

Modelling and Forecasting Interest Rates during Stages of the Economic Cycle: A Knowledge-Discovery Approach

David Diaz
Escuela de Economía y
Negocios
Departamento de
Administración
Universidad de Chile
Santiago,
Chile

ddiaz@unegocios.cl

Babis Theodoulidis
Manchester Business School
The University of Manchester
Manchester,
United Kingdom

b.theodoulidis@manchester.ac.uk

Carlos Dupouy
Central Bank of Chile
Santiago,
Chile

cdupouy@bcentral.cl

Corresponding Author

David Diaz
Escuela de Economía y Negocios
Departamento de Administración
Universidad de Chile
Santiago, Chile

Diagonal Paraguay 256, Oficina 1103
Telephone: +56229783708
ddiaz@unegocios.cl

The opinions and results of this work are the sole responsibility of the authors. They do not represent in any way the institutional views or policies of their affiliations.

Abstract

Modelling the structure of risk-free rates and their relation to other economic and financial variables during different stages of the economic cycles has attracted much interest from both the theoretical and practical perspectives. The previous literature has emphasized the deployment of expert systems and knowledge-discovery approaches motivated by the need to address the limitations of the econometric models. However, it has failed to address the interpretability aspects and, more importantly, the need to provide methodological support that allows the deployment of such techniques in a more systematic way. This approach entails the definition of a process that includes the usual steps taken by experts to address similar problems and allows the relative merits of different techniques in relation to common goals and objectives to be gauged.

This paper addresses the interpretability and the lack of methodological support by proposing a knowledge-discovery methodology that includes a minimal common number of steps to model, analyse, evaluate and deploy different non-linear techniques and models. Furthermore, the interpretability is addressed through the use of open-box techniques, such as decision trees.

The proposed methodology helps to discover and describe hidden patterns, allowing for the study and characterization of economic cycles, and economic cycle stages, as well as the description of the historic relationships between interest rates and other relevant economic variables. These patterns can also be used in the forecasting of economic cycle stages, interest rates and other related variables of concern. The output of the methodology can provide actionable information for market agents, such as monetary authorities, financial institutions, and individual investors, as well as for the academic community, to increase further the knowledge and understanding of financial markets, thus enriching and complementing existing financial theories.

Keywords: interest rates, yield curve, economic cycles, forecasting, data mining, decision trees

1. Introduction

In finance, expert systems are used to help practitioners with their decision making in relation to stock market, securities and commodities prediction. This task has been dealt with two main approaches, one that explicitly considers theory in relation to market characteristics (Cox, Ingersoll, & Ross, 1985; Diebold & Li, 2006; Estrella & Hardouvelis, 1991; Estrella & Mishkin, 1997, 1998; Svensson, 1994; Vasicek, 1977) and another that explores the existing patterns and relationships in the data without making explicit reference to financial theories (Araújo, Oliveira, & Meira, 2015; Cervelló-Royo, Guijarro, & Michniuk, 2015; Chang, Liu, Lin, Fan, & Ng, 2009; Guresen, Kayakutlu, & Daim, 2011; Kamo & Dagli, 2009; Svalina, Galzina, Lujić, & Šimunović, 2013),

One of the most significant problems in finance that relies on expert systems involves forecasting the future movements of interest rates (Enke & Thawornwong, 2005; Hong & Han, 2002; Ju, Kim, & Shim, 1997; Oh & Han, 2000). Part of the challenge has to do with the nonlinear and dynamic nature of interest rates, as well as their relationship to economic cycles and other economic variables. The latter relates to the fact that interest rates are part of the monetary policy of governments in order to control inflation and price stability.

Typically, authorities make use of the Monetary Policy Rate (MPR), which serves as an anchor and level for all other interest rates in the local market, as well as issuing risk-free debt at different maturities and denominations (Bank of England, 2015). The implicit zero-coupon rates on these emissions form the risk-free term structure that is usually represented as a graph known as the *Yield Curve* (Alexander, Sharpe & Bailey, 2003), which corresponds to one of the most important economic indicators in existence. Its importance lies in the number of factors that are summarized on its shape, such as risk premium levels, investment expectations, real sector activity, and the balance between supply and demand for debt of different maturities in the future (Estrella & Hardouvelis, 1991; Estrella & Trubin, 2006; Estrella, 2005).

In order to determine the right levels of MPR and other risk-free debt rates, the monetary authorities adopt a continuous monitoring and review process of market conditions, during which they characterize the economic cycles through the analysis of several factors and the mechanisms by which monetary decisions impact inflation and economic activity (see Figure 1). In a similar way, secondary market participants monitor economic cycles and the interactions between variables, and take trading decisions based on the expected evolution of such factors (Resnick & Shoesmith, 2002).

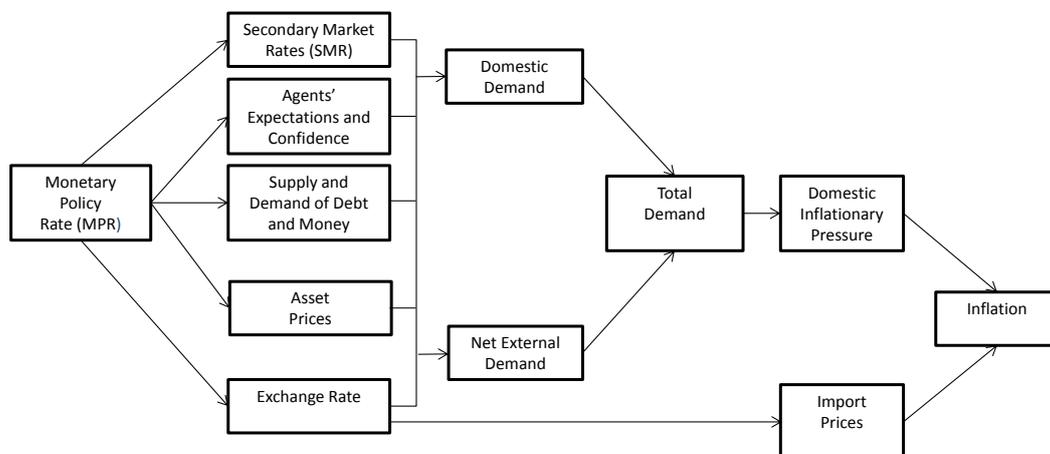


Figure 1: The transmission mechanism of monetary policy [adapted from Bank of England (2015)]

Understanding the dynamics of interest rates, and the relationship between the shape of the yield curve and the future economic cycle evolution is a problem that holds great interest for academia and market participants. In the literature, two main streams of research can be identified: theoretical and empirical. The theoretical stream attempts to explain the characteristics (such as the slope, level and curvature) of the yield curve on the basis of various financial theories (Alexander et al., 2003; Brown & Dybvig, 1986; Cox, Ingersoll Jr & Ross, 1981; Heath, Jarrow & Morton, 1992; Hicks, 1939; Ho & Lee, 1986; Nelson & Siegel, 1987; Svensson, 1994; Vasicek, 1977).

The empirical stream deploys the use of different econometric methods and knowledge-discovery techniques in order to model the structure and dynamics of the yield curve (Enke & Thawornwong, 2005; Gogas, Chionis & Pragkidis, 2009; Gogas, Papadimitriou, Matthaiou & Chrysanthidou, 2014; Jacovides, 2008; Ju, Kim & Shim, 1997; Kim & Noh, 1997; Oh & Han, 2000; Vela, 2013; Zimmermann, Tietz & Grothmann, 2002). Recently, the use of knowledge-discovery techniques has gained popularity, mainly due to the fact that these are able to handle complex, non-linear relationships between variables, seasonality and the presence of structural breaks (Gogas et al., 2014; Jacovides, 2008; Ju et al., 1997; Kim & Noh, 1997; Oh & Han, 2000; Vela, 2013; Zimmermann et al., 2002). However, these are not exempt from limitations. Most noticeable is the difficulty of interpretation, bearing in mind that the deployed techniques can be considered “black-boxes”, in which it is extremely hard to understand the abundant non-linear patterns that are taken into account for each prediction.

Furthermore, the deployment of knowledge-discovery techniques reported in the literature is carried out on an ad hoc basis that considers specific definitions of the problem, specific tools and special characteristics of the data sets. The task of understanding relationships in the context of interest rates and related economic variables would greatly benefit from the use of a methodological approach, as is the norm in other domains. This approach entails the definition of a process that includes the usual steps taken by experts to address similar problems, and allows for comparison between the different techniques in order to be able to gauge their relative merits in relation to common goals and objectives.

This paper addresses the lack of methodological support by proposing a knowledge-discovery methodology that includes a minimal common number of steps to model, analyse, and evaluate non-linear relationships between interest rates and the relevant related economic variables. In addition, the use of decision trees (J R Quinlan, 1993; J. R. Quinlan, 1990) when modelling interest rates is examined and compared with a wide range of statistical and knowledge-discovery techniques commonly mentioned in the literature. The comparison is carried out in terms of model accuracy, fitness and interpretability of its results. The proposed methodology is used to model and forecast risk-free interest rates, taking into consideration the stages of the economic cycles. Through the use of data-mining techniques such as clustering and decision trees, the proposed methodology addresses the limitations of previous work. To the best of our knowledge this is the first attempt to use decision trees as the tool to investigate and present in human, understandable form the non-linear patterns found in the relationships between interest rates and related economic variables.

The proposed methodology covers the complete lifecycle of a typical knowledge-discovery process, including the deployment of the resulting models in a simulated setting. Its use is demonstrated and evaluated through the characterization and analysis of the economic cycles and interest rate forecasting in relation to a specific secondary market bond. In addition to the novel application of these techniques in this context, the results show their appropriateness and benefits compared to existing econometric models, in relation to both the selection and ranking of variables at each stage of the cycle and to the prediction of movement of the interest rate.

The structure of the paper is as follows. Section 2 summarizes the relevant literature in terms of the different streams of research, highlighting the gaps and limitations addressed in this paper. Section 3 presents an overview of the proposed methodology in relation to its phases and the objectives pursued in each phase. Section 4 delivers a further description of the modelling and evaluation phases, giving details of the models and their results and how these are evaluated. Section 5 discusses a possible scenario for the deployment of selected models in order to provide the organizational context for the exploitation of the results. Finally, Section 6 summarizes the work, discusses its limitations and identifies areas for future work.

2. Literature Review

From a theoretical perspective, previous research attempts to explain the forms generally displayed on the term structure of interest rate curves, four of which have been the most widely accepted in the financial literature. These are: the *Theory of Unbiased Expectations* (Hicks, 1939) that seeks to explain the shape of the yield curve as the result of investor expectations about interest rates and future inflation; the *Liquidity Preference Theory* (Hicks, 1939) that is a modification of the theory of expectations through the inclusion of a premium for liquidity of short-term rates; the *Market Segmentation Theory* (Alexander et al., 2003) which assumes that investors and borrowers are restricted by law, preference or custom to certain maturities, and that they will not leave their market or enter a different market, even when there are substantial changes to market conditions and the *Preferred Habitat Theory* according to which investors and borrowers still have preferred segments of the market in which they operate but they are willing to move between segments when there are significant differences in yield between segments (Alexander et al., 2003).

From an empirical perspective, four different streams of literature are identified. The first sub-stream includes models that use the slope of the yield curve as the main economic variable to forecast inflation and GDP growth (Estrella & Mishkin, 1997). Under this, a number of studies have adopted linear and non-linear econometric models and consider the spread between long-term and short-term government bond rates as a significant variable to explain inflation and other variables related to economic cycles such as real economy activity growth, industrial production and recession indices (Ang, Piazzesi & Wei, 2004) (Estrella & Mishkin, 1997) (Bernanke, Boivin & Elias, 2005) (Bernanke et al., 2005) (Diebold & Li, 2006) (Estrella & Hardouvelis, 1991) (Estrella & Trubin, 2006). These models exhibit strong predictive power over time and across different countries (Estrella, Rodrigues, & Schich, 2003).

The second sub-stream includes the so-called *data-driven models*, which deploy methods that try to fit a mathematical function to spot rates such as *splines* (McCulloch, 1971, 1975) and *parsimonious models* (Nelson & Siegel, 1987) (Svensson, 1994). Overall, data-driven and parsimonious models are not intended to model the future behaviour of interest rates. They focus on interpolating the current shape of the term structure. As such they have limited applicability in the estimations of future volatilities, option pricing or simple simulation, as they do not attempt to extract such information from the cross-sectional data for which the adjustments are generated.

The third sub-stream of research are *dynamic models*. These include the equilibrium models (Cox et al., 1985; Vasicek, 1977) and the arbitrage-free models (Ho & Lee, 1986) (Hull & White, 1990) (Heath et al., 1992). There is a long list of variations to these models; however, the detailed review of this family of models is outside the scope of this paper.

The fourth sub-stream of research under the empirical line of investigation refers to data-driven approaches that rely on knowledge-discovery techniques to model interest rates. The main motivation

behind this approach is to address the limitations of previous streams to handle complex non-linear relationships between variables, in addition to seasonality and the presence of structural breaks. The majority of research in this sub-stream relates to predicting interest rates, that is, modelling the time dynamics of one specific point on the yield curve (Jacovides, 2008; Kim & Noh, 1997; Oh & Han, 2000; Vela, 2013; Zimmermann et al., 2002). Work has also been published on analysing the characteristics of the economic cycles on the basis of the slope of yield curves attempting to forecast recessions using machine-learning techniques, such as Support Vector Machines (SVM) (Gogas et al., 2014).

Several other studies have used different knowledge-discovery and machine-learning techniques for related tasks: in Kim and Noh (1997), neural networks and case-based reasoning (CBR) are used both separately and together to forecast interest rates, and their results are compared. CBR produces cases (groups) of variables in which the result is the same (or very similar). When used in combination with neural networks, CBR is used to reduce the combinations of variables that are taken as input and also provide an explanation based on the patterns found. For the US, the performance of the neural networks exceeds CBR together with a random walk. For Korea, neither model produces statistically significant results.

In Ju et al. (1997) fuzzy logic and genetic algorithms are used to estimate interest rates and are applied to the Korean economy. The authors argue that the proposed approach shows fuzzy logic to be a viable alternative to neural networks and that it is worth examining hybrid models that combine both. In Oh and Han (2000), neural networks are used to estimate the rate of interest in the US. They define as “turning points” any structural change that is the result of the implementation of monetary policy by the government. The authors argue that the combination of the detection of turning points together with neural networks is statistically superior to pure neural networks. In Zimmermann et al. (2002) three variations of neural networks are deployed to estimate the yield curve based on data from Germany in three and six months. The results show that it is possible to optimize the performance of neural networks, especially for error correction. In Hong and Han (2002), an expert system consisting of a series of rules is used which discriminates whether news items from the internet can be considered “positive” or “negative”. This is used as input to a neural network that also includes economic and financial variables such as the amount of money, economic growth, monetary policy rate, commodity prices and unemployment rate. The results show a performance improvement compared to a conventional neural network and argue the importance of presenting a causal explanation for each news item on the market. In Enke and Thawornwong (2005) a two-stage neuro-hybrid rolling forecasting model, called fuzzy inference neural network, is proposed to deal with the dynamic, non-linear, volatile and complex nature of interest rates. It is shown that the model achieves a good performance in relation to neural networks. In Jacovides (2008) neural networks and SVM are compared in estimating the six-month interest rate for the United Kingdom with different maturities. The results show that the performance of SVM is superior to neural networks. In Joseph, Larrain and Singh (2011) neural networks are used to predict recessions based on interest rate spreads – in other words, the difference between the average yields on ten-year US Treasury bonds and on three-month US Treasury bills. It is shown that neural network models outperform regression models, as evidenced by the R-squared and mean square error performance metrics. In Vela (2013) the predictive power of the yield curves for different countries in Latin America is analysed based on variables from previous research such as Nelson and Siegel (1987) and Svensson (1994). In comparative terms, the neural networks had a higher yield than the other models, but for some curves, the neural networks did not exceed the random walk model. In Gogas et al. (2014) SVM was used to analyse the predictive ability of the yield curve in terms

of the US real GDP cycles. The results show that a high overall accuracy -between 66 and 100 per cent- in forecasting recessions can be achieved.

Overall, for the fourth sub-stream of research, the increased popularity of expert systems and knowledge-discovery techniques, motivated by the need to address the limitations of the econometric models, is evident. Powerful non-linear techniques, such as neural networks and SVM, are becoming the customary tools when modelling complex relationships between interest rates and related economic variables. However, these are not exempt from limitations. Most noticeable is the difficulty of interpretation, considering that these techniques can be considered “black-boxes” in which it is extremely hard to understand the abundant non-linear patterns that are taken into consideration for each prediction. Furthermore, the deployment of knowledge-discovery techniques reported in this sub-stream of research are customized for the specific definition of the problem addressed and cannot be deployed in a wider context, i.e. alternative datasets, granularities and data-mining techniques. This implies the lack of a broader methodological approach, common steps, benchmarks and evaluation standards, as is the norm in other domains (Chapman et al., 2000).

3. Methodology Overview

The main objective of the methodology is to be able not only to model interest rate and interest rate term structures together with their relationships with other relevant economic variables, but also to uncover hidden patterns and information embedded in the financial data sets using knowledge-discovery techniques. Another objective of the methodology is the study and characterization of economic cycles, and economic cycle stages, as well as the description of the historic relationships between interest rates and other relevant economic variables. Finally, another specific objective of the proposed methodology is the use of all previously gathered knowledge in the forecasting of a given target variable. In this paper, the methodology is exemplified through the forecasting of one specific secondary market interest rate, which is represented by the rates at which five-year nominal zero-coupon bonds should be traded. It is important to note that this particular variable of interest is just one of the possible financial variables that could be chosen.

Figure 2 represents the specific objectives of the methodology and the order in which they are addressed.

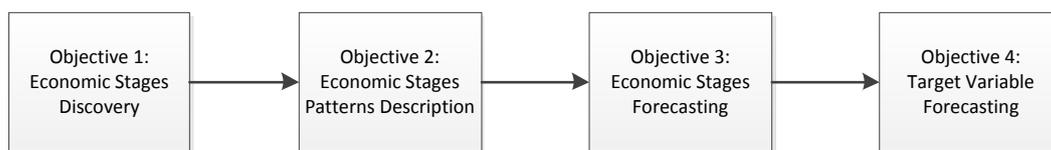


Figure 2: Objectives of the proposed methodology

The proposed methodology uses the Cross-Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al., 2000) as its basis, adapting and extending it in order to support the precise modelling and forecasting of interest rates and related variables of interest during the different stages of the economic cycle. It consists of a number of sequential phases, namely data understanding, data preparation, modelling, evaluation and deployment. Each phase contains a number of sub-phases and as with the main phases, sub-phases also consider reverse and backward flows in order to accommodate continuous improvement and special requirements for modelling using specific knowledge-discovery and statistical techniques.

This work considers the definition of an “economic cycle” as the cyclical process whereby the economy fluctuates between stages of expansion and contraction (Investopedia, 2015). Typically, two stages or

“turning points” are defined within each economic cycle, namely growth and recession. However, economic theories have also defined alternative characterizations of stages such as the four-stage economic cycle (full recession, early recovery, late recovery and early recession) or the three-stage economic cycle (recovery, recession, flat (or transition)) (Bhaumin, 2011).

In order to work with a common denominator of what is to be considered an economic cycle, we adopt the definition that identifies three main states of the economic indicators, specifically the recovery or *growth* stage, the *recession* stage and the *transition* stage.

In the remainder of this section, the different phases of the proposed methodology are discussed.

Data understanding

The first phase involves understanding the requirements and the data available for analysis. It also involves the collection and evaluation of the available data in relation to the objective of the analysis (fit-for-purpose) and includes an initial analysis from a theoretical and statistical (descriptive) perspective. For example, statistical analysis includes analysis of the distribution of the key variables and quality issues, for example, missing data and errors. In relation to interest rates, this step concerns the collection and understanding of real-sector and market data that could be considered for characterization of the economic cycles and forecasting of the interest rates or any other related variable of interest (see, for example, Figure 1).

In this study, a data set for the Chilean markets for the period spanning January 2006 to December 2012 is used for the purpose of demonstrating the methodology and its implementation. This particular market and the period were chosen to demonstrate the deployment of the proposed methodology based on findings that the Chilean yield curve is a good candidate for modelling using knowledge-discovery techniques such as neural networks (Vela, 2013). Data was obtained from Bloomberg data services. The set of variables used in the study covers the range and type of variables that are usually mentioned in the literature, including economic and financial variables, such as monetary policy rates, short- and long-term interest rates, the spread between long- and short-term rates, interest rate swaps, foreign interest rates, inflation rates, exchange rates, commodity prices, industrial production indexes, stock indexes and volatility indexes, among others. The relevant data for the analysis in this paper include the variables, as shown in Table 1 together with their summary statistics. The granularity and their type are given in the Description column in Table 1. The number of observations is 1736 for all variables under consideration except for CLP_5Y_BLP which is 1647 that includes some missing values from the source.

The detailed analysis of the descriptive statistics shown in Table 1 is an important component of the data understanding process allowing us to identify stylized facts of the variables of concern. For instance, the mean and standard deviation of the variables give us an estimation of the long term equilibrium value and its average historic fluctuations which can be used as a reference for future interpretation of the models. Table 2 illustrates the direction and magnitude of the linear relationships of the data sets. Considering that Chile has a small, open economy, it is no surprise that copper prices (COPPER) correlate negatively with Country Risk variables, as copper is one of the most important exports of the country fact that also explain the negative correlation with the exchange rate variable (CLPDOLAR). A high correlation between swap rates for different maturities (SWAPCAMERA) and medium-term interest rate (CLP_5y) variables is also observed.

Economic Variable	Min	Max	Range	Mean	Std Dev	Skewness	Kurtosis	Median	Description (Granularity/Type)
CLP_5Y	3.599	8.068	4.469	5.719	0.685	0.455	1.15	5.73	Five-year zero-coupon interest rates (Daily/ Interest Rate)
CLP_5Y_BLP	3.18	8.45	5.27	5.732	0.824	0.286	1.217	5.69	Five-year zero-coupon interest rate. Source Bloomberg. (Daily/ Interest Rate)
CDS_CL_5Y	12.498	322.962	310.464	79.775	57.354	1.374	2.22	73.668	Five-year credit default swap Chile (Daily/ Credit Default Swap)
CLIMMOMS	-3.77	6.28	10.05	0.325	1.1	0.793	11.711	0.42	Economic activity indicator seasonally adjusted mom% (Monthly/ Economic Activity Indicator)
CLIMSA	90.24	118.14	27.9	102.509	7.539	0.472	-0.856	100.7	Economic activity indicator seasonally adjusted index (2008=100) (Monthly/ Economic Activity Indicator)
IMACEC	87.8	124.2	36.4	102.832	8.34	0.461	-0.564	101.1	Monthly economic activity indicator (Monthly/ Economic Activity Indicator)
CLPDOLAR	429.55	682.75	253.2	517.127	42.62	1.166	2.042	515.43	Parity peso dollar (Daily/ Exchange Rate)
COMP_INF	0	5.667	5.667	3.182	0.862	-1.396	5.719	3.16	Compensatory inflation (Daily/ Inflation)
COPPER	124.75	462.85	338.1	328.369	69.125	-0.898	0.673	340.025	Copper prices (Daily/ Commodity Price)
EMBI_CL_CORP	144.558	241.896	97.338	185.842	27.836	0.269	-1.229	179.364	Corporate Chile emerging market bond index (Daily/ Country Corporate Risk)
EMBI_CL_SOB	71	411	340	149.471	69.273	1.884	3.558	136	Sovereign Chile emerging market bond index (Daily/ Country Sovereign Risk)
FPD	0	4225700	4225700	783241.07	708066.71	1.1	1.189	590900	Draining permanent facility (Daily/ Bank System Liquidity)
FPL	0	585429	585429	42905.33	77127.54	3.09	11.732	11552	Injection permanent facility (Daily/ Bank System Liquidity)
IPSA	1988.24	5040.97	3052.73	3474.993	853.238	0.111	-1.226	3302.33	Chilean stock market selective price index (Daily/ Stock Market Indicator)
MPR	0.5	8.25	7.75	4.342	2.178	-0.498	-0.575	5	Monetary policy rate (Daily/ Monetary Policy Rate)
PETROLEUM	33.87	145.29	111.42	81.632	19.981	0.309	0.233	80.21	Petroleum prices (Daily/ Commodity Price)
SPREAD_CLP5_MPR	-2.841	5.409	8.25	1.377	1.98	0.592	-0.559	0.617	Difference between medium-term rate (CLP_5y) and short-term rate of the Chilean yield curve (MPR) (Daily/ Yield Curve Slope)
SWAPCAMERA_1y	0.76	9.01	8.25	4.699	1.895	-0.427	-0.332	5.03	One-year swap rate (Daily/ Interest Rate)
SWAPCAMERA_5y	3.05	8.15	5.1	5.632	0.882	0.285	-0.054	5.52	Five-year swap rate (Daily/ Interest Rate)
SWAPCAMERA_10y	3.84	8.11	4.27	5.996	0.721	0.164	-0.116	5.96	Ten-year swap rate (Daily/ Interest Rate)
T10_USD	1.388	5.295	3.907	3.448	1.025	-0.154	-0.951	3.5	Treasury rate USA ten years (Daily/ USA Treasury Ten year rate)
TIB	0.3	8.38	8.08	4.339	2.199	-0.528	-0.571	5.02	Interbank interest rates (Daily/ Interbank Interest Rate)
UF	17911.62	22732.79	4821.17	20587.36	1446.42	-0.453	-1.043	20982.4	Unidad de Fomento (Daily/ Inflation Index)
VIX	9.89	80.86	70.97	23.12	10.83	1.896	4.644	20.655	Volatility index (Daily/ Volatility Index)

Table 1: List of the economic variables and their summary statistics

Correlation Matrix	CLP_5Y	CLP_5Y_BLP	CDS_CL_5Y	CLIMMOMS	CLIMSA	IMACEC	CLP-DOLAR	COMP_INF	COPPER	EMBI_CL_CORP	EMBI_CL_SOB	FPD	FPL	IPSA	MPR	PETROLEUM	Spread_CLP5_MPR	Swap camera_1y	Swap camera_5y	Swap camera_10y	T10_USD	TIB	UF	VIX
CLP_5Y	1.00	0.98	-0.37	-0.10	-0.20	-0.19	-0.20	0.53	0.36	-0.30	-0.17	-0.23	0.10	-0.14	0.43	0.53	-0.13	0.69	0.91	0.89	0.41	0.43	-0.27	-0.11
CLP_5Y_BLP		1.00	-0.35	-0.10	-0.14	-0.13	-0.23	0.72	0.38	-0.26	-0.15	-0.22	0.10	-0.13	0.54	0.55	-0.25	0.77	0.92	0.87	0.36	0.54	-0.24	-0.11
CDS_CL_5Y			1.00	-0.12	0.30	0.27	0.51	-0.02	-0.54	0.30	0.90	0.39	0.01	0.06	0.04	-0.22	-0.17	-0.20	-0.47	-0.54	-0.62	0.04	0.63	0.82
CLIMMOMS				1.00	0.10	0.03	-0.08	-0.03	0.06	0.05	-0.14	0.11	0.01	0.08	-0.16	0.01	0.15	-0.11	-0.09	-0.07	0.01	-0.16	-0.04	-0.04
CLIMSA					1.00	0.89	-0.48	0.04	0.44	0.96	0.09	0.49	0.05	0.85	-0.03	0.50	-0.04	-0.05	-0.35	-0.47	-0.87	-0.03	0.88	0.02
IMACEC						1.00	-0.43	0.01	0.39	0.87	0.06	0.48	0.06	0.77	-0.03	0.46	-0.03	-0.06	-0.32	-0.43	-0.79	-0.03	0.79	0.03
CLP-DOLAR							1.00	-0.15	-0.87	-0.47	0.58	-0.01	-0.04	-0.58	0.05	-0.72	-0.13	-0.11	-0.12	-0.08	0.18	0.05	-0.18	0.60
COMP_INF								1.00	0.18	-0.02	0.09	0.15	0.13	0.02	0.21	0.44	-0.05	0.33	0.43	0.39	0.04	0.20	0.07	0.09
COPPER									1.00	0.41	-0.63	0.03	0.05	0.58	0.02	0.70	0.11	0.22	0.28	0.25	-0.05	0.02	0.12	-0.59
EMBI_CL_CORP										1.00	0.04	0.52	0.03	0.89	-0.25	0.42	0.17	-0.27	-0.49	-0.55	-0.87	-0.26	0.90	-0.01
EMBI_CL_SOB											1.00	0.23	0.02	-0.17	0.25	-0.21	-0.33	0.02	-0.23	-0.34	-0.44	0.24	0.43	0.86
FPD												1.00	-0.13	0.42	-0.18	0.12	0.12	-0.24	-0.33	-0.34	-0.49	-0.19	0.57	0.21
FPL													1.00	0.06	0.08	0.06	-0.05	0.11	0.06	0.03	-0.03	0.08	0.03	0.03
IPSA														1.00	-0.31	0.44	0.29	-0.24	-0.37	-0.40	-0.63	-0.31	0.72	-0.18
MPR															1.00	0.22	-0.95	0.93	0.61	0.38	0.06	1.00	-0.17	0.16
PETROLEUM																1.00	-0.05	0.37	0.40	0.32	-0.23	0.21	0.34	-0.25
Spread_CLP5_MPR																	1.00	-0.78	-0.36	-0.11	0.07	-0.95	0.09	-0.21
Swap camera_1y																		1.00	0.81	0.62	0.21	0.93	-0.26	-0.01
Swap camera_5y																			1.00	0.95	0.56	0.61	-0.49	-0.18
Swap camera_10y																				1.00	0.69	0.38	-0.57	-0.24
T10_USD																					1.00	0.07	-0.93	-0.32
TIB																						1.00	-0.17	0.15
UF																							1.00	0.34
VIX																								1.00

Table 2: Correlations between the economic variables

Swap contracts have the characteristic of representing net interchange of interest rate payments with near-zero credit risk, correlating with risk-free rates of fixed income instruments issued by the Monetary Authority. A positive correlation between stock market indicators and economic activity indicators is expected, as the stock market should move closely to economic expectations. A negative correlation between economic activity indicators and US Treasury bills also reflects the openness of the Chilean economy and the trend of global investors to move towards emerging markets, configuring different fly-to-risk or fly-to-quality net inflows depending on the expected performance of the given markets. Finally, a high correlation between the monetary policy rate (MPR) and the interbank rate (TIB) reflects the accomplishment of one of the most important operative objectives of the Central Bank, maintaining a minimum spread through the injection or drain of liquidity to banks in order to ensure the effectiveness of its monetary policy.

Data preparation

The second phase relates to the preparation of the data, in which the variables are transformed and modified from the initial raw data for use with the modelling step. This is an iterative process that helps with the improvement of the analysis (modelling phase) and can involve adding or removing variables or lags and modifications to the transformation of the variables. A number of data preparation steps have been carried out and these are discussed in more detail in Section 4.2.

Modelling

During this phase, modelling techniques are identified and applied to selected variables, and their outputs are examined in relation to the analysis task. There are several available modelling techniques, some of which require data to be in certain formats, which necessitates iterations with the data preparation tasks. Moreover, set-up and calibration parameters might need to be specified for each modelling technique, and these are examined and optimized where possible during this step. For the purposes of the analysis in this paper, a number of modelling techniques available as part of the IBM SPSS Modeller (IBM, 2014) were used. Furthermore, the modelling phase was separated into four steps for the purposes of the overall analysis (as shown in Figure 3) and it is discussed in detail in Section 4.

Evaluation

This phase involves evaluation of the modelling outputs according to the degree to which they meet the organizational objectives for which the whole analysis is carried out, and the extent to which they can be useful in informing decision-making. The expectation is that following the modelling phase a number of high-quality models will have been built. For the purposes of the evaluation phase in this paper, the outputs of the different modelling steps were evaluated considering both statistical fitness and domain expert validation. In particular, Silhouette measures (Rousseeuw, 1987) were used when evaluating clustering techniques. Classification power or accuracy (precision, recall and F-metric) was measured using a confusion matrix and/or hit ratio counts when evaluating classifiers techniques; and paired sample t-tests, MSE, MAPE, MASE (Hyndman & Koehler, 2006) and directional accuracy (Greer, 2003) were utilized to examine if the mean difference in forecasting ability (error rates) was different for the set of models that included the economic stage variable and those that did not. These are discussed in detail in Section 4.

Deployment

This phase takes the evaluation results in terms of the selected models for deployment and determines a plan for implementation. In this paper, deployment assumes a scenario in which the best models are used for informing trading decisions, measuring the economic performance of a market-timing strategy (Resnick & Shoemith, 2002). As such, the return on investment of trading bonds is calculated,

considering costs and profits, and following the investment decisions that are informed from the forecast CLP_5y rates. Within this scenario, the results of the analysis are measured in relation to whether they result in profit for the organization, and a number of related measurements of performance are calculated, namely total profit, number of trades, average return per trade, and risk-return measurements, including standard deviation of the portfolio valuation and Sharpe ratios (Sharpe, 1994). Recommendations are also made for how best the models can be used for such a task. Furthermore, based on the deployment scenario, there might be additional ways in which knowledge from the decision rules and forecasting model can be used. This is discussed in detail in Section 5. An overview of the proposed methodology is shown in Figure 3.

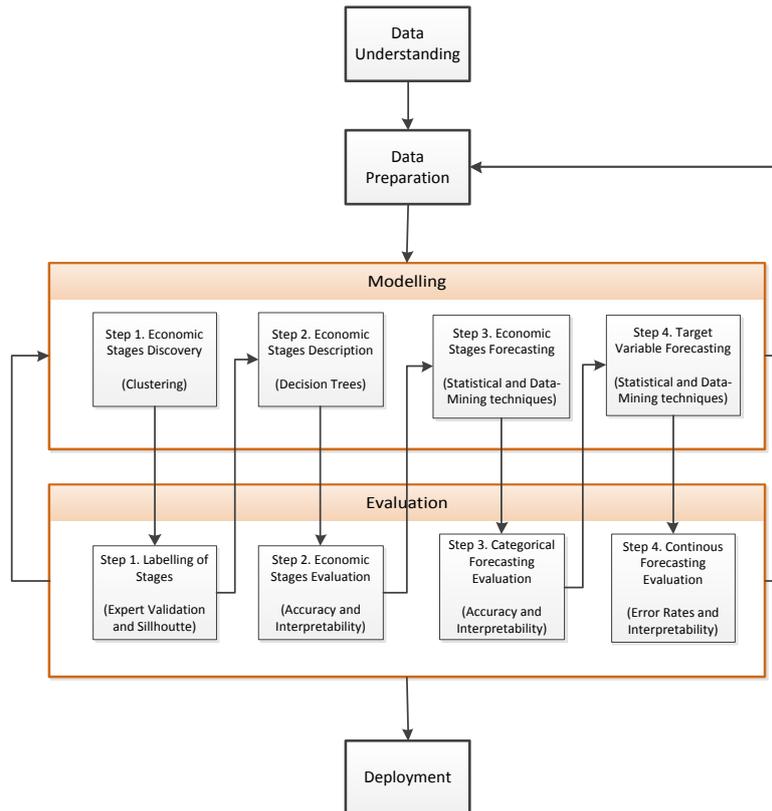


Figure 3: Detailed steps and methodology

4. Modelling and Evaluation

4.1. Overview

This section discusses in detail the proposed methodology in terms of the steps required for the data preparation, modelling, and evaluation phases (see Figure 3). This section is structured as follows: Section 4.2 discusses the transformation of the variables selected in the data understanding phase in order to be used for the modelling phase. Section 4.3 elaborates upon the details of the modelling phase and the first step, the *Economic Stages Discovery*, using unsupervised data-mining techniques. This section also discusses the labelling of the economic stages that is carried out as part of the evaluation phase of this step. Section 4.4 discusses the *Economic Stages Description*, which involves the characterization of the different stages of the economic cycles in terms of the relationships between the variables involved at each stage. This is defined as a classification task for a cross-sectional sample and addressed using supervised data-mining techniques. This section also includes evaluation of the output

of the classification task, which is part of the evaluation phase of the methodology. Section 4.5 discusses the *Economic Stages Forecasting*, using a range of data-mining and statistical techniques. This is also defined as a supervised classification task, but for a time-series sample. As such, the outputs are evaluated in terms of the accuracy and interpretability of the results, for both an in-sample and out-of-sample partition, and their relative merits are elaborated in relation to their economic and organizational implications, as well as their statistical properties. Section 4.6 discusses the *Target Variable Forecasting* step of the methodology, using a specific target variable, namely the CLP_5y rate. This is defined as a time-series supervised forecasting exercise and its outputs are evaluated for the previously defined in-sample and out-of-sample partition. This process includes comparison of the forecasting error rates with or without explicitly including the 'Economic Stage Label' as an optional input variable, and their forecasting power is evaluated in terms of model fitness and organizational and economic interpretability of the results.

4.2. Data Preparation

As part of the data preparation phase, data transformations include changes in the granularity, scale and creation of lagged variables. First, daily observations were used to create weekly averages for each of the 24 input variables. Similarly, monthly averages were also created. A total of twenty-four (24) monthly (calendar month) variables and twenty-four (24) weekly variables were created (a total of 48 variables). The weekly columns were joined with the monthly columns using the weekly label as the key. Monthly averages were joined considering their lagged values in order to avoid superposition of monthly values that had still not occurred in the time sequence. For instance, the weekly average value of week 3 on month 2 was joined with the monthly average of month 1. The weekly average value of week 4 on month 2 was also joined with the monthly average of month 1, and so on.

Second, all 48 variables were transformed into a standardized (same scale) form. This process is recommended when working with knowledge-discovery techniques that calculate multi-dimensional distances between pairs of fields, such as the Clustering and Support Vector Machines. All variables were transformed by subtracting the mean and dividing by the standard deviation of each field – in other words, using a z-score transformation approach. In relation to the z-score values interpretation, z-scores close to zero represent states of the variable in which its value was close to its historical mean (see Table 2). Moreover, z-score values up to ± 1 represent states of the variable in which its value was one standard deviation away from the mean (positive or negative). This is still inside its ordinary range of fluctuation. Z-score values up to ± 2 , ± 3 , or more represent states of the variable that are not so common and can be considered high or even extreme in terms of its historical behaviour.

Third, several lagged variables of these already transformed fields were created. This process is required for modelling in the forecasting steps, assuming that only historical values can be used to forecast the future. Ten lags for each of the weekly variables (240 variables in total) were created. In addition, 18 lags for each of the monthly variables (432 variables in total) were created. After all this data preparation, a total of 720 new variables were created in addition to the 24 original variables and the CPLy5 target variable.

4.3. Economic Stages Discovery

During step 1, *Economic Stages Discovery*, clusters were created for the weekly variables in order to uncover periods in which the economic variables behaved similarly and, as such, could be mapped to

stages in the economic cycle. For this task, all the weekly z-score variables were considered as input data.

Three different clustering techniques were used: *TwoSteps (x-means)*, *k-means* and *Kohonen Networks (Self-Organizing Maps, SOMs)* (Alpaydin, 2010). For the k-means technique, five alternative models were created, considering the number of clusters as k=2, 3, 4, 5, 6. Both TwoSteps and Kohonen Networks are unsupervised techniques that automatically suggest the number of natural groupings or clusters discovered in the data set. Furthermore, to evaluate the various clustering models, the *silhouette measure* (Rousseeuw, 1987) was calculated.

The summary of results that includes the silhouette values and other measures obtained in this step can be seen in Table 3. The highest silhouette value was obtained with k-means A (k=2), followed by k-means B (k=3) and automatic detection of k=3 using the TwoSteps technique. Considering the overall results, it was decided to keep for follow-up analysis the “k-means B” clustering results, as this corresponds better to the three stages of the economic cycle approach (Bhaumin, 2011) compared to the “TwoStep A” model that also results in three clusters. The reason is that the “k-means B” model includes a cluster that includes a larger proportion of the weekly values and could be labelled as the most often-encountered stage of the economic cycle, the *transition stage*.

	Model	Silhouette	Number of Clusters	Smallest Cluster (N)	Smallest Cluster (%)	Largest Cluster (N)	Largest Cluster (%)	Smallest/Largest
1	K-means A	0.356	2	120	34	228	65	0.526
2	K-means B	0.338	3	98	28	149	42	0.658
3	TwoStep A	0.331	3	86	25	144	42	0.597
4	K-means C	0.328	4	33	9	147	42	0.224
5	K-means D	0.297	6	10	2	95	27	0.105
6	K-means E	0.236	5	10	2	158	45	0.063
7	Kohonen A	0.174	12	7	2	62	17	0.113

Table 3: Summary of results for various clustering techniques

In relation to the importance of each input variable to act as a discriminator between clusters, the 24 variables were ranked according to their F-test measure. The results are shown in Table 4. Interpretation of the F-test results indicates that the top five variables are all characteristics closely related to the yield curve, followed by variables that relate to real activity and inflation (COPPER is the main export commodity of the Chilean economy, whereas PETROLEUM is one of the most important import commodities). It is worth noticing that the means of each variable for each cluster are statistically different from one another at 99 per cent confidence (sig. 0.00%), indicating that on average all clusters group weeks in which the values of the respective variables are different, thus validating the presence of structural breaks and inflection points in the data set.

Following the first step of modelling (*Economic Stages Discovery*), and examination of the statistical properties from the various clustering techniques, the clusters identified by the “k-means B” model were selected for evaluation and labelling by a human expert in order to examine whether the resulting clusters map easily to humanly identifiable economic stages. Furthermore, to facilitate this task, a list of key economic events was used as a reference for expected inflexion points that could delimit the beginning and end of the different stages of the economic cycles according to the suggestions of the domain experts. The events selected were: Event A - 23 October 2008, Lehman Brothers bankruptcy; Event B - 21 May 2010, Greek Bailout; Event C- 19 March 2011, Earthquake in Japan; Event D - 8 August 2011, Downgrade of USA rating.

Field	statistic	cluster-1*	cluster-2*	cluster-3*	F-Test	df	Sig.
		CL1_Recession(Down)	CL2_Growth(Up)	CL3_Transition(Flat)			
CLP_5Y_BLP_z-score	mean	0.219	-1.466	1.455	355.661	2, 345	0.00%
	standard deviation	0.855	0.819	0.867			
	standard error	0.086	0.081	0.071			
	count	98	101	149			
Swap camera_1y_z-score	mean	-0.272	-1.093	1.748	324.587	2, 345	0.00%
	standard deviation	0.958	1.1	0.719			
	standard error	0.097	0.109	0.059			
	count	98	101	149			
CLP_5Y_z-score	mean	0.272	-1.483	1.293	282.258	2, 345	0.00%
	standard deviation	0.84	0.815	1.004			
	standard error	0.085	0.081	0.082			
	count	98	101	149			
Swap camera_5y_z-score	mean	0.098	-1.502	1.095	223.996	2, 345	0.00%
	standard deviation	0.713	0.787	1.168			
	standard error	0.072	0.078	0.096			
	count	98	101	149			
MPR_z-score	mean	-0.957	0.267	1.91	216.786	2, 339	0.00%
	standard deviation	0.544	1.342	1.094			
	standard error	0.056	0.134	0.09			
	count	93	101	148			
PETROLEUM_z-score	mean	0.55	-1.02	1.3	208.59	2, 345	0.00%
	standard deviation	0.782	0.895	0.94			
	standard error	0.079	0.089	0.077			
	count	98	101	149			
TIB_z-score	mean	-0.823	0.325	1.82	194.103	2, 345	0.00%
	standard deviation	0.661	1.326	1.044			
	standard error	0.067	0.132	0.086			
	count	98	101	149			
COPPER_z-score	mean	0.743	-1.16	0.837	168.751	2, 345	0.00%
	standard deviation	0.804	1	0.898			
	standard error	0.081	0.099	0.074			
	count	98	101	149			
UF_z-score	mean	0.273	1.405	2.003	162.955	2, 345	0.00%
	standard deviation	1.115	0.51	0.528			
	standard error	0.113	0.051	0.043			
	count	98	101	149			
Swap camera_10y_z-score	mean	0.25	-1.507	0.559	134.948	2, 345	0.00%
	standard deviation	0.744	0.83	1.249			
	standard error	0.075	0.083	0.102			
	count	98	101	149			
COMP_INF_z-score	mean	0.042	-0.581	0.462	88.741	2, 345	0.00%
	standard deviation	0.627	0.543	0.635			
	standard error	0.063	0.054	0.052			
	count	98	101	149			

Table 4: Predictor importance for clustering results

Figure 4 shows the behaviour of the CLP_5y variable during the period of analysis, overlaid with the corresponding weekly label of the economic stage resulting from the “k-means B” clustering model. This was given to a domain expert, who examined the clusters (economic stages) and the list of key economic events and provided the labels for each of them in relation to one of the possible stages of the economic cycle. The labels were assigned as follows:

- Cluster 1 assigned to *recession*, shown as dark-blue circles in Figure 4
- Cluster 2 assigned to *growth*, shown as red squares in Figure 4
- Cluster 3 assigned to *transition*, shown as light-blue crosses in Figure 4

The labels for each stage were incorporated as a new variable in the data set so that in the following step of the methodology they could be taken into consideration. This new variable was named ‘Economic Stage Label’.

Alternative ways to detect structural breaks and turning points to characterize the economic stages could have considered other qualitative methods such as polls of economic agents’ expectations, the Delphi technique, and focus groups involving panels of experts. In addition, other quantitative methods such as the Pettitt test (Oh & Han, 2000) and the Hodrick–Prescott filter (Gogas et al., 2014) could also have been used. A complete analysis of alternative ways of discovering structural breaks and labelling economic stages is outside the scope of this paper; however, it is important to mention that the knowledge-discovery approach taken in this paper, that is, the clustering technique, presents some important advantages, such as the automatic recognition of natural groupings considering the non-linear relationships between the variables and high dimensionality in the data sets. Moreover, the suggested technique also has the ability to incorporate and compare alternative ways, as well as requiring relatively low input from domain experts.

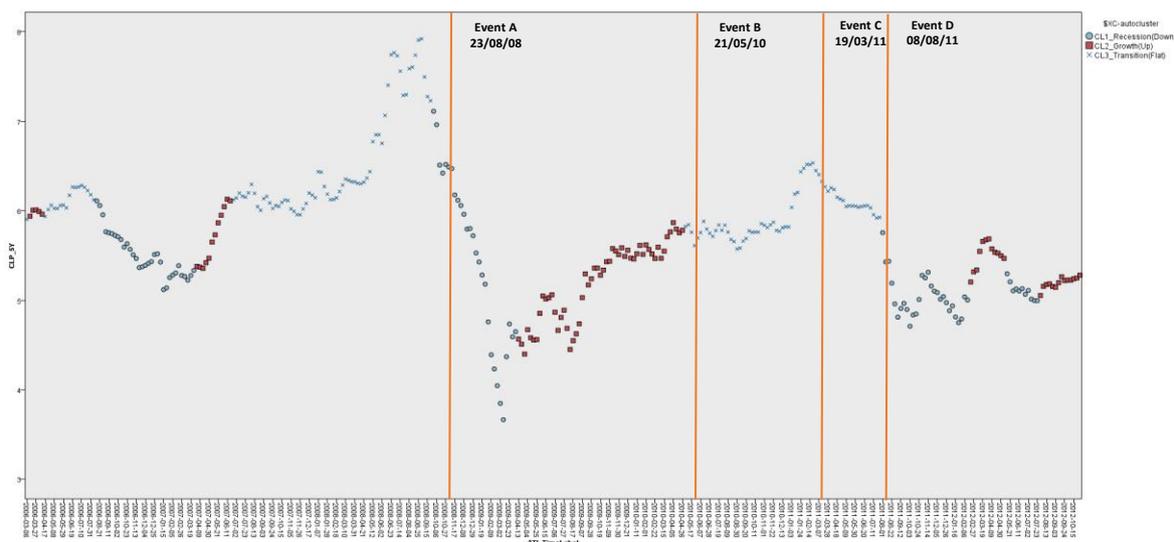


Figure 4: Clustering results for the CLP_5y and labelling of economic stages

4.4. Economic Stages Description

The objective of the second modelling step (*Economic Stages Description*) is to describe the characteristics of the financial and economic variables at each stage of the economic cycle. Furthermore, the aim is to present the information to a human expert in a way that is easy to understand and interpret. To achieve this aim, a decision-tree classifier was used as the supervised modelling technique. As a result of the use of decision trees, uncovered patterns are presented using an “*if-then*” structure, in which the “*then*” component corresponds to one of the three economic stages, namely *recession*, *growth* and *transition*, and the “*if*” component includes logical tests for values of the different economic and financial variables.

In order to apply the decision-tree technique, the data set was divided into two randomly selected samples: 50 per cent of the available data was allocated to the training sample and the remaining 50 per cent to the testing data set. The training data set was used to build a decision tree (C5.0) using ‘*Economic Stage Label*’ as the target variable (“*then*” part of the rule). The resulting tree and its rule-set are shown in Figure 5. It is worth mentioning that the technique incorporates a feature selection process that discards variables not considered relevant for discrimination between the classes. In this case, the set of possible input variables included all the economic and financial variables in their z-score transformation, that is, the decision tree selected only 5 out of 24 available sets of variables.

From the patterns presented in Figure 5, it is interesting to note that in most of the generated rules the variable SPREAD_CLP5_MPR is selected. This is in accordance with the literature that argues that the predictability of economic growth, or the lack thereof, namely recessions, is a function of the slope of the yield curve (Estrella & Mishkin, 1998; Estrella et al., 2003). Furthermore, both the short-term rates (SWAPCAMERA_1y), and SPREAD_CLP5_MPR are examined continuously by the Central Bank to predict economic growth, and the value of the SWAPCAMERA_1y and MPR could be adjusted as these are relevant to assess market expectations.

From looking at Figure 5, it is also possible to appreciate that the first variable used to discriminate (split condition) between stages of the economic cycle is SWAPCAMERA_1y. This is also in accordance with the literature, and it presents supporting evidence for the explanation that the term

structure contains information about future expectations of spot rates, which is linked to the unbiased expectations theory and the liquidity preference theory, as liquidity premiums captured in the swap rates seem to play an important role (McCulloch, 1987; Patton & Timmermann, 2010; Richardson, Richardson, & Smith, 1992).

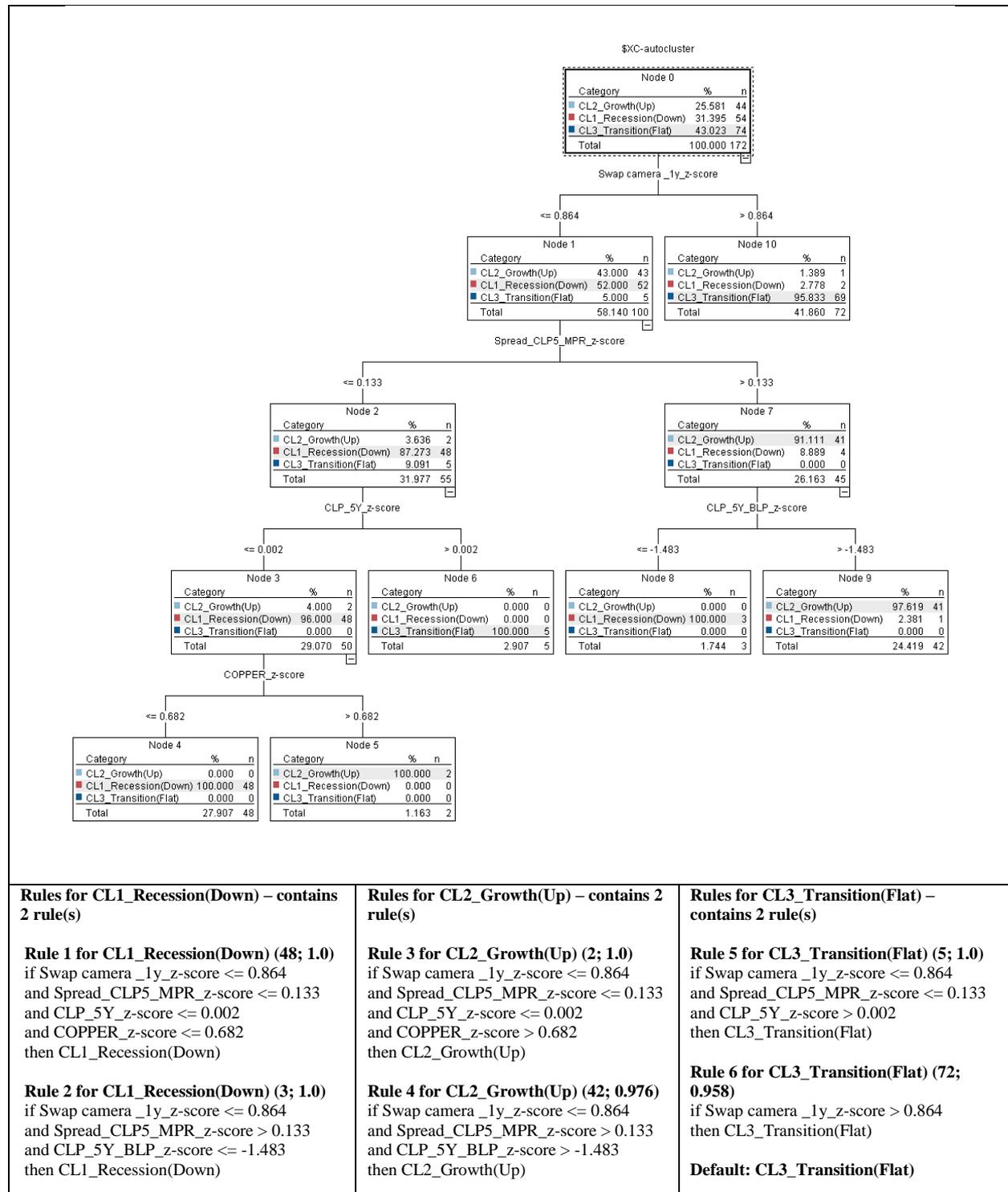


Figure 5: The C5.0 decision-tree classifier for the *Economic Stages Description* step

The importance of the copper prices, particularly for the economy in study, is reflected in the fact that the COPPER z-score is present in both Rule 1 for recession and Rule 3 for growth. This kind of patterns

is characteristic of the open economies that depend to a large extent on the exports of commodities, such as copper in the case of Chile.

Other variables that were selected are also consistent with the existing literature. For example, the role of the yield curve and its slope is again reflected in the fact that all variables (with the exception of the COPPER) relate in some way to either the current shape of the yield curve (SPREAD_CLP5_MPR, CLP_5Y, CLP_5Y_BLP) or its expected future shape (SWAPCAMERA_1y). In particular, *Rule 6* states that whenever the SWAPCAMERA_1y_z-score is higher than 0.86 then the expected economic stage is *transition*. This can be interpreted as signifying that when market agents expect short-term swap rates to be higher than the recent average, then the economic stage will be in transition. Moreover, when there are low levels of COPPER prices and low interest rates, then rules such as *Rule 1* and *Rule 2* describe the economic stage as *recession*. Alternatively, *Rule 3* and *Rule 4* state that higher COPPER prices and high interest rates are associated with the *growth* stage. COPPER is one of the main exports of the Chilean economy, and the results show that following its behaviour is key to describing the different stages of the economic cycles.

The classification accuracy of the resulting rule-set is measured through precision, recall and f-measure (Alpaydin, 2010). Overall, as shown in Figure 6, the rule-set produced by C5.0 is validated by the high recall and precision measures. In particular, in relation to the recall measure which represents the proportion of the relevant instances that are characterized, the average recall for the training partition was higher than the testing partition reaching a 0.97 and 0.92 recall accuracy respectively. Furthermore, in relation to f-measure the *growth* stage presents the lowest score (0.88 in testing sample). This result may be explained from the different proportions of each class in the training and testing partitions. A higher number of *growth* instances is present in the testing partition, and a balancing strategy (Alpaydin, 2010) may be needed if the importance between the classes is an issue for the user. This strategy is exemplified in the next step.

Panel A: 'Partition' = 1_Training

Coincidence Matrix (rows show actuals)	CL1_Recession(Down)	CL2_Growth(Up)	CL3_Transition(Flat)		
CL1_Recession(Down)	50	2	2		
CL2_Growth(Up)	0	43	1		
CL3_Transition(Flat)	0	0	74		
Unequal measures of importance	Recall (R)	Precision (P)	Effectiveness (E)	F-measure	
CL1_Recession(Down)	0.93	1.00	0.04	0.96	
CL2_Growth(Up)	0.98	0.96	0.03	0.97	
CL3_Transition(Flat)	1.00	0.96	0.02	0.98	
Average	0.97	0.97	0.03	0.97	

Panel B: 'Partition' = 2_Testing

Coincidence Matrix (rows show actuals)	CL1_Recession(Down)	CL2_Growth(Up)	CL3_Transition(Flat)		
CL1_Recession(Down)	45	1	1		
CL2_Growth(Up)	6	45	3		
CL3_Transition(Flat)	0	2	73		
Unequal measures of importance	Recall (R)	Precision (P)	Effectiveness (E)	F-measure	
CL1_Recession(Down)	0.96	0.88	0.08	0.92	
CL2_Growth(Up)	0.83	0.94	0.12	0.88	
CL3_Transition(Flat)	0.97	0.95	0.04	0.96	
Average	0.92	0.92	0.08	0.92	

Figure 6: Training and testing results for the *Economic Description Step*

4.5. Economic Stages Forecasting

The objective of this step is to forecast the state of the economic stage using a one-week-ahead horizon. As such, the target variable was a categorical variable and alternative models were built considering the data set as a longitudinal or time-series sample. Bearing in mind that there were three possible values

of the target variable ‘Economic Stage Label’, and given the fact that the number of instances in each class was very dissimilar, it was considered to be an unbalanced classification problem and therefore oversampling (Alpaydin, 2010) was used as the corresponding data-mining strategy. Figure 7 shows the distribution of examples in each class before applying the oversampling balance technique.

Value	Proportion	%	Count
CL1_Recession(Down)		25.21	61
CL2_Growth(Up)		27.69	67
CL3_Transition(Flat)		47.11	114

Figure 7: Number of instances per class available in the training sample

Assuming that only historical values could be used to forecast the future, lagged weekly z-score variables were used as input and were complemented by monthly lagged variables. The available data was divided into two main sections: the first section included data from January 2005 to December 2010; and the second included data from January 2011 to December 2012.

Fifty per cent of the first section of the data set was used to train nine different classification models. These models were also tested for the remaining 50 per cent of the first section of the data, namely the testing sample. Figure 8 shows unequal measures of importance for both training and testing samples, and also the number of variables (fields) used by each of them.

	Model	No. Fields Used	Training Sample				Testing Sample			
			Recall (R)	Precision (P)	Effectiveness (E)	F-measure	Recall (R)	Precision (P)	Effectiveness (E)	F-measure
1	CHAID 1	6	0.982	0.983	0.018	0.982	0.954	0.958	0.045	0.955
2	C5 1	3	0.982	0.983	0.018	0.982	0.949	0.950	0.051	0.949
3	C&R Tree 1	18	0.906	0.918	0.098	0.902	0.925	0.933	0.075	0.925
4	Quest 1	16	0.960	0.962	0.041	0.959	0.957	0.956	0.045	0.955
5	SVM 1	700	0.885	0.939	0.103	0.897	0.878	0.937	0.110	0.890
6	Neural Net 1	700	0.646	0.967	0.241	0.759	0.684	0.986	0.203	0.797
7	KNN Algorithm 1	700	0.646	0.967	0.241	0.759	0.684	0.986	0.203	0.797
8	Logistic regression 1	700	0.505	0.783	0.414	0.586	0.494	0.724	0.418	0.582
9	Bayesian Network 1*	700	0.155	1.000	0.674	0.551	0.149	1.000	0.675	0.325

* Precision and Recall of the Bayesian Network Model is the average of the performance of two classes. Model did not predict any instance as Recession (down) class

Figure 8: Training and Testing sample results of the *Economic Stages Forecasting* step.

Recall (R), Precision (P), Effectiveness (E), and F-measure columns presents the average performance in the three available classes.

As mentioned previously, the decision trees incorporate a feature selection process that automatically scans and selects from the available inputs the best combination of variables to carry out the forecasting task. In Figure 8 it can be appreciated that the best four models correspond to variations of decision trees and that from the 700 available input lagged variables, between 3 and 18 variables were automatically selected. The best performing decision tree was constructed using the CHAID (Alpaydin, 2010) technique and achieved an F-measure of 0.955 in the testing sample, using six variables as input. Figure 9 presents the decision tree and the corresponding rule-set.

From the patterns presented in Figure 9, it is interesting to note that none of the selected variables correspond to one of the available monthly lagged variables (these variables include in their label the character “#” and none of these are present). This could relate to the fact that the forecasting horizon is relatively short, one-week-ahead, and thus, the long-term trends of the available variables are less relevant for short-term forecasting. Moreover, most of the rules include a form of the inflation rate variable (UF_z-score lag 6) and particularly, rules for recession include this inflation variable in high ranges (z-scores need to be at least 0.615 standard deviations away from the inflation mean). The presence of a form of the exchange rate related variable (CLP-DOLAR_z-score lag 8) in several rules

indicates that foreign inflation rates, and/or foreign economic performance, are also important for discriminating between economic stages.

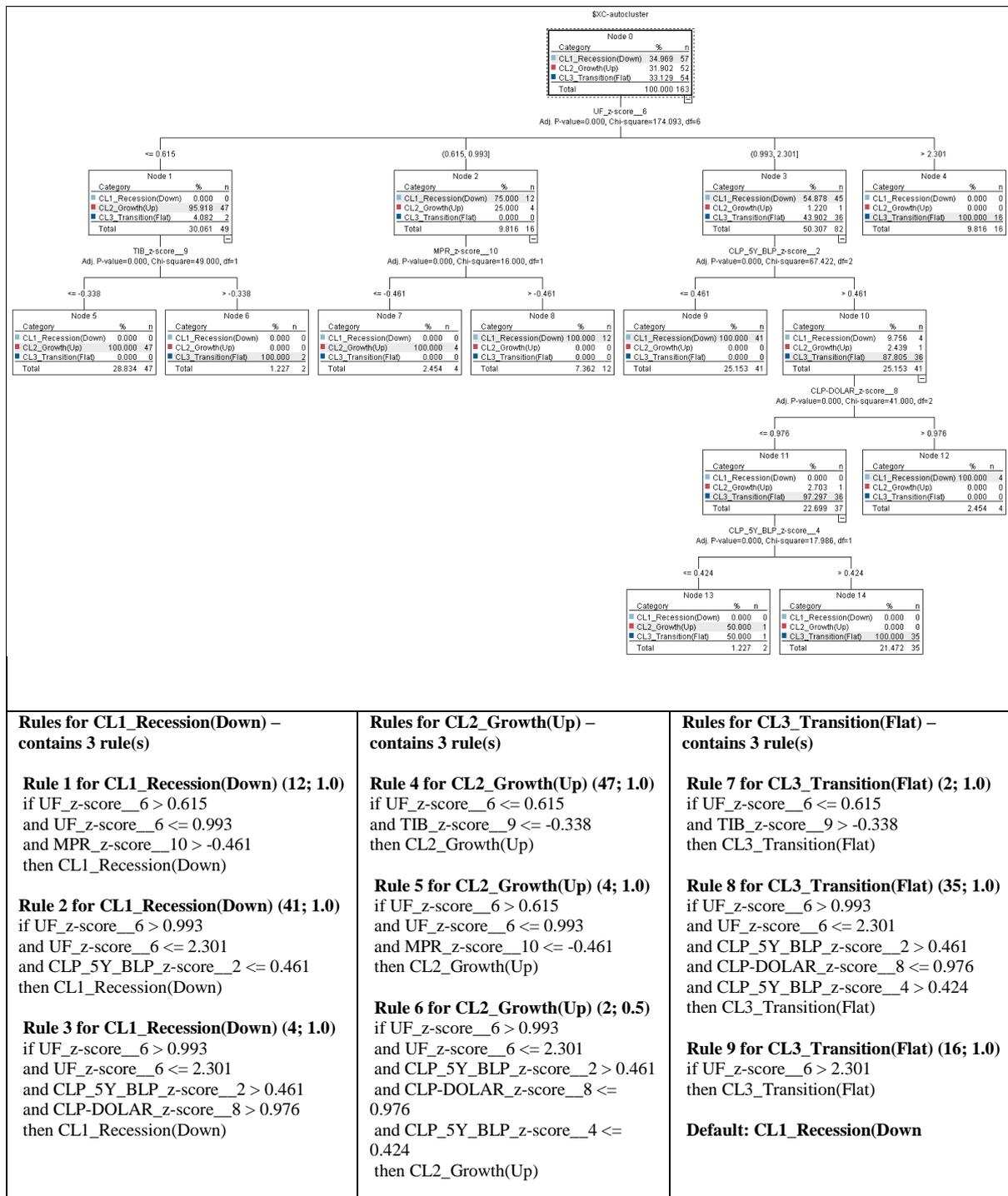


Figure 9: Best decision tree (CHAID) for the *Economic Stages Forecasting* step

Interestingly, rules for recession are in line with what is expected from literature, in other words, rules contain patterns that reflect high levels of exchange rates (depreciation of the Peso against the Dollar), high inflation rate, together with high short term rates and low long term rates, this is, inverted yield curves. For instance, Rule 3 indicates that on top of high domestic inflation rates (UF z-score lag 6 > 0.99 and UF z-score lag 6 <= 2.3) high exchange rates (CLP-DOLAR z-score lag 8 > 0.97) and

relatively high medium-term interest rates (CLP_5y_z-score lag 2) are precursors of a recession. In relation to rules for growth, the opposite pattern is also captured, and low short term rates can be evidence of monetary authorities giving stimuli for growth. For instance, Rule 4 indicates that low levels of interbank rates (TIB_z-score_9<=-0.338) together with low levels of inflation (UF_z-score lag 6<=0.615) are associated with growth stages in 47 instances. A full analysis of all the available rules and patterns is not within the scope of this paper; however, these are presented in extended form in Figure 9.

Figure 10 shows the detailed results of the CHAID tree. In particular, these compare favourably with the ones of the economic description stage, possibly through the use of balancing that addressed the issue of the different proportions of each class in two partitions. In relation to the unequal measures of importance, the recall accuracy of the *transition stage* has the lowest score (0.88). This means that the model generates a higher number of false positives when describing transition stages. This may be due to the fact that it is difficult to distinguish between transition stages that follow a growth stage and transition stages that follow recession stages. Perhaps, more segments or types of structural break should be considered in order to improve the recall accuracy.

Panel A: 'Partition' = 1 Training

Coincidence Matrix (rows show actuals)	CL1_Recession(Down)	CL2_Growth(Up)	CL3_Transition(Flat)		
CL1_Recession(Down)	57	0	0		
CL2_Growth(Up)	0	57	0		
CL3_Transition(Flat)	1	1	56		
Unequal measures of importance	Recall (R)	Precision (P)	Effectiveness (E)	F-measure	
CL1_Recession(Down)	1.00	0.98	0.01	0.99	
CL2_Growth(Up)	1.00	0.98	0.01	0.99	
CL3_Transition(Flat)	0.97	1.00	0.02	0.98	
Average	0.982	0.983	0.018	0.982	

Panel B: 'Partition' = 2 Testing

Coincidence Matrix (rows show actuals)	CL1_Recession(Down)	CL2_Growth(Up)	CL3_Transition(Flat)		
CL1_Recession(Down)	62	2	0		
CL2_Growth(Up)	0	56	0		
CL3_Transition(Flat)	6	1	49		
Unequal measures of importance	Recall (R)	Precision (P)	Effectiveness (E)	F-measure	
CL1_Recession(Down)	0.97	0.91	0.06	0.94	
CL2_Growth(Up)	1.00	0.95	0.03	0.97	
CL3_Transition(Flat)	0.88	1.00	0.07	0.93	
Average	0.954	0.958	0.045	0.955	

Figure 10: Training and testing results for the best decision tree (CHAID) for the *Economic Stages Forecasting* step

In relation to the out-of-sample partition, as shown in Figure 11, the classification accuracy of the resulting rule-set is lower as expected (average recall 0.77, average precision 0.89). In particular, the recall accuracy for the *growth stage* has the lowest measure (0.38) and this is the main reason for the overall lower classification accuracy. The high number of false positives of the growth stages may be due to the process that was followed to characterize the clusters (economic stages) in relation to the key economic events as part of the economic stages discovery step. The labelling results in the out-of-sample partition may be fed back to the domain experts in order to reconsider the ways in which economic stages are characterized.

1.1.Target Variable Forecasting

The aim of this stage is the forecast of a continuous target variable using a longitudinal or time-series approach. Thus, the dataset was divided into two main sections: the first section included data from January 2005 to December 2010; and the second included data from January 2011 to December 2012.

In particular, the chosen target variable is the medium-term rate represented in the CLP_5y variable. Corresponding lagged weekly z-scores were used as inputs and were complemented by monthly lagged variables to address the medium-to-long-term perspective. This variable was selected for a one-week-ahead forecast exercise, as it is one of the most liquid bond rates in the Chilean market. Bonds in the five-year-maturity range are issued periodically by the monetary authority, the Central Bank of Chile, and as such are among the most traded securities in the secondary fixed-income market. Longer maturity instruments are usually held by pension funds to hedge positions reducing the available liquidity of such securities, making five-year maturity an important habitat for other types of institutional investor, mainly banks and mutual funds.

Partition=Out-of-sample

Coincidence Matrix (rows show actuals)	CL1_Recession(Down)	CL2_Growth(Up)	CL3_Transition(Flat)	
CL1_Recession(Down)	40	0	0	
CL2_Growth(Up)	15	10	1	
CL3_Transition(Flat)	2	0	28	
Unequal measures of importance	Recall (R)	Precision (P)	Effectiveness (E)	F-measure
CL1_Recession(Down)	1.00	0.70	0.18	0.82
CL2_Growth(Up)	0.38	1.00	0.44	0.56
CL3_Transition(Flat)	0.93	0.97	0.05	0.95
Average	0.77	0.89	0.22	0.78

Figure 11: Out-of-sample partition results for the *Economic Stages Forecasting* step

Various statistical and data-mining techniques were applied which are available as part of the IBM SPSS Modeller toolset (IBM, 2014). For all these techniques, the default settings proposed by the toolset were used, as the exhaustive optimization of parameters and model's architectures was out of the scope of the paper. In detail, the techniques used were: k-nearest neighbour (KNN), artificial neural networks (ANN), support vector machines (SVM), decision trees (CHAID, C&R Tree), generalised linear models (GENLIN), and regression models (linear with/without feature selection). Moreover, an ensemble strategy was used averaging the prediction of each of the aforementioned modelling techniques to obtain more accurate predictions. In addition, a naïve model was also estimated for comparison purposes.

The k-nearest neighbour (KNN) is a method for classifying cases based on their similarity to other cases. The average or median target value of the nearest neighbours was used to obtain the predicted value for the new case. The k parameter was set to 3 and the distance computation metric used was Euclidean. ANN can be used to approximate automatically for nonlinear relationships between the inputs and the target variable. The trade-off for this flexibility is that the neural network is not easily interpretable. Specifically, the multilayer perceptron architecture was used and the number of neurons and hidden layers was automatically selected by the tool. SVM can be used to improve the predictive accuracy of a model without over fitting the training data and can be used to categorise a high-dimensional feature space. Specifically, SVM was used with the regularisation parameter C= 10, regression precision E = 0.1 and as kernel type, the radial basis function. Two specific decision tree techniques were used in the form of the Chi-squared Automatic Interaction Detection (CHAID) and the Classification and Regression (C&R) Tree. Both of these techniques can be applied to forecast a continuous variable. CHAID is a classification method that uses chi-square statistics to identify optimal splits. C&R tree uses recursive partitioning to split the training records into segments with similar output target values. For both of these techniques, the configuration was to build a single decision tree with up to 5 depth levels. Linear regression is a statistical technique used for predictions based on the values of numeric input fields. Regression models fit a line that minimizes the discrepancies between predicted and actual output target values whereas the generalised linear models (GENLIN), in addition, use a specified link function to covariate the relationship between the input variables. The predictions for the naïve model were done simply using the previous week rate as the forecasted value.

The aim of the evaluation in this step was twofold: first, to identify accurately the best models i.e., the best combinations of model-variable types; and, second, to assert statistically whether considering the previously forecast economic stage in the form of an input variable improved the forecasting ability of the models. The strategy followed was to try a large collection of models and combinations of input variables. Equation 1 shows the first type, which includes input variables $Var_{i,t}$ where subscript i represents one of the original economic variables and t is a subscript for time that takes negative values from -1 to -10 to represent corresponding lagged versions of the given economic input variable. D_1 is a dummy variable that takes the value of 1 if the ‘Economic Stage Label’ variable ($Stage_t$) is included in the model and 0 if not. The variable CLP_{5y_t} represents the target variable. Equation 2 shows the second type which considers as input auto-regressive components of the target variable CLP_{5y_t} at times $t-1$ to $t-10$, and also, a D_1Stage_t component. Equation 3 shows the third kind of input types, which can be considered a combination of the previous two equations, incorporating auto-regressive with $t=t-1$, $Var_{i,t}$ with $t=t-1$ to $t-10$ and D_1Stage_t components.

$$CLP_{5y_t} = f(Var_{i,t-1}, \dots, Var_{i,t-10}, D_1Stage_t) \quad \text{Equation 1}$$

$$CLP_{5y_t} = f(CLP_{5y_{t-1}}, \dots, CLP_{5y_{t-10}}, D_1Stage_t) \quad \text{Equation 2}$$

$$CLP_{5y_t} = f(CLP_{5y_{t-1}}, Var_{i,t-1}, D_1Stage_t) \quad \text{Equation 3}$$

A particular case was the use of Principal Components Analysis (PCA) for the creation of factors that summarized the available input variables. Firstly, eight additional $Var_{i,t}$ components were included representing extra lagged monthly z-score values. These were noted as $mVar_{i,t}$ as shown in Equation 4 and 5 respectively.

$$CLP_{5y_t} = f(Var_{i,t-1}, \dots, Var_{i,t-10}, mVar_{i,t-1}, \dots, mVar_{i,t-10}, D_1Stage_t) \quad \text{Equation 4}$$

$$CLP_{5y_t} = f(CLP_{5y_{t-1}}, Var_{i,t-1}, \dots, Var_{i,t-10}, mVar_{i,t-1}, \dots, mVar_{i,t-10}, D_1Stage_t) \quad \text{Equation 5}$$

Later on, eight factors were built for the lagged weekly z-score values, and eight factors for the lagged monthly z-score values. These were noted $F1$ to $F8$ for weekly lagged factors, and $mF1$ to $mF8$ for monthly factors. Equation 6, and 7 show the final structure of these models.

$$CLP_{5y_t} = f(F1, \dots, F8, mF1, \dots, mF8, D_1Stage_t) \quad \text{Equation 6}$$

$$CLP_{5y_t} = f(CLP_{5y_{t-1}}, F1, \dots, F8, mF1, \dots, mF8, D_1Stage_t) \quad \text{Equation 7}$$

The different inputs shown on Equations 1, 2, 3, 6 and 7 were used to estimate different models by means of the 9 available data-mining techniques plus their ensemble to generate a total of 828 estimated models. Furthermore, a naïve model was also estimated as mentioned previously. To measure the models’ fitness, error rates – measured as mean squared errors (MSE), mean absolute percentage error (MAPE), mean absolute scaled error (MASE) (Hyndman & Koehler, 2006) and Directional Accuracy (Greer, 2003) were calculated for each available combination of model-inputs-variables and their results were ranked by their MASE in the testing sample. To assert whether it was best to consider as an input variable the economic stage, paired sample t-tests were utilized to examine if the mean difference in forecasting ability (MASE) was different for the set of models that included the mentioned variable and those that did not. Table 5 shows the MASE of all the combinations of variables and models that included, and did not include, the stage as an input variable. It can be seen that in the testing sample the MASE was smaller (0.221 difference) for the models that did include the stage as an input. This difference was statistically significant (sig.=0.000%)

A selection of the best model structures was later evaluated for their out-of-sample performance using the previously defined partitions of the data sets. In particular, the best five performing models for each

structure presented on Equations 1, 2, 3, 6 and 7 (25 in total) were selected. To assert whether it was best to consider as an input variable the economic stage, paired sample t-tests were again utilized to examine if the mean difference in forecasting ability (MASE) was different for the set of models that included the mentioned variable and the ones that did not. Figure 12 summarizes these results.

Testing Sample Performance MASE – All Models*		
t-Test: Paired Two Sample for Means	With stage	Without stage
	<i>MASE</i>	<i>MASE</i>
Mean	2.508	2.730
Variance	2.236	3.063
Observations	347	347
Pearson Correlation	0.970	
Hypothesized Mean Difference	0.000	
Df	346	
t Stat	-8.731	
SIG. P(T<=t) one-tail	0.000	
t Critical one-tail	1.649	
SIG. P(T<=t) two-tail	0.000	
t Critical two-tail	1.967	

Table 5: Paired sample t-test of MASE in testing sample for all combinations of input variables and models.

*134 models were not included in this test due feature selection option that drop the Stage variable automatically.

From Figure 12, Panel A, it is possible to notice that, in general, models that include the stage were better, or at least had no difference, in terms of MASE, than the models that did not include the stage. This difference is statistically significant (MASE mean dif.= -0.147, 3.9% one-tail sig.). In terms of the size of the errors (MASE), as shown in Figure 12, Panel A, models that used the auto-regressive components and PCA factors as inputs exhibited better in-testing performance than models that included other variables. In general, ensemble, GENLIN and linear models with automatic feature selection were the best performers. From Figure 12, Panel B, it also possible to notice that, at least in statistical terms, it was not worse to consider models that included the ‘Economic Stage Label’ as input (MASE mean dif.= +0.124, 29.5% one-tail sig.). As expected in general models worsen their performance in terms of MAPE, MSE, and MASE accuracy, however in terms of Directional Accuracy models continued to perform remarkably well (average DA in testing-sample=0.58, sd.=0.09; average DA in out-of-sample=0.65, sd.=0.10). A similar behaviour is shown in terms of correlations of the forecasted values with the real data (average ρ in testing-sample=0.95, sd.=0.07; average ρ in out-of-sample=0.89, sd.=0.15).

Given the considerable number of competing models that were evaluated, the interpretability of the results is an important concern and, thus, preference was given to models that use a reduced number of input features and that are relatively easy to interpret by domain experts. Figure 13 presents the best performing decision Tree models for each input structure (Equations 1, 2, 3, 6 and 7) in terms of testing and out-of-sample performance. CART using auto-regressive or PETROLEUM variables and ‘Economic Stage Label’ as input variables rank top. The same model achieves average performance when comparing MASE with models presented in Figure 11, but surpass them in terms of Directional Accuracy (DA). This implies that non-linear black-box models, such as GENLIN, KNN, SVM and other types of neural networks that use PCA factors and ‘Economic Stage Label’ as input variables, could be preferable when interpretability is not a concern and accurate continuous estimations are needed. However, Decision Trees could be preferable when understanding of the patterns that originate each prediction is required and when directional forecasting is deemed more important.

Panel A: Testing Sample Performance – Selected Models

ID	Structure	Inputs	Technique	MASE_w	MASE_wo	difference	Best?	ro_w	ro_wo	DA_w	DA_wo	M2w_MSE	M2wo_MSE	M2w_MAPE	M2wo_MAPE
0	Naïve	AR1 alpha=1	Naïve	1.000	1.000	0.00		0.973	0.973	0.585	0.585	0.016	0.016	0.016	0.016
1	Eq. 7	AR1, PCA, Stage	ensemble	0.932	0.921	0.01	0	0.975	0.975	0.593	0.577	0.015	0.015	0.015	0.015
2	Eq. 3	AR1, COMPINF_1, Stage	genlin	0.945	0.933	0.01	0	0.975	0.975	0.593	0.602	0.014	0.014	0.015	0.015
3	Eq. 3	AR1, CLIMSA_1, Stage	genlin	0.957	0.966	0.01	1	0.975	0.974	0.561	0.569	0.015	0.015	0.015	0.015
4	Eq. 3	AR1, VIX_1, Stage	genlin	0.958	0.960	0.00	1	0.975	0.974	0.593	0.593	0.014	0.015	0.015	0.015
5	Eq. 3	AR1, IPSA_1, Stage	genlin	0.959	0.974	0.01	1	0.974	0.974	0.585	0.585	0.015	0.015	0.015	0.015
6	Eq. 3	AR1, CDS5y_1, Stage	genlin	0.960	0.962	0.00	1	0.975	0.974	0.585	0.585	0.015	0.015	0.015	0.015
7	Eq. 2	AR1, Stage	genlin	0.973	0.989	0.02	1	0.974	0.973	0.577	0.585	0.015	0.016	0.015	0.016
8	Eq. 2	AR1 to AR10, Stage	ensemble	0.985	0.996	0.01	1	0.974	0.973	0.602	0.634	0.016	0.016	0.016	0.016
9	Eq. 2	AR1, Stage	linear	1.001	0.989	0.01	0	0.973	0.973	0.585	0.585	0.016	0.016	0.016	0.016
10	Eq. 2	AR1, Stage	svm	1.001	0.981	0.02	0	0.973	0.973	0.577	0.585	0.016	0.016	0.016	0.016
11	Eq. 2	AR1, Stage	ensemble	1.031	1.034	0.00	1	0.974	0.971	0.577	0.585	0.018	0.018	0.016	0.017
12	Eq. 7	AR1, PCA, Stage	linear	1.036	1.036	0.00	1	0.991	0.991	0.553	0.553	0.016	0.016	0.016	0.016
13	Eq. 7	AR1, PCA, Stage	regression	1.073	1.073	0.00	1	0.999	0.999	0.472	0.472	0.016	0.016	0.017	0.017
14	Eq. 7	AR1, PCA, Stage	genlin	1.097	1.073	0.02	0	0.999	0.999	0.480	0.472	0.017	0.016	0.017	0.017
15	Eq. 7	AR1, PCA, Stage	knn	1.397	1.446	0.05	1	0.999	0.999	0.634	0.642	0.034	0.043	0.023	0.023
16	Eq. 6	PCA, Stage	ensemble	1.425	1.454	0.03	1	0.959	0.957	0.618	0.610	0.030	0.031	0.023	0.023
17	Eq. 6	PCA, Stage	knn	1.500	1.561	0.06	1	1.000	1.000	0.626	0.626	0.039	0.051	0.024	0.025
18	Eq. 6	PCA, Stage	genlin	1.608	1.635	0.03	1	1.000	1.000	0.439	0.431	0.037	0.038	0.026	0.026
19	Eq. 6	PCA, Stage	regression	1.635	1.635	0.00	1	1.000	1.000	0.431	0.431	0.038	0.038	0.026	0.026
20	Eq. 6	PCA, Stage	neuralnetwork	1.723	1.911	0.19	1	1.000	1.000	0.431	0.407	0.041	0.059	0.027	0.031
21	Eq. 1	SPREAD_CLP5_MPR, Stage	knn	2.263	4.102	1.84	1	0.849	0.580	0.667	0.626	0.118	0.297	0.034	0.064
22	Eq. 1	MPR, Stage	knn	2.414	3.029	0.61	1	0.843	0.807	0.772	0.732	0.134	0.166	0.037	0.046
23	Eq. 1	SWAPCAMERA_5y, Stage	knn	2.415	3.071	0.66	1	0.873	0.838	0.667	0.602	0.098	0.127	0.037	0.047
24	Eq. 1	UF, Stage	knn	2.549	2.445	0.10	0	0.843	0.893	0.724	0.699	0.123	0.086	0.040	0.039
25	Eq. 1	CLP_5y_BLP, Stage	knn	2.557	2.889	0.33	1	0.871	0.853	0.740	0.699	0.104	0.122	0.039	0.044
		average		1.42	1.56	-0.15	0.76	0.96	0.95	0.59	0.58	0.04	0.05	0.02	0.02
		sdev		0.58	0.87	0.40	0.44	0.05	0.09	0.09	0.08	0.04	0.07	0.01	0.01
		max		2.56	4.10	0.10	1.00	1.00	1.00	0.77	0.73	0.13	0.30	0.04	0.06
		min		0.93	0.92	-1.84	0.00	0.84	0.58	0.43	0.41	0.01	0.01	0.01	0.01

Panel B: Out-of-Sample Performance – Selected Models

ID	Structure	Inputs	Technique	MASE_w	MASE_wo	difference	Best?	ro_w	ro_wo	DA_w	DA_wo	M2w_MSE	M2wo_MSE	M2w_MAPE	M2wo_MAPE
0	Naïve	AR1 alpha=1	Naïve	1.000	1.000	0.00		0.972	0.972	0.579	0.579	0.010	0.010	0.013	0.013
1	Eq. 7	AR1, PCA, Stage	ensemble	2.353	1.846	0.51	0	0.957	0.962	0.600	0.632	0.045	0.027	0.028	0.023
2	Eq. 3	AR1, COMPINF_1, Stage	genlin	1.028	0.979	0.05	0	0.974	0.974	0.589	0.579	0.009	0.009	0.013	0.013
3	Eq. 3	AR1, CLIMSA_1, Stage	genlin	1.050	1.064	0.01	1	0.973	0.972	0.611	0.611	0.009	0.010	0.013	0.014
4	Eq. 3	AR1, VIX_1, Stage	genlin	1.082	1.004	0.08	0	0.974	0.974	0.589	0.579	0.010	0.009	0.014	0.013
5	Eq. 3	AR1, IPSA_1, Stage	genlin	1.077	0.988	0.09	0	0.974	0.973	0.589	0.579	0.010	0.009	0.014	0.013
6	Eq. 3	AR1, CDS5y_1, Stage	genlin	1.112	1.021	0.09	0	0.974	0.973	0.589	0.579	0.010	0.010	0.014	0.013
7	Eq. 2	AR1, Stage	genlin	1.075	0.999	0.08	0	0.973	0.972	0.589	0.579	0.010	0.010	0.014	0.013
8	Eq. 2	AR1 to AR10, Stage	ensemble	1.109	1.116	0.01	1	0.971	0.969	0.611	0.537	0.010	0.012	0.014	0.014
9	Eq. 2	AR1, Stage	linear	1.031	0.999	0.03	0	0.973	0.972	0.568	0.579	0.009	0.010	0.013	0.013
10	Eq. 2	AR1, Stage	svm	1.082	1.000	0.08	0	0.973	0.972	0.589	0.579	0.010	0.010	0.014	0.013
11	Eq. 2	AR1, Stage	ensemble	1.109	1.086	0.02	0	0.970	0.970	0.589	0.579	0.011	0.011	0.014	0.014
12	Eq. 7	AR1, PCA, Stage	linear	1.375	1.375	0.00	1	0.971	0.971	0.600	0.600	0.014	0.014	0.017	0.017
13	Eq. 7	AR1, PCA, Stage	regression	2.190	2.190	0.00	1	0.961	0.961	0.632	0.632	0.035	0.035	0.027	0.027
14	Eq. 7	AR1, PCA, Stage	genlin	2.409	2.190	0.22	0	0.955	0.961	0.632	0.632	0.044	0.035	0.030	0.027
15	Eq. 7	AR1, PCA, Stage	knn	8.187	6.551	1.64	0	0.830	0.850	0.842	0.863	0.480	0.315	0.100	0.079
16	Eq. 6	PCA, Stage	ensemble	1.832	1.442	0.39	0	0.940	0.961	0.789	0.716	0.028	0.016	0.024	0.018
17	Eq. 6	PCA, Stage	knn	1.031	0.999	0.03	0	0.973	0.972	0.568	0.579	0.009	0.010	0.013	0.013
18	Eq. 6	PCA, Stage	genlin	1.392	1.032	0.36	0	0.973	0.973	0.579	0.579	0.015	0.010	0.018	0.013
19	Eq. 6	PCA, Stage	regression	1.082	1.000	0.08	0	0.973	0.972	0.589	0.579	0.010	0.010	0.014	0.013
20	Eq. 6	PCA, Stage	neuralnetwork	1.835	1.837	0.00	1	0.951	0.949	0.947	0.947	0.023	0.024	0.023	0.023
21	Eq. 1	SPREAD_CLP5_MPR, Stage	knn	1.084	1.084	0.00	1	0.972	0.972	0.589	0.589	0.010	0.010	0.014	0.014
22	Eq. 1	MPR, Stage	knn	1.031	0.999	0.03	0	0.973	0.972	0.568	0.579	0.009	0.010	0.013	0.013
23	Eq. 1	SWAPCAMERA_5y, Stage	knn	1.201	1.087	0.11	0	0.1010	1.008	0.495	0.505	0.011	0.010	0.015	0.014
24	Eq. 1	UF, Stage	knn	1.051	0.999	0.05	0	0.973	0.972	0.558	0.579	0.010	0.010	0.013	0.013
25	Eq. 1	CLP_5y_BLP, Stage	knn	1.122	0.999	0.12	0	0.974	0.972	0.579	0.579	0.010	0.010	0.014	0.013
		average		1.60	1.44	0.16	0.24	0.96	0.97	0.62	0.61	0.03	0.03	0.02	0.02
		sdev		1.44	1.13	0.33	0.44	0.03	0.03	0.10	0.10	0.09	0.06	0.02	0.01
		max		8.19	6.55	1.64	1.00	1.01	1.01	0.95	0.95	0.48	0.31	0.10	0.08
		min		1.03	0.98	-0.01	0.00	0.83	0.85	0.49	0.51	0.01	0.01	0.01	0.01

Figure 12: Testing and out-of-sample performance for selected Target Variable Forecasting models

2. Deployment

The creation of good-quality models is generally not the end of a knowledge-discovery process. Models can certainly increase knowledge about the data but in order to achieve organizational and business value, the knowledge gained will need to be organized and presented in such a way that can be used and exploited within a real context. This typically involves deploying “live” models to support the decision-making processes of an organization.

ID	Structure	Inputs	Technique	MASE_w	MASE_wo	diference	Best?	ro_w	ro_wo	DA_w	DA_wo	M2w_MSE	M2wo_MSE	M2w_MAPE	M2wo_MAPE	Partition
26	Eq. 1	MPR, Stage	chaid	3.189	3.818	-0.63	1	0.800	0.754	0.976	0.984	0.183	0.217	0.048	0.058	Testing
27	Eq. 2	AR1 to AR10, Stage	cart	1.325	1.278	0.05	0	0.954	0.959	0.935	0.935	0.032	0.029	0.021	0.021	Testing
28	Eq. 3	AR1, PETROLEUM_1, Stage	cart	1.310	1.310	0.00	1	0.955	0.955	0.951	0.951	0.033	0.033	0.021	0.021	Testing
29	Eq. 6	PCA, Stage	chaid	1.983	1.983	0.00	1	0.881	0.881	0.627	0.967	0.101	0.101	0.030	0.030	Testing
30	Eq. 7	AR1, PCA, Stage	cart	1.421	1.373	0.05	0	0.950	0.955	0.927	0.927	0.038	0.034	0.023	0.022	Testing
26	Eq. 1	MPR, Stage	chaid	5.791	7.011	-1.22	1	0.758	0.551	0.926	0.968	0.213	0.319	0.074	0.090	Out_of_Sample
27	Eq. 2	AR1 to AR10, Stage	cart	2.628	2.628	0.00	1	0.912	0.912	0.916	0.916	0.059	0.059	0.035	0.035	Out_of_Sample
28	Eq. 3	AR1, PETROLEUM_1, Stage	cart	3.050	3.050	0.00	1	0.916	0.916	0.895	0.895	0.075	0.075	0.039	0.039	Out_of_Sample
29	Eq. 6	PCA, Stage	chaid	7.728	7.728	0.00	1	0.627	0.627	0.947	0.947	0.362	0.362	0.100	0.100	Out_of_Sample
30	Eq. 7	AR1, PCA, Stage	cart	2.475	2.475	0.00	1	0.920	0.920	0.916	0.916	0.055	0.055	0.033	0.033	Out_of_Sample
			average	3.09	3.27	-0.18	0.80	0.87	0.84	0.90	0.94	0.12	0.13	0.04	0.04	
			sdev	2.10	2.32	0.42	0.42	0.11	0.15	0.10	0.03	0.11	0.12	0.03	0.03	
			max	7.73	7.73	0.05	1.00	0.95	0.96	0.98	0.98	0.36	0.36	0.10	0.10	
			min	1.31	1.28	-1.22	0.00	0.63	0.55	0.63	0.89	0.03	0.03	0.02	0.02	

Figure 13: Testing and out-of-sample performance for selected *Decision Trees Forecasting* models

The proposed methodology could be applied within different organizational contexts, having different types of user and decision-making processes in mind. For instance, it could be applied in a monetary authority context such as a central bank, which could use, for example, the uncovered patterns learned from the *Economic Stages Description* step to further understand the variables that characterize the economic stages and their relationships. In addition, it could be applied by both institutional and individual investors, which could use, for instance, the forecast interest rates resulting from the *Target Variable Forecasting* stage, informing their trading decisions. It could also be applied within an academic or financial research context, where the knowledge gained at the different stages could be analysed, evaluated and compared with outputs from other econometric approaches, thus enriching existing theories or helping to create new ones.

Only one scenario is assumed for the purposes of the discussion of the deployment phase in this paper, but several others could be of interest. In particular, it is assumed that the organizational context for the selected scenario is that of an institutional investor. The models to be deployed will be used to inform weekly trading decisions on the basis of the forecast CLP_5y rates. The economic performance of alternative market-timing strategies (Resnick & Shoemsmith, 2002) will also be monitored and evaluated on a weekly basis.

The deployment phase consists of two steps: *planning* and *monitoring and maintenance* (Chapman et al., 2000). In terms of the planning step for the selected scenario, it is assumed that the institutional investor prefers models that exhibit high levels of accuracy even as a trade-off from interpretability. Thus, out of all the models that have been evaluated, it has been decided to prioritize models that include the economic stage “economic stage label” as input variable, and which also consider the PCA factors as input, given the fact that such models presented smaller error rates in the testing sample. The forecast CLP_5y rates of the selected models will be used to generate *trade* or *hold* signals on a weekly basis for the given bond. Trading signals will be generated according to two alternative strategies, namely *active* and *risk-averse* strategies. The active strategy generates *trade (hold)* signals when the forecast interest rate decreases (increases) from its previous weekly value. The risk-averse strategy generates *trade (hold)* signals following a similar logic but imposes a minimum change in the forecast interest rate of ± 10 basis points before actually generating a signal. As a consequence, the risk-averse strategy will generate fewer trading signals, thus increasing the number of weeks in which the position of the investor is to “hold”.

In order to calculate the return on investment of different strategies, it is assumed that for *trade* signals all the available cash is invested at the beginning of the week in the respective bond, which is then sold at the closing prices of the last day of the week. For *hold* signals, no trading is executed and, therefore, only cash is held during that week while waiting for the appearance of the next signal. In the case of the risk-averse strategy, if the minimum change threshold is not reached, the available cash is also kept

until the next signal appears. To exemplify the deployment scenario, a reference figure of 1m US\$ was used as the starting available cash for the purposes of the return-on-investment estimations. It is assumed that the following information is provided in the form of a simple report to traders on a weekly basis that includes the forecast CLP_5y rate for the following week, and the two signals generated based on the active and risk-averse strategies.

In terms of the monitoring step of the deployment phase, it is assumed that information regarding the back-testing (historical) performance of the selected forecasting models, and the simulated trading results of the alternative strategies, is also presented to traders on a weekly basis. In particular, metrics for monitoring include the average historical weekly results of the total profit, the number of trades, average return per trade, and risk-return metrics such as the standard deviation of the portfolio valuation and the Sharpe ratios (Sharpe, 1994). A positive *total profit* and gain over “Buy and Hold” strategy (Resnick & Shoosmith, 2002) implies successful deployment of the selected models. Figure 14 presents an example of the dashboard of the metrics defined as part of the monitoring phase. For this purpose, simulation data from the second partition, namely, from January 2011 to December 2012, was used. Thus, this corresponds to an example of the view that was available to traders in the first week of January 2013. From its inspection, one can see that the *active* strategy, informed by SVM and linear model forecasts, generated better values for the risk and return metrics. Possible maintenance actions that could be taken include the withdrawal of the *risk-averse* strategy and/or the recalibration of alternative models, including KNN, neural networks, CHAID and C&R Tree.

Active Strategy	Buy & Hold	KNN Algorithm 1	Neural Net 1	SVM 1	CHAID 1	C&R Tree 1	Generalized Linear 1	Linear 1	Regression 1
Total Cumulative Return %	2.41%	1.79%	4.57%	5.89%	1.52%	1.68%	7.25%	7.53%	7.46%
Weekly Average Return %	0.03%	0.02%	0.05%	0.06%	0.02%	0.02%	0.08%	0.08%	0.08%
Yearly Average Return %	1.26%	0.93%	2.38%	3.07%	0.79%	0.87%	3.77%	3.92%	3.89%
Average Return per Trade %	2.41%	0.08%	0.16%	0.19%	0.12%	0.15%	0.25%	0.30%	0.26%
Number of Trades	1	23	29	31	13	11	29	25	29
Percentage of Time with an active Position	1.04%	23.96%	30.21%	32.29%	13.54%	11.46%	30.21%	26.04%	30.21%
Adj Sharpe Ratio	0.054	0.092	0.181	0.180	0.148	0.146	0.232	0.239	0.240
Risk Averse Strategy	Buy & Hold	KNN Algorithm 1	Neural Net 1	SVM 1	CHAID 1	C&R Tree 1	Generalized Linear 1	Linear 1	Regression 1
Total Cumulative Return %	2.41%	1.95%	0.91%	1.14%	1.52%	1.68%	2.12%	1.35%	1.81%
Weekly Average Return %	0.03%	0.02%	0.01%	0.01%	0.02%	0.02%	0.02%	0.01%	0.02%
Yearly Average Return %	1.26%	1.01%	0.47%	0.60%	0.79%	0.87%	1.10%	0.71%	0.94%
Average Return per Trade %	2.41%	0.14%	0.09%	0.10%	0.22%	0.17%	0.35%	0.12%	0.30%
Number of Trades	1	14	10	11	7	10	6	11	6
Percentage of Time with an active Position	1.04%	14.58%	10.42%	11.46%	7.29%	10.42%	6.25%	11.46%	6.25%
Adj Sharpe Ratio	0.054	0.105	0.067	0.098	0.148	0.146	0.095	0.114	0.081

Figure 14: Dashboard of the historical performance of *active* and *risk-averse* strategies

In the longer term, the organization could follow up with additional monitoring activities that relate to the accuracy of the models or changes to the organizational context, and revise the deployment scenario accordingly.

3. Summary and Conclusions

This paper introduces a knowledge-discovery methodology to model and forecast interest rates, taking into consideration the interplay of related economic and financial variables at different stages of the economic cycle. This problem is of great interest to academia and market participants, as reflected in the literature. The evidence reported favours the theoretical explanation that the term structure contains information about future expectations of spot rates, which is linked to the unbiased expectations theory and the liquidity preference theory. In addition, the preferred habitat theory seems to be linked to the liquidity preference theory and presents a better fit for certain features in the term structure.

Nonetheless, it is possible to identify some caveats in previous work. Recent literature has placed an emphasis on the need to address the limitations of the econometric models. However, it has failed to address the interpretability aspects and, more importantly, the need to provide methodological support that allows the deployment of such advanced techniques in a more systematic way.

This paper addresses the interpretability aspects and the lack of methodological support, making a contribution to the field by helping stakeholders in finance that use expert systems with their decision making by proposing a knowledge-discovery methodology that explicitly considers theory in relation to market characteristics (Cox et al., 1985; Diebold & Li, 2006; Estrella & Hardouvelis, 1991; Estrella & Mishkin, 1997, 1998; Svensson, 1994; Vasicek, 1977) that explores existing patterns and relationships in the data (Araújo et al., 2015; Cervelló-Royo et al., 2015; Chang et al., 2009; Guresen et al., 2011; Kamo & Dagli, 2009; Svalina et al., 2013) and, that includes minimal common number of steps to model, analyse and evaluate non-linear relationships between interest rates and the relevant related economic variables using open-box techniques. To the best of our knowledge, this is the first attempt to use decision trees as the tool to investigate and present in human understandable form the non-linear patterns found in the relationships between interest rates and related economic variables.

The proposed methodology defines four objectives: i) discover, and ii) describe hidden patterns, characterizing economic cycles, and economic cycle stages; iii) use these patterns for forecasting economic cycle stages, and iv) forecasting interest rates and other related variables of interest. These are achieved following the five phases of the proposed methodology, considering reverse and backward flows, in order to accommodate continuous improvement and special requirements for modelling. The proposed methodology is sufficiently general to cater for the modelling of any other related target variable.

Regarding the instantiation of the methodology, the results of the models developed as part of the *Economic Stages Discovery* and *Economic Stages Description* steps present evidence in accordance with literature that argues that the predictability of economic growth, or the lack thereof, namely recessions, is a function of the slope of the yield curve, supporting the theoretical explanation that the term structure contains information about future expectations of spot rates. This is because the variable representing the spread between long- and short-term rates, and the swap rates variable, form an important part of the decision rules created with the decision-tree technique. The results also show patterns that are specific to the country from which the data sets were selected. For example, these indicate that, in the case of Chile, low COPPER price levels are associated with a higher probability of recessions.

The results of the *Economic Stages Forecasting* step indicate the prevalence of short-term lagged variables over medium–long-term lagged variables as inputs, at least for the selected forecasting horizon. This could relate to the fact that the chosen forecasting horizon is relatively short, one-week-ahead, and, thus, the long-term trends may be less relevant for the short-term instantiation of the forecasting task. Furthermore, results indicate that inflation rates, both domestic and foreign, and exchange rates are extremely important for discriminating between economic stages. The latter variables place the emphasis on Chile's sensitivity to external events, which also corroborates the role of COPPER prices and their relation to expected recessions. The forecasting power of the resulting decision trees was validated through precision and accuracy ratios on the in- and out-of-sample performance.

Regarding the *Target Variable Forecasting* step, undoubtedly a good selection of input variables is crucial to improving the performance of the model. This is demonstrated by the improvement of

accuracy when previously gathered knowledge in the form of an input variable that reflects the expected economic stage is included. Moreover, models that used auto-regressive components and PCA factors as inputs exhibited better in-testing performance (MASE) than models that included other variables. In general, ensemble, GENLIN and linear models with automatic feature selection were the best performers. These results are valid for the testing sample and were statistically significant at the 99 per cent confidence level. However, Decision Trees could be preferable when understanding of the patterns that originate each prediction is required and when directional forecasting (DA) is deemed more important.

Finally, in the deployment phase, the outputs of the methodology can provide actionable information for market agents. The deployment phase was exemplified using a scenario in which the models deployed inform weekly trading decisions as part of an institutional investor context. In terms of the planning step for the selected scenario, it is assumed that the institutional investor prefers models that exhibit high levels of accuracy, even as a trade-off from interpretability. The economic performance of alternative market-timing strategies was monitored and evaluated following this assumption. From inspection of the results, the scenario presents a case in which the *active* strategy, informed by SVM and linear model forecasts, generated better values for risk and return metrics. Later on, possible maintenance actions discussed in these phases included the withdrawal of other market-timing strategies and/or the recalibration of alternative models. Moreover, in the longer term, the organization could follow up with additional monitoring activities and revise the deployment scenario.

The limitations of the work include the sequential steps followed in the proposed methodology which implies that any errors during one of the stages is propagated to the subsequent ones. In the proposed methodology this is especially true for the characterisation of the economic cycles (Economic Stage Discovery, Description and Forecasting) because the output is used as an input to the subsequent stages. To address this the proposed methodology includes feedback loops that help to optimise the overall process and reduce the impact of error propagation. Another limitation relates to the fact that the proposed methodology helps to discover relationships between the economic variables but not explicitly examines causality. For this, additional techniques such as Wiener-Granger causality test (Granger, 1980) should be used. Finally, a further limitation relates to the relatively short span of the dataset (ten years). In general, the longer the period the better the analysis. However, the purpose of the work here is to propose a generic methodology rather than examine a particular forecasting case.

The work reported in this paper could be extended in a number of different directions. For example, future work could involve further investigation of the descriptive variables and study of the dynamics of the relationships between them. Another future direction related to the previous point is the study of the upcoming implementation of regulations based on Basel III. New liquidity requirements for banks should theoretically affect the stability of the relationships between the economic and financial variables in periods of crisis. Future work could also study how to optimize the out-of-sample performance of the forecasting models, perhaps considering a forecast-and-retrain or “rolling” modelling strategy (Parisi, Parisi, & Díaz, 2008). Furthermore, in relation to relatively low recall accuracy when forecasting *transition* stages, perhaps alternative characterizations of stages and/or alternative ways to detect structural breaks could also be researched. The work could also be extended by further validating the use of methodology in a broader set of countries and considering new sets of economic and financial variables. Moreover, the work could be validated using different forecasting horizons, spans of data sets, and complementary machine-learning strategies such as bagging and boosting techniques (Alpaydin, 2010; Mahato & Attar, 2014). Also, other econometric and data-driven models could be evaluated such as Vector Auto Regressions (VARs) and variants (such as FAVAR) (Bernanke et al.,

2005) that are expected to perform very well in this type of task. In particular, VAR have the advantage of modelling the target variable not only as a function of explanatory variable, but as part of a simultaneously determined system of equations. Other data-driven techniques which are extensions of the Moving Average forecasting technique, including exponential smoothing, double smoothing, and Holt-Winters (Chatfield, 1978) which include components for seasonal trends, could also be tested.

Acknowledgements

The authors would like to acknowledge the contributions from Karen Moreno and Raúl Muñoz, (MSc students, MSc in Finance, Universidad de Chile), for their help with the literature review, collection of the data sets, and preliminary versions of the modelling.

References

- Alexander, G. J., Sharpe, W. F., & Bailey, J. V. (2003). *Fundamentals of Investments/ Fundamentos de Inversiones*. Pearson Educación.
- Alpaydin, E. (2010). *Introduction to Machine Learning*. MIT Press.
- Ang, A., Piazzesi, M., & Wei, M. (2004). What Does the Yield Curve Tell us about GDP Growth? *NBER Working Papers*. Retrieved from <http://ideas.repec.org/p/nbr/nberwo/10672.html>
- Araújo, R. de A., Oliveira, A. L. I., & Meira, S. (2015). A hybrid model for high-frequency stock market forecasting. *Expert Systems with Applications*, 42(8), 4081–4096. doi:10.1016/j.eswa.2015.01.004
- Bank of England. (2015). The transmission mechanism of monetary policy.
- Bernanke, B. S., Boivin, J., & Eliasch, P. (2005). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *The Quarterly Journal of Economics*, 120(1), 387–422.
- Bhaumin, S. (2011). Productivity and the Economic Cycle. Retrieved February 5, 2015, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32108/11-772-productivity-and-the-economic-cycle.pdf
- Brown, S., & Dybvig, P. (1986). The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates. *The Journal of Finance*, 41, 617–630. doi:10.1111/j.1540-6261.1986.tb04523.x
- Cervelló-Royo, R., Guijarro, F., & Michniuk, K. (2015). Stock market trading rule based on pattern recognition and technical analysis: Forecasting the DJIA index with intraday data. *Expert Systems with Applications*, 42(14), 5963–5975. doi:10.1016/j.eswa.2015.03.017
- Chang, P.-C., Liu, C.-H., Lin, J.-L., Fan, C.-Y., & Ng, C. S. P. (2009). A neural network with a case based dynamic window for stock trading prediction. *Expert Systems with Applications*, 36(3), 6889–6898. doi:10.1016/j.eswa.2008.08.077
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0 Step-by-step data mining guide. Retrieved from <http://www.crisp-dm.org/CRISPWP-0800.pdf>
- Chatfield, C. (1978). The Holt-Winters Forecasting Procedure on JSTOR. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 27(3), 264–279.
- Cox, J., Ingersoll, J., & Ross, S. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385. doi:10.2307/1911242
- Cox, J., Ingersoll Jr, J., & Ross, S. (1981). A Re-examination of Traditional Hypotheses about the Term Structure of Interest Rates. *The Journal of Finance*, 36(4), 769–799.
- Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2), 337–364.

- Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications*, 29(4), 927–940.
- Estrella, A. (2005). Why Does the Yield Curve Predict Output and Inflation? *The Economic Journal*, 115(505), 722–744.
- Estrella, A., & Hardouvelis, G. (1991). The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance*, 46(2), 555–576.
- Estrella, A., & Mishkin, F. S. (1997). The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review*, 41(7), 1375–1401.
- Estrella, A., & Mishkin, F. S. (1998). Predicting U.S. Recessions: Financial Variables as Leading Indicators. *Review of Economics and Statistics*. doi:10.1162/003465398557320
- Estrella, A., Rodrigues, A. P., & Schich, S. (2003). How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States. *The Review of Economics and Statistics*, 85(3), 629–644. doi:10.2139/ssrn.249965
- Estrella, A., & Trubin, M. R. (2006). The Yield Curve as a Leading Indicator : Some Practical Issues. *Finance*, 12, 1–7.
- Gogas, P., Chionis, D., & Pragkidis, I. (2009). Predicting European Union recessions in the euro era: The yield curve as a forecasting tool of economic activity. *International Advances in Economic Research*, 16(1).
- Gogas, P., Papadimitriou, T., Matthaïou, M., & Chrysanthidou, E. (2014). Yield Curve and Recession Forecasting in a Machine Learning Framework. *Computational Economics*.
- Granger, C. W. J. (1980). Testing for causality. *Journal of Economic Dynamics and Control*, 2, 329–352.
- Greer, M. (2003). Directional accuracy tests of long-term interest rate forecasts. *International Journal of Forecasting*, 19(2), 291–298. doi:10.1016/S0169-2070(01)00141-8
- Guresen, E., Kayakutlu, G., & Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 38(8), 10389–10397. doi:10.1016/j.eswa.2011.02.068
- Heath, D., Jarrow, R., & Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60, 77–105. doi:10.2307/2951677
- Hicks, J. R. (1939). *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. Oxford: Clarendon Press.
- Ho, T. S. Y., & Lee, S. B. (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41, 1011–1029. doi:10.2307/2328161
- Hong, T., & Han, I. (2002). Knowledge-based data mining of news information on the Internet using cognitive maps and neural networks. *Expert Systems with Applications*, 23(1), 1–8. doi:10.1016/S0957-4174(02)00022-2

- Hull, J., & White, A. (1990). Pricing Interest-Rate-Derivative Securities. *Review of Financial Studies*, 3, 573–592. doi:10.2307/2962116
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. 9
- IBM. (2014, December 17). Retrieved December 17, 2014, from <http://www-01.ibm.com/software/analytics/spss/products/modeler/>
- Investopedia. (2015). Retrieved February 5, 2015, from <http://www.investopedia.com/terms/e/economic-cycle.asp>
- Jacovides, A. (2008). Forecasting Interest Rates from the Term Structure: Support Vector Machines Vs Neural Networks.
- Joseph, A., Larrain, M., & Singh, E. (2011). Predictive Ability of the Interest Rate Spread Using Neural Networks. *Procedia Computer Science*, 6, 207–212.
- Ju, Y. J., Kim, C. E., & Shim, J. C. (1997). Genetic-based fuzzy models: Interest rate forecasting problem. *Computers & Industrial Engineering*. doi:10.1016/S0360-8352(97)00193-9
- Kamo, T., & Dagli, C. (2009). Hybrid approach to the Japanese candlestick method for financial forecasting. *Expert Systems with Applications*, 36(3), 5023–5030. doi:10.1016/j.eswa.2008.06.050
- Kim, S. H., & Noh, H. J. (1997). Predictability of interest rates using data mining tools: A comparative analysis of Korea and the US. *Expert Systems with Applications*, 13(2), 85–95.
- Mahato, P. K., & Attar, V. (2014). Prediction of gold and silver stock price using ensemble models. In *2014 International Conference on Advances in Engineering & Technology Research (ICAETR - 2014)* (pp. 1–4).
- McCulloch, J. H. (1971). Measuring the Term Structure of Interest Rates. *The Journal of Business*, 44(1), 19–31.
- McCulloch, J. H. (1975). The Tax-adjusted Yield Curve. *The Journal of Finance*, 30(3), 811–830. doi:10.1111/j.1540-6261.1975.tb01852.x
- McCulloch, J. H. (1987). The monotonicity of the term premium : A closer look. *Journal of Financial Economics*, 18, 185–192. doi:DOI: 10.1016/0304-405X(87)90068-7
- Nelson, C., & Siegel, A. (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*. doi:10.1086/296409
- Oh, K. J., & Han, I. (2000). Using change-point detection to support artificial neural networks for interest rates forecasting. *Expert Systems with Applications*, 19(2), 105–115.
- Parisi, A., Parisi, F., & Díaz, D. (2008). Forecasting gold price changes: Rolling and recursive neural network models. *Journal of Multinational Financial Management*, 18(5), 477–487.
- Patton, A. J., & Timmermann, A. (2010). Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts. *Journal of Financial Economics*, 98, 605–625. doi:10.1016/j.jfineco.2010.06.006

- Quinlan, J. R. (1990). Learning logical definitions from relations. *Machine Learning*, 5, 239–266. doi:10.1007/BF00117105
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann San Mateo California (Vol. 1). doi:10.1016/S0019-9958(62)90649-6
- Resnick, B. G., & Shoesmith, G. L. (2002). Using the Yield Curve to Time the Stock Market. *Financial Analyst Journal*, 58(3), 82–90.
- Richardson, M., Richardson, P., & Smith, T. (1992). The monotonicity of the term premium. *Journal of Financial Economics*. doi:10.1016/0304-405X(92)90013-N
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65.
- Sharpe, W. F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*, 21(1), 49–58. doi:10.3905/jpm.1994.409501
- Svalina, I., Galzina, V., Lujčić, R., & Šimunović, G. (2013). An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. *Expert Systems with Applications*, 40(15), 6055–6063. doi:10.1016/j.eswa.2013.05.029
- Svensson, L. E. O. (1994). Estimating and Interpreting Forward Interest Rates. *IMF Working Paper*, 94, 1–29.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*. doi:10.1016/0304-405X(77)90016-2
- Vela, D. (2013). Forecasting Latin-American yield curves: An artificial neural network approach. Retrieved November 15, 2014, from http://www.banrep.gov.co/sites/default/files/publicaciones/archivos/be_761.pdf
- Zimmermann, H.-G., Tietz, C., & Grothmann, R. (2002). Yield curve forecasting by error correction neural networks and partial learning. In *ESANN 2002, 10th Eurorean Symposium on Artificial Neural Networks, Bruges, Belgium, April 24-26, 2002, Proceedings* (pp. 407–412).