REFERENCES


Marshall, P. and Walker, E. 2002. Asymmetric reaction to information and serial


Masulis, R.W. and Ng, V.K. 1995. Overnight and daytime stock-return dynamics on
the London stock exchange: the impact of the “Big Bang” and the 1987 stock-market


2003: 16.


McQueen, G. and Roley, V.V. 1993. Stock Prices, news and business conditions.


Morrison, D.G. 1972. Upper bounds for correlations between binary outcomes and

Neftci, S. 1984. Are economic time series asymmetric over the business cycle?

Nel, H. 1996. The term structure of interest rates and economic activity in South

illusion? Working paper, *Economics Department, University of Washington, Seattle*. 


APPENDIX 1

PREDICTING TURNING POINTS IN THE SOUTH AFRICAN ECONOMY

A1.1 INTRODUCTION

Following the recent trend in the literature, the term structure was used as explanatory variable in the Markov switching regime model of the South African business cycle (see chapter five). Theoretically the term structure can be used as leading indicator of turning points in the economy, but it has to be established whether it is superior to other indicators in practice as well. The appendix is organized as follows: The next section gives a brief overview of the relevant literature. Section A1.3 describes the econometric technique, and section A1.4 describes the leading indicators used in the empirical analysis. Section A1.5 presents the results of the empirical analysis, while section A1.6 provides the conclusion.

A1.2 LITERATURE REVIEW

Estrella and Mishkin (1998) compared the performance of various financial variables, including four term structures of interest rates, stock prices, monetary aggregates, indices of leading indicators and other economic variables such as GDP, CPI and exchange rates, as predictors of US recessions. They estimated probit models with quarterly data for the period 1959 to 1995, and evaluated the performance of the leading indicators by using the pseudo-$R^2$ value developed for dichotomous models by Estrella (1998). Their results indicated that the interest rate spread outperforms the other indicators for forecasting beyond one quarter ahead. They also tested the performance of all the possible models that includes both the interest rate spread and one other indicator as explanatory variables.

Several studies confirmed the result of Estrella and Mishkin (1998) that the interest rate spread is successful with predicting business cycle turning points. Estrella and Hardouvelis (1991) were the first to empirically analyze the term structure as a
predictor of real economic activity. Regressions of future GNP growth on the slope of the yield curve and other information variables showed that a steeper (flatter) slope implies faster (slower) future growth in real output. The forecasting accuracy in predicting cumulative changes is highest 5 to 7 quarters ahead. In addition, they also used a probit model to analyze the predictive power of the term structure on a binary variable that simply indicates the presence or absence of a recession.

Bernard and Gerlach (1996) tested the ability of both the domestic and foreign term structures to predict business cycle turning points in eight industrial countries for the period 1972 to 1993. Using probit models, they show that the domestic term spreads are statistically significant in explaining business cycle turning points in all eight countries. The period over which the domestic term spread successfully forecasts the turning points vary across countries, but the optimal forecast period ranges from two to five quarters. Nel (1996) studied the relationship between the term structure of interest rates and the South African business cycle. He found that they were cointegrated, in other words a contemporaneous relationship, despite a poor overall fit.

Cook and Smith (2001) assessed the effectiveness of transplanting a forecasting method based on a probabilistic approach in the South African context. They tested the ability of some of the components of the composite index of leading indicators to predict both the official Reserve Bank turning points as well as the mechanistic turning points of the composite index of coincident indicators. This is done by estimating a probit model with all the chosen leading indicators simultaneously as explanatory variables. Their results indicate an ability of the model to accurately forecast business cycle turning points in the 1980s. However, in the 1990s, the model displays a diminished capacity to forecast the turning points. The present analysis differs from their study in several ways. Instead of evaluating the joint performance of the leading indicators, we are evaluating the performance of the leading indicators individually to find the individual leading indicator that most accurately predicts business cycle turning points. Methodologically, we use the pseudo R² developed by Estrella (1998) for models with dichotomous dependent variables to evaluate the models, unlike their qualitative evaluation.
A1.3 THE TECHNIQUES

A1.3.1 The Probit Model

Several authors have used probit models to model business cycle turning points (see e.g. Estrella and Hardouvelis, 1991; Dueker, 1997; Dotsey, 1998; Estrella and Mishkin, 1998; Bernard and Gerlach, 1996). The probit form is dictated by the fact that the variable being predicted takes on only two possible values – whether the economy is in a recession or not. The model is defined in reference to a theoretical linear relationship of the form:

\[ Y_{t+k} = \alpha + \beta^* x_t + \varepsilon_t \]  \hspace{1cm} (A1.1)

where \( Y_t \) is an unobserved variable that determines the occurrence of a recession at time \( t \), \( k \) is the length of the forecast horizon, \( \varepsilon_t \) is a normally distributed error term, and \( x_t \) the value of the explanatory variable at time \( t \). The parameters \( \alpha \) and \( \beta \) are estimated with maximum likelihood. The observable recession indicator \( R_t \) is related to this model by

\[ R_t = 1 \text{ if } Y_t > 0, \text{ and } 0 \text{ otherwise.} \]  \hspace{1cm} (A1.2)

The form of the estimated equation is

\[ P(R_{t+k} = 1) = F(\alpha + \beta^* x_t) \]  \hspace{1cm} (A1.3)

where \( F \) is the cumulative normal distribution function.

The model is estimated by maximum likelihood. The recession indicator is obtained from the South African Reserve Bank, that is, \( R_t = 1 \) if they classify the economy to be in a downward phase at time \( t \), and 0 otherwise (see table A1.1).
Table A1.1  Business Cycle Phases According to SARB since 1978

<table>
<thead>
<tr>
<th>Upward phase</th>
<th>Downward phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 1978</td>
<td>August 1981</td>
</tr>
<tr>
<td>April 1983</td>
<td>June 1984</td>
</tr>
<tr>
<td>April 1986</td>
<td>February 1989</td>
</tr>
<tr>
<td>June 1993</td>
<td>November 1996</td>
</tr>
<tr>
<td>September 1981</td>
<td>March 1983</td>
</tr>
<tr>
<td>July 1984</td>
<td>March 1986</td>
</tr>
<tr>
<td>March 1989</td>
<td>May 1993</td>
</tr>
<tr>
<td>December 1996</td>
<td>August 1999</td>
</tr>
</tbody>
</table>

A1.3.2 Pseudo-\(R^2\) for Models with Dichotomous Dependent Variables

Estrella (1998) developed a pseudo \(R^2\) that is a simple measure of goodness of fit in the context of a dichotomous dependent variable, which corresponds intuitively to the widely used coefficient of determination \((R^2)\) in a standard linear regression\(^1\). Models for dichotomous dependent variables, such as probit and logit models, are usually estimated by maximizing the likelihood function, which is defined as:

\[
L = \prod_{y_i=0} F(\beta x_i) \prod_{y_i=1} F(1 - \beta x_i). \quad (A1.4)
\]

Let the unconstrained maximum value of the likelihood function \((L)\) be \(L_U\), and its maximum value under the constraint that all coefficients are zero except for the constant as \(L_C\). Denote the number of observations with \(n\). Then

\(^1\) Estrella (1998) suggest the following three requirements for an \(R^2\) analog for models with dichotomous dependent variables: (i) It has to be contained by the interval \([0,1]\), where zero represents no fit and one represents a perfect fit. (ii) It has to be based on a valid test statistic for the hypothesis that all the coefficients, except the constant, are zero. (iii) Its derivative with respect to the test statistic should be consistent with the corresponding derivative in the linear case. Estrella (1998) shows that most previous measures of fit, specifically McFadden (1974), Cragg and Uhler (1970), Aldrich and Nelson (1989), Veall and Zimmermann (1992), Morisson (1972), Goldberger (1973) and Davidson and McKinnon (1993), lacks at least one of the three abovementioned properties that an \(R^2\) should have.
The form of this function ensures that the values 0 and 1 correspond to “no fit” and “perfect fit” respectively, and that intermediate values have roughly the same interpretations as their analogues in the linear case.

Estrella’s pseudo $R^2$ is easy to apply. First, a probit model with only a constant as explanatory variable is estimated to calculate the maximum value of the restricted likelihood function ($L_C$). Next, a probit model is estimated with the appropriate number of months ahead of the explanatory variable in order to calculate the unconstrained maximum likelihood ($L_U$). These two values are simply substituted into the formula of the pseudo $R^2$. These $R^2$-values are comparable, and the model with the highest is the best model.

A1.4 INDICATORS EXAMINED AND DATA USED

The primary focus of this analysis is to compare the performance of different individual economic indicators in predicting business cycle turning points. Variables such as interest rates, international indicators, stock price indices and monetary aggregates are examined. The performance of these individual indicators will also be compared with the performance of the composite index of leading indicators compiled by the South African Reserve Bank. Most of the components of the composite index of leading indicators for example share prices, money supply and the number of residential building plans passed are also tested individually.

It should be kept in mind that the objective of the composite index of leading indicators is not solely to predict the turning points of the business cycle, but also to provide information regarding the levels of economic growth. It is therefore possible that an individual indicator, even a single component of the composite index, can outperform the index in terms of predicting turning points, even though the index itself is better at predicting the course of the business cycle or the business cycle.
turning points. All the variables included in the analysis are well-established leading economic indicators, and the selection is based on that of Estrella and Mishkin (1998).

Financial variables such as different stock indices are commonly associated with the expectations of future economic events. According to the dividend model of Williams (1938), stock prices are the sum of expected future dividends discounted by future interest or discounting rates. This means that stock indices are forward-looking indicators of expected economic conditions and interest rates and should therefore be good leading economic indicators. Following Estrella and Mishkin (1998), the overall stock index as well as the financial, mining and commercial share indices and the price-earnings ratio were included in the analysis.

Two monetary policy variables, namely short-term interest rates and (different definitions of) money supply, were also included in the analysis. In addition, the long-term interest rate was included since it should reflect expected future short-term interest rates according to the expectations hypothesis.

Recently the yield spread, defined as the difference between the long-term interest rate and the short-term interest rate, as leading indicator has received considerable attention in the literature (see e.g Estrella and Hardouvelis (1991), Bernard and Gerlach (1996), and Estrella and Mishkin (1998)). Assume that the country is currently enjoying high growth, so that there is a general agreement among investors that the country is heading for a slow-down or recession in the future. Consumers want to hedge against the recession, and therefore purchase financial instruments (e.g. long-term bonds) that will deliver pay-offs during the economic slowdown. The increased demand for long-term bonds causes an increase in the price of long-term bonds, in other words a decrease in the yield on long-term bonds. In order to finance these purchases, investors sell their shorter-term assets, which results in a decline in the price of short-term assets, and an increase in the yield on short-term assets. In other words, if a recession is expected, long-term interest rates will fall and short-term interest rates will rise. Consequently, prior to the recession, the slope of the term structure of interest rates will become flat (or even inverted), which means that the yield spread declines. Similarly, long-term interest rates rises while short-term interest
Table A1.2  List of Variables

<table>
<thead>
<tr>
<th>Series</th>
<th>Description</th>
<th>Transformation Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS</td>
<td>Short-term nominal interest rate (3 month BA rate)</td>
<td></td>
</tr>
<tr>
<td>RL</td>
<td>Long-term nominal interest rate (10-year government bond yield)</td>
<td></td>
</tr>
<tr>
<td>SPR</td>
<td>Yield spread, defined as the long-term minus the short-term interest rate (RL-RS)</td>
<td></td>
</tr>
<tr>
<td>M3 (RM3)</td>
<td>Nominal (real) M3 money supply</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>M2 (RM2)</td>
<td>Nominal (real) M2 money supply</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>M1 (RM1)</td>
<td>Nominal (real) M1 money supply</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>JSE</td>
<td>All-share index</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>FS</td>
<td>Financial shares</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>MS</td>
<td>Mining shares</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>CS</td>
<td>Commercial shares</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>PE</td>
<td>Price-earnings ratio</td>
<td></td>
</tr>
<tr>
<td>NEE</td>
<td>Nominal effective exchange rate</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>REE</td>
<td>Real effective exchange rate</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>R$</td>
<td>Rand-US$ exchange rate</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>US</td>
<td>US composite index of leading indicators</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>TR</td>
<td>Composite index of leading indicators of trading partners</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>BP</td>
<td>Building plans passed</td>
<td></td>
</tr>
<tr>
<td>INF</td>
<td>CPI inflation rate</td>
<td></td>
</tr>
<tr>
<td>UO</td>
<td>Manufacturing, unfilled orders</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>NO</td>
<td>Manufacturing, new orders</td>
<td>Year on year growth</td>
</tr>
<tr>
<td>CIL</td>
<td>Composite index of leading indicators</td>
<td>Year on year growth</td>
</tr>
</tbody>
</table>

rates falls when an expansion is expected, so that an upward-sloping yield curve predicts an expansion.

South Africa is a small, open economy and is therefore extremely vulnerable to changes in economies in the rest of the world, especially those of our trading partners and the dominant economies such as the US and Europe. This is increasingly the case since the early 1990s when South Africa re-entered the international economy after
economic sanctions were lifted and globalization generally increased interdependence amongst countries. This motivated the inclusion of the composite index of leading indicators of South Africa’s trading partners as well as that of the US. Since South Africa is such an open economy, exchange rates have a significant influence on the performance of the economy, and since it takes time for changes in the exchange rate to affect domestic prices and hence economic growth, the exchange rate could be a leading indicator of the economy, especially when using high frequency data.

Lastly some macroeconomic indicators such as building plans passed, and unfilled and new manufacturing orders are included on the basis that they reflect the expectations of economic agents.

A1.5 EMPIRICAL ANALYSIS

Monthly data for the period March 1978 to March 2001 was used in the empirical analysis. Forecasts for 1 to 18 months ahead, in other words up to a year and a half, were considered.

A1.5.1 Performance of Individual Leading Indicators

The pseudo $R^2$ developed by Estrella (1998) (see section A1.3.2) is used to compare the forecast performance of each individual leading indicator in forecasting business cycle turning points for 1 to 18 months ahead. The pseudo $R^2$ values of the models are given in table A1.3. Three different transformations of each variable were tested, namely the series in levels, in first differenced from, and the year on year growth in the series. Only the transformation of each series that performed best is reported, the rest of the results are omitted for brevity and available from the author upon request. The transformation of each series that was used is reported in table A1.1. The highest $R^2$ value of each series is indicated in bold print.
Table A1.3  Pseudo $R^2$-values of Leading Indicators

<table>
<thead>
<tr>
<th>Months ahead</th>
<th>SPR</th>
<th>RL</th>
<th>R$$</th>
<th>M3</th>
<th>RM3</th>
<th>TR</th>
<th>INF</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.409</td>
<td>0.158</td>
<td>0.231</td>
<td>0.008</td>
<td>0.083</td>
<td>0.017</td>
<td>0.016</td>
<td>0.141</td>
</tr>
<tr>
<td>2</td>
<td>0.478</td>
<td>0.160</td>
<td>0.266</td>
<td>0.010</td>
<td>0.077</td>
<td>0.023</td>
<td>0.016</td>
<td>0.131</td>
</tr>
<tr>
<td>3</td>
<td>0.540</td>
<td>0.163</td>
<td>0.287</td>
<td>0.018</td>
<td>0.074</td>
<td>0.028</td>
<td>0.016</td>
<td>0.119</td>
</tr>
<tr>
<td>4</td>
<td>0.587</td>
<td>0.163</td>
<td>0.283</td>
<td>0.033</td>
<td>0.079</td>
<td>0.031</td>
<td>0.017</td>
<td>0.108</td>
</tr>
<tr>
<td>5</td>
<td>0.618</td>
<td>0.162</td>
<td>0.268</td>
<td>0.059</td>
<td>0.091</td>
<td>0.033</td>
<td>0.018</td>
<td>0.099</td>
</tr>
<tr>
<td>6</td>
<td>0.635</td>
<td>0.160</td>
<td>0.253</td>
<td>0.087</td>
<td>0.108</td>
<td>0.037</td>
<td>0.018</td>
<td>0.093</td>
</tr>
<tr>
<td>7</td>
<td>0.643</td>
<td>0.160</td>
<td>0.234</td>
<td>0.122</td>
<td>0.131</td>
<td>0.040</td>
<td>0.017</td>
<td>0.089</td>
</tr>
<tr>
<td>8</td>
<td>0.627</td>
<td>0.160</td>
<td>0.214</td>
<td>0.155</td>
<td>0.158</td>
<td>0.043</td>
<td>0.016</td>
<td>0.082</td>
</tr>
<tr>
<td>9</td>
<td>0.578</td>
<td>0.152</td>
<td>0.196</td>
<td>0.194</td>
<td>0.190</td>
<td>0.047</td>
<td>0.016</td>
<td>0.076</td>
</tr>
<tr>
<td>10</td>
<td>0.536</td>
<td>0.144</td>
<td>0.173</td>
<td>0.238</td>
<td>0.230</td>
<td>0.052</td>
<td>0.015</td>
<td>0.071</td>
</tr>
<tr>
<td>11</td>
<td>0.483</td>
<td>0.132</td>
<td>0.152</td>
<td>0.283</td>
<td>0.270</td>
<td>0.058</td>
<td>0.015</td>
<td>0.068</td>
</tr>
<tr>
<td>12</td>
<td>0.424</td>
<td>0.118</td>
<td>0.134</td>
<td>0.340</td>
<td>0.324</td>
<td>0.065</td>
<td>0.014</td>
<td>0.066</td>
</tr>
<tr>
<td>13</td>
<td>0.358</td>
<td>0.106</td>
<td>0.120</td>
<td>0.383</td>
<td>0.368</td>
<td>0.073</td>
<td>0.014</td>
<td>0.068</td>
</tr>
<tr>
<td>14</td>
<td>0.297</td>
<td>0.096</td>
<td>0.116</td>
<td>0.421</td>
<td>0.406</td>
<td>0.079</td>
<td>0.013</td>
<td>0.071</td>
</tr>
<tr>
<td>15</td>
<td>0.245</td>
<td>0.090</td>
<td>0.120</td>
<td>0.452</td>
<td>0.439</td>
<td>0.084</td>
<td>0.012</td>
<td>0.074</td>
</tr>
<tr>
<td>16</td>
<td>0.203</td>
<td>0.088</td>
<td>0.131</td>
<td>0.466</td>
<td>0.398</td>
<td>0.088</td>
<td>0.010</td>
<td>0.078</td>
</tr>
<tr>
<td>17</td>
<td>0.172</td>
<td>0.087</td>
<td>0.148</td>
<td>0.452</td>
<td>0.446</td>
<td>0.091</td>
<td>0.009</td>
<td>0.082</td>
</tr>
<tr>
<td>18</td>
<td>0.150</td>
<td>0.087</td>
<td>0.170</td>
<td>0.455</td>
<td>0.450</td>
<td>0.096</td>
<td>0.008</td>
<td>0.088</td>
</tr>
</tbody>
</table>
From the results in table A1.3 it is clear that the year on year change in the Reserve Bank’s composite index of leading indicators leading 3 months has the highest $R^2$ value, followed by the yield spread leading 7 months. These three models explain 71.2595 percent and 64.3182 percent respectively of the variation in the dependent variable. However, the composite index of leading indicators is only available with a four to five month lag, and is subject to revision. In other words, the optimal number of months ahead is not available in time for forecasting. The months that are available yield lower $R^2$ values than the yield spread, which is immediately available and not subject to revision.
A1.5.2 Probit Models

Table A1.4 presents the results of the probit models with the composite index of leading indicators and the yield spread. Each of the models was estimated with only one explanatory variable and a constant, with the leading time chosen on the basis of the pseudo R² values in table A1.3. The parameters were estimated with maximum likelihood.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Lead (months)</th>
<th>Constant</th>
<th>Standard error</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>Pseudo R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPR</td>
<td>7</td>
<td>0.246</td>
<td>0.107</td>
<td>-0.493</td>
<td>0.050</td>
<td>64%</td>
</tr>
<tr>
<td>CIL</td>
<td>3</td>
<td>0.361</td>
<td>0.119</td>
<td>-0.273</td>
<td>0.030</td>
<td>71%</td>
</tr>
</tbody>
</table>

The results in table A1.4 are interpreted as follows:

\[ P(R_{t+7} = 1) = F(0.246 - 0.493*SPR_t) \]  
\[ P(R_{t+3} = 1) = F(0.361 - 0.273*CLI_t) \]

where F is the cumulative normal distribution, \( R_t \) is a dummy variable that takes on the values one if the economy is in a recession in period t, and \( P(R_{t+i} = 1) \) is the probability that the economy is in a recession in period \( t+i \).

These results are consistent with a priori expectations. According to the results in equation A1.7 there is a negative relationship between the composite index of leading indicators and the probability of a recession, which means that an increase in the composite index of leading indicators predicts a decline in the probability of a future recession. In other words, an increase in the composite index of leading indicators...
indicates a higher probability of an economic upswing, which is consistent with the construction of the composite index of leading indicators. According to equation A6 there is a negative relationship between the interest rate spread and the probability of a recession in future, which means that increases in the interest rate spread lowers the probability of a future recession. This is consistent with the theoretical relationship between the interest rate spread and economic activity, according to which the interest rate spread will decline prior to a recession (see section A1.4).

Table A1.5  Probability of a Recession Two Quarters Ahead as a Function of the Short-Term Interest Rate, the Interest Rate Spread and the Composite Index of Leading Indicators

<table>
<thead>
<tr>
<th>SPR&lt;sub&gt;t&lt;/sub&gt;</th>
<th>P(R&lt;sub&gt;t+7&lt;/sub&gt; = 1)</th>
<th>CLI&lt;sub&gt;t&lt;/sub&gt;</th>
<th>P(R&lt;sub&gt;t+3&lt;/sub&gt; = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6</td>
<td>1.00</td>
<td>-13.00</td>
<td>1.00</td>
</tr>
<tr>
<td>-5</td>
<td>1.00</td>
<td>-10.00</td>
<td>1.00</td>
</tr>
<tr>
<td>-4</td>
<td>0.99</td>
<td>-7.00</td>
<td>0.99</td>
</tr>
<tr>
<td>-3</td>
<td>0.96</td>
<td>-4.00</td>
<td>0.93</td>
</tr>
<tr>
<td>-2</td>
<td>0.89</td>
<td>-1.00</td>
<td>0.74</td>
</tr>
<tr>
<td>-1</td>
<td>0.77</td>
<td>2.00</td>
<td>0.43</td>
</tr>
<tr>
<td>0</td>
<td>0.60</td>
<td>5.00</td>
<td>0.16</td>
</tr>
<tr>
<td>1</td>
<td>0.40</td>
<td>8.00</td>
<td>0.03</td>
</tr>
<tr>
<td>2</td>
<td>0.23</td>
<td>11.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>0.11</td>
<td>14.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td>17.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>20.00</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>0.00</td>
<td>26.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.499</td>
<td>0.5</td>
<td>13.322</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Given these formulas, the probability of a recession associated with certain values of the explanatory variables can be calculated easily. For example, a yield spread of 0.6 percent in a certain period indicates that the probability that the economy will be in a
recession seven periods ahead is 25 percent. The recession probabilities of some of the possible values of the explanatory variables are given in table A1.5. The last row in table A1.5 presents the values of the three economic indicators associated with the probability of a recession of exactly 50 percent. In other words, values of the interest rate spread and composite index of leading indicators below that value predicts that the economy is more likely to be in a recession than an expansion seven or three months ahead respectively, while a short-term interest rate above the value predicts that the economy is more likely to be in an expansion than a recession seven months ahead.

Figures A1.1 and A1.2 plot the estimated probability of a recession derived from each model. The shaded areas denote periods of actual recessions as classified by the South African Reserve Bank, and the lines indicate the probability that the economy is in a recession in that period.

**Figure A1.1  Recession Probability Predicted by Interest Rate Spread**

![Graph showing recession probability predicted by interest rate spread](image)

Source: Own calculations
Figure A1.2  Recession Probability Predicted by Composite Index of Leading Indicators

The lines in figures A1.1 and A1.2 represent the probability that the economy will be in a recession in a particular period as calculated by the three different probit models using the interest rate spread and the composite index of leading indicators respectively as explanatory variables. If the probability of a recession is greater (lower) than 50 percent, it will be regarded as a predicted recession (expansion). These predicted recessions can be compared with the official dates of the South African Reserve Bank presented by the shaded areas. For example, the composite index of leading indicators predicted a recession early in 1981 (when the probability of a recession exceeded 50 percent) compared with the actual recession that occurred at the end of 1981.

None of the two models missed any cycle. However, the model with the composite index of leading indicators gave a false signal of a downswing in January 1996 and an upswing in January 1997. In addition, the model with the composite index of leading indicators gave a false signal of a downswing in January 2001. In general, all three models performed fairly well. The model with the yield spread seems to have performed somewhat worse at the beginning of the sample with the 1983-1984 upswing, while they performed quite well for the rest of the period. On the other hand,
the performance of the model with the composite index of leading indicators seemed to have deteriorated over the sample period.

The deteriorating performance of the composite index and the improving performance of the interest rate model might be the result of important structural change in the economy. And, unlike the composite index, neither of the interest rate models gave any false signals. In addition, the optimal forecast period of the yield spread model is seven months compared to three months in the case of the composite index, and the interest rate variables are available in time and are not revised. Therefore, the yield spread model is preferred to the model with the composite index of leading indicators.

A1.6 CONCLUSION

The objective of this analysis was to compare the performances of difference leading indicators in terms of predicting turning points of the South African business cycle. The pseudo $R^2$ indicated that two best individual indicators are the yield spread and the composite index of leading indicators compiled by the South African Reserve Bank. They led the turning points with seven and three months respectively. A close inspection of the probit models of these two individual indicators as explanatory variables indicated that the yield spread model is preferred to the model with the composite index. Data availability is better in the case of the yield spread, and unlike the composite index, it did not give any false signals. In general, the yield spread model’s performance seemed to have improved over the course of the sample period, while the performance of the composite index seemed to have deteriorated over the course of the sample period. Performance at the end of the sample is obviously more important for forecasting purposes, but these trends might also be reflecting an underlying structural change in the economy, which makes the interest rate models even more desirable since it seems as if they are better at predicting the new structure than the composite index.
APPENDIX 2

MODEL EVALUATION FOR DIFFERENT LOSS FUNCTIONS

A2.1 INTRODUCTION

In chapter seven the forecasting and modeling accuracy of different stock market models were compared using the RMSE, RMSPE and Theil’s inequality coefficient U. In addition, the sign and signed rank tests of Diebold and Mariano (1995) for testing whether the forecasting accuracy of two models are statistically different, were used. These tests require that a loss function be specified. In chapter seven the results of these tests are presented for loss functions that minimize the error terms and the squared error terms. In addition, asymmetric linex loss functions were used since the theory presented in chapter three suggested that investors may behave asymmetrically. The linex loss function is specified as follows:

\[ g(e_t) = \frac{B}{\alpha^2} \left\{ \exp(\alpha e_t) + \alpha e_t - 1 \right\} \quad (A2.1) \]

where \( e \) is the error term of the estimated model. The parameter \( \alpha \) determines the degree of asymmetry. If \( \alpha > 0 \), then the losses are approximately linear for negative error terms and approximately exponential for positive error terms. By defining the error \( e \) as the actual value less the simulated value, positive values of \( \alpha \) corresponds to the case in which underpredictions are more costly than overpredictions. Negative values, on the other hand, corresponds to the case where the function is exponential to the left of the origin and linear to the right. Furthermore, the closer \( \alpha \) is to zero, the closer the function approximates the standard quadratic case.

As explained in chapter six, overpredictions are more dangerous to investors than underpredictions, and therefore negative values of \( \alpha \) are used in this study so that overpredictions are more costly than underpredictions. In chapter seven the results of
the sign test\textsuperscript{2} was already given for different negative values of $\alpha$, which are consistent with the case where overpredictions are more costly than underpredictions. In this appendix, the results will be presented for different positive values of $\alpha$, in other words where overpredictions are less costly than underpredictions. In addition, the influence of different values of $\beta$ on the results will also be illustrated.

A2.2 ESTIMATION RESULTS

In tables A2.1 and A2.2 the models are compared for the sample and forecast periods respectively using the sign test with linex loss functions with different positive values of $\alpha$. In other words, overpredictions are assumed to be less costly than underpredictions\textsuperscript{3}. The null hypothesis of equal modeling accuracy of the random walk and cointegration models during the sample period is rejected for all the loss functions except the first two, which are the closest to being symmetric loss functions. In none of the cases is the null hypothesis of equal forecast accuracy rejected for any pair of models. In other words, using loss functions for which overpredictions are assumed to be less costly than underpredictions, the only statistically significant difference in accuracy is between the random walk and the cointegration model during the sample period.

According to the results in table A2.2, the null hypothesis of equal forecasting accuracy is not rejected for any pair of models. The results in table A2.3 illustrate the impact of the parameter $\beta$ in the linex loss function (see equation A2.1). According to these results $\beta$ does not influence the conclusion of the sign test since the outcome remains constant for a given value of $\alpha$.

\textsuperscript{2} The signed rank test requires a symmetric loss function and is hence not relevant in this case.

\textsuperscript{3} Theoretically overpredictions will be more costly to investors than underpredictions (see chapter seven). The comparisons of the models have been presented in chapter seven. However, the counter-intuitive counterpart, where overpredictions are less costly than underpredictions, are presented in this appendix for completeness.
Table A2.1  Equal Accuracy Tests for Modelling Performance with Different $\alpha$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.1</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Ct}))=0$</td>
<td>1.48</td>
<td>1.69</td>
<td>2.9*</td>
<td>2.7*</td>
<td>3.2*</td>
<td>2.9*</td>
<td>2.7*</td>
<td>2.7*</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Ct}))\neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Vt}))=0$</td>
<td>1.05</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.48</td>
<td>1.26</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Ct})-g(e_{Vt}))=0$</td>
<td>-0.6</td>
<td>-1.5</td>
<td>-1.9</td>
<td>-1.7</td>
<td>-1.5</td>
<td>-1.9</td>
<td>-2.1</td>
<td>-2.1</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Ct})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant on a 10% level of significance.

Table A2.2  Equal Forecast Accuracy Tests with Different $\alpha$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.1</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Ct}))=0$</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Ct})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Vt}))=0$</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Ct})-g(e_{Vt}))=0$</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Ct})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Significant on a 10% level of significance.
Table A2.3  Equal Modelling Accuracy Tests with Different $\beta$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>0.1</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>10</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha=1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Ct}))=0$</td>
<td>2.6*</td>
<td>2.6*</td>
<td>2.6*</td>
<td>2.6*</td>
<td>2.6*</td>
<td>2.6*</td>
<td>2.6*</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Ct})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Vt}))=0$</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Ct})-g(e_{Vt}))=0$</td>
<td>-1.9</td>
<td>-1.9</td>
<td>-1.9</td>
<td>-1.9</td>
<td>-1.9</td>
<td>-1.9</td>
<td>-1.9</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Ct})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha=-1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Ct}))=0$</td>
<td>-1.48</td>
<td>-1.48</td>
<td>-1.48</td>
<td>-1.48</td>
<td>-1.48</td>
<td>-1.48</td>
<td>-1.48</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Ct})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Vt}))=0$</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Ct})-g(e_{Vt}))=0$</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Ct})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha=2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Ct}))=0$</td>
<td>3.2*</td>
<td>3.2*</td>
<td>3.2*</td>
<td>3.2*</td>
<td>3.2*</td>
<td>3.2*</td>
<td>3.2*</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Ct})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Vt}))=0$</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
<td>1.26</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Ct})-g(e_{Vt}))=0$</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Ct})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha=-2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Ct}))=0$</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-1.5</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Ct})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Rt})-g(e_{Vt}))=0$</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Rt})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_0$: $\text{med}(g(e_{Ct})-g(e_{Vt}))=0$</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
<td>1.90</td>
</tr>
<tr>
<td>$H_A$: $\text{med}(g(e_{Ct})-g(e_{Vt})) \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A2.3 CONCLUSION

In this appendix the null hypothesis of equal forecasting accuracy was tested using the sign test suggested by Diebold and Mariano (1995). An asymmetric linex loss function was used. The influence of the parameter $\beta$ in the linex function was shown to be insignificant. In addition, the case in which underpredictions of the stock market is more costly than overpredictions was illustrated. The results showed that the null hypothesis of equal modeling accuracy of the random walk and cointegration models during the sample period is rejected for all the loss functions except the first two, which are the closest to being symmetric loss functions. In all the other cases the models are equally accurate. In other words, using loss functions for which overpredictions are assumed to be less costly than underpredictions, the only statistically significant difference in accuracy is between the random walk and the cointegration model during the sample period. All the models are equally accurate in forecasting the stock market when overpredictions are less costly than underpredictions.