The Term Structure Of Interest Rates And Economic Activity: An Empirical Critique

Clinton Watkins

Department of Economics
University of Western Australia

Abstract Intertemporal asset pricing models, relating macroeconomic variables to asset returns and the term structure of interest rates, suggest that the term structure (yield curve) contains forecast-enhancing information about the future growth rate of real economic activity. Empirical evidence that the term structure contains information regarding a variety of macroeconomic variables has been presented in the literature over the last two decades. Recent empirical studies on the information content of the yield spread have been conducted using simple linear univariate forecasting models. U.S. and G7 data have been found to support a leading association between the term structure of interest rates and changes in the level of future real economic activity. However, many of the econometric models presented in the literature are far from adequate; in particular, several empirical studies of the term structure have not adequately accounted for important explanatory variables. The purpose of the paper is to evaluate the significance of these empirical models. Published empirical research is evaluated in the light of the data used, choice of both dependent and explanatory variables, important omitted explanatory variables, type of model chosen, economic hypotheses tested, methods of estimation and calculation of standard errors for inference (including generated regressors), reported descriptive statistics, use of diagnostic tests of auxiliary assumptions, use of nested and non-nested tests and information criteria, and empirical implications.

1. INTRODUCTION

Interest rates have long been thought to contain and reflect information about the expectations of agents in the economy. Intertemporal general equilibrium models, such as those of Cox, Ingersoll and Ross (Econometrica, 53, 1985, pp363-407), relate macro-economic variables to asset returns and the term structure of interest rates.

The term structure of interest rates represents the relationship among the yields on default-free securities that differ only in their term to maturity. Graphically, the term structure is depicted using the yield curve. Term (or yield) spread defines the difference between a long- and a short-term interest rate for an equivalent security.

In theory, a positively sloped yield curve precedes a future increase in real output, while a flattening or inversion precedes a fall in real output. The mechanism by which intertemporal asset pricing models imply this result may be summarised by the following. Agents are risk averse and desire to hedge their income. Agents purchase bonds of a maturity that will deliver payoffs to smooth expected fluctuations in their level of income. When agents expect a future downturn in economic activity, they prefer long term bonds to short term bonds as the former deliver a payoff during the expected recession. The demand for long term bonds drives up the price of long term bonds, and the long term interest rate falls. The short term interest rate increases as the short term bond price falls. If an increase in economic activity is expected, short term bonds are preferred as they deliver a payoff before the expected expansion. Short term interest rates fall, and long term interest rates rise.

A generic specification for the term structure-based economic growth forecasting model is that of equation (1) below:

$$Y_{t,t+k} = f(TS_t; I_t; V_t; OV_t) \qquad (1)$$

where:
i) Growth in output, $Y_{t,t+k}$, is over period t to t+k,
ii) TS is the current term structure of interest rates,
iii) I represents particular interest rate levels,
iv) V is the variability of interest rates,
v) OV represents other relevant variables.

If the slope of the yield curve accurately forecasts economic growth, and the relationship is stable over both time and monetary regimes, then the term structure of interest rates is a valuable leading economic indicator.

2. EMPIRICAL ISSUES

Five recent, important papers examining term structure based economic growth forecasting models are divided into twenty cases, depending on the type of model in question. The studies examined are Chen (1991), Estrella and Hardouvelis (1991), Harvey (1989, 1991a,b).

The economic hypotheses tested can be classified into three main types (see Table 1). The first type relate to the term structure, the second and third types relate to causality. There are only two applications of the causality hypotheses. One for U.S. or world yield spreads leading or lagging those of the G7 nations, and one for U.S. or world business cycles leading or lagging those of the G7 nations. The bulk of the hypotheses tested relate to the term structure. There are seven aspects of hypotheses relating to the term structure. The most frequently considered such example, where there are eleven cases, has to do with explanations of the future growth in real GNP. Other components of the term structure hypothesis examined involve explanation of a greater proportion of future growth in real GNP relative to the use of stock market variables, macroeconomic variables, and survey data or institutional forecast models. There are two examples where the probability of a recession is forecasted, and the success or otherwise of such forecasts are used to evaluate the term structure hypothesis indirectly. Finally there is one example for each of the explanation of individual components of GNP and explanation of the conditional variance of the growth rate of GNP, as methods for testing the term structure hypothesis.

Type of Hypothesis	Freq.
The term structure:	
explains future growth in real GNP	11
explains a greater proportion of future growth in real GNP than do stock market variables	5
explains a greater proportion of future growth in real GNP than do macroeconomic variables	4
explains a greater proportion of future growth in real GNP than do survey data or institutional forecast models	3
forecasts the probability of a recession	2
explains individual components of GNP	1
explains the conditional variance of the growth rate of GNP	1
US or world yield spreads lead or lag those of the G7 nations	1
US or world business cycles lead or lag those of the G7 nations	1

Table 1: Economic Hypotheses Tested

In the context of the above, no mention of non-stationarity is made in any of the reviewed papers. Therefore, it is implicitly assumed that variables are stationary over the sample period. Moreover, the correlations observed are assumed not to be spurious, and merely resulting from some trend properties of the data.

2.1. Sample Data

In Table 2, number of quarterly observations is decomposed into three broad categories. In terms of the frequency of use of such quarterly observations, consideration is given to both modelling as well as out-of-sample forecasting. It should be noted, however, that each attempt at out-of-sample forecasting did not use or consider the issue of virgin data, but used data that was part of the modelling process. In the majority of instances, modelling was conducted over what would be considered small samples. For what would be regarded as a moderately sized sample for quarterly observations, namely 101-145 observations, there are 10 examples where data is used for modelling, but no studies retained any observations for out-of-sample forecasting.

No. of Quarterly Observations	Frequency Of Use	
	Modelling	Out-of-sample Forecasting
1 - 20	1	0
21 - 30	1	2
31 - 50	2	0
51 - 100	14	2
101 - 145	10	0

Table 2: Number Of Quarterly Observations

2.2. Dependent Variables

Measures of the dependent variable used in the term structure-based economic growth forecasting model are the marginal percentage change in real GNP used in eleven models, and the cumulative percentage change in real GNP used in seven applications. A dichotomous dependent variable, the probability of a recession, was used in a variant of the model used to forecast the probability of recession. The percentage change in real GNP components, conditional variance of the real GNP growth rate, and the yield spread occurred in one model each.

Type	Number
Marginal percentage change in real GNP	11
Cumulative percentage change in real GNP	7
Dichotomous (forecast of the probability of a recession)	2
Percentage change in real GNP components	1
Conditional variance of the real GNP growth rate	1
Yield spread	1

Table 3: Choice of Dependent Variable

2.3. Explanatory Variables

Table 4 indicates how frequently particular explanatory variables are included in the surveyed empirical models. Each model uses some term spread variable as a proxy for the term structure (or slope of the term structure) since the term structure is not a precisely measurable variable. The term structure is represented by either the 5-year, 10-year, or lagged 10-year nominal term spread variables. Interest rate theory suggests that the information content of the term structure is captured by three measurable factors[1]: (i) the short term rate of interest; (ii) the long term rate of interest; and (iii) interest rate variability. As the term spread is the difference between a long- and short-term interest rate, the first two measurable factors suggested by interest rate theory are, to some extent, represented by this measure. Some models also include the level of a short-term interest rate. None of the models include any variable capturing the variability of interest rates. Three of the surveyed models contain an expected real interest rate regressor (see Table 7) that is generated from another model. Other explanatory variables listed in Table 4 are used in nested or non-nested forecasting models which are evaluated relative to the term structure-based model.

Type	Number
5-year nominal term spread[a,b]	10
10-year nominal term spread	11
Lagged 10-year nominal term spread	1
1-month Treasury Bill rate	3
Ex ante real federal funds rate[c]	3
1-quarter stock return	3
4-quarter stock return	3
Dividend yield	3
Default spread	3
Lagged growth in real GNP[b,c]	3
Lagged industrial production growth	3
Lagged growth of an index of leading indicators[c]	2
Lagged rate of inflation[c]	2
Survey forecast of cumulative percentage change in real GNP	1

[a] Includes one entry specified for 4 leads, current and 4 lagged values of the variable.
[b] Includes one combined 3- and 4-year term spread.
[c] Includes one entry for the variable lagged by 4 quarters.

Table 4: Choice of Explanatory Variables

A proxy variable in a regression necessarily implies the presence of measurement error since the correct variable is not used. Measurement error in an explanatory variable causes serious problems for ordinary least squares estimation, essentially biased and inconsistent estimates[2]. One solution is to use instrumental variable estimation, which yields consistent but inefficient estimators. Frequently, a suitable instrument correlated with the proxy variable and uncorrelated with the error term is difficult to obtain.

Models that use the nominal term spread as a proxy for the term structure assume either (i) the real interest rate is constant and the term spread reflects expected inflation, or (ii) the expected rate of inflation is constant and models forecasting growth in real GNP are extracting information from the real term structure. Neither assumption is reasonable, and both are at odds with the empirical evidence[3].

Harvey (1989, 1991a,b) assumes inflation follows a first order integrated moving average process[4] (IMA(1,1)). He continues that this assumption means the *intercept* contains another variable, the expected real short-term rate of interest. The idea that the intercept (as a constant) can contain a variable, unless it is constant, is bizarre. Further, this variable in the intercept is shown not to contribute to the explanatory power of the model for one- to three-quarter forecasting horizons! This should not be an unexpected result! An omitted variable manifests itself in the error term, resulting in biased estimates of the coefficients.

The nominal yield spread between two points on the yield curve is a restrictive proxy for the term structure. A two step OLS estimation procedure that estimates the steepness of a specific non-linear transformation of the yield curve using the entire length of the yield curve has been proposed by Frankel and Lown (1994)[5]. Improved measurement of the term structure must be beneficial for estimation where a model contains only one proxy variable.

2.4. Excluded Explanatory Variables.

Most studies recognised the possibility of omitted variables, although none discussed the consequences of this for their models. In one case, no excluded explanatory variables were acknowledged, although in each other case, more than one excluded explanatory variable was recognised. In Table 5, excluded explanatory variables, as recognised through self reporting in the various papers examined, are tabulated. Apart from the one case that did not recognise the exclusion of any explanatory variables, three broad categories are considered. These are bond market variables, stock market variables, and other variables.

Although in several cases it was recognised that the "correct" slope of the yield curve had been omitted, no authors recognised the fact that their measure of the term structure, the term spread, is a proxy variable.

[1]After empirically evaluating various theoretical models of the term structure, Stambaugh (1988) concludes at most, three variables appear to be sufficient to represent the information content of the term structure of interest rates.
[2]Garber and Kepper (1980) conduct a detailed analysis of this errors-in-variable problem.

[3] See, for example, Mishkin (1990).
[4]If the inflation rate is assumed to follow an IMA(1,1) process, it will, in effect, be constant, and the term spread will reflect the real term structure.
[5]In this case the second step of the estimation procedure is to forecast future inflation. However, the first step may be adapted to be used in a similar two step procedure to forecast the level future economic activity.

Further more, there is no recognition of measurement error or proxy variable problems with OLS estimation.

Type	Number
Bond Market Variables[1]	
"Correct" slope of the yield curve	7
Interest rate volatility	7
Expected short-term real interest rate	9
Real federal funds rate	3
Real yield spread	9
Nominal yield spread[2]	9
Additional yields on bonds of different maturities	7
Stock Market Variables	
Risk of portfolio[3]	1
Cash flows from stock market investments[3]	1
Investors' time horizon[3]	1
Discount rate[3]	2
Share price index	3
Other variables	
Other information variables	10
Other macroeconomic variables	2
Some measure of relative price change	3
Rate of inflation	3
None	1

[1] Long term interest rates have not been included separately in any models.
[2] Nominal yield spread as defined by the difference between the yield on a bond with 5 quarters to maturity and a bond with 1 quarter to maturity.
[3] Variables recognised as being excluded specifically from models alternative to the term structure model.

Table 5: Recognition of Excluded Explanatory Variables

2.5. Model Specification

Three types of regression model are considered in the empirical applications, namely, a simple linear regression model, the multiple linear regression model, and the probit model. Of the twenty cases of empirical applications considered, twelve cases were in the context of the simple linear regression model, six were multiple linear regression models, and on two occasions, the probit model was estimated.

Type	Number
Simple linear regression model	12
Multiple linear regression model	6
Probit	2
Total	20

Table 6: Model Specification

2.6. Method of Estimation

Ordinary least squares estimation was used for almost all models (see Table 7). Heteroscedasticity and/or serial correlation was expected in a number of models. These were estimated by OLS and the standard errors were given by the Newey West heteroscedasticity consistent method. The method of estimation was not reported for three regression models, which were presumed to be OLS or OLS with Newey-West standard errors. Where out of sample forecasts were desired, the models were estimated by OLS on a sub-sample. Recursive least squares is presumed to be the method by which out of sample forecasts are generated recursively since the method of forecast generation was not reported. Two Probit models explaining the probability of recession were estimated by maximum likelihood.

Type	Number		
	No Generated Regressor	Generated Regressor	Total
OLS	2	0	2
OLS (presumably)	2	0	2
OLS, with Newey-West SE's	7	2	9
OLS (presumably), with Newey-West SE's	1	0	1
OLS / Recursive Least Squares (presumably)	4	0	4
Maximum likelihood	1	1	2
Total	17	3	20

Table 7: Methods of Estimation and Inference

Generated regressors were present in three models. Two were estimated by two-step ordinary least squares with Newey-West standard errors. This method produces standard errors that have been shown to be no better than those produced by the (incorrect) two-step ordinary least squares procedure[6]. The expected real interest rate used in the probit model of Estrella and Hardouvelis (1991) was generated by OLS, the structural equation was estimated by maximum likelihood. Correct two-step estimation was not used for any model containing a generated regressor.

No models acknowledged the measurement error problems (bias and inconsistency of OLS estimates) implied by the use of proxy variables. Instrumental variable estimation was not used for any model.

2.7. Diagnostic Tests

No diagnostic tests were conducted in any of the surveyed papers even though authors discussed the possibility of structural change, serial correlation, moving average errors, heteroscedasticity, omission of relevant variables, and predictive failure.

[6]McAleer and Smith (1994) provide evidence from Monte Carlo experiments showing that the Newey-West procedure performs no better than two-step OLS.

Type[a]	Number		
	Reported	Not Reported	Total
R^2 / adjusted R^2 / pseudo-R^2	15	5	20
Standard error	5	15	20
ME / MAE / RMSE	4	16	20
Diagnostics	0	20	20
Nested / Non-nested tests	2[b]	18	20

[a] R^2 is the coefficient of multiple determination (or squared multiple correlation coefficient), ME is the mean error of the regression, MAE is the mean absolute error, and RMSE is the root mean squared error.
[b] Two entries represent non-nested F-tests reported, but not recognised by the authors.

Table 8: Reporting of Descriptive Statistics, Diagnostics and Nested / Non-nested Tests

Diagnostic testing plays a critical role in assessing the adequacy of empirical economic models. Given the theoretical model, an intermediate specification is needed to obtain an equation suitable for estimation and testing. This involves making numerous auxiliary assumptions about functional form, relevant variables and that they approximate the "true" variables of interest, stability of the model etc. Pesaran and Smith (1985, p138) noted that "A consequence of this procedure is that one cannot know whether the results of the statistical analyses reflect inferentially on the economic theory or on the auxiliary assumptions".

In every paper surveyed, it was assumed the error term was heteroscedastic. No diagnostic test for heteroscedasticity was conducted, however, Newey-West heteroscedasticity consistent standard errors were frequently used. Estrella and Hardouvelis (1991) use Newey-West standard errors to correct for assumed moving average errors of unknown order and conditional heteroscedasticity, without testing these assumptions.

2.8. Descriptive Statistics

Discrimination criteria refer to goodness of fit measures such as the coefficient of multiple determination and information criteria. These descriptive statistics assess how well different models fit the data, with some adjustment for parsimony. The philosophy behind using discrimination criteria to choose between models is that the best predicting model is the closest approximation to the "true" specification. The adjusted coefficient of multiple determination was the most commonly reported descriptive statistic (see Table 8). In some cases, the adjusted coefficient of multiple determination was the only statistic, of any kind, reported for a regression model. The adjusted R^2 was most used in the evaluation of competing nested or non-nested models, or predictive ability. A number of measures of the regression error were used to evaluate and compare competing regressor models. In five cases, the standard error of the regression was reported as a measure of the statistical adequacy of an estimated regression model. In four cases, the out-of-sample forecasting ability of competing models was compared on the basis of the models' mean absolute error and root mean squared error[7]. The principal disadvantage of discriminating between models on the basis of goodness of fit measures is that each model is evaluated only in terms of its own performance. One model will always be chosen, regardless of whether or not it can predict the consequences of separate alternatives.

2.9. Non-Nested Testing

Non-nested tests are specification tests with specific alternatives. The purpose of separate tests is to achieve high power against the specified alternative. In the assessment of separate specific alternatives, an appropriate philosophy is to test whether the null model can predict the performance of an alternative model "significantly well". The essential difference between discriminating on the basis of descriptive statistics or information criteria and on the basis of testing is that the latter enables the classical inferential procedures to be applied.

Given that a substantial part of the empirical research surveyed aims to test hypotheses (see Table 1) of the superiority of the term structure-based forecasting model over alternative models, or other variables relevant to alternative models, it is surprising that no nested or non-nested tests were (intentionally) used (see Table 8). Estrella and Hardouvelis (1991) conduct non-nested F-tests between the term structure-based model and a survey information model without recognition. Being unaware of the non-nested test used, the authors misinterpret the test statistics, and draw mistaken inference[8].

To conduct the non-nested F-test, the separate models are nested within a general model. The standard F-test is then used to test the competing models as special cases. However, this test does not use information regarding the competing separate models in an optimal fashion. There is the possibility of high multicollinearity

[7]Mean error was also used in two cases.
[8]McAleer (1995) cites a common problem in empirical publications dealing with non-nested models in that the authors are unaware of which non-nested tests are actually being used, and misinterpret published test statistics.

between the regressors, and low degrees of freedom due to the number of regressors, in the general model.

Estrella and Hardouvelis (1991, p572) conduct a test between alternative non-nested models for 1, 2 and 3 quarters ahead forecasting horizons and interpret the results to show "that spread is a better predictor of future output growth than the median survey forecast". Seemingly oblivious to the fact that they have conducted a non-nested test, the authors compare the models the basis of each equation's coefficient of multiple determination. In addition, they observe that the R^2 does not increase when an additional variable SURVEYF is included. A correct analysis of the non-nested F-test shows that, for one-quarter ahead, both models are rejected. For two- and three-quarters ahead, the survey information model is rejected while the term structure based model is not.

3. CONCLUSION

With regard to the empirical conclusions, in almost every case, the empirical papers find that the term structure is a statistically adequate forecasting model, and, indeed, is a superior model when compared with other alternatives. The ability to draw such a conclusion from the broad range of empirical papers considered rests very strongly on the models being interpreted as adequate from a statistical point of view. The published empirical research was evaluated in the light of the data used, choice of both the dependent and set of explanatory variables, self reporting of important omitted explanatory variables, the type of model actually chosen for estimation, the economic hypotheses tested therein, the various methods of estimation conducted and the calculation of correct standard errors for statistical inference (some of which were effected by the presence of generated regressors), the reported descriptive statistics, the use of diagnostic tests of the set of auxiliary assumptions, and the use of both nested and non-nested tests and information criteria.

On the basis of a serious analysis of the empirical implications of each of these estimation papers, an indubitably correct inference would be that the empirical conclusions drawn by each and every one of these authors, with regard to each and every one of their estimated models, would be that the models themselves are statistically inadequate. Consequently any empirical conclusion drawn on the basis of such inadequate models would need to be interpreted very cautiously indeed.

4. ACKNOWLEDGMENTS

I would like to thank Michael McAleer for his invaluable advice, many helpful comments and discussions; Mike Veall for helpful comments and suggestions; and participants at the Fourth General Meeting of the International Environmetrics Society and the Fifth International Conference on Statistical Methods for the Environmental Sciences, Ontario, Canada, August 1994.

5. REFERENCES

Chen, Nai-Fu., Financial Investment Opportunities and the Macroeconomy, *Journal of Finance*, 46, 529-554, 1991.

Cox, J.C., J.E. Ingersoll, JR., and S.A. Ross, An Intertemporal General Equilibrium Model Of Asset Prices, *Econometrica*, 53, 363-384, 1985a.

Cox, J.C., J.E. Ingersoll, JR., and S.A. Ross, A Theory Of The Term Structure Of Interest Rates, *Econometrica*, 53, 385-407, 1985b.

Estrella, A. and G.A. Hardouvelis, The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance*, 46, 555-576, 1991.

Frankel, J.A. and C.S. Lown, An Indicator Of Future Inflation Extracted From The Steepness Of The Interest Rate Yield Curve Along Its Entire Length, *Quarterly Journal of Economics*, 109(2), 517-530, 1994.

Garber, S. and S. Kepper, Extending the classical normal errors-in-variable model, *Econometrica*, 48, 1541-1546, 1980.

Harvey, C.R., Forecasts of Economic Growth from the Bond and Stock Markets, *Financial Analysts Journal*, September/October, 38-45, 1989

Harvey, C.R., The Term Structure And World Economic Growth, *Journal of Fixed Income*, June, 7-19, 1991a.

Harvey, C.R., Interest Rate Based Forecasts of German Economic Growth, *Weltwirtschaftliches Archiv*, 124, 701-718, 1991b.

McAleer, M., The significance of testing empirical non-nested models, *Journal of Econometrics*, 67, 149-171, 1995.

McAleer, M., and J. Smith, Newey-West covariance matrix estimates for models with generated regressors, *Applied Economics*, 26, 635-640, 1994.

Mishkin, F.S. What does the Term Structure tell us about Future Inflation?, *Journal of Monetary Economics*, 25, 77-95, 1990.

Pagan, A. R. Econometric Issues in the Analysis of Regressions with Generated Regressors, *International Economic Review*, 25, 221-47, 1984.

Pesaran, H. And R. Smith, Keynes on Econometrics, in T. Lawson and H. Pesaran (eds), *Keynes' Economics: Methodological Issues*, Croom Helm, London, 134-150, 1985.

Stambaugh, R.F., The Information in Forward Rates: Implications For Models of Term Structure, *Journal of Financial Economics*, 21, 41-70, 1988.