves are poor proxies for the systematic influences is suggested by literature that reports relatively low $\bar{R}^2$s in the absence of a residual market factor or a market index (Connor, 1995; Van Rensburg; 2000; Szczygielski & Chipeta, 2015). This attests to the complexity of the return generating process. Also, investors may not be rational and will therefore respond irrationally (or will not respond at all) to macroeconomic news with sentiment being more important (Malkiel, 2003). Lin, Rahman and Yung (2009) provide support for this in a study of the determinants real estate investment trust (REIT) returns, finding that investor sentiment is more important relative to macroeconomic factors. Finally, as suggested by Panetta (2002) and Spyridis *et al*. (2012), the return generating process is unstable and therefore the same macroeconomic factors may not be continuous proxies for unspecified systematic influences over time. Furthermore, the response to macroeconomic news embodied by macroeconomic factors may differ between crises and non-crises periods (McQueen & Roley, 1993; De Lint, 2002). Notably, the residual market factor fails to substantially improve the approximation of systematic influences in returns. This is perhaps to be expected and highlights the limitation of this approach in previous studies that rely upon some "broad market index," to derive the residual market factor (McElroy & Burmeister, 1988: 33). Unless this broad market index is the all-inclusive but completely elusive market portfolio, then any residual market factor will not reflect all return generating factors and will not be a proxy for omitted influences (Born & Moser, 1988; Gadzinski *et al.,* 2018).

The inclusion of characteristic-based factors significantly improves the approximation of F_{1t}, with the $\bar{R}^2$ increasing from 0.294 in Panel C in Table 2 to 0.450 (see Table 1D in the supplementary appendix). The improvement for F_{2t} is lower, but still noticeable with an increase in the $\bar{R}^2$ to 0.306 from 0.255. The increase in the $\bar{R}^2$ for F_{3t} is negligible. These observations suggest a number of implications for our analysis. Firstly, all macroeconomic factors remain statistically significant across factor scores. This implies that results are generally robust. The loss of significance for $IM\varepsilon_t$ can be explained by its weak significance in the first place. Secondly, for F_{1t}, SMB_t, HML_t and RMW_t, are statistically significant and one additional macroeconomic factor is now statistically significant, MET_t. This latter finding can be attributed to underspecification of the original model (equation (8)). For F_{2t}, SMB_t and UMD_t are statistically significant. Overall, this and the significance of characteristic-based factors from F_{1t}, suggests that characteristic-based factors contribute to approximating systematic influences over and above the macroeconomic and residual market factors. This is further evidence that the residual market factors are not adequate proxies for systematic influences. Thirdly, our findings are similar to that of Aretz *et al.* (2010) in that they show that the SMB_t and HML_t factors reflect the macroeconomic environment and are therefore proxies for systematic influences. Where our findings differ is that that we also consider profitability and investment factors, RMW_t and CMA_t. This, together with a consideration of the ability of characteristic-based factors to approximate measures of systematic influences, constitutes a novel finding within the South African context.

4.3. Model stability and rolling regression results

Table 3 reports the results of equation (8) re-estimated with breakpoints (Section 3.5.). Each factor regression is characterized by five breaks, translating into six distinct periods. This suggests that any linear factor model will be characterized by instability in the relationship between returns and macroeconomic factors (see Panetta, 2002). For example, while there is a significant relationship between F_{1t} and BP_{t-1} and MET_t during the 2001M01 to 2003M04 period in Panel A, these factors are no longer significant proxies until the period 2011M05 to 2014M06. This suggests that the response of stock prices will differ according to period and that

Table 3. Factor Score Regressions with Breakpoints

Panel A: F_{1t}

Period	Obs.	Sig. Factors
2001M01- 2003M04	28	$-\alpha$ **, BP_{t-1} **, $-USD\varepsilon_t$ ***, $-MET_t$ **, $-LTY_t$ ***, $M\varepsilon_t$ **
2003M05 - 2006M05	37	α *, BUS_t ***, $-USD\varepsilon_t$ ***, $-LTY_t$ ***, $TLI\varepsilon_t$ ***, $M\varepsilon_t$ **
2006M06 - 2008M11	30	$LEAD_{t-1}$ ***, BUS_t ***, $-LTY_t$ ***, $-TLI\varepsilon_t$ ***, $M\varepsilon_t$ **
2008M12 - 2011M04	29	α***, $-LEAD_{t-1}$ *, $-LTY_t$ **, $M\varepsilon_t$ ***
2011M05 - 2014M06	38	BP_{t-1} **, $LEAD_{t-1}$ ***, $-USD\varepsilon_t$**, $-MET_t$ ***, $-LTY_t$ **
2014M07 - 2016M12	30	BUS_t ***, $TLI\varepsilon_t$ ***
Break Dates	2003M05, 2006M06, 2008M12, 2011M05, 2014M07	
$\bar{R}^2$	0.432	
F-statistic	3.461***	
AIC	2.653	
BIC	3.671	

Panel B: F_{2t}

Period	Obs.	Sig. Factors
2001M01 - 2003M04	28	α ***, $-BP_{t-1}$ **, $-BUS_t$ **, $USD\varepsilon_t$ ***, $TLI\varepsilon_t$ **, $M\varepsilon_t$ **
2003M05 - 2005M11	31	$M\varepsilon_t$ **
2005M12 - 2009M02	39	MET_t ***, $TLI\varepsilon_t$ ***, $M\varepsilon_t$ ***, $IM\varepsilon_t$ ***
2009M03 - 2012M03	37	BUS_t ***, $-USD\varepsilon_t$ ***, MET_t ***, $-LTY_t$ **, $TLI\varepsilon_t$ ***
2012M04 - 2014M07	28	$-\alpha$ ***, $-LEAD_{t-1}$ *, $USD\varepsilon_t$ ***, MET_t ***, LTY_t ***, $M\varepsilon_t$ *
2014M08 - 2016M12	29	BP_{t-1} ***, $-USD\varepsilon_t$ ***, MET_t ***, $-LTY_t$ **
Break Dates	2003M05, 2005M12, 2009M03, 2012M04, 2014M08	
$\bar{R}^2$	0.578	
F-statistic	5.436***	
AIC	2.520	
BIC	3.538	

Panel C: F_{3t}

Period	Obs.	Sig. Factors
2001M01 - 2003M04	28	$-\alpha$ ***, BUS_t ***, $M\varepsilon_t$ ***, $IM\varepsilon_t$ ***
2003M05 - 2006M09	41	$USD\varepsilon_t$ ***, MET_t **, $M\varepsilon_t$ ***
2006M10 - 2009M01	28	$-\alpha$ **, BP_{t-1} ***, $-USD\varepsilon_t$ **, $-MET_t$ ***, $M\varepsilon_t$ **, $IM\varepsilon_t$ *
2009M02 - 2012M04	39	$USD\varepsilon_t$ ***, MET_t ***, LTY_t ***, $M\varepsilon_t$ ***
2012M05 - 2014M08	28	α *, BUS_t ***, $TLI\varepsilon_t$ **, $M\varepsilon_t$ ***, $IM\varepsilon_t$ ***
2014M09 - 2016M12	28	α **, $USD\varepsilon_t$ **, $-MET_t$ **, $M\varepsilon_t$ ***
Break Dates	2003M05, 2006M10, 2009M02, 2012M05, 2014M09	
$\bar{R}^2$	0.600	
F-statistic	5.861***	
AIC	2.489	
BIC	3.507	

The asterisks, ***, ** and *, indicate statistical significance at the respective 1%, 5% and 10% levels of significance. Obs. is the number of observations for each segment. Estimation settings: Bai-Perron test with 1 to *m* globally determined breaks. Break selection based upon sequential evaluation with 0.15 trimming and a significance level of 10%. Based upon the trimming settings and the sample size, the maximum permissible number of breaks is 5. Least squares with Newey-West heteroscedasticity and autocorrelation consistent (HAC) standard errors used, permitting heterogeneous error distributions across breaks. All factors are in innovations, defined as follows: BP_{t-1}- number of building plans passed, $LEAD_{t-1}$- domestic composite cyclical leading indicator, BUS_t - business activity, $USD\varepsilon_t$- (orthogonalized) fluctuations in the Rand-Dollar exchange rate, MET_t- world metal prices, LTY_t- long-term government bond yields, $TLI\varepsilon_t$- (orthogonalized) leading indicator for South Africa's trading partners, $M\varepsilon_t$- the residual market factor orthogonal to the macroeconomic factor set, derived from returns on the JSE All Share Index, $IM\varepsilon_t$- a second residual market factor orthogonal to the macroeconomic factor set and $M\varepsilon_t$.

macroeconomic factors may be better proxies for systematic influences during certain periods, relative to others.

Not only is the ability of macroeconomic factors to proxy for pervasive influences unstable, the direction of the relationships changes between periods. For example, for F_{3t} in Panel C, during the 2003M05 to 2006M09 period, the relationship between factor scores and MET_t is positive, negative for the 2006M10 to 2009M01 period, positive for the 2009M02 to 2012M04 period and negative for the 2014M09 to 2016M12 period. The ability of macroeconomic factors to proxy for systematic influences may be limited to specific subperiods, rendering macroeconomic factors ineffective at explaining returns over the entire sample period or over longer periods in general (Panetta, 2002; Spyridis *et al.,* 2012).

It can be argued that instability in relationships arises as a result of the length of the sample period chosen (totalling a 192 months) and that a shorter sample period may yield more desirable results. For example, Berry, Burmeister and McElroy's (1988) sample spans a period of 11 years, Van Rensburg's (1996) sample spans almost 10 years and Panetta (2002) subdivides his sample into four non-overlapping subperiods of 4 years each. The Northfield Macroeconomic Equity Risk Model uses 60 months of data (NIS, 2015). In contrast, Sadorsky and Henriques (2001) use a sample that spans 26 years whereas Szczygielski and Chipeta (2015) use a sample that spans almost 16 years. However, the results in Table 3 suggest that structural changes take place frequently and subperiods during which relationships between factor scores and macroeconomic factors are stable and shorter than any of the sample periods set out above. The implication is that macroeconomic factors proxy for systematic influences over (very) short periods of time and that proxies differ over time.

Another notable observation in Table 3 is that the residual market factor is insignificant during certain periods. For example, for F_{1t}, $M\varepsilon_t$ is not statistically significant for the 2011M05 to 2014M06 and 2014M07 to 2016M12 periods. In contrast, $M\varepsilon_t$ is highly statistically significant for F_{3t} across all subperiods. This provides a potential explanation for the substantial increase in the $\bar{R}^2$ from 0.131 in Panel A to 0.431 for F_{3t} in Panel B in Table 2. This implies that $M\varepsilon_t$ continues to be a proxy for unspecified influences in returns represented by the factor scores of F_{3t} across all segments. In contrast, the lack of significance for $M\varepsilon_t$ during certain periods for F_{1t} and F_{2t} in Table 3 presents a potential explanation as to why the $\bar{R}^2$ improves in Table 2 but not by as much as for F_{3t} in Panel B.

Next, we re-estimate equations (6), (7), (8) and (9) as rolling regressions (Section 3.5). The respective $\bar{R}^2$s derived from these rolling regressions are reported in Figures 2, 3 and 4.

Figure 2. Rolling Factor Score Regressions for Factor 1

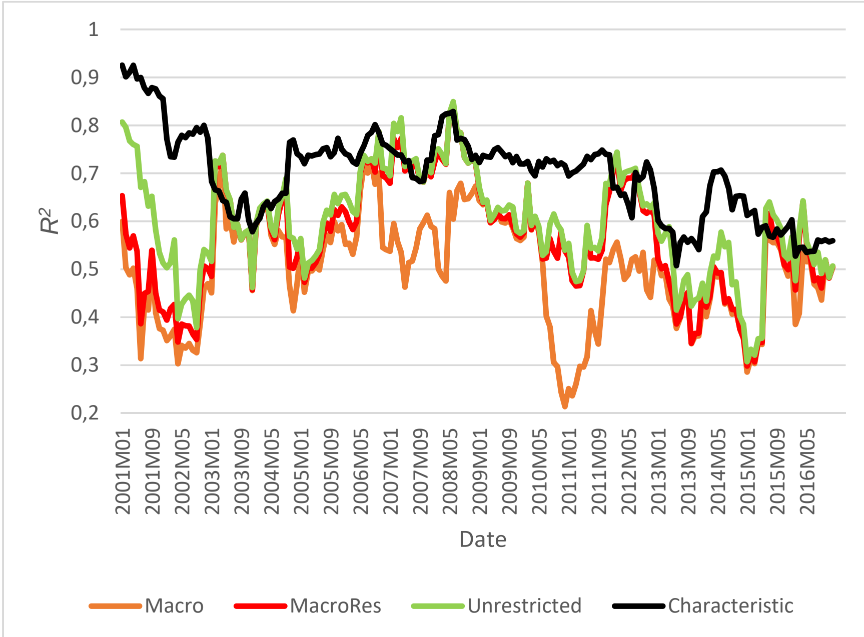


Figure 3. Rolling Factor Score Regressions for Factor 2

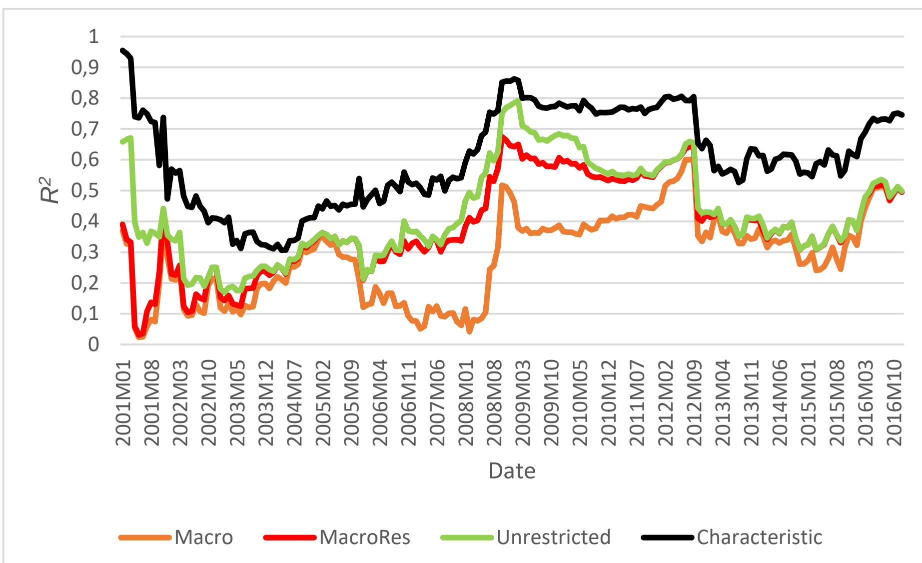
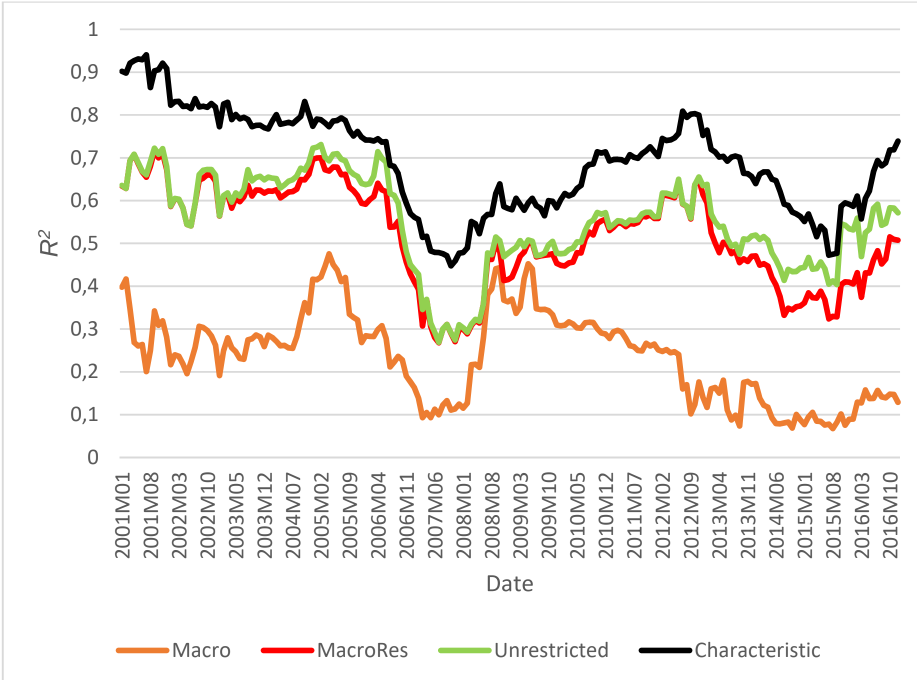


Figure 4. Rolling Factor Score Regressions for Factor 3



Notes: Figure 2, 3 and 4 report the rolling $\bar{R}^2$s for equation (6), (7), (8) and (9) relating each series of factor scores to factor set. The $\bar{R}^2$for equation (6) is termed as the "macro" and is the $\bar{R}^2$for a purely macroeconomic factor set. The $\bar{R}^2$for equation (7) is reported as "MacroRes" and is the $\bar{R}^2$for the macroeconomic factors and the residual market factor, $M\varepsilon_t$. The $\bar{R}^2$for equation (8) is reported as "Unrestricted" and is the $\bar{R}^2$for the macroeconomic factors and both residual market factors, $M\varepsilon_t$and $IM\varepsilon_t$. The $\bar{R}^2$for equation (9) is reported as "Characteristic" and is the $\bar{R}^2$ for the macroeconomic factors and both residual market factors, $M\varepsilon_t$ and $IM\varepsilon_t$, and the characteristic based factors, SMB_t, HML_t, UMD_t, RMW_tandCMA_t.

The most noticeable aspect in Figures 2, 3 and 4 is that the ability of the macroeconomic factors to approximate factor scores is highly unstable. Also, the ability of macroeconomic factors to approximate factor scores increases around crises periods and decreases afterwards. An increase in the $\bar{R}^2$, most pronounced for F_{2t} and F_{3t} and coinciding with the global financial crises, is observed around the 2008 to 2009 period in Figures 3 and 4. This suggests that during global periods of turbulence, macroeconomic fundamentals become more important (see McQueen & Roley, 1993; De Lint, 2002). The inclusion of $M\varepsilon_t$ offsets the drop in the ability of macroeconomic factors to proxy for systematic influences. This is especially evident for F_{1t} and F_{2t} for the 2006 to 2008 period in Figures 2 and 3. This also suggests that markets may be driven by sentiment during favourable economic periods, reflected by the general market indices used to derive the residual market factors. For F_{3t}, the residual market factor compensates for the loss of the ability of macroeconomic factors to proxy for systematic influences after 2008. Finally, the characteristic-based factors contribute to approximating factor scores; for each factor, the inclusion of the characteristic-based factors produces an $\bar{R}^2$ over and above that of the model that comprises the macroeconomic and residual market factors implying that the former factors are insufficient proxies.

4.4. Implications

Our findings have implications for South African researchers, investors, econometricians and economists. For researchers, they confirm the findings of prior studies on the South African stock market and developed markets. The return generating process is characterised by multiple factors (see for example Yli-Olli & Virtanen, 1993; Chimanga & Kotze, 2009). Systematic influences can be proxied for by a set of macroeconomic factors; namely BP_{t-1}, $LEAD_{t-1}$, BUS_t, USD_t, MET_t, LTY_t and TLI_t, in this study. However, identifying such factors as interpretable proxies is a challenge. Out of an extensive set of factors, totalling 52 unique factors and 208 factors considered contemporaneously and with up to three lags, only seven qualify. As in other studies (Connor, 1995; Middleton & Satchell, 2001), macroeconomic factors are poor proxies for systematic influences. From a purely econometric standpoint, specifications that comprise only macroeconomic factors are likely to be underspecified and will result in misleading inferences (Dominguez, 1992; Brauer & Gómez-Sorzano, 2004; Van Rensburg, 2002; Bucevska, 2011). The failure to adequately reflect systematic influences and the resultant underspecification also has implications for investors and researchers who apply asset pricing models such as the CAPM or the APT that require the estimation of risk premia (see Brennan, Chordia & Subrahmanyam, 1998: 349; Zaremba, Czapkiewicz, Szczygielski & Kaganov, 2019).

The results in Section 4.3 indicate that macroeconomic factors are more important during times of crises suggesting that investors pay closer attention to macroeconomic news during such periods. De Lint (2002) makes a similar finding for Mexico and six Asian countries, finding that local factors, which are more likely to be captured by the macroeconomic factor set in this study, are more important during times of crises. This is a potential explanation as to why the macroeconomic factors are better proxies for systematic influences during some periods relative to others. Relatedly, our findings also bring into question the sample lengths used in prior research of a similar nature, ranging between 4 and 26 years (Sadorsky & Henriques, 2001; Panetta, 2002). Our findings indicate that periods of structural stability range between 28 and 41 months. This suggests that studies that use longer periods may report that a macroeconomic factor is statistically significant (insignificant) over an extended period of time when, in reality, it is statistically significant (insignificant) for a limited number of short periods. Observed significance (or insignificance) over an extended period of time is the result of a bias arising from structural instability. The recommendation is that consideration is given to the ability of macroeconomic factors to exert a continuous influence on stock returns and their importance over the entire economic cycle.

An important implication is that a commonly used proxy for omitted factors that is often seen as sufficiently addressing factor omission in the literature (Deetz *et al.,* 2009; Czaja *et al.,* 2010) is ineffective or only partially effective. The results in Panel B of Table 3 show that the $\bar{R}^2$ is still far below 1, even with the residual market factor included. Moreover and worryingly, the results in Table 3 show that the residual market factor is not a stable proxy for systematic influences. Relatedly, Figures 2, 3 and 4 show that the residual market factor offsets the decrease in the ability of macroeconomic factors to proxy for systematic influences during certain periods. The explanatory power of the model comprising the macroeconomic factors and the residual market factors diverges from that of the macroeconomic factors only. This can be potential explained in the emergence of different macroeconomic proxies during these periods or the rising importance of sentiment during periods of favourable activity, prior to economic downturns. However, its contribution is unstable and at times, non-existent. The contribution of the second residual market factor is marginal at best. The recommendation to researchers and econometricians is that the use of a residual market factor should not be the default solution to underspecification. We suggest that the solution may instead lie in combining macroeconomic *and* statistical factors in such specifications to address underspecification. Such an approach is also suggested by Van Rensburg (1997) and applied in the Northfield Macroeconomic Equity Risk Model (NIS, 2015).

Finally, through an extension and robustness test, we show that characteristic-based factors are proxies for systematic influences (see Table 1C in the supplementary appendix). Although not our main result, this is an advance in the South African context. These findings are similar to those of Aretz *et al.* (2010). Where our approach differs is that we relate factor scores to these factors as opposed to relating characteristic-based factors to macroeconomic factors. The time-series (as opposed to cross-sectional) explanatory power of these factors has not been tested in the South African context. The implication is that macroeconomic factors and characteristic-based factors can be combined to arrive at improved specifications of return behaviour (Brown, Hiraki, Arakwa & Ohno, 2009).

5. Conclusion

We investigate the ability of macroeconomic factors to proxy for pervasive influences in stock returns. Where our study differs is in that we separate the systematic components of returns from idiosyncratic components in returns. By doing so, we are able to directly assess the ability of macroeconomic factors to proxy for systematic influences. We also consider the ability of a commonly used proxy for omitted factors, the residual market factor, to address and resolve underspecification.

In line with prior literature on developed markets, we find that macroeconomic factors are poor proxies for systematic influences and are unlikely to yield an adequate description of the return generating process. This will continue to be the case if idiosyncratic components are fully diversifiable and the only remaining drivers of returns are systematic in nature. This will also be the case if a seemingly generous description of the return generating process is specified. Worryingly, we find that the commonly used residual market factor does not substantially improve the approximation of factor scores. A failure to adequately specify a model can lead to misleading inferences and results. Interestingly and although this is not the focus of our study, we find that characteristic-based factors contribute to approximating the systematic drivers of returns in South Africa. Consequently, a further avenue for research is to investigate the nature of the mechanism through which this occurs. Also, a further avenue of research is the ability of a factor analytic augmentation derived from the residuals of a specification applied across a number of series to proxy for remaining uncaptured systematic influences.

Our findings serve as a reference for researchers who are interested in modelling the relationship between asset returns and factor sets, especially macroeconomic factors.

References

Al-Tamimi, H.A.J, Alwan, A.A. & Rahman, A.A.A. (2011). Factors affecting stock prices in the UAE financial markets. *Journal of Transnational Management,* 16(3), 3-19.

Amenc, N., Esakia, M., Goltz, F., & Luyten, B. (2019). Macroeconomic risks in equity factor investing. *Journal of Portfolio Management,* 45(6), 39-60.

Aretz, K., Bartram, S. M., & Pope, P. F. (2010). Macroeconomic risks and characteristic-based factor models. *Journal of Banking and Finance, 34*(6), 1383-1399.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*(1), 47–78.

Baldacci, E., Gupta, S., & Mati, A. (2011). Political and fiscal risk determinants of sovereign spreads in emerging markets. *Review of Development Economics*, *15*(2), 251-263.

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A, 160*(901), 268-282.

Berry, M. A., Burmeister, E., & McElroy, M. B. (1988). Sorting out risks using known APT factors. *Financial Analysts Journal*, *44*(2), 29-42.

Bessler, W., & Kurmann, P. (2014). Bank risk factors and changing risk exposures: Capital market evidence before and during the financial crisis. *Journal of Financial Stability*, *13*, 151-166.

Bessler, W., Kurmann, P., & Nohel, T. (2015). Time-varying systematic and idiosyncratic risk exposures of US bank holding companies. *Journal of International Financial Markets, Institutions and Money*, *35*, 45-68.

Biger, N., & Page, M. J., (1993) Unit trusts' performance: Does the yardstick matter?, *Journal of Studies in Economics and Econometrics, 17*(1), 1-15.

Bilson, C. M., Brailsford, T. J., & Hooper, V. J. (2001). Selecting macroeconomic variables as explanatory factors of emerging stock market returns. *Pacific-Basin Finance Journal*, *9*(4), 401-426.

Birz, G., & Lott, J. R. (2011). The effect of macroeconomic news on stock returns: New evidence from newspaper coverage. *Journal of Banking and Finance*, *35*(11), 2791-2800.

Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods*, *17*(3), 399-417.

Born, J. A., & Moser, J. T. (1988). An investigation into the role of the market portfolio in the arbitrage pricing theory. *Financial Review,* 23(3), 287-299.

Brauer, J., & Gómez-Sorzano, A. G. (2004). Homicide cycles in Colombia, 1950-1999. *International Journal of Applied Econometrics and Quantitative Studies*, *1*(1), 1-22.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, *49*(3), 345-373.

Brown, S. J., Hiraki, T., Arakawa, K., & Ohno, S. (2009). Risk premia in international equity markets revisited. *Pacific-Basin Finance Journal*, *17*(3), 295-318.

Bucevska, V. (2011). Heteroscedasticity. In M. Lovric (Ed.). *International Encyclopedia of Statistical Science.* 630-633. Berlin, Heidelberg: Springer.

Burmeister, E., & Wall, K. D. (1986). The arbitrage pricing theory and macroeconomic factor measures. *Financial Review*, *21*(1), 1-20.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance 52*(1), 57–82.

Carlson, J. B., Craig, B., & Schwarz, J. C. (2000). Structural uncertainty and breakpoint tests: an application to equilibrium velocity. *Journal of Economics and Business*, *52*(1-2), 101-115.

Chan, K.C., Chen, N., & Hsieh, D.A. (1985). An exploratory investigation of the firm size effect. *Journal of Financial Economics,* 14(3), 451-471.

Chang, S. J. (1991). A study of empirical return generating models: A market model, a multifactor model, and a unified model. *Journal of Business Finance & Accounting*, *18*(3), 377-391.

Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, *59*(3), 383-403.

Chen, S. J., & Jordan, B. D. (1993). Some empirical tests in the arbitrage pricing theory: Macro variables vs. derived factors. *Journal of Banking and Finance, 17*(1), 65-89.

Chimanga, A., & Kotze, D. (2009). A multivariate analysis of factors affecting stock returns on the JSE. *African Finance Journal*, *11*(2), 80-96.

Connor, G. (1995). The three types of factor models: A comparison of their explanatory power. *Financial Analysts Journal*, *51*(3), 42-46.

Czaja, M. G., Scholz, H., & Wilkens, M. (2010). Interest rate risk rewards in stock returns of financial corporations: Evidence from Germany. *European Financial Management*, *16*(1), 124-154.

De Lint, C. R. (2002). Risk profiles: How do they change when stock markets collapse?. *Journal of International Financial Markets, Institutions and Money*, *12*(1), 59-80.

Deetz, M., Poddig, T., Sidorovitch, I., & Varmaz, A. (2009). An evaluation of conditional multi-factor models in active asset allocation strategies: an empirical study for the German stock market. *Financial Markets and Portfolio Management, 23*(3), 285-313.

DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation, 14*(20), 1-11.

Dominguez, K.M. (1992). *Exchange rate efficiency and the behaviour of international asset markets.* New York, NY: Routledge

Elton, E. J., & Gruber, M. J. (1988). A multi-index risk model of the Japanese stock market. *Japan and the World Economy*, *1*(1), 21-44.

Elton, E.J. & Gruber, M.J. (2018). The impact of Ross's exploration of APT on our research. *Journal of Portfolio Management, 44*(6), 98-107.

Elton, E.J., Gruber, M.J., Brown, S.J. & Goetzmann, W.N. (2014). *Modern portfolio theory and investment analysis* (9th ed.). New York, NY: Wiley.

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

Fama, E. F., & K. R. French. (2015). A five-factor asset pricing model. *Journal of Financial Economics, 116*(1), 1–22.

French, J. (2017). Macroeconomic forces and arbitrage pricing theory. *Journal of Comparative Asian Development,* 16(1), *1-20.*

Gadzinski, G., Schuller, M., & Vacchino, A. (2018). The global capital stock: Finding a proxy for the unobservable global market portfolio. *Journal of Portfolio Management,* 44(7), 12-23.

Greene, W.J. (2012). *Econometric analysis* (7th ed.). Harlow, Essex: Pearson Education Limited.

Hamao, Y. (1988). An empirical examination of the arbitrage pricing theory. *Japan and the World Economy,* 1(1), 45-61.

Hansen, B.E. (2012, July). *Advanced time series and forecasting - Lecture 5: structural breaks.* Lecture presented at the Summer School in Economics and Econometrics, University of Crete, Crete.

Kryzanowski, L., Lalancette, S., & To, M. C. (1994). Some tests of APT mispricing using mimicking portfolios. *Financial Review, 29*(2), 153-192.

Lin, C. Y., Rahman, H., & Yung, K. (2009). Investor sentiment and REIT returns. *Journal of Real Estate Finance and Economics*, 39(4), 450-471.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives,* 17(1), 59-82.

Merville, L.J., Hayes-Yelken, L.J., Xu, Y. (2001). Identifying the factor structure of equity returns. *Journal of Portfolio Management, 27*(4), 51-61.

McElroy, M. B., & Burmeister, E. (1988). Arbitrage pricing theory as a restricted nonlinear multivariate regression model iterated nonlinear seemingly unrelated regression estimates. *Journal of Business and Economic Statistics,* 6(1), 29-42.

McQueen, G., & Roley, V. V. (1993). Stock prices, news, and business conditions. *Review of Financial Studies*, *6*(3), 683-707.

Middleton, L. P., & Satchell, S. E. (2001). Deriving the arbitrage pricing theory when the number of factors is unknown. *Quantitative Finance*, *1*(5), 502-508.

Nel, W. S. (2011). The application of the capital asset pricing model (CAPM): A South African perspective. *African Journal of Business Management*, *5*(13), 5336.

Newey, W. K., & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review, 28*(3), 777-787.

Northfield Information Services (NIS) (2015). *U.S. Macroeconomic Equity Risk Model.* Available from the Northfield Information Services, Inc. website: http://www.northinfo.com/documents/7.pdf/ Accessed 11 October 2018.

Page, M.J. (1989). Model selection for measuring security price performance. *South African Journal of Business Management, 17*(1), 78-81.

Panetta, F. (2002). The stability of the relation between the stock market and macroeconomic forces. *Economic Notes, 31*(3), 417-450.

Priestley, R. (1996). The arbitrage pricing theory, macroeconomic and financial factors, and expectations generating processes. *Journal of Banking and Finance*, *20*(5), 869-890.

PWC (2015). Africa: A closer look at value. Retrieved December 1, 2017, from the PWC website at https://www.pwc.co.za/en/assets/pdf/valuation-methodology-survey-2015.pdf

Roll, R., & Ross, S. A. (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance*, *35*(5), 1073-1103.

Sadorsky, P. (2008). Assessing the impact of oil prices on firms of different sizes: Its tough being in the middle. *Energy Policy,* 36(10), 3854-3861.

Sadorsky, P., & Henriques, I. (2001). Multifactor risk and the stock returns of Canadian paper and forest products companies. *Forest Policy and Economics*, *3*(3-4), 199-208.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(3), 485-493.

Spyridis, T., Sevic, Z., & Theriou, N. (2012). Macroeconomic vs. statistical APT approach in the Athens Stock Exchange. *International Journal of Business*, *17*(1), 39-64.

Szczygielski, J. J., & Chipeta, C. (2015). Risk factors in returns of the South African stock market. *Studies in Economics and Econometrics*, *39*(1), 47-70.

Van Rensburg, P. (1995). Macroeconomic variables and the Johannesburg Stock Exchange: A multifactor approach. *De Ratione, 9*(2), 45-63.

Van Rensburg, P. (1996). Macroeconomic identification of the priced APT factors on the Johannesburg Stock Exchange. *South African Journal of Business Management, 27*(4), 104-112.

Van Rensburg, P. (1997). Employing the prespecified variable approach to APT factor identification on the segmented Johannesburg Stock Exchange. *South African Journal of Accounting Research, 11*(1), 57-74.

Van Rensburg, P. (2000). Macroeconomic variables and the cross-section of Johannesburg Stock Exchange returns. *South African Journal of Business Management, 31(*1), 31-43.

Van Rensburg, P. (2002). Market segmentation on the Johannesburg stock exchange II. *Studies in Economics and Econometrics, 26*(1), 83-99.

Van Rensburg, P., & Slaney, K. (1997). Market segmentation on the Johannesburg stock exchange. *Journal for Studies in Economic and Econometrics, 21*(3), 1-23.

Wei, K. J. (1988). An asset-pricing theory unifying the capm and apt. *Journal of Finance*, *43*(4), 881-892.

Yli-Olli, P., & Virtanen, I. (1993). Some empirical tests of the arbitrage pricing theory using transformation analysis. *Empirical Economics, 17*(4), 507-522.

Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 9*(2), 79-94.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432-442.

Zaremba, A., Czapkiewicz, A., Szczygielski, J. J., & Kaganov, V. (2019). An application of factor pricing models to the Polish stock market. *Emerging Markets Finance and Trade*, *55*(9), 2039-2056.

Endnotes follow:

[i] The sectors comprising the sample together with the respective descriptive statistics for the return data are reported in Table 1A of the Appendix 1.

[ii] Sampling adequacy is confirmed using the Kaiser-Meyer-Olkin (KMO) test, producing a value of 0.927.

[iii] Following preliminary analysis, USD_t is orthogonalised against LTY_t and MET_t, and TLI_t is orthogonalised against MET_t to mitigate the consequences of multicollinearity. The resultant residual series, $USD\varepsilon_t$ and $TLI\varepsilon_t$, are used in place of the original innovation series. Correlation between the original innovation series and the orthogonalised series are over 0.8, indicating that orthogonalisation does not substantially change the nature of these factors.