Doctoral theses at NTNU, 2021:272

Ranik Raaen Wahlstrøm

Financial data science for exploring and explaining the ever-increasing amount of data

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Economics and Management
NTNU Business School

NTNU
Norwegian University of
Science and Technology

Ranik Raaen Wahlstrøm

Financial data science for exploring and explaining the ever-increasing amount of data

Thesis for the Degree of Philosophiae Doctor

Trondheim, September 2021

Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Economics and Management
NTNU Business School

ISBN 978-82-326-5700-1 (printed ver.)
ISBN 978-82-326-6261-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:272

Printed by NTNU Grafisk senter

Acknowledgments

I am very grateful to my main supervisor Florentina Paraschiv for exceptionally good guidance. Her enthusiasm, encouragements, and continuous academic and moral support have made me passionate about my work. I am also grateful to my co-supervisor Roland Füss and my co-authors Akarsh Kainth, Markus Schmid, and Michael Schürle for their valuable feedback and productive collaborations. I thank NTNU Business School and the Faculty of Economics and Management at the Norwegian University of Science and Technology for the opportunity to pursue a doctoral degree within the field of my interest. Further, I am very appreciative for the valuable support from my colleagues at NTNU Business School and co-members at the Centre for Banking and Finance. My profound gratitude goes to my family for the continuous encouragements and help. I am especially grateful to Martha and our son Fredrik for all the support and happiness they have offered me.

Trondheim, May 2021

Ranik Raaen Wahlstrøm

Summary

Financial data science is an interdisciplinary emerging new research paradigm which intersects econometrics and data science. It is considered to be a discipline in its own right and is not only concerned with statistical inference but also with exploring and explaining data for advancing financial decision making.

This thesis contributes to data-driven financial studies by arguing for two potential improvements of financial data science over pure financial econometrics. First, this thesis argues for letting a data-driven process guide the selection of model variables in cases of data sets with many observations or many competing variables available. This is addressed in Articles 2 and 3 in this thesis, in applications to corporate finance with focus on company bankruptcy prediction. Second, this thesis argues for evaluating models not solely based on goodness of fit criteria and standard statistical metrics, but also on the real economic implications of their predictions and the stability of their estimated parameter values when these have an economic interpretation. Articles 1 and 3 of this thesis address this in applications to yield curve modeling and company bankruptcy prediction, respectively.

For central banks, this thesis makes recommendations on relevant modeling and data choices when fitting parsimonious yield curve models for monetary policy decisions. The recommendations have a particular emphasis on the stability of parameter estimates over time, as these have an intrinsic economic meaning. Further, this thesis shows that feature selection methods improve bankruptcy prediction models commonly used by banks and financial regulators. Moreover, it proposes an improved bankruptcy prediction model for small and medium-sized enterprises (SMEs) compared to the benchmark model employed by the Financial Supervisory Authority of Norway. Finally, this thesis documents evidence for financial regulators concerning the benefits of aligning national accounting standards towards International Financial Reporting Standards (IFRS).

Table of contents

Research articles:

Article 1: Wahlstrøm, R.R., Paraschiv, F., Schürle, M., 2021. A Comparative Analysis of Parsimonious Yield Curve Models with Focus on the Nelson-Siegel, Svensson and Bliss Versions. Computational Economics. doi:10.1007/s10614-021-10113-w.

Article 2: Kainth, A., Wahlstrøm, R.R., 2021. Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies. Journal of Risk and Financial Management 14, 123. doi:10.3390/jrfm14030123.

Article 3: Paraschiv, F., Schmid, M., Wahlstrøm, R.R., 2021. Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods. Working Paper, Norwegian University of Science and Technology and University of St. Gallen, to be submitted to the Review of Finance.

1. Introduction

Financial econometrics is the application of statistical methods to problems in finance (Brooks, 2019). The underlying platform of most econometric modeling consists of linear regression, parameter estimation, and hypothesis testing with statistical significance levels (Greene, 2012; Varian, 2014a; Mullainathan and Spiess, 2017; De Prado, 2018; Brooks, 2019; Simonian and Fabozzi, 2019; Khraisha, 2020). However, using linear regression when solving problems in finance can be problematic, as it relies on strong assumptions that are often false, e.g., assumptions of multivariate normal distributions and linear relationships. Further, practices of multiple hypothesis testing have produced potentially false findings due to selection bias and the pressure to produce significant results (Kim and Ji, 2015; Harvey et al., 2016; Harvey, 2017; Khraisha, 2020). Moreover, the emphasis on statistical significance levels is contrary to the American Statistical Association, which states that using statistical significance for justifying scientific claims can lead to erroneous beliefs and poor decision making (Wasserstein and Lazar, 2016; Wasserstein et al., 2019). The lack of confidence in statistical significance levels has even led to their use being banned in scientific journals (Trafimow and Marks, 2015). In addition, given the increasing amount of data employed in financial applications over the last years, the likelihood of favorable statistical significance levels for the model parameters also increases, making them less suitable for interpretations of results and evaluation. (Harvey, 2017; Brooks et al., 2019).

1.1. Financial data science and flexible machine learning techniques

Financial data science is an emerging new research paradigm which expands the scope of financial econometrics to cope with these problems (Brooks et al., 2019; Simonian and Fabozzi, 2019; Khraisha, 2020). Rather than being one area of applied data science, which is the study of extracting knowledge and insights from data (Dhar, 2013), financial data science is considered to be a discipline in its own right, at the intersection between data science and econometrics (Simonian and Fabozzi, 2019). Thus, while econometrics is concerned with statistical inference, financial data science is also concerned with how the exploration and explanation of data can advance financial decision making (Brooks et al., 2019). Particularly, compared to regression analysis, which is commonly used in econometrics, financial data science also makes use of more flexible machine learning techniques, e.g., artificial neural networks. Such techniques can capture multivariate non-linear relations and rely on few or no assumptions about the data or the error terms. Thus, they can discover complex structures that are not specified in advance, making them suitable for harnessing the new opportunities for financial and economic research emerging due to the ever-increasing amount of data available (Einav and Levin, 2014; Varian, 2014b; Mullainathan and Spiess, 2017).[1]

An example of a research field that has shifted from traditional regression analysis to more flexible machine learning techniques is company bankruptcy prediction.[2] Early studies on this topic typically use discriminant analysis (e.g., Altman, 1968; Altman et al., 1977; Taffler, 1984). However, when solving problems in economics and finance, including bankruptcy prediction, it is problematic to use discriminant analysis, as it makes several assumptions that do not hold for economic and financial data (Joy and Tollefson, 1975; Deakin, 1976; Eisenbeis, 1977). This includes the assumption of equal variance-covariance matrices across the classes of data, as well as the assumption of multivariate normal distribution of input variables. Consequently, Martin (1977) and Ohlson (1980) use logistic regressions for bankruptcy prediction, which rely on less restrictive assumptions and produce more intuitive outputs. However, logistic regressions are sensitive to outliers, missing values, and multicollinearity (Balcaen and Ooghe, 2006). The latter is particularly problematic for bankruptcy prediction, as input variables often are financial ratios which frequently share the same accounting numbers in the numerators

[1]A profound introduction to machine learning techniques in general can be found in James et al. (2013). As highlighted by Hoepner et al. (2021), regression analysis commonly used in econometrics also constitute machine learning techniques as machine learning occurs whenever a "computer program is said to learn from experience *E* with respect to some class of tasks *T* and performance measure *P* if its performance at tasks in *T*, as measured by *P*, improves with experience *E*." (Mitchell, 1997, p.2). Clearly, this often applies to regression analysis when solving problems in economics and finance. However, among all machine learning techniques, regressions are the least flexible.

[2]I refer to Jones et al. (2015, 2017) and Næss et al. (2017) for empirical evaluations of techniques used for bankruptcy prediction.

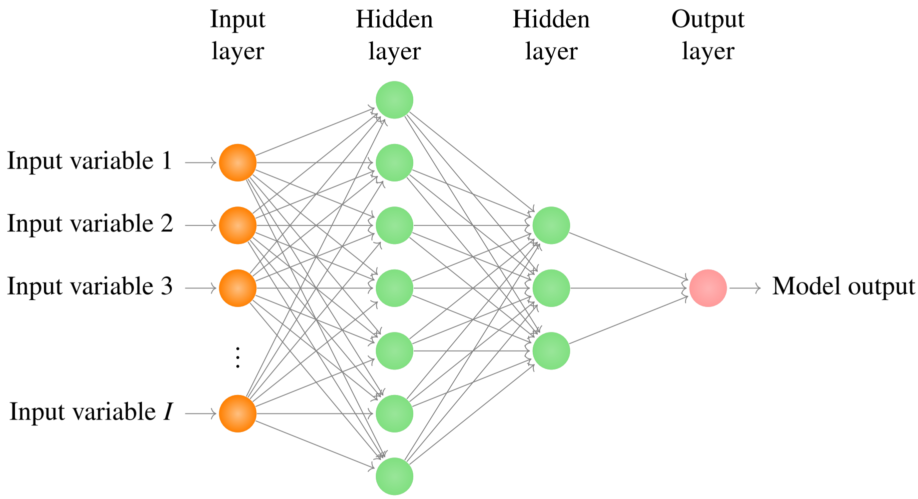
or denominators. Another serious drawback of logistic regressions is that they can be sensitive to extreme non-normality of input variables (Mcleay and Omar, 2000)

Recently, researchers and practitioners started to use more flexible machine learning techniques for bankruptcy prediction. Among these techniques, artificial neural networks are the most widely used since the 1990s (Bellovary et al., 2007; Kumar and Ravi, 2007). Their use for bankruptcy prediction is found, among others, in Tam and Kiang (1992), Zhang et al. (1999), Geng et al. (2015), du Jardin and Séverin (2012), and du Jardin (2015). Typically, bankruptcy prediction studies applying this technique use a subgroup called feedforward artificial neural networks which consist of several layers in the following order: First, they consist of an input layer which contains a number of nodes that corresponds to the number of model input variables. Each node in this layer has a value that is the same as the value of one of the input variables, respectively. Second, feedforward artificial neural networks consist of a predefined number of hidden layers, each containing a predefined number of nodes. Finally, they consist of an output layer containing one or more nodes which represent the model output. Figure 1 illustrates a feedforward artificial neural network that has two hidden layers with seven and three nodes, respectively. Further, each node in the hidden and output layers of feedforward artificial neural networks has a value which is computed by a predefined transfer function as illustrated in Figure 2. The input of this function is the sum of a bias value and the values of the nodes in the previous layer each multiplied with an associated weight. Feedforward artificial neural networks are trained by estimating all weights and bias values of all its nodes in the hidden and output layers.[3]

Another example of flexible machine learning techniques being applied for bankruptcy prediction include decision trees (e.g., Marais et al., 1984; Frydman et al., 1985; Cielen et al., 2004; Gepp et al., 2010; Tsai and Hsu, 2013). These perform model training by dividing the input variable space into distinct and non-overlapping regions, each falling

[3]A thorough description of a feedforward artificial neural network and how it is trained is given in Appendix B in Article 3 of this thesis.

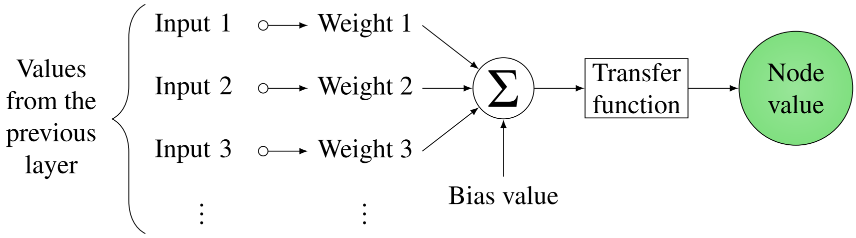
Figure 1: Illustration of a feedforward artificial neural network.



This feedforward artificial neural network has two hidden layers with seven and three nodes, respectively. The input layer has I nodes whose values are given by the I input variables. Each node in the hidden and output layers has a value which is computed as illustrated in Figure 2. The node of the output layer represents the model output.

into one of the possible prediction classes, e.g., bankrupt or non-bankrupt. After the input variable space is divided, any new observation is predicted by a decision tree to the class of the region in which the observation is located in. Further, support vector machines introduced by Vapnik (1998) are also flexible machine learning techniques used in several bankruptcy prediction studies (e.g., Min and Lee, 2005; Shin et al., 2005; Härdle et al., 2009). Support vector machines aim to solve a binary classification problem by dividing the input variable space by a linear hyperplane into two regions. Each region is assigned with one of the two classes, and any new observation is predicted to the class of the region in which the observation is located in. For managing the non-linearities in the data, support vector machines first map the original input variable space into a higher dimensional space using a kernel function. This makes it more likely to obtain a satisfactory separation by the linear hyperplane. Moreover, the k-nearest neighbor has also been used for bankruptcy prediction (e.g., Park and Han, 2002). This technique is non-parametric and predicts any new observation based on the class affiliation of the

Figure 2: The computation of the value of a node in a hidden or output layer of a feedforward artificial neural network.



The values of each node in the previous layer is multiplied by an associated weight. Further, the sum of all these products and a bias value is the input of a transfer function. The output of this transfer function is the computed value of the node. Feedforward artificial neural networks are trained by estimating all weights and biases of all its nodes in the hidden and output layers.

observations nearest in terms of distance in the input variable space.

1.2. Motivation and implications of this thesis

The purpose of this thesis is to contribute to data-driven financial studies by arguing for two potential improvements of financial data science over pure financial econometrics. These potential improvements are emphasized in three research articles, as applications to yield curve modeling and corporate finance, with focus on company bankruptcy prediction.

First, this thesis argues that in the case of data sets with many observations, or in the event that a data set is extensive in the number of competing explanatory variables, the choice of model variables should be guided by data-driven processes. This can be done by using feature selection methods, which systematically and empirically choose a predetermined number of input variables to be used for modeling. Indeed, feature selection methods can enhance generalization, improve explanatory power, reduce the computation time, and give a better understanding of the data (Guyon and Elisseeff, 2003; Chandrashekar and Sahin, 2014; Tian et al., 2015). Further, they are classified into filter, wrapper, and embedded methods (Chandrashekar and Sahin, 2014). *Filter methods* use a predefined criterion, e.g., the Pearson correlation coefficient, to measure the relationships between the values of the single variables and the classifications of the observations in the data. The variables that rank highest according to this predefined criterion are those selected by the filter methods. *Wrapper methods* select input variables heuristically (John et al., 1994; Kohavi and John, 1997). They start with all variables available or no variables at all, before iteratively removing or adding one or more variables until a predetermined number of input variables is reached. This procedure follows an algorithm which chooses variables to remove or add in accordance to their performance when used in a model trained and evaluated with the data. *Embedded feature selection methods* incorporate feature selection as part of model parameter estimation. One of these methods is the least absolute shrinkage and selection operator (LASSO) method popularized by Tibshirani (1996). This method includes a penalty term in the objective function used when estimating parameter values. The weighting of this penalty term determines the number of estimated parameter values that become zero. Initially, the model parameters are estimated with a weighting of the penalty term so high that all the estimated parameter values of all variables become zero. After this, the model parameters are re-estimated repeatedly, each time with a gradually lower weighting of the penalty term, such that one by one the estimated parameter values become non-zero. This process stops when a predetermined number of estimated parameter values becomes non-zero, and the variables associated with these non-zero estimated parameter values are those selected by the LASSO method.

As opposed to using feature selection methods, previous studies within fields in finance, including bankruptcy prediction, often choose variables ad-hoc from a large list of competing variables based on subjective criteria, e.g., their frequency of use in other studies (Balcaen and Ooghe, 2006; Appiah et al., 2015; Tian et al., 2015; Tian and Yu, 2017; Gupta et al., 2018). Article 3 of this thesis addresses the importance of variable selection by using feature selection methods in an application to bankruptcy prediction and shows major implications for the decision making in credit risk management. Further, the data also guide the input variable selection in Article 2 of this thesis. This article is also an application to bankruptcy prediction, and in particular the variables are selected such that they do not violate the assumptions about the data made by the technique used for model estimation.

Second, this thesis argues that models should be evaluated not only by goodness of fit criteria and standard statistical metrics, but also by *i)* the real economic implications of the models' predictions and *ii)* the stability of estimated parameter values. The latter is highly relevant, especially when model parameters build on decision making processes in the financial sector. One example in this case are the parametric parsimonious yield curve models, where estimated model parameters are used by central banks for monetary policy decision making. Yield curves, and thus the parameters of models used for constructing them, reflect both changes in future economic activity and response actions taken by the monetary authorities (Bretscher et al., 2018). Indeed, the short end of the yield curve reflects monetary policy decisions by central banks in response to changes in inflation, economic activity, and other economic conditions (Taylor, 1993). Furthermore, the medium range and long end of the yield curve reflect the central banks' inflation targets, credibility, and communication about the intended future course of action (Lengwiler and Lenz, 2010). Article 1 of this thesis estimates parsimonious yield curve models and reveals that the stability over time of their estimated parameter values are highly affected by different modeling and data choices. These choices include model configuration, parameter constraining, data selection, and approaches for selecting initial parameter values for the numerical estimation procedure. Moreover, Article 1 provides details on these choices, as well as recommendations to promote the stability of parameter estimates and thus their financial interpretation.

Further, a major problem with using solely statistical metrics for model evaluation is that they assume equal costs of the different types of prediction errors, which is an assumption that often is false in real-world applications (Altman et al., 1977; Zmijewski, 1984; Stein, 2005; Balcaen and Ooghe, 2006; Agarwal and Taffler, 2007, 2008; Bauer and Agarwal, 2014; De Bock et al., 2020). For example, models for automatic fraud detection flag potential fraudsters for manual investigation. In this case, not flagging a fraudster is often much costlier than flagging a non-fraudster. However, statistical metrics for model evaluation treat the costs as equal. In this thesis, the weaknesses of statistical metrics for model evaluation are addressed in Article 3 in applications to company bankruptcy prediction. Rather than using statistical metrics, Article 3 proposes to evaluate models based on the economic implications of their predictions for the users. Particularly, Article 3 evaluates bankruptcy prediction models by their effects on the profitability of banks in a simulation of a competitive credit market based on actual market data. The banks in this simulation use bankruptcy prediction models for credit decisions and pricing, which are also something real banks use such models for.

The rest of this thesis is organized as follows: Section 2 deals with variable selection guided by the data, which is showed empirically in Articles 2 and 3 in applications to company bankruptcy prediction. Further, Section 3 shows the value of evaluating models based on the real economic implications of their predictions as well as the stability of their parameter estimates. This is investigated empirically in Articles 1 and 3 in applications to yield curve modeling and company bankruptcy prediction, respectively. A summary of the three research articles included in this thesis, as well as their scientific contributions, are presented in Section 4. Section 5 lists some additional academic activities I have been involved in during my time as PhD student at NTNU Business School. Finally, the three research articles included in this thesis follow at the end.

2. Data guided variable selection

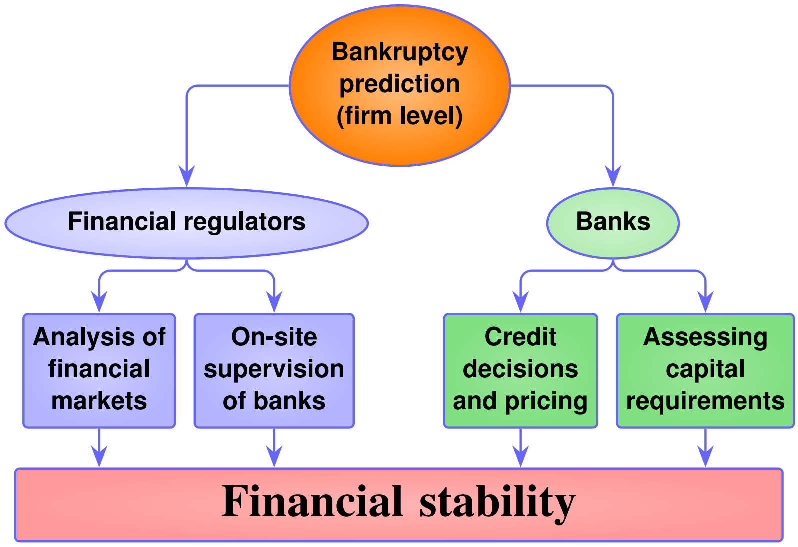
This thesis argues that in the light of the ever-increasing amount of observations and potential model variables, the selection of variables should be guided by the data. This is addressed in Articles 2 and 3 of this thesis in applications to bankruptcy prediction.

2.1. Relevance of variable selection for bankruptcy prediction

The bankruptcy of a company has significant negative economic consequences, such as loss of jobs, loans, equity, future earnings, and future tax revenues. Thus, accurately predicting bankruptcy is of critical importance for many actors, either for company recovery processes, or for reducing the negative effects for stakeholders and the economy if bankruptcy is unavoidable. Also, accurate bankruptcy predictions at the firm level by financial regulators and banks are a precondition for managing systematic risk and promoting financial stability, as outlined in Figure 3. First, for financial regulators, bankruptcy prediction is a key element for the analysis of financial markets and for the on-site supervision of banks (Bernhardsen and Larsen, 2007). Second, for banks, bankruptcy prediction is considered to be the core of credit risk management and has become even more relevant after the Basel regulatory framework introduced the rating of borrowers as a central criterion for minimum capital requirements (Härdle et al., 2009; BIS, 2017). In particular, the Basel framework allows banks to use statistical prediction models for calculating borrower rating used for setting minimum capital requirements. The predictions made by such models are furthermore used by banks for evaluating the risks associated with new and existing customers when making credit decisions and pricing, i.e., when deciding on whether to grant loans and on what terms. Even small improvements in the models' prediction abilities can lead to significant economic benefits for banks by avoiding charging borrowers incorrectly (Stein, 2005).

The first use of accounting numbers for assessing the creditworthiness of companies is found in Rosendale (1908) who use the current ratio for this purpose. This is followed by Smith and Winakor (1930, 1935), FitzPatrick (1932), and Merwin (1942) who investigate how the values of also other individual accounting-based ratios are related to company failure. The prediction of company failure based on individual accounting-based ratios is first analyzed by Beaver (1966, 1968). He also suggests that further research should investigate whether even more precise predictions are possible if multiple accounting-based ratios are considered simultaneously. This is followed up by Altman (1968) who introduces the first multivariate model for bankruptcy prediction. His model uses discriminant analysis and five accounting-based ratios categorized into the main aspects of a company's financial profile: liquidity, profitability, leverage, solvency (coverage), and activity. More recent studies, e.g., Shumway (2001), argue that the accuracy

Figure 3: Bankruptcy prediction for managing systematic risk and promoting financial stability.



Financial regulators perform bankruptcy prediction on firm level for analysis of financial markets and for on-site supervision of banks. Further, banks use bankruptcy prediction for making credit decisions and pricing, as well as for setting minimum capital requirements in accordance with financial regulatory frameworks. In sum, this makes bankruptcy prediction a precondition for managing systematic risk and promoting financial stability.

of bankruptcy prediction models is improved if using market-based input variables collected from market data in combination with accounting-based variables collected from financial statements. Following this, many studies consider bankruptcy prediction models with both market-based and accounting-based variables (e.g., Chava and Jarrow, 2004; Campbell et al., 2008; Beaver et al., 2012; Tian et al., 2015; Blöchlinger and Leippold, 2018). However, other studies argue that models with solely accounting-based variables are sufficient or even outperform models with both groups of variables (e.g., Reisz and Perlich, 2007; Agarwal and Taffler, 2008).

Overall, the literature has introduced hundreds of input variables for bankruptcy prediction models, which can be observed, e.g., in the overview of variables found in Table 2 in the review study of Kumar and Ravi (2007). Thus, the selection of variables requires careful consideration. However, existing studies often choose variables based on subjective criteria such as the researchers' own discretion, data availability, or frequency of use in other studies (Balcaen and Ooghe, 2006; Appiah et al., 2015; Tian et al., 2015; Tian and Yu, 2017; Gupta et al., 2018). Indeed, previous literature has not given appropriate attention to the selection of variables (Härdle et al., 2009; Tian et al., 2015), which is also still actively debated among both academics and practitioners (see, e.g., Fitzgerald, 2009; Toplensky, 2020). This thesis addresses this by arguing for letting the data guide the selection of model variables.

2.2. Variable selection methods applied in this thesis

In Article 2 of this thesis, model variables for bankruptcy prediction are selected by starting with an initial set of variables, before subsequently removing several of them to avoid violation of the no multicollinearity assumption about the data made by the estimation technique, logistic regressions. After this selection, Article 2 compares the abilities of the selected model variables to predict bankruptcy when applied to financial statements derived under International Financial Reporting Standards (IFRS) and local Generally Accepted Accounting Principles (GAAP), respectively. This is done to investigate any differences in the quality of financial reporting caused by using IFRS compared to local GAAP.

Further, Article 3 of this thesis tests a filter, a wrapper, and an embedded feature selection method, respectively, for selecting input variables for bankruptcy prediction models. The methods are allowed to select from a total of 155 accounting-based variables retrieved from prior studies, derived from a comprehensive dataset of privately held Norwegian small and medium-sized enterprises (SMEs) in 2006-2014. The article shows that an embedded LASSO feature selection method yields the best model performance across different time periods and across two different estimation techniques – an artificial neural network and logistic regressions. Further, Article 3 confirms the superiority of the LASSO method when evaluating the effects of the different variable sets given by the feature selection methods on bank profitability in a simulation of a competitive credit market. This simulation employs real-world data and is detailed in Section 3.1.

During my time as PhD student at NTNU Business School, I have also contributed to the literature with the research presented in the article of Pelja and Wahlstrøm (2021).[4] This article is not part of this thesis. Nevertheless, it is relevant, as it emphasizes that the development of models, e.g., the selection of their variables, should be guided by the characteristics of the underlying data. Particularly, this article assesses the performance of bankruptcy prediction models on different subsets of data. It uses a data set of 992,369 financial statements and reveals that the bankruptcy prediction models under consideration perform better when applied to medium-sized companies compared to smaller and larger companies. This finding is robust across three variable sets and two estimation techniques.

3. Value based model evaluation

This thesis argues for that model evaluation should not be done solely based on statistical metrics commonly used in the existing literature, e.g., accuracy, Brier score, and decile rankings. Rather, models should also be evaluated based on the stability of their parameter estimates when these have a specific financial interpretation. This is addressed in applications to yield curve modeling in Section 3.2. Further, model evaluation should also be based on the real economic implications of the models' predictions for the users. This is addressed next in applications to company bankruptcy prediction.

3.1. Economic implications

The costs for a bank are typically much higher when *i)* predicting *low* probabilities of bankruptcy for potential new borrowers that actually *go bankrupt* compared to *ii)* predicting *high* probabilities of bankruptcy for potential new borrowers that actually *do not go bankrupt*. The former generally causes severe costs as it often results in the bank granting loans to bad borrowers that eventually default on their loans. The latter, on the other hand, generally causes relatively smaller costs of not receiving the potential interest profits of lending to good borrowers. However, statistical metrics commonly used

[4]The current version of this article is available at http://pelja2021.ranik.no. It is submitted to the Norwegian scientific journal Magma, ISSN 1500-0788. Details about this submission is given in Section 5.

for evaluating bankruptcy prediction models assume that the different types of prediction errors carry equal costs.

This drawback of statistical metrics when evaluating bankruptcy prediction models is addressed in Article 3 of this thesis. It proposes to rather evaluate bankruptcy prediction models based on the real economic implications of their predictions for banks. This is done in a simulation of a competitive credit market that builds on the frameworks of Stein (2005) and Blöchlinger and Leippold (2006) for credit decision making and credit risk pricing. Article 3 follows the simulations in Agarwal and Taffler (2007, 2008) and Bauer and Agarwal (2014), yet extends them by employing real-world data from the effective size of the whole Norwegian SME loan market. In particular, the simulation in Article 3 includes all companies in this loan market, and lets each company be a potential borrower that wants to borrow an amount equivalent to that of the interest-bearing debt from its financial statement. Additionally, the simulation includes several hypothetical banks, each using one of the bankruptcy prediction models to be evaluated, respectively, to derive a credit spread for each potential borrower. This credit spread is used to decide whether to grant a loan to the potential borrower and, if the loan is granted, on what terms. In cases where a potential borrower is granted a loan from several banks, it borrows solely from the bank offering the best terms. Further, the simulation in Article 3 computes the profits of each hypothetical bank based on the revenues from their lending and the losses from their bankrupted borrowers. Finally, the profits of each hypothetical bank are used for evaluating the bankruptcy prediction models they apply.

3.2. Parameter stability

For several problems in finance, the estimated model parameter values have a specific financial meaning. When this is the case, the stability of these parameter values over time becomes a key consideration. This is addressed in Article 1 of this thesis, which makes recommendations concerning relevant modeling choices for central banks when using parametric parsimonious yield curve models for monetary policy decisions.

Yield curves describe the spot rates, forward rates, or discount factors for different times to maturity (BIS, 2005).[5] They are considered to be the most basic building block of finance and are used for many applications among academics, practitioners, and central bankers (BIS, 2005; Gürkaynak et al., 2007; Diebold and Rudebusch, 2013; Duffee, 2013). These applications include managing financial risk, allocating portfolios, structuring debt, valuating capital goods, pricing financial assets and derivatives, making monetary policy decisions, and predicting or explaining related variables, e.g., macroeconomic activity, real rates, inflation, and the dynamics of risk premia. Yield curves are constructed from spot rates, forward rates, or discount factors derived from the observed market prices of fixed-income instruments and their future cash flows, i.e., coupon payments and face value repayments, as well as their time to maturity (James and Webber, 2000; Diebold and Rudebusch, 2013). Fama and Bliss (1987) provide an approach for constructing forward rates or spot rates at maturities other than those of the observed future cash flows of the instruments. This approach first considers forward rates at the cash flows' different maturities before sequentially constructing "unsmoothed Fama-Bliss" forward rates or spot rates of synthetic instruments at other maturities.[6] In any case, as yield curves are continuous, they require a functional form to be fitted to the spot rates, forward rates, or discount factors derived from either the observed or synthetic instruments.

One option for constructing continuous yield curves is to use linear non-parametric spline-based methods (James and Webber, 2000; BIS, 2005). For example, McCulloch (1971, 1975b,a) uses cubic splines, i.e., splines of order three, to construct yield curves to observed discount factors. However, splines may not produce a good curve at short and long maturities, due to their tendency to oscillate excessively at the outer ranges of the curve. This results in yield curves that tend to diverge at long maturity where the yields typically flatten, i.e., do not change with increasing time to maturity (Shea, 1984).

---

[5]The spot rate $s(m)$ for times to maturity $m \in [0, \infty)$ is the annualized percentage return for a fixed-income instrument which pays no coupons. It relates to the discount factor $\delta(m)$ by $s(m) = -\frac{\log(\delta(m))}{m}$. Further, the spot rate relates to the forward rate $f(m)$ by $f(m) = s(m) + m\dot{s}(m)$ where $\dot{s}(m)$ is the derivative of $s(m)$ with respect to m.

[6]See Bliss (1997b) for details on "unsmoothed Fama-Bliss" forward rates and spot rates.

Vasicek and Fong (1982) and Fisher et al. (1995) address this by using exponential splines and smoothing splines, respectively, which ensures that the curve converges to a fixed limit with increasing time to maturity.

Another option for constructing continuous yield curves is to use parsimonious parametric models consisting of few factors driven by a set of parameters (e.g., Vasicek, 1977; Litterman and Scheinkman, 1991; Bliss, 1997a). Such parsimonious models are appealing for several reasons (Diebold and Rudebusch, 2013; Duffee, 2013). First, they are more manageable and interpretable than splines because they effectively collapse a high-dimensional modeling situation into a low-dimensional one. Second, financial theory suggests the factor structure of the parsimonious yield curve models (Diebold and Rudebusch, 2013). Third, parsimonious yield curve models provide a good fit to the data as it appears it is possible to explain almost all the variation over time in observed yields with only a few principal components. Finally, parsimonious yield curve models are flexible enough to capture a range of monotonic, humped and S-type shapes typically found in observed yields (De Pooter, 2007).

A frequently used parsimonious yield curve model is the Nelson-Siegel model proposed by Nelson and Siegel (1987). It gives the spot rate $s(m)$ as a function of time to maturity $m \in [0, \infty)$ given by

$$s(m) = \beta_0 + \beta_1 \frac{1 - e^{\frac{-m}{\tau}}}{\frac{m}{\tau}} + \beta_2 \left(\frac{1 - e^{\frac{-m}{\tau}}}{\frac{m}{\tau}} - e^{\frac{-m}{\tau}} \right) \tag{1}$$

where β_0, β_1, β_2, and $\tau > 0$ are parameters to be estimated. The first, second, and third factors of Equation (1) control the long, short, and medium segments of the yield curve, respectively, and may therefore be interpreted as the level, slope, and curvature factors (Nelson and Siegel, 1987; Diebold and Li, 2006). The magnitudes of these three factors are given by β_0, β_1, and β_2, respectively. The decay parameter τ determines the exponential decay rate of the slope and curvature factors, as well as the location of the hump or trough associated with the curvature factor.

While many extensions of the Nelson-Siegel model have been proposed (e.g., Björk and Christensen, 1999; Diebold et al., 2005), the original Nelson-Siegel model and the extension by Svensson (1994, 1995) are those most used by central banks (BIS, 2005; Gürkaynak et al., 2007; Nymand-Andersen, 2018). The latter model extends the former with an additional curvature factor, which is considered beneficial since it allows for an extra curvature in the yield curve at longer maturities (Svensson, 1994, 1995; Diebold and Li, 2006; Gürkaynak et al., 2007). Some studies further derive dynamic versions of the parsimonious models (e.g., Diebold and Li, 2006; Diebold et al., 2006; De Pooter, 2007; Koopman et al., 2010). However, since the maturity of fixed-income instruments in the market varies over time, the dynamic model versions require the use of synthetic instruments with maturity dates that are fixed over time.

Some previous studies fix the decay parameters of parsimonious yield curve models, e.g., τ in Equation (1), to estimate them simply by ordinary least squares regression (e.g., Diebold and Li, 2006). However, fixing any parameters is not the practice of central banks (BIS, 2005; Gürkaynak et al., 2007; Nymand-Andersen, 2018). Indeed, not fixing any parameters of parsimonious yield curve models often results in a better fit with the data as the location of humps or troughs in the curvature factor(s) are allowed to vary over time (Koopman et al., 2010; Diebold and Rudebusch, 2013). When not fixing any parameters, the parsimonious models need to be estimated by using an iterative algorithm to numerically solve a non-convex optimization problem with many local minima. This poses difficulties as different initial parameter values for the algorithm may lead to different local minima, i.e., different final estimated parameter values (Gimeno and Nave, 2009; Manousopoulos and Michalopoulos, 2009; Gilli et al., 2010). Still, the different parameter values may result in similar yield curve shapes, and thus, similar goodness of fit (Gürkaynak et al., 2007). However, as these parameters have an economic meaning, any unstable behavior can make them hard to interpret.

Because of this, Article 1 of this thesis emphasizes that, in addition to the goodness of fit, the stability of estimated parameter values of parsimonious yield curve models over time becomes relevant when they are used for economic interpretations. The article is the first to compare the stability of estimated model parameters among different parsimonious yield curve models and different approaches for predefining initial parameter values for the model estimation. In addition, Article 1 examines the robustness of the findings when constraining model parameters that define the location of the yield curve humps and troughs, as well as applying filter criteria for the selection of instruments in the sample.

4. Research articles

This section introduces the three research articles included in this thesis, as well as their scientific contributions.

Article 1: A Comparative Analysis of Parsimonious Yield Curve Models with Focus on the Nelson-Siegel, Svensson and Bliss Versions

This article is co-authored with Florentina Paraschiv at NTNU and Michael Schürle at the University of St.Gallen, in Switzerland. It is published in Computational Economics, ISSN 0927-7099, and is available at https://doi.org/10.1007/s10614-021-10113-w

In this article, we fit the Nelson-Siegel, Bliss, and Svensson parsimonious yield curve models for every trading day between 2000 and 2019 to observed market prices of US Treasury bills, notes, and bonds. Following the practice of central banks, we estimate all model parameters by solving a non-convex optimization problem numerically, which requires predefining initial parameter values for the estimation. We evaluate different modeling and data choices, including model configuration, parameter constraining, data selection, and approaches for selecting initial parameter values for the estimation procedure. Our study reveals that the different choices result in negligible differences in the goodness of fit. However, they result in significant differences in the stability of model parameter estimates over time. An unstable behavior over time of parameter estimates can make them hard to interpret, which is a serious drawback given that central banks use their intrinsic financial interpretation for monetary policy decision making. We recommend using the Nelson-Siegel model while deriving initial values for the parameter estimation procedure from the observed yields themselves. This ensures the most stable parameter estimates. Further, we find that the extra flexibility of the Svensson model is superfluous due to confounding effects. Moreover, to achieve better stability, we recommend neither excluding instruments with maturities above ten years, as often done in previous empirical studies, nor constraining the location of the humps or troughs of the curvature factors of the Svensson model as suggested by De Pooter (2007), Ferstl and Hayden (2010), and Sasongko et al. (2019).

Article 2: Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies

This article is published in Journal of Risk and Financial Management, ISSN 1911-8066, and is available at https://doi.org/10.3390/jrfm14030123. It is co-authored with Akarsh Kainth at NTNU and will also be included in his doctoral thesis.

In this article, we assess International Financial Reporting Standards (IFRS). These were introduced as a replacement for local Generally Accepted Accounting Principles (GAAP) to contribute to more transparency and cross-country comparability through the use of fair values and more disclosure requirements (De George et al., 2016). Particularly, we investigate any differences caused by the alleged benefits of IFRS over local GAAP on the quality of financial reporting. We do this by comparing the performance of bankruptcy prediction models when applied to financial statements derived under IFRS and local GAAP, respectively. For this purpose, we use a comprehensive dataset of 2,290,551 financial statements of privately held companies over the period 2006-2018 from Sweden and Norway, based on IFRS and Norwegian GAAP, respectively. Our findings suggest that IFRS result in better bankruptcy prediction models compared to Norwegian GAAP. This indicates that the transparency and cross-country comparability promoted by IFRS prevent the management of companies facing insolvency from hiding the company's true situation by engaging in window dressing of the accounts or creative accounting practices. As a result, investors, creditors, financial regulators, and other stakeholders can expect a more accurate picture of companies based on their financial statements when these are derived under IFRS. Thus, our findings provide empirical evidence of the benefits for financial regulators of aligning national accounting standards towards IFRS.

Article 3: Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods

This article is co-authored with Florentina Paraschiv at NTNU and Markus Schmid at the University of St.Gallen, in Switzerland. The article is complete, and we are currently getting good suggestions for improvements from recognized experts in the field. All authors agree to submit this article to the Review of Finance (RoF), ISSN 1572-3097, by the summer of 2021. The article has also been presented at international conferences with peer reviews. These include the 4th Shanghai-Edinburgh Fintech Conference and the 6th Fintech International Conference, the 2021 Winter Research Conference on Machine Learning and Business at the University of Miami, and the 2020 FIBE Conference in Bergen.

In the context of this thesis, we make two main contributions in this article. First, we show that variables for bankruptcy prediction models chosen by alternative feature selection methods are superior to variables chosen ad-hoc based on subjective criteria. This is shown in applications to bankruptcy predict of privately held SMEs using a comprehensive dataset of financial statements from such companies in Norway over the period 2006-2014. For each financial statement, we extract a total of 155 accounting-based input variables derived from prior literature. We test several feature selection methods for choosing among these and find that the best model performance is achieved when using the variables chosen by an embedded LASSO feature selection method. This finding is robust over different time periods and across the two employed estimation techniques – an artificial neural network and logistic regressions.

Second, we contribute to the literature in this article by proposing to evaluate bankruptcy prediction models based on the real economic implications of their predictions for banks in term of bank profitability. This is done in a simulation of a competitive credit market that employs real-world data from the whole Norwegian SME loan market. Our approach is an improvement over the common practice of using solely statistical metrics to assess bankruptcy prediction models as these incorrectly assume equal costs of different types of prediction errors. The superiority of alternative feature selection methods over ad-hoc chosen variables is confirmed in our simulation. In addition, our simulation confirms that the preferred feature selection method is the LASSO method.

We also contribute in this article to the bankruptcy prediction literature by giving insights about privately held SMEs, as opposed to most prior studies which are limited to larger and listed companies (e.g., Campbell et al., 2008; Tian et al., 2015; Liang et al., 2016). Further, since the analyses in our study are done using Norwegian data, we offer an improved bankruptcy prediction model for SMEs compared to the benchmark employed by the Financial Supervisory Authority of Norway.[7]

[7] See Eklund et al. (2001) and Bernhardsen and Larsen (2007) for details about the model used by the Financial Supervisory Authority of Norway. Its variables are used as benchmark in our study. They are also considered by our feature selection methods.

5. Other contributions

As a PhD student at NTNU Business School, I have also been involved in other academic activities. These include:

- Conducting the study presented in the research article Pelja and Wahlstrøm (2021).[8] The current version of this article is available at: http://pelja2021.ranik.no. It is submitted to the Norwegian scientific journal Magma, ISSN 1500-0788, and has undergone a peer review. The main conclusion from both the reviewer and the editorial staff after having undergone peer review is that publication is recommended following a revision. We have submitted an updated version of this article, where all comments from the peer review are answered, and we are now awaiting a decision from the editors.
- Participating in EU COST Action "Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry" (FinAI) CA19130 funded by the Horizon 2020 Framework Programme of the European Union.
- Being peer reviewer for the following journals:
  - Computational Economics, ISSN 0927-7099
  - Computational Management Science, ISSN 1619-697X
  - Financial Markets and Portfolio Management, ISSN 1934-4554
- Presenting my research in an internal seminar for practitioners at the Central Bank of Norway, as well as in several internal and external conferences and seminars, e.g., the 4$^{\text{th}}$ Shanghai-Edinburgh Fintech Conference and the 6$^{\text{th}}$ Fintech International Conference, the 2021 Winter Research Conference on Machine Learning and Business at the University of Miami, the 2020 FIBE Conference in Bergen, and the 2$^{\text{nd}}$ Yushan Conference.
- Performing many tasks related to teaching and supervising students, e.g.,:

[8] See the last paragraph in Section 2.2 of this thesis for a brief description of the content of this research article.

- Being course coordinator for BBAN4001 “Data Science” (second degree level) for the autumn semester 2020, which included lecturing, exam and assignments design and grading.
- Creating and grading the re-sit examination May 2020 for BMRR4015 “Advanced Data- and Transaction Analysis” (second degree level).
- Conducting lectures in IF440 “Capital Markets and Uncertainty” (second degree level) and MET1002 “Statistics for Business” (Foundation courses, level I).
- Co-supervising master thesis groups at NTNU Business School and the Department of Mathematical Sciences at NTNU.

References

Agarwal, V., Taffler, R., 2008. Comparing the performance of market-based and accounting-based bankruptcy prediction models. Journal of Banking & Finance 32, 1541–1551.

Agarwal, V., Taffler, R.J., 2007. Twenty-five years of the Taffler z-score model: Does it really have predictive ability? Accounting and Business Research 37, 285–300.

Altman, E.I., 1968. Financial ratios, discriminant analysis and prediction of corporate bankruptcy. Journal of Finance 23, 589–609.

Altman, E.I., Haldeman, R.G., Narayanan, P., 1977. ZETA analysis, a new model to identify bankruptcy risk of corporations. Journal of Banking & Finance 1, 29–54.

Appiah, K.O., Chizema, A., Arthur, J., 2015. Predicting corporate failure: a systematic literature review of methodological issues. International Journal of Law and Management 57, 461–485.

Balcaen, S., Ooghe, H., 2006. 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. The British Accounting Review 38, 63–93.

Bauer, J., Agarwal, V., 2014. Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. Journal of Banking & Finance 40, 432–442.

Beaver, W.H., 1966. Financial ratios as predictors of failure. Journal of Accounting Research 4, 71–111.

Beaver, W.H., 1968. Market Prices, Financial Ratios, and the Prediction of Failure. Journal of Accounting Research 6, 179–192.

Beaver, W.H., Correia, M., McNichols, M.F., 2012. Do differences in financial reporting attributes impair the predictive ability of financial ratios for bankruptcy? Rev Account Stud 17, 969–1010.

Bellovary, J., Giacomino, D., Akers, M., 2007. A Review of Bankruptcy Prediction Studies: 1930 to Present. Journal of Financial Education 33, 1–42.

Bernhardsen, E., Larsen, K., 2007. Modelling credit risk in the enterprise sector – further development of the SEBRA model. Economic Bulletin 2/2007.

BIS, 2005. Zero-coupon yield curves: technical documentation. BIS Papers No. 25. Bank of International Settlements.

BIS, 2017. Basel III: Finalising post-crisis reforms. Report. Bank of International Settlements.

Björk, T., Christensen, B.J., 1999. Interest rate dynamics and consistent forward rate curves. Mathematical Finance 9, 323–348.

Blöchlinger, A., Leippold, M., 2006. Economic benefit of powerful credit scoring. Journal of Banking & Finance 30, 851–873.

Blöchlinger, A., Leippold, M., 2018. Are ratings the worst form of credit assessment except for all the others? Journal of Financial and Quantitative Analysis 53, 299–334.

Bliss, R.R., 1997a. Movements in the term structure of interest rates. Economic Review - Federal Reserve Bank of Atlanta 82, 16–33.

Bliss, R.R., 1997b. Testing term structure estimation methods. Advances in Futures Options Research 9, 197–232.

Bretscher, L., Schmid, L., Vedolin, A., 2018. Interest Rate Risk Management in Uncertain Times. The Review of Financial Studies 31, 3019–3060.

Brooks, C., 2019. Introductory econometrics for finance. Fourth edition. ed., Cambridge University Press, Cambridge.

Brooks, C., Hoepner, A.G.F., McMillan, D., Vivian, A., Simen, C.W., 2019. Financial data science: the birth of a new financial research paradigm complementing econometrics? The European Journal of Finance 25, 1627–1636.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In Search of Distress Risk. The Journal of Finance 63, 2899–2939.

Chandrashekar, G., Sahin, F., 2014. A survey on feature selection methods. Computers & Electrical Engineering 40, 16–28.

Chava, S., Jarrow, R.A., 2004. Bankruptcy prediction with industry effects. Review of Finance 8, 537–569.

Cielen, A., Peeters, L., Vanhoof, K., 2004. Bankruptcy prediction using a data envelopment analysis. European Journal of Operational Research 154, 526–532.

De Bock, K.W., Coussement, K., Lessmann, S., 2020. Cost-sensitive business failure prediction when misclassification costs are uncertain: A heterogeneous ensemble selection approach. European Journal of Operational Research 285, 612–630.

De George, E.T., Li, X., Shivakumar, L., 2016. A review of the IFRS adoption literature. Rev Account Stud 21, 898–1004.

De Pooter, M., 2007. Examining the Nelson-Siegel Class of Term Structure Models: In-Sample Fit versus Out-of-Sample Forecasting Performance. Working Paper, Tinbergen Institute .

De Prado, M.L., 2018. Advances in financial machine learning. John Wiley & Sons.

Deakin, E.B., 1976. Distributions of financial accounting ratios: Some empirical evidence. The Accounting Review 51, 90–96.

Dhar, V., 2013. Data science and prediction. Commun. ACM 56, 64–73.

Diebold, F.X., Li, C., 2006. Forecasting the term structure of government bond yields. Journal of Econometrics 130, 337–364.

Diebold, F.X., Piazzesi, M., Rudebusch, G.D., 2005. Modeling bond yields in finance and macroeconomics. The American Economic Review 95, 415–420.

Diebold, F.X., Rudebusch, G.D., 2013. Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach. The Econometric and Tinbergen Institutes Lectures, Princeton University Press, Princeton.

Diebold, F.X., Rudebusch, G.D., Boragan Aruoba, S., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. Journal of Econometrics 131, 309–338.

Duffee, G., 2013. Chapter 7 - forecasting interest rates. Handbook of Economic Forecasting 2, 385–426.

Einav, L., Levin, J., 2014. Economics in the age of big data. Science 346.

Eisenbeis, R.A., 1977. Pitfalls in the application of discriminant analysis in business, finance, and economics. Journal of Finance 32, 875–900.

Eklund, T., Larsen, K., Bernhardsen, E., 2001. Model for analysing credit risk in the enterprise sector. Economic Bulletin 3/2001.

Fama, E.F., Bliss, R.R., 1987. The information in long-maturity forward rates. American Economic Review 77, 680–692.

Ferstl, R., Hayden, J., 2010. Zero-coupon yield curve estimation with the package termstrc. Journal of Statistical Software 36, 1–34.

Fisher, M., Nychka, D.W., Zervos, D., 1995. Fitting the Term Structure of Interest Rates with Smoothing Splines. Federal Reserve System Working Paper No. 95-1 .

Fitzgerald, P., 2009. New bankruptcy-risk model looks to topple Altman's Z-score. The Wall Street Journal .

FitzPatrick, P.J., 1932. A comparison of the ratios of successful industrial enterprises with those of failed companies. The Certified Public Accountant , 598–605; 656–662; 727–731.

Frydman, H., Altman, E.I., Kao, D.L., 1985. Introducing recursive partitioning for financial classification: The case of financial distress. The Journal of Finance 40, 269–291.

Geng, R., Bose, I., Chen, X., 2015. Prediction of financial distress: An empirical study of listed Chinese companies using data mining. European Journal of Operational Research 241, 236–247.

Gepp, A., Kumar, K., Bhattacharya, S., 2010. Business failure prediction using decision trees. Journal of Forecasting 29, 536–555.

Gilli, M., Große, S., Schumann, E., 2010. Calibrating the Nelson–Siegel–Svensson model. Report. Computational Optimization Methods in Statistics, Econometrics and Finance.

Gimeno, R., Nave, J.M., 2009. A genetic algorithm estimation of the term structure of interest rates. Computational Statistics & Data Analysis 53, 2236–2250.

Greene, W.H., 2012. Econometric analysis. 7th ed., international ed. ed., Pearson, Boston.

Gürkaynak, R.S., Sack, B., Wright, J.H., 2007. The U.S. treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304.

Gupta, J., Gregoriou, A., Ebrahimi, T., 2018. Empirical comparison of hazard models in predicting SMEs failure. Quantitative Finance 18, 437–466.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. Journal of Machine Learning Research 3, 1157–1182.

Harvey, C.R., 2017. Presidential Address: The Scientific Outlook in Financial Economics. The Journal of Finance 72, 1399–1440.

Harvey, C.R., Liu, Y., Zhu, H., 2016. … and the Cross-Section of Expected Returns. Rev Financ Stud 29, 5–68.

Hoepner, A.G.F., McMillan, D., Vivian, A., Simen, C.W., 2021. Significance, relevance and explainability in the machine learning age: an econometrics and financial data science perspective. The European Journal of Finance 27, 1–7.

Härdle, W., Lee, Y.J., Schäfer, D., Yeh, Y.R., 2009. Variable selection and oversampling in the use of smooth support vector machines for predicting the default risk of companies. Journal of Forecasting 28, 512–534.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning: with Applications in R. 1st ed. 2013, corr. 7th printing 2017 edition ed., Springer, New York.

James, J., Webber, N., 2000. Interest rate modelling. Wiley series in financial engineering, Wiley, Chichester.

du Jardin, P., 2015. Bankruptcy prediction using terminal failure processes. European Journal of Operational Research 242, 286–303.

du Jardin, P., Séverin, E., 2012. Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. European Journal of Operational Research 221, 378–396.

John, G.H., Kohavi, R., Pfleger, K., 1994. Irrelevant features and the subset selection problem. Machine Learning Proceedings 1994 , 121–129.

Jones, S., Johnstone, D., Wilson, R., 2015. An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. Journal of Banking & Finance 56, 72–85.

Jones, S., Johnstone, D., Wilson, R., 2017. Predicting corporate bankruptcy: An evaluation of alternative statistical frameworks. Journal of Business Finance & Accounting 44, 3–34.

Joy, O.M., Tollefson, J.O., 1975. On the financial applications of discriminant analysis. Journal of Financial and Quantitative Analysis 10, 723–739.

Kainth, A., Wahlstrøm, R.R., 2021. Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies. Journal of Risk and Financial Management 14, 123.

Khraisha, T., 2020. A Holistic Approach to Financial Data Science: Data, Technology, and Analytics. The Journal of Financial Data Science 2, 64–84.

Kim, J.H., Ji, P.I., 2015. Significance testing in empirical finance: A critical review and assessment. Journal of Empirical Finance 34, 1–14.

Kohavi, R., John, G.H., 1997. Wrappers for feature subset selection. Artificial Intelligence 97, 273–324.

Koopman, S.J., Mallee, M.I.P., Wel, M.V.d., 2010. Analyzing the Term Structure of Interest Rates Using the Dynamic Nelson–Siegel Model With Time-Varying Parameters. Journal of Business & Economic Statistics 28, 329–343.

Kumar, P.R., Ravi, V., 2007. Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. European Journal of Operational Research 180, 1–28.

Lengwiler, Y., Lenz, C., 2010. Intelligible factors for the yield curve. Journal of Econometrics 157, 481–491.

Liang, D., Lu, C.C., Tsai, C.F., Shih, G.A., 2016. Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. European Journal of Operational Research 252, 561–572.

Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. Journal of fixed income 1, 54–61.

Manousopoulos, P., Michalopoulos, M., 2009. Comparison of non-linear optimization algorithms for yield curve estimation. European Journal of Operational Research 192, 594–602.

Marais, M.L., Patell, J.M., Wolfson, M.A., 1984. The experimental design of classification models: an application of recursive partitioning and bootstrapping to commercial bank loan classifications. Journal of Accounting Research 22, 87–114.

Martin, D., 1977. Early warning of bank failure: A logit regression approach. Journal of Banking & Finance 1, 249–276.

McCulloch, J.H., 1971. Measuring the term structure of interest rates. The Journal of Business 44, 19–31.

McCulloch, J.H., 1975a. Money and Inflation: A Monetarist Approach. Academic Press.

McCulloch, J.H., 1975b. The tax-adjusted yield curve. Journal of Finance 30, 811–830.

Mcleay, S., Omar, A., 2000. The sensitivity of prediction models to the non-normality of bounded and unbounded financial ratios. The British Accounting Review 32, 213–230.

Merwin, C.L., 1942. Financing small corporations in five manufacturing industries, 1926-1936. National Bureau of Economic Research, New York.

Min, J.H., Lee, Y.C., 2005. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Systems with Applications 28, 603–614.

Mitchell, T., 1997. Machine Learning. McGraw Hill.

Mullainathan, S., Spiess, J., 2017. Machine Learning: An Applied Econometric Approach. Journal of Economic Perspectives 31, 87–106.

Nelson, C.R., Siegel, A.F., 1987. Parsimonious modeling of yield curves. The Journal of Business 60, 473–489.

Næss, A.B., Wahlstrøm, R.R., Helland, F.F., Kjærland, F., 2017. Konkursprediksjon for norske selskaper - en sammenligning av regresjonsmodeller og maskinlæringsteknikker. Bred og spiss! NTNU Handelshøyskolen 50 år: En Vitenskapelig Jubileumsantologi , 313–330.

Nymand-Andersen, P., 2018. Yield curve modelling and a conceptual framework for estimating yield curves: evidence from the European Central Bank's yield curves. Report 9289933569. ECB Statistics Paper.

Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research 18, 109–131.

Paraschiv, F., Schmid, M., Wahlstrøm, R.R., 2021. Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods. Working Paper, Norwegian University of Science and Technology and University of St. Gallen, to be submitted to the Review of Finance .

Park, C.S., Han, I., 2002. A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. Expert Systems With Applications 23, 255–264.

Pelja, I., Wahlstrøm, R.R., 2021. Hvordan påvirker bedriftens størrelse predikering av konkurs? Working Paper, NHH Norwegian School of Economics and Norwegian University of Science and Technology, submitted to Magma .

Reisz, A.S., Perlich, C., 2007. A market-based framework for bankruptcy prediction. Journal of Financial Stability 3, 85–131.

Rosendale, W.M., 1908. Credit Department Methods. Bankers' Magazine .

Sasongko, A., Utama, C.A., Wibowo, B., Husodo, Z.A., 2019. Modifying hybrid optimisation algorithms to construct spot term structure of interest rates and proposing a standardised assessment. Computational Economics 54, 957–1003.

Shea, G.S., 1984. Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximations. Journal of Financial and Quantitative Analysis 19, 253–269.

Shin, K.S., Lee, T.S., Kim, H.j., 2005. An application of support vector machines in bankruptcy prediction model. Expert Systems with Applications 28, 127–135.

Shumway, T., 2001. Forecasting bankruptcy more accurately: A simple hazard model. The Journal of Business 74, 101–124.

Simonian, J., Fabozzi, F.J., 2019. Triumph of the Empiricists: The Birth of Financial Data Science. The Journal of Financial Data Science 1, 10–13.

Smith, R.F., Winakor, A.H., 1930. A Test Analysis of Unsuccessful Industrial Companies. Bureau of Business Research, Bulletin 31. Urbana, IL: University of Illinois Press .

Smith, R.F., Winakor, A.H., 1935. Changes in Financial Structure of Unsuccessful Industrial Corporations. Bureau of Business Research, Bulletin 51. Urbana, IL: University of Illinois Press .

Stein, R.M., 2005. The relationship between default prediction and lending profits: Integrating ROC analysis and loan pricing. Journal of Banking & Finance 29, 1213–1236.

Svensson, L.E.O., 1994. Estimating and interpreting forward interest rates: Sweden 1992 - 1994. NBER Working Paper Series , 4871.

Svensson, L.E.O., 1995. Estimating forward interest rates with the extended Nelson–Siegel method. Sveriges Riksbank Quarterly Review .

Taffler, R.J., 1984. Empirical models for the monitoring of UK corporations. Journal of Banking & Finance 8, 199–227.

Tam, K.Y., Kiang, M.Y., 1992. Managerial applications of neural networks: The case of bank failure predictions. Management Science 38, 926–947.

Taylor, J.B., 1993. Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy 39, 195–214.

Tian, S., Yu, Y., 2017. Financial ratios and bankruptcy predictions: An international evidence. International Review of Economics & Finance 51, 510–526.

Tian, S., Yu, Y., Guo, H., 2015. Variable selection and corporate bankruptcy forecasts. Journal of Banking & Finance 52, 89–100.

Tibshirani, R., 1996. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288.

Toplensky, R., 2020. Banks are only as sound as their models. The Wall Street Journal .

Trafimow, D., Marks, M., 2015. Editorial. Basic and Applied Social Psychology 37, 1–2.

Tsai, C.F., Hsu, Y.F., 2013. A meta-learning framework for bankruptcy prediction. Journal of Forecasting 32, 167–179.

Vapnik, V.N., 1998. Statistical Learning Theory. Adaptive and learning systems for signal processing, communications, and control, Wiley, New York.

Varian, H.R., 2014a. Beyond Big Data. Bus Econ 49, 27–31.

Varian, H.R., 2014b. Big Data: New Tricks for Econometrics. The Journal of Economic Perspectives; Nashville 28, 3–28.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177–188.

Vasicek, O.A., Fong, H.G., 1982. Term structure modeling using exponential splines. Journal of Finance 37, 339–348.

Wahlstrøm, R.R., Paraschiv, F., Schürle, M., 2021. A Comparative Analysis of Parsimonious Yield Curve Models with Focus on the Nelson-Siegel, Svensson and Bliss Versions. Computational Economics .

Wasserstein, R.L., Lazar, N.A., 2016. The ASA Statement on p-Values: Context, Process, and Purpose. The American Statistician 70, 129–133.

Wasserstein, R.L., Schirm, A.L., Lazar, N.A., 2019. Moving to a world beyond "p< 0.05". null 73, 1–19.

Zhang, G., Y. Hu, M., Eddy Patuwo, B., C. Indro, D., 1999. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. European Journal of Operational Research 116, 16–32.

Zmijewski, M.E., 1984. Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research 22, 59–82.

Article 1

Title:
A Comparative Analysis of Parsimonious Yield Curve Models with Focus on the Nelson-Siegel, Svensson and Bliss Versions

Authors:
Ranik Raaen Wahlstrøm
Florentina Paraschiv
Michael Schürle

Reference:
Wahlstrøm, R.R., Paraschiv, F., Schürle, M., 2021. A Comparative Analysis of Parsimonious Yield Curve Models with Focus on the Nelson-Siegel, Svensson and Bliss Versions. Computational Economics. doi:10.1007/s10614-021-10113-w.

Computational Economics
https://doi.org/10.1007/s10614-021-10113-w

A Comparative Analysis of Parsimonious Yield Curve Models with Focus on the Nelson-Siegel, Svensson and Bliss Versions

Ranik Raaen Wahlstrøm[1] · Florentina Paraschiv[1,2] · Michael Schürle[2]

Accepted: 13 March 2021

Abstract
We shed light on computational challenges when fitting the Nelson-Siegel, Bliss and Svensson parsimonious yield curve models to observed US Treasury securities with maturities up to 30 years. As model parameters have a specific financial meaning, the stability of their estimated values over time becomes relevant when their dynamic behavior is interpreted in risk-return models. Our study is the first in the literature that compares the stability of estimated model parameters among different parsimonious models and for different approaches for predefining initial parameter values. We find that the Nelson-Siegel parameter estimates are more stable and conserve their intrinsic economical interpretation. Results reveal in addition the patterns of confounding effects in the Svensson model. To obtain the most stable and intuitive parameter estimates over time, we recommend the use of the Nelson-Siegel model by taking initial parameter values derived from the observed yields. The implications of excluding Treasury bills, constraining parameters and reducing clusters across time to maturity are also investigated.

Keywords Parsimonious yield curve models · Term structure · Monetary policy decisions · Non-linear least squares · Initial values

Mathematics Subject Classification E430 · G120

We are grateful to four anonymous referees, Fred Espen Benth, Martin Hellmich, and Paul Söderlind for very useful feedbacks, as well as our colleagues from Norges Bank for very useful comments during the participation at their scientific seminar. Florentina Paraschiv thanks the funding from *Adolf Øiens Donasjonsfond* Energizing New Computational Frontiers and the Isaac Newton Institute for Mathematical Sciences for its hospitality during the programme "The mathematics of energy systems" which was supported by EPSRC Grant Number EP/R014604/1. This work acknowledges research support by COST Action "Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry" (FinAI) CA19130.

Extended author information available on the last page of the article

Published online: 15 April 2021

1 Introduction

The term structure of interest rates describes the relationship between yields and time to maturity of fixed-income instruments. Another name, which is often connected with the graphical representation of this relation, is yield curve. The discount function, which is considered the most basic building block of finance, can be inferred directly from it (Gürkaynak et al., 2007). Both financial market participants, policymakers and academics are concerned with modeling the yield curve (Duffee, 2013). From the perspective of a central bank, the yield curve can be used for drawing correct inferences regarding the appropriateness of its monetary policy stance (BIS, 2005; Cœuré, 2017). Many central banks use parsimonious data-driven models for this purpose.

In this paper, we empirically investigate implications of relevant modelling choices for central banks when using such models. We investigate the implications on both the goodness of fit and the stability of estimated model parameter values over time. The latter becomes relevant as parameters of parsimonious models used by (central) banks have a specific financial meaning, e.g., when their dynamic behavior is interpreted in bond risk-return models (Gimeno & Nave, 2009). We perform our analysis using data of US Treasury bills, notes and bonds for all 4996 trading days between 2000 and 2019.

Some previous studies estimate model parameters in monthly steps using synthetic zero bond yields for constant maturities up to 10 years. These must be derived in a preliminary step from prices of coupon-bearing bonds by other approaches. In this case, after fixing certain parameters the model under consideration can be estimated simply by ordinary least squares (OLS) regression. By further assuming stochastic processes for the non-fixed parameters, some authors then derive dynamic versions of parsimonious models. We instead follow the common practice of central banks of estimating all parameters of the original static models directly to the daily observed market prices of the above mentioned Treasury instruments with maturities up to 30 years. As no parameters are fixed, the full set of model parameters must be obtained by solving a non-convex optimization problem by means of a non-linear least squares method, which requires the specification of a set of initial values. As Gimeno & Nave (2009) point out, the latter is crucial for the stability of estimated parameters. Using daily data gives us more observations to fit the models, lowers the influence of any month-end effects and is consistent with the practice of central banks (BIS, 2005; Gürkaynak et al., 2007; Nymand-Andersen, 2018). Our study complements the existing literature on the following points: We offer a comprehensive picture of the robustness of parsimonious models with respect to different approaches for selecting initial values for the fitting procedure, constraints on certain parameters in relation to confounding effects, as well as filter criteria for the selection of instruments considered in the estimation.

Our results support previous evidence suggesting that the magnitudes of the first two factors of the parsimonious models represent the level of the yield curve. However, we show that one of the two curvature factors of the parsimonious

Svensson model is superfluous due to confounding effects. Furthermore, our tests of yield curve models as well as different approaches for the selection of initial parameter values for the non-linear fitting procedure imply that central banks, when using the yield curve for monetary policy decisions, should prefer the less flexible Nelson-Siegel model, as well as initial values that are derived from observed yields. These suggestions lead to the most stable and intuitive parameter estimates over time, which makes it easier to give them a financial interpretation, without compromising the goodness of fit. Finally, we test the implications on our findings when preimposing restrictions on the distance between the locations of humps or troughs in the yield curve (like in De Pooter, 2007; Ferstl & Hayden, 2010), excluding Treasury bills (like in Gürkaynak et al., 2007) and controlling for clustering of instruments across time to maturity. Overall, we observe persisting confounding effects in the curvature factors of the Svensson model and an insignificant effect on the goodness of fit. In the cases of controlling for clustering of instruments across time to maturity or preimposing restrictions on the distance between the locations of the humps or troughs in the yield curve, we observe a significant increase in the variation in parameter values. In particular, we observe more variation in the level factor of the yield curve when instruments with more than 10 years are excluded, meaning that the inclusion of longer maturities leads to a better approximation for the long end of the yield curve.

The rest of this paper is organized as follows. Section 2 introduces formally the relevant parsimonious yield curve models that are investigated in this study, and reviews earlier related empirical work. Section 3 explains the data and the fitting procedure applied here, including the different approaches for selecting initial values. Results are presented and interpreted in Sect. 4. Finally, conclusions are given in Sect. 5.

2 Theoretical Background

Let us first introduce important definitions related to the construction of discount factors, spot rates and yields to maturity. Suppose that $\mathbf{C} = \{c_{(i,j)}\}_{i=1,\ldots,N, j=1,\ldots,L}$ is a matrix of cash flows from all coupon payments and the repayment of the face value from government securities i at times j, and that $\mathbf{p} = \{p_i\}_{i=1,\ldots,N}$ is the corresponding price vector. Then it is possible to find a vector $\delta = \{\delta_j\}_{j=1,\ldots,L}$ of discount factors from the following equation (James and Webber 2000):

$$\mathbf{p} = \mathbf{C}\delta + \epsilon \tag{1}$$

where $\epsilon = \{\epsilon_i\}_{i=1,\ldots,N}$ is a vector of errors. Finding δ directly by solving (1) using OLS regression does not work very well, because $\mathbf{C}$ has too many columns compared to the length of $\mathbf{p}$, and too many zeros since the cash flows of government instruments rarely occur on the same date (James & Webber 2000). A better way is to define the discount factor as a function $\delta(m)$ of time to maturity $m \in [0, \infty)$, and then let $\delta = (\delta(m_1), \ldots, \delta(m_L))'$ be the vector of discount factors for all cash flow dates $\{m_j\}_{j=1,\ldots,L}$. $\delta(m)$ is an example of a term structure, which links time to maturity and discount factors.

The term structure may also be represented by the *spot rate* $s(m)$ (Müller, 2002; BIS, 2005), which is the annualized percentage return for an instrument which pays no coupons.[1] It relates to the discount factor by

$$s(m) = -\frac{1}{m}\log(\delta(m)). \tag{2}$$

The *yield to maturity* y_i is the internal rate of return that sets the present value of a instrument's cash flows (coupon payments and repayment of face value) equal to its market price p_i:

$$p_i = \sum_{j=1}^{L} c_{ij}\, e^{-y_i \cdot m_j} \tag{3}$$

2.1 Models for Estimating the Term Structure

There exist many types of models for estimating the term structure. Some models are concerned with using the spread between long- and short-term interest rates to forecast inflation and real activity of a country or region (Fama & Bliss, 1987; Mishkin, 1990b, a; Shiller & Campbell, 1991; Estrella et al., 2003; Bernanke et al., 2005; Ang et al., 2006; Estrella & Trubin, 2006; Rudebusch & Williams, 2009). Such models require as input yields of specific maturities. However, since usually we do not observe the yields of arbitrary maturities directly, other models are needed that derive them from the prices of traded instruments. Often these models describe the term structure by a continuous function, whose parameters are found by fitting the resulting yield curve to observed market data. Furthermore, there are dynamic models which focus mainly on pricing fixed-income derivatives, and less on forecasting or interpolating the yield curve. Such models include equilibrium models (Vasicek, 1977; Cox et al., 1985; Duffie & Kan, 1996; Bianchi & Cleur, 1996; De Rossi, 2010), no-arbitrage models (Ho & Lee, 1986; Hull & White, 1990; Heath et al., 1992; Eydeland, 1996) and models stating that the interest rates depend on macroeconomic variables (Ang & Piazzesi, 2003; Moench, 2008; Rudebusch & Wu, 2008; Audrino, 2012). Other models rely on machine learning techniques that are capable of incorporating non-linear relationships between economic variables to predict interest rates. These techniques include support vector machines (Gogas et al., 2015), fuzzy logic and genetic algorithms (Ju et al., 1997), neural networks (Kim & Noh, 1997; Oh & Han, 2000; Hong & Han, 2002; Bianchi et al. 2020b, a) and case-based reasoning (Kim & Noh 1997). However, the financial literature has been slow to adapt such methods (Bianchi et al. 2020b), possibly because it is not necessary straightforward to understand their abundant non-linear patterns (Diaz et al., 2016) and it is claimed that they are not suitable for parameter inference (see Mullainathan & Spiess, 2017). Finally, data-driven yield curve models fit

[1] From the spot rate, which is based on the price of a transaction that takes place immediately, one may also derive forward rates which is the settlement price of a transaction at a predetermined date in the future. See BIS (2005) for details.

mathematical functions, including spline-based and parsimonious functions, to discount factors, spot rates, forward rates or par yields (Müller, 2002; BIS, 2005).

Many central banks use parsimonious data-driven models for the interpolation of yield curves and the assessment of monetary policy measures (BIS, 2005). Indeed, such models have an economic interpretation and provide a good fit of the resulting term structures to observed yields or prices, respectively, of fixed income instruments. This also makes them ideal as basis for measuring risk in fixed income portfolios (Caldeira et al., 2015). The parsimonious Nelson-Siegel model of Nelson & Siegel (1987) and its extensions by Svensson (1994, 1995) and Bliss (1997) use a single exponential function over the entire maturity range. The popularity of these models stems from the fact that – unlike for example spline models – they provide a parsimonious approximation of the yield curve and use only a small number of parameters, yet are flexible enough to capture a range of monotonic, humped and S-type shapes observed in yield data (De Pooter, 2007).

2.2 Specification of Parsimonious Yield Curve Models

The Nelson-Siegel model was proposed by Nelson & Siegel (1987) to interpolate the yield curve (in terms of spot rates) by the following function:

$$s(m) = \beta_0 + \beta_1 \frac{1 - e^{\frac{-m}{\tau_1}}}{\frac{m}{\tau_1}} + \beta_2 \left(\frac{1 - e^{\frac{-m}{\tau_1}}}{\frac{m}{\tau_1}} - e^{\frac{-m}{\tau_1}} \right) \tag{4}$$

where $s(m)$ is the spot rate at any given time to maturity m, and β_0, β_1, β_2 and τ_1 are parameters whose specific values result from the fitting procedure. The first, second and third factors of Equation (4) may be interpreted as the level, slope and curvature factors, respectively, as they control the long, short and medium segments of the yield curve (Nelson & Siegel, 1987; Diebold & Li, 2006). This is due to the characteristics of the factor loadings for different times to maturity, which we illustrate in Fig. 1.

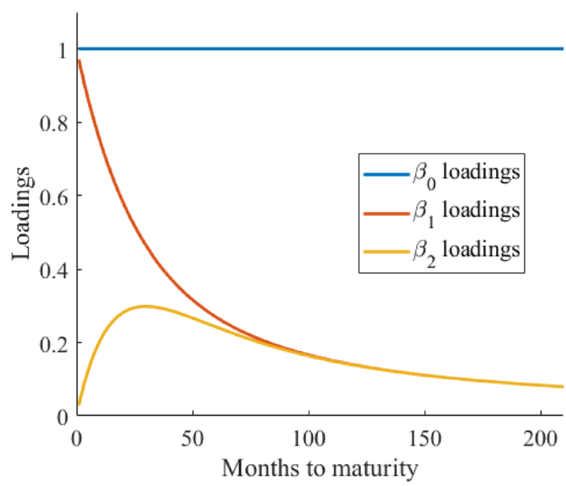


Fig. 1 Illustration of the factor loadings over time to maturity in months of the Nelson-Siegel model as given in Eq. (4)

The level factor β_0 represents the limit value of the spot rate when the maturity m goes to infinity and must be strictly positive. The assumption that its loading is constantly one reflects a market where participants have no information to distinguish expectations for different times to maturity far into the future (Dahlquist & Svensson, 1996). The loading of the slope factor β_1 starts at one when $m = 0$ and monotonically decreases towards zero as time to maturity increases. The loading of the curvature factor β_2 starts at zero, its absolute value attains a certain maximum as time to maturity increases, and then decays to zero with further increasing time to maturity. Its sign controls if a hump-shape ($\beta_2 > 0$) or a trough-shape ($\beta_2 < 0$) is generated. The decay parameter $\tau_1 > 0$ determines the exponential decay rate (in years to maturity) of the slope and curvature factors. In addition, its value controls the location of the hump or trough, respectively, associated with the curvature factor. The sum $\beta_0 + \beta_1$ determines the level of the short end, i.e., the starting value of the yield curve for $m = 0$.

Diebold et al., (2005) propsed a reduced Nelson–Siegel model without the curvature factor. They argued the level and slope factors explain almost all variation, but acknowledged that for shaping the entire yield curve two factors are most likely not enough. This was confirmed by De Pooter (2007), who found that this reduced two-factor Nelson-Siegel model performed poorly in yield curve fitting because of the lack of the curvature factor.

As the slope and curvature factors of the Nelson–Siegel model rapidly approach zero (see Diebold & Li, 2006), only the level factor is left to fit the yield curve at longer maturities (Diebold & Rudebusch, 2013). To address this, Svensson (1994, 1995) extended the Nelson-Siegel model to a four-factor model by adding a second curvature factor, which allows to reflect a second hump or trough in the yield curve and increases the flexibility to fit it to observed market data:

$$s(m) = \beta_0 + \beta_1 \frac{1 - e^{\frac{-m}{\tau_1}}}{\frac{m}{\tau_1}} + \beta_2 \left(\frac{1 - e^{\frac{-m}{\tau_1}}}{\frac{m}{\tau_1}} - e^{\frac{-m}{\tau_1}} \right) + \beta_3 \left(\frac{1 - e^{\frac{-m}{\tau_2}}}{\frac{m}{\tau_2}} - e^{\frac{-m}{\tau_2}} \right) \tag{5}$$

where β_3 determines the magnitude of the second curvature factor, while τ_2 determines the location of the second hump (if $\beta_3 > 0$) or trough (if $\beta_3 < 0$). Gürkaynak et al. (2007) argue that the Svensson model should be preferred to the Nelson-Siegel model since the yield curve slopes down at the very long end, and thus the second curvature factor of the Svensson model is needed to model a second hump at longer maturities. Using government bonds from the Euro zone, Nymand-Andersen (2018) also found that the Svensson model performs slightly better than the Nelson-Siegel model with respect to flexibility and goodness of fit. He also compared both models with spline-based approaches and concluded that the latter are sensitive to the applied optimization algorithm, the fixing of smoothing parameters, the selection of penalty functions and the location of knot points.

Björk & Christensen (1999) extended the original Nelson–Siegel model to a four-factor model by adding a second slope factor, as opposed to the Svensson model which adds a second curvature factor. Furthermore, they constructed a five factor model by extending the latter by a fifth factor, which increases linearly with time to maturity. Diebold et al. (2006) found that these two extensions provide only negligible improvement in the model fit, suggesting that fewer factors are sufficient. De Pooter (2007) argued that the fifth factor is problematic since it implies a linear increase in yields with maturity.

While in (4) the loadings of the slope and the curvature factor are governed by the same decay parameter τ_1, Nelson & Siegel (1987) discussed already in their original paper a generalization where this restriction is relaxed by introduction of an individual decay parameter $\tau_2 > 0$ in the last term:

$$s(m) = \beta_0 + \beta_1 \frac{1 - e^{\frac{-m}{\tau_1}}}{\frac{m}{\tau_1}} + \beta_2 \left(\frac{1 - e^{\frac{-m}{\tau_2}}}{\frac{m}{\tau_2}} - e^{\frac{-m}{\tau_2}} \right). \tag{6}$$

Here, τ_1 determines again the exponential decay rate of the slope factor, while τ_2 controls the decay rate of the curvature factor as well as the location of the hump or trough. Nelson & Siegel (1987) found in tests that the model variant in equation (6) with individual decay parameters was overparameterized. Therefore they proposed the more parsimonious formulation in equation (4). However, Bliss (1997) remarked that their finding of overparameterization resulted from using a sample of instruments with maturity of up to one year only, and that overparameterization should not pose any problem when also longer maturities were considered. Thus, we will also consider the generalized version in equation (6) in the sequel and refer to it as Bliss model. By comparison of (5) and (6), it is obvious that the Bliss model may also be seen as a special case of the Svensson model with its $\beta_2 = 0$.

Any model that is an extension of the Nelson-Siegel model can be used to obtain a fit that is at least as good as the one obtained with the Nelson–Siegel model, since it includes the latter as a special case. However, a lower number of factors in the yield curve model is typically adequate (Diebold & Rudebusch, 2013). Dahlquist & Svensson (1996) compared the Nelson-Siegel model with the dynamic Longstaff & Schwartz (1992) term structure model and found that the former is well above what is needed for monetary policy analysis. Söderlind & Svensson (1997) stated that the original Nelson-Siegel model gives a satisfactory fit in many cases, but in some cases, when the term structure is very complex, the Svensson model improves the fit considerably. Both studies used data for Swedish government bonds denoted in Swedish Krona. Similarly, De Pooter (2007) found that the parsimonious Nelson-Siegel model offers a satisfactory fit, while the more elaborate models with multiple decay parameters (the Bliss model) or additional factors (the Svensson model) lead to an improvement for specific time points when the yield curve exhibits more complex shapes.

2.3 Challenges with the Estimation of Parsimonious Yield Curve Models

Since the parameters β_0, β_1 and β_2 of the Nelson–Siegel model can be associated with the level, slope and curvature of the yield curve, Diebold & Li (2006) recognized that they must vary over time along with the curve's changing shape. However, the authors assumed that the fourth parameter τ_1 can be fixed at a specific value such that the loading of the curvature factor in (4) achieves its maximum for a maturity of 2.5 years, which is commonly seen as "medium-term". By fixing the value of τ_1 and fitting the model in (4) directly to spot rates, the remaining parameters on each observation date can be estimated simply by OLS regression as then the factor loadings only depend on the maturity. In a subsequent step, Diebold & Li (2006) fit autoregressive models to the obtained series of β_0, β_1 and β_2, which leads to a dynamic version of the Nelson-Siegel model. This approach has been extended by Koopman et al. (2010), who treated also τ_1 in (4) as a fourth latent factor and modeled its dynamics jointly with the other parameters by a vector autoregressive process. The corresponding non-linear model was estimated with an extended Kalman filter.

Not fixing the value of τ_1 (and τ_2) leads generally to a better fit of the yield curve since it allows the location of humps or troughs in the curve to vary over time (Koopman et al., 2010; Diebold & Rudebusch, 2013). If the non-dynamic yield curve models in (4), (5) and (6) were fitted to spot rates, one could also perform a grid search over different values of τ_1 (and τ_2), estimate for each grid point the remaining parameters by OLS and select the solution with the best goodness of fit. However, as spot rates are usually not directly observable, this requires to derive them first from prices of traded instruments with another term structure estimation method like, e.g., unsmoothed Fama-Bliss rates (Fama & Bliss, 1987) or bootstrapping (Hagan & West, 2006). Yet, such approaches suffer from a lack of available instruments with very long maturities. Therefore, the above-mentioned papers consider only spot rates up to 10 years.

As central banks usually estimate the yield curve up to maturities of 30 years, their common practice is to fit parsimonious models directly to observed market prices of the relevant instruments (BIS, 2005; Gürkaynak et al., 2007; Nymand-Andersen, 2018). Estimating the full parameter set $\beta_0, \beta_1, \beta_2, \tau_1$ (and β_3, τ_2) then leads to a non-linear optimization problem due to the specific form of equations (4), (5) and (6), where the non-linearity is introduced by τ_1 (and τ_2, respectively). In practice, the estimation task is further complicated by the fact that the corresponding non-linear problem is also non-convex and has many local minima, and small changes in instrument prices as well as different initial values for the optimization algorithm may lead to different solutions (Gimeno & Nave, 2009; Manousopoulos & Michalopoulos, 2009; Gilli et al., 2010). As a result, the empirically observed model parameter values become instable and occasionally jump discretely from one day to the next. Gürkaynak et al. (2007) pointed out that although the jumps in parameters can be large, the changes in fitted yields over most of the considered maturity range are quite muted. Indeed, the estimation may arrive at similar yield curve shapes for very different combinations of parameters.

However, parameter instability poses difficulties when giving them an economic interpretation. Lengwiler & Lenz (2010) highlighted that the three factors in the Nelson-Siegel model are not mutually orthogonal, which means that each of them has innovations that are dependent on the other two factors. The authors argued that this results in difficulties in forming expectations about each factor. To address this issue, the authors demonstrated how to construct mutually orthogonal factors. Furthermore, they constructed their own three factors, which can be identified as the long, short and curvature factors. To our knowledge, this approach has not become widely accepted among academics and practitioners, and therefore we do not consider it in this paper.

Due to the similar factor loading structure for the third and fourth factors of the Svensson model, a specific potential problem arises when the decay parameters τ_1 and τ_2 assume similar values. In this case, the Svensson model reduces to the three-factor Nelson-Siegel model with a magnitude of the curvature factor equal to the sum of β_2 and β_3, and the parameters cannot be identified individually but only by their sum (De Pooter, 2007). This effect can be observed in Gürkaynak et al. (2007), where the estimates of β_2 and β_3 take large absolute values up to 10^5, but with opposite signs when the values of τ_1 and τ_2 coincide.[2] To make sure that the second curvature factor of the Svensson model increases the flexibility at other times to maturity than the first curvature factor, i.e., in order to prevent confounding effects, previous studies have suggested to preimpose restrictions on the distance between the values of τ_1 and τ_2. De Pooter (2007), who used instruments with maturities up to 10 years, preimposed the restriction of $\tau_1 \geq \tau_2 + 6.69$ to ensure that the maximum loading of the second curvature factor is at least twelve months shorter than the maximum loading of the first curvature factor. This effectively adds the extra flexibility gained from the fourth factor of the Svensson model at maturities shorter than that of the third factor, which is counterintuitive if the motivation for the second curvature factor is a better fit for the long end of the yield curve. On the other hand, Sasongko et al. (2019) preimposed the restriction $\tau_2 > \tau_1$, which implies that the maximum loading of the second curvature factor is at longer maturities than the maximum loading of the first curvature factor. This is in accordance with Ferstl & Hayden (2010) who introduced the R package *termstrc* for fitting yield curves. The authors proposed the restriction of $\tau_2 > \tau_1 + \Delta\tau$, where $\Delta\tau$ is predefined and has the default value of 0.5 in their package.[3] Furthermore, the authors also use $\Delta\tau = 0.5$ in one of their examples of using the package.

2.4 Data Choices when Estimating Parsimonious Yield Curve Models

Bolder & Stréliski (1999) emphasized that besides the optimization problem, a second key issue in the application of yield curve models is the data problem, i.e., the selection of instruments to be considered. This aspect is particularly important for parsimonious models where a single instrument can have a large impact on the shape of the whole curve and not only near its maturity (Manousopoulos & Michalopoulos, 2009).

The earlier cited papers by Diebold et al. (2006), De Pooter (2007) and Koopman et al. (2010) use Kalman filter-based estimation methods to identify the evolution of the latent factors in the context of a dynamic Nelson-Siegel model or one of its extensions. This requires the use of spot rates with constant maturities to model the measurement equation, which links observations with latent factors over time. With the exception of Treasury bills, which are essentially zero bonds with maturities up

[2] See data posted on www.federalreserve.gov/econres/feds/2006.htm, accessed 6th of January 2021.

[3] This default value was found in the R package *termstrc* downloaded from github.com/datarob/termstrc at 9th of March 2020.

to one year at the time of issue, spot rates are not directly observable. Therefore, the authors use monthly updated unsmoothed Fama-Bliss (Fama & Bliss, 1987) rates of synthetic instruments with constant maturities that are derived from prices of coupon-bearing Treasury notes and bonds by an iterative procedure. Due to the unavailability of long-term bonds, the above-mentioned papers restrict themselves to set of constant maturities up to 10 years. Only Christensen et al. (2007, 2009) considered maturities up to 30 years, taking into account a specific sample period in which Treasury bonds with the corresponding maturities were actually issued, and found clear evidence that models with more than three factors provide a better fit to the long end of the yield curve. Details on the derivation of unsmoothed Fama-Bliss rates are described in Bliss (1997), where the method is tested against other approaches, among them the Nelson-Siegel curve. However, the practice of central banks is to fit the models directly to observed prices of government securities instead of spot rates of synthetic instruments (BIS, 2005; Gürkaynak et al., 2007; Nymand-Andersen, 2018).

When selecting instruments for fitting the models, securities with special features such as being callable, variable coupon or perpetual bonds should be excluded (Nymand-Andersen, 2018). There are also reasons for excluding standard "plain-vanilla" instruments. For example, the trading volume of bonds often decreases considerably close to the maturity date, and thus the quoted prices may not accurately reflect the theoretically correct ones (BIS, 2005). Gürkaynak et al., (2007) excluded all Treasury bills and consider only notes and bonds for the purpose of yield curve fitting. This was motivated by the observation that bills are priced differently from notes and bonds with less than one year to maturity due to liquidity, taxes, and other effects. The authors also referred to Duffee (1996), who found that movements in bill yields are often disconnected from yields of notes and bonds. They also excluded the two most recently issued securities of each original term to maturity because these instruments often trade at a premium due to demand from the repurchase agreement (Repo) market and higher liquidity.

The overview in BIS (2005) showed that most central banks, which either use the Nelson-Siegel or the Svensson models to derive yield curves, follow different approaches in excluding securities, often because of country-specific reasons. The Bank of Canada excludes instruments that trade at a premium or discount of more than 500 basis points from their coupon because the price of these instruments may be distorted by tax effects (BIS, 2005). Several central banks exclude securities close to their maturity, among them the Federal Reserve (maturities below 30 days), the European Central Bank (ECB, maturities below three months), the Bank of Japan (below six months with the exception of some short-term instruments), the Bank of France (depending on the type of instrument) as well as the Swiss National Bank (below one year).

The Bundesbank found for their data set that excluding treasuries with maturities between three and twelve months implies imprecise estimates for the one-year rate, which is of particular interest for policy makers. Therefore, they exclude only instruments with less than three months time to maturity. Other central banks reflect the short end of the term structure by replacing bonds with other, more liquid instruments such as repo rates (England, Spain) or money market rates (Norway, Switzerland). In order to consider only instruments with sufficient liquidity, the European Central Bank requires a minimum daily trading volume of EUR 1 million and a maximum bid-ask spread of 3 basis points, while Canada applies a minimum outstanding amount as filter. For an extended overview of the various approaches applied by different central banks, we refer to the report by the BIS (2005).

2.5 Parsimonious Models for Forecasting

Some authors investigate also the use of parsimonious models for forecasting future interest rates. Diebold & Li (2006) reported a good forecasting performance of their dynamic extension of the Nelson-Siegel model for US Treasury yields between January 1985 and December 2000. Carriero (2011) found that the out-of-sample performance deteriorates if the sample period is extended to 2009. Duffee (2011) reported that the model is inferior to random walk forecasts when the data sample is expanded with more recent observations. Moench (2008) concluded on the basis of a subsample analysis that the strong forecasting performance documented by Diebold & Li (2006) might be due to their specific choice of the forecasting period. De Pooter (2007) found that only the four-factor model by Björk & Christensen (1999) could compete with Moench's favorite model, which uses several macroeconomic variables and parameter restrictions implied by no-arbitrage constraints. Doshi et al. (2020) proposed to use horizon-specific forecasting loss functions when estimating term structure models, instead of traditional loss functions like mean-squared error, and found that this improves out-of-sample forecasting performance. However, a further assessment of forecasting capabilities of yield curve models is beyond the scope of this paper. We refer to Duffee (2013) for a profound examination of yield curve models used for forecasting and to Carriero et al. (2012) for an extensive comparison of different modelling approaches that are estimated with Bayesian vector autoregression. It should be emphasized that parsimonious yield curve models were originally not intended for forecasting since they do not contain information on the dynamics of the yield curve (Lengwiler & Lenz, 2010; Diaz et al., 2016), unless further assumptions are made on the evolution of the factors as, e.g., in the extension by Diebold & Li (2006).

3 Data and Methodology

We fit the Nelson–Siegel, the Svensson and the Bliss models to mid prices of US Treasury securities for each of the 4996 trading days between 1st January 2000 and 31st December 2019, calculated as average of the closing bid and ask price for non-callable US bills, notes and bonds retrieved from the database of the Center for Research in Security Prices (CRSP). Following the procedures applied by several central banks, we exclude instruments with a remaining time to maturity of less than three months, as suggested by Gürkaynak et al. (2007). As mentioned earlier, they also proposed to exclude Treasury bills motivated by the findings in Duffee (1996). We test the effect of excluding vs. including the T-bills in Section 4.4.

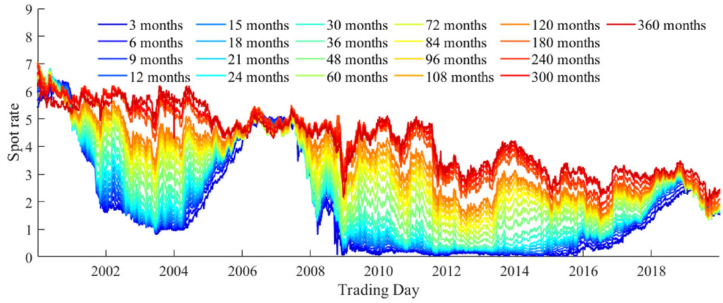


Fig. 2 Evolution of daily spot rates for fixed maturities from 3 to 360 months (30 years). The lines have unique colors from blue shades for the shortest maturities to red shades for the longest maturities. The spot rates shown are yields of synthetic instruments and are derived from market prices of Treasury instruments by bootstrapping

Figure 2 shows the evolution of daily spot rates for fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120, 180, 240, 300 and 360 months. Based on the distances between the spot rates of shorter and longer maturities, we observe that the period of investigation covers times with normal, flat and inverted yield curves. Further, the investigation period covers the shocks on the global markets after the 9/11 terror attacks in 2001, the Financial Crisis of 2007–2008, as well as rising and falling interest rates. Note that the spot rates shown are yields of synthetic instruments derived from the market prices of Treasury bills, notes and bonds by bootstrapping. They are displayed here to illustrate the different yield curve regimes during the investigation period, while the parsimonious yield curve models considered in this paper are directly fitted to prices of traded instruments.

3.1 Optimization Problem

As outlined previously, fitting a yield curve model to market data requires the minimization of an error measure χ, which is based on the differences between observed and fitted (i.e., obtained from the model) yields or prices. The choice between yield or price error minimization is not definite and depends on the intended use of the yield curve. When the purpose is deriving interest rates for monetary policy decisions, it suggests itself to minimize yield errors. By contrast, if the purpose is pricing of bonds, minimizing price errors appears more suitable. In both cases, a discount function is calculated from the yield curve obtained for the current choice of parameters and used to calculate the bond prices implied by the model. In the case of price error minimization, observed prices can be compared directly with estimated prices. A beneficial feature from a computational point of view is that analytical gradients for the error measure χ can be derived (Ferstl & Hayden, 2010), which facilitates the numerical solution of the fitting procedure. In the case of yield error minimization, in addition Eq. (3) must be solved for each instrument i to obtain its estimated yield to maturity from the corresponding model-implied price. Since this requires an iterative procedure for all coupon-bearing bonds in each step of the optimization algorithm, minimizing yield errors is computationally more demanding than price error minimization. Furthermore, gradients of the error measure must be estimated numerically.

Svensson (1994) pointed out that bond prices are rather insensitive to changes in yields for short maturities and, thus, a minimization of price errors may lead to large yield errors for short-term securities. Since a change in the yield results in a small (large) change in the price of a bond with a short (long) maturity, minimizing price errors would lead to an over-fitting of the long end of the term structure at the expense of the short end (BIS, 2005). This may be corrected by weighting the price errors of each individual bond by the inverse of its (modified) duration. In this way, yields for short maturities may be captured more accurately with less computational effort. Among the nine central banks in the overview of the BIS (2005) that adopted the Nelson-Siegel or the Svensson model, five apply a minimization of duration-weighted prices, while four use yield error minimization.

Formally, let y_i be the yield to maturity and p_i the price of security i observed on a specific trading day. For ease of notation, the time indices will be dropped in the sequel. The corresponding values derived from one of the parsimonious yield curve models (4), (5) or (6) are denoted by $\hat{y}_i(\gamma)$ and $\hat{p}_i(\gamma)$, respectively, where γ is the vector of parameters. The error for instrument i is the difference between observed and fitted value, i.e., $\epsilon_i(\gamma) = y_i - \hat{y}_i(\gamma)$ if yield errors are minimized or $\epsilon_i(\gamma) = (p_i - \hat{p}_i(\gamma))/dur_i$ for minimization of duration-weighted price errors, where dur_i is the modified duration of security i. Thus, with N securities (after filtering) considered in the estimation, the error measure to be minimized is

$$\chi(\gamma) = \sum_{i=1}^{N} [\epsilon_i(\gamma)]^2. \tag{7}$$

The resulting optimization problem

$$\min_{\mathbf{l} \leq \gamma \leq \mathbf{u}} \chi(\gamma) \tag{8}$$

is a (bound-constrained) non-linear least squares problem with lower and upper bounds $\mathbf{l}$ and $\mathbf{u}$ on the values of the parameters. If additional restrictions on the distance between the parameters τ_1 and τ_2 for the Svensson model are taken into account, problem (8) becomes a constrained non-linear optimization problem. Depending on the setting, we apply different solution algorithms. Details are described in Appendix A.

3.2 Bounds, Restrictions and Initial Values

The lower and upper bounds $\mathbf{l}$ and $\mathbf{u}$ defined above help to avoid that the fitting procedure results in a local minimum where the yield curve model parameters have (too) extreme values without any intuitive financial interpretation. As mentioned earlier, such extreme values can be observed, for example, from the data of Gürkaynak et al. (2007), where no bounds were defined and the estimated parameters assume extreme magnitudes up to absolute values above 10^5. We apply the same values for the bounds as in section 2 of Gilli et al. (2010), which are listed

Table 1 Initial values derived from observed yields in accordance with the financial interpretation of parameters (Manousopoulos & Michalopoulos, 2009) as well as lower and upper bounds (Gilli et al., 2010) used when fitting model parameters

Parameter	Initial value	Lower bound	Upper bound
β_0	See Equation (9)	0	15
β_1	See Equation (10)	-15	30
β_2	0.0	-30	30
β_3	0.0	-30	30
τ_1	1.0	0	30
τ_2	1.0	0	30

in Table 1. τ_1 and τ_2 must be strictly positive since they control the location of the first and, in case of the Svensson model, second hump (trough). We allow for values up to 30 which permits the model to take into account potential humps (troughs) at the very long end of the yield curve.

For the time being, we choose not to preimpose any restrictions on the distance between τ_1 and τ_2, but rather aim at understanding the behavior of the original model specification. However, in Sect. 4.3 we present the implications of our findings when preimposing constraints on the distance between τ_1 and τ_2, and conclude that such restrictions are disadvantageous when using the yield curve for monetary policy decisions.

Any non-linear fitting procedure requires the specification of an initial choice of the parameters and then tries to improve the fit by updating γ iteratively until it converges to a (local) minimum. Due to the existence of many local minima, the resulting goodness of fit depends largely on the choice of the starting values (Gimeno & Nave, 2009; Manousopoulos & Michalopoulos, 2009). For fitting the Svensson model, we consider six different approaches to determine these initial values.[4]

Approach #1 uses the initial values listed in Table 1, which are directly derived from observed yields and consistent with the financial interpretation of the parameters as in Manousopoulos & Michalopoulos (2009). The initial values of the magnitudes of the long-term (level) factor β_0 and the short-term (slope) factor β_1 are approximated for each trading day by

$$initial\ \beta_0 = \frac{y_1 + y_2 + y_3}{3} \tag{9}$$

$$initial\ \beta_1 = y_s - initial\ \beta_0 \tag{10}$$

where y_1, y_2 and y_3 are the observed yield to maturity in percent of the three instruments with the longest time to maturity and y_s is the observed yield to maturity in percent of the instrument with the shortest time to maturity observed on that day.[5]

In *approach #2* we fit first the less flexible Nelson-Siegel model to the data, where the initial values for the corresponding parameters are set as in the first

[4] Whenever fitting the Nelson-Siegel & Bliss models, we use approach #1 for initial values.

[5] y_1, y_2, y_3 and y_s are retrieved after any filtering of the data set, including the exclusion of instruments with a remaining time to maturity of less than three months as discussed above.

approach. In a second step, the obtained values of β_0, β_1, β_2 and τ_1 for the Nelson-Siegel model are used as initial values for fitting the Svensson model, together with the values for β_3 and τ_2 from Table 1. According to BIS (2005), a similar approach is applied by the Bank of France. *Approach #3* works analogously to approach #2, but uses the Bliss model to find values for β_0, β_1, β_2, τ_1 and τ_2, which are then used as initial values for fitting the Svensson model.

Approach #4 is inspired by the Swiss National Bank (Müller, 2002). It uses the Nelder-Mead or downhill simplex algorithm (Nelder & Mead, 1965; Box, 1965) with initial values from Table 1 to obtain a full set of all six parameters of the Svensson model by solving problem (8). In order to further improve the goodness of fit, the obtained six parameters are used again as initial values for the non-linear optimization described before.

The assumption that the yield curve should usually not change much from one day to the next is the motivation for *approach #5*, which uses as initial values for any trading day the parameters found from the non-linear optimization on the previous trading day.[6] However, we observed in preliminary tests that using only this approach might lead to extreme parameter values that tend to persist over longer time periods as the optimization algorithm gets trapped in a far from optimal local minimum. A remedy for this problem is to choose randomly alternative initial values that are uniformly distributed between the specified bounds (Gilli & Schumann, 2010).

This leads to the last *approach #6*, in which we compare for each trading day the goodness of fit obtained from solving the non-convex optimization problem for 105 different sets of initial values for the six parameters. These include 100 randomly selected sets drawn from intervals defined by the bounds in Table 1, the four sets of starting values used also by approaches #1 to #4, as well as the set of parameter estimates identified by approach #6 for the previous trading day. By selecting the parameter set with the best goodness of fit among all alternatives, approach #6 always results in the best fit according to the chosen error measure. The consideration of many sets of randomly chosen starting values in addition to those of the other approaches reduces significantly the risk that the algorithm gets trapped in a "bad" local minimum.

4 Results

In this section, we present and discuss the results obtained through the methodology described in the previous section. Section 4.1 shows comparatively the implications of approaches for selecting initial parameter values. Section 4.2 presents a comparative examination of parsimonious yield curve models and sheds light on confounding effects in the Svensson model. Section 4.3 shows the implications when preimposing restrictions on the distance between τ_1 and τ_2, while Section 4.4 presents robustness checks performed by considering different subsets of the data.

[6] We use approach #1 for initial values for the very first trading day in our data set, as data for the previous trading day in this case is not given.

Table 2 Proportion of all trading days between 2000 and 2019 when different approaches for initial values lead to the best goodness of fit

	Proportions when minimizing yield errors	Proportions when minimizing duration-weighted price errors
(a) Proportions when approach #6 selects as initial values any of the four sets of starting values used also by approaches #1 to #4, the set of parameter estimates identified by approach #6 for the previous trading day, or one of the 100 randomly selected sets drawn from intervals defined by the bounds in Table 1		
Approach #1	2.1 %	2.7 %
Approach #2	4.7 %	4.0 %
Approach #3	1.7 %	2.3 %
Approach #4	1.9 %	2.3 %
Parameters found with approach #6 on the previous day	37.7 %	35.2 %
One of the 100 randomly selected sets	52.0 %	53.5 %
Proportions when approaches #1 to #5 lead to the best goodness of fit		
Approach #1	12.0 %	14.0 %
Approach #2	25.7 %	25.7 %
Approach #3	16.2 %	14.5 %
Approach #4	18.7 %	20.5 %
Approach #5	27.3 %	25.3 %

Results are obtained when using the Svensson model fitted by minimizing yield errors and duration-weighted price errors, respectively. The approaches #1 to #6 are defined in Section 3.2

4.1 Implications of Approaches for Selecting Initial Parameter Values

Tables 2a and 2b show the proportion of all trading days (between 2000 and 2019) on which the various approaches for initial values lead to the best goodness of fit in terms of the lowest sum of squared errors when the Svensson model is fitted. The tables have two columns for the proportions when minimizing yield errors vs. duration-weighted price errors, i.e., price errors are divided by the modified duration of the corresponding bonds to avoid an overweighting of instruments with high duration. Table 2a shows how often approach #6 selects a solution in which one of the 100 combinations of random numbers was chosen to initialize the fitting procedure, compared to a parameter set obtained from one of the other approaches. We observe that in most cases one of the randomly selected sets of initial values leads to the best goodness of fit, followed by using the parameter values found with approach #6 on the previous day. Table 2b shows how often approaches #1 to #5 lead to the best goodness of fit. In this case, the proportions of the different approaches among the best solutions are more balanced as none of them are based on the comparison of several sets of initial values. Overall, without consideration of approach #6, using the initial values from the fitted Nelson-Siegel model (approach #2) or always using the values identified on the previous day (approach #5) result in the best goodness of fit.

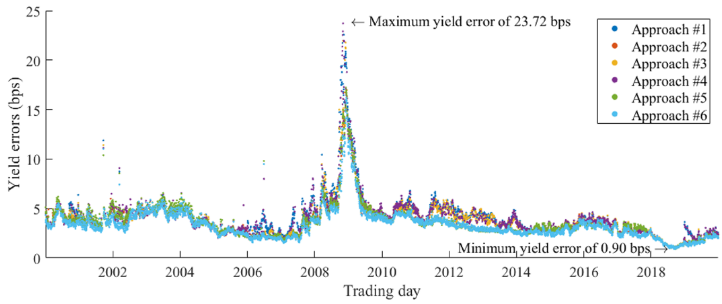


Fig. 3 Evolution of average absolute yield errors $\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i(\gamma))|$ in basis points (bps) of the N instruments taken into account on each trading day between 2000 and 2019, when the yield curve is fitted with the Svensson model by minimizing yield errors and using different approaches for initial values. The approaches #1 to #6 are defined in Sect. 3.2

Figure 3 summarizes the goodness of fit when the yield curve is fitted with the Svensson model by minimizing yield errors using the different approaches for initial values. To assess the magnitude of the mispricing of individual instruments in terms of yield to maturity, we report here the average absolute yield error $\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i(\gamma)|$ in basis points (bps) of the N instruments taken into account on each trading day between 2000 and 2019. We observe a maximum and minimum value of 23.72 bps and 0.90 bps, respectively, as well as a mean of 3.67 bps regardless of which approach for initial values is chosen. Further, we observe a worse goodness of fit from late 2007 to mid 2009, which corresponds to the Financial Crisis of 2007–2008. However, this is the same for all approaches for initial values. No significant deterioration in the goodness of fit can be found during the shocks on the global markets after the 9/11 terror attacks in 2001. Further, we observe that the times of normal, flat and inverted yield curves, as well as rising and falling interest rates, are not indicators for the choice of a specific approach for initial values. Overall, we observe rather small differences (of a few basis points) in the goodness of fit between the various approaches for the selection of initial values.[7]

Yet, the choice of the initial values has significant implications on the stability of the resulting Svensson model parameter estimates and their interpretability. Figures 4 and 5 display the evolution of β_0 and β_1 across all trading days between 2000 and 2019 when yield errors are minimized. Obviously, the estimated parameters exhibit a more stable and intuitive pattern when initial values are derived from observed yields, as illustrated in the top and middle panels of Fig. 4 for approach #1 and #2, respectively. Also, for approach #5 we observe in the middle panels of Figure 5 a more stable pattern, but there is tendency of getting trapped in local minima with extreme parameter values. The top and bottom panels of Fig. 5 imply that the variation increases significantly when approaches #4 and #6 for initial values are applied. In particular, parameters can take very different values over consecutive trading days. This is counterintuitive, since market conditions under normal circumstances persist. Thus, the financial interpretation of parameters drops

[7] Similarly, insignificant changes in the goodness of fit across different approaches result when duration-weighted price errors are minimized instead of yield errors.

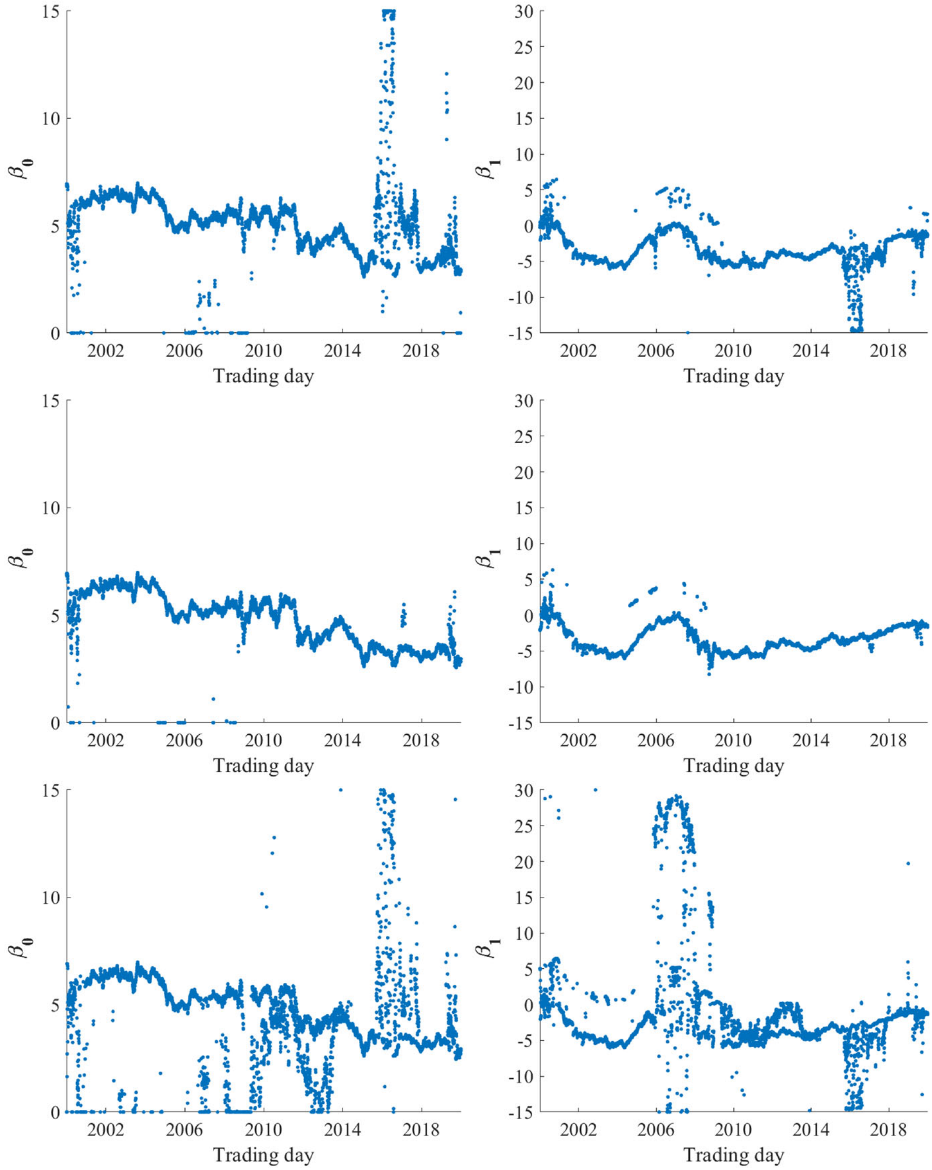


Fig. 4 Values of β_0 and β_1 across trading days derived from the Svensson model fitted by minimizing yield errors and using different approaches for initial values. Top panels show values when using approach #1 for initial values. Middle panels display values when using approach #2 for initial values. Bottom panels present values when using approach #3 for initial values. The approaches are defined in Sect. 3.2

for both approaches. The optimization with the downhill simplex algorithm in approach #4 and the random sampling in approach #6 lead to larger deviations compared to the use of initial values derived directly from data. Based on these insights, approaches #4 and #6 are not recommended if the goal is to interpret parameter values for monetary policy decisions.

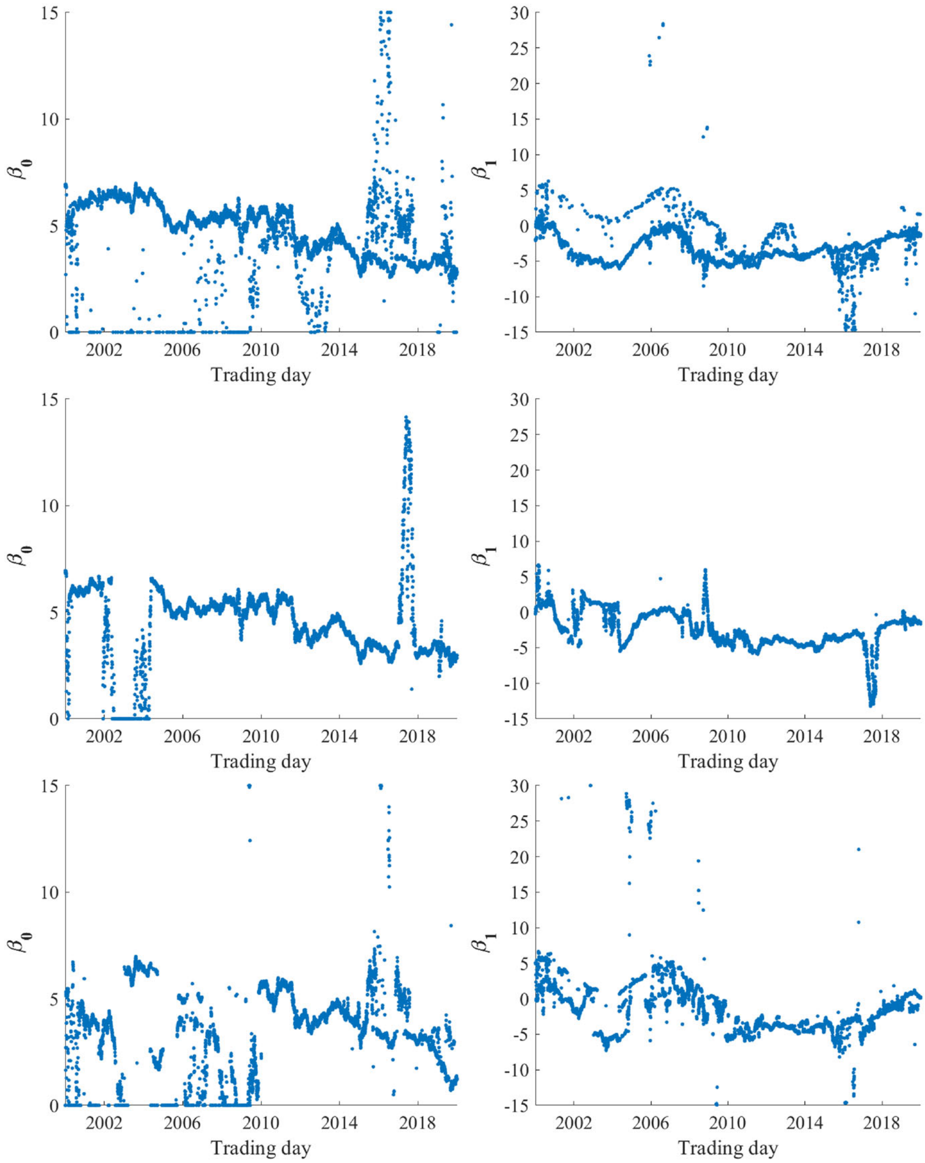


Fig. 5 Values of β_0 and β_1 across trading days derived from the Svensson model fitted by minimizing yield errors and using different approaches for initial values. Top panels show values when using approach #4 for initial values. Middle panels display values when using approach #5 for initial values. Bottom panels present values when using approach #6 for initial values. The approaches are defined in Sect. 3.2

For reasons of space we have limited ourselves to the presentation of evolution of the first two parameters β_0 and β_1 since we focus on these in subsequent discussions. However, our findings concerning the stability of parameter values applies also to β_2, β_3, τ_1 and τ_2. This becomes evident in Table 3, which exhibits the standard

Table 3 Standard deviation across all trading days between 2000 and 2019 of estimated parameter values derived from the Svensson model fitted by minimizing yield errors and using different approaches for initial values, which are defined in Section 3.2

Approach for initial values	β_0	β_1	β_2	β_3	τ_1	τ_2
Approach #1	1.76	2.48	8.09	9.33	3.57	4.38
Approach #2	1.34	1.98	13.73	11.61	3.13	1.17
Approach #3	2.18	7.01	15.12	15.00	4.72	4.26
Approach #4	1.75	2.61	13.94	14.50	3.98	3.80
Approach #5	1.97	2.53	12.52	12.28	5.72	2.72
Approach #6	2.19	4.30	8.54	8.73	6.55	7.32

deviations of all estimated parameters of the Svensson model over the entire sample period.

In conclusion, we suggest using initial values derived from observed yields (approaches #1 and #2) since this leads to the most stable and intuitive parameter estimates. However, we achieve a slightly better goodness of fit by using many combinations of initial values (approach #6), but at the expense of large variations in the estimated values of model parameters. Thus, this approach should rather be avoided when the interpretability of the estimated parameter values is important. In addition, simultaneously testing many initial values is computationally expensive. Using the parameter values obtained from fitting the model on the previous trading day as initial values (approach #5) provides a compromise between parameter stability and goodness of fit. However, this approach gets too often trapped in a local minimum with extreme parameter values and, thus, alternative initial values should be considered as well.

4.2 Comparative Examination of Parsimonious Yield Curve Models and Confounding Effects in the Svensson Model

This section presents a comparative examination of the Nelson-Siegel, Bliss and Svensson models. First, we compare the evolution of the level and the slope factors with a short- and a long-term spot rate. Second, we investigate the curvature factors, and find confounding effects in the two curvature factors of the Svensson model, which suggests that one of them is superfluous. Finally, we compare the models with respect to their goodness of fit and the behavior of the estimated parameter values.

The two top panels of Fig. 6 show the values of the magnitudes of the level and slope factors over time, derived from the Nelson–Siegel model fitted by minimizing yield errors and using approach #1 for initial values. The left panel shows the evolution of β_0 together with the 30 year spot rate, while the right panel illustrates the evolution of the sum $\beta_0 + \beta_1$ together with the 3 month spot rate. Both market rates are given in percent and were derived from the bond price data set by bootstrapping. We observe that β_0 matches the spot rates for longer times to

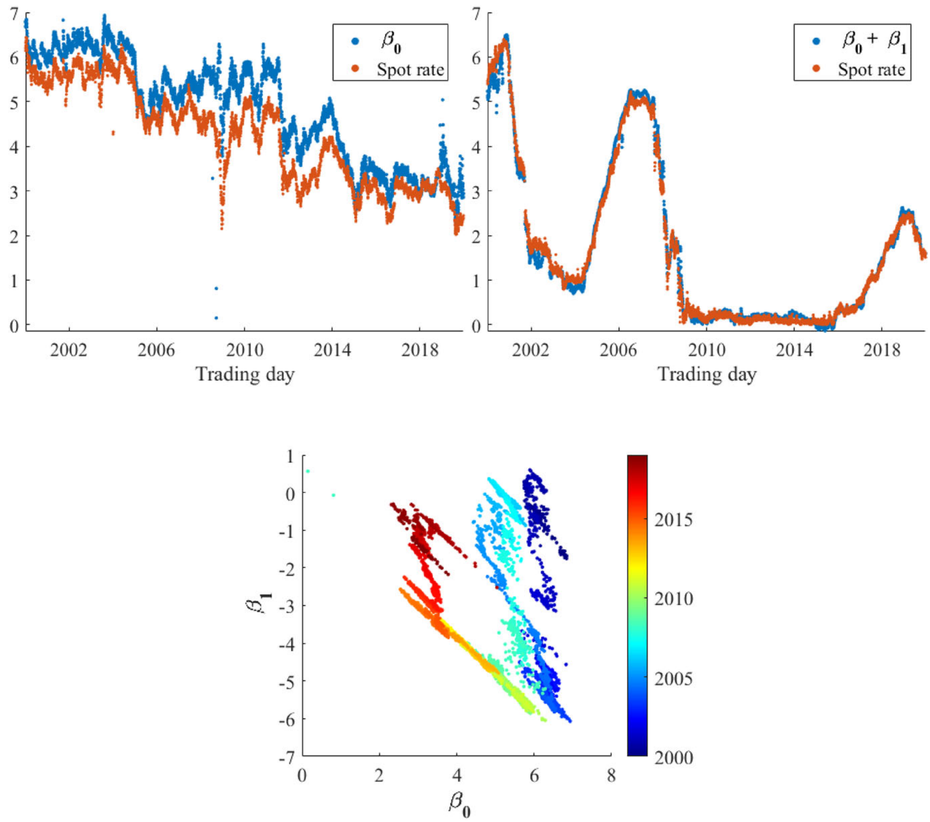


Fig. 6 Top left panel shows daily values of β_0 and spot rates for 360 months in percent derived from bootstrapping. The correlation between β_0 and the spot rates is 0.95 for the whole period of 2000–2019. Top right panel displays daily values of $\beta_0 + \beta_1$ and spot rates for 3 months in percent derived from bootstrapping. The correlation between $\beta_0 + \beta_1$ and the spot rates is 1.00 for the complete investigation period. Bottom panel presents joint evolution of β_0 and β_1 values for all trading days between 2000 and 2019. Each plot in the bottom panel has an unique color representing the trading day, which goes from blue for 1st of January 2000 to red for 31st of December 2019. All values of β_0 and β_1 in the three panels are derived from the Nelson-Siegel model fitted by minimizing yield errors and using approach #1 for initial values

maturity (360 months), with a correlation of 0.95 during 2000–2019. Further, we observe that $\beta_0 + \beta_1$ matches the spot rates for shorter times to maturity (3 months), with a correlation of 1.00 during 2000–2019. This is an empirical evidence that the magnitudes of the first two factors of the Nelson–Siegel model represent the level of the yield curve, as discussed in Sect. 2.2. We find the same evidence when using the Bliss and Svensson models and other approaches for initial values.[8] Further, we observe an almost perfect negative correlation between β_0 and β_1 over consecutive trading days. This is illustrated in the bottom panel of Fig. 6, which shows the joint evolution of β_0 and β_1 for all trading days derived from the Nelson–Siegel model

8 The empirical evidence is not necessary as obvious as in Figure 6. This because the fluctuation of parameter values is changing with different models and approaches for initial values, as discussed below and above, respectively. Results are available upon request.

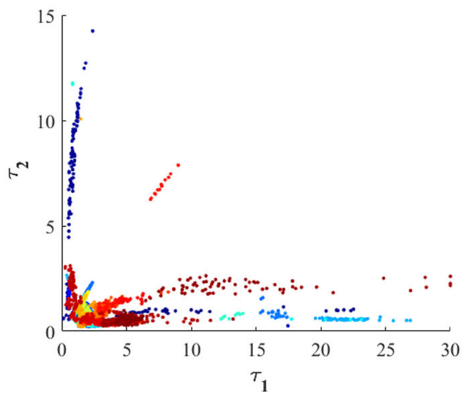


Fig. 7 Joint evolution of parameter values for all trading days between 2000 and 2019, derived from the Svensson model fitted by minimizing yield errors and using approach #2 for initial values. Each plot in the figure has an unique color representing the trading day, which goes from blue for 1st of January 2000 to red for 31st of December 2019

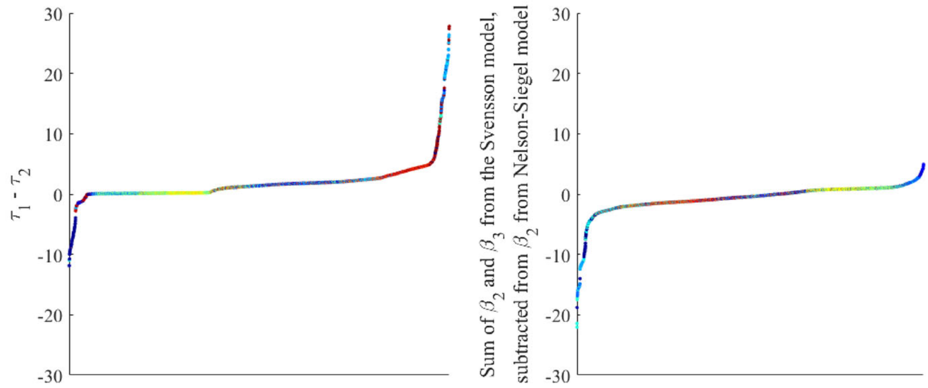
fitted by minimizing yield errors and using approach #1 for initial values. To illustrate different patterns across different trading day intervals, each plot in the panel has a unique color representing the trading day, which goes from blue for 1st of January 2000 to red for 31st of December 2019, as shown in the color bar on the right. The same colors are also used in subsequent figures. The observed high negative correlation means that the starting value of the yield curve at zero maturity ($\beta_0 + \beta_1$) remains almost constant in the corresponding trading day intervals. That is, investors' expectations for the near future remain practically constant over consecutive trading days, even if their expectations far into the future (represented by β_0) vary. We find the same evidence when using the Bliss and Svensson models and other approaches for initial values.[9] To sum up, the level and slope factors have a high degree of financial interpretation, which make them well suited for monetary policy decisions.

For the curvature factors, however, we observe confounding effects. Figure 7 shows exemplary the joint evolution of daily parameter values derived from the Svensson model fitted by minimizing yield errors and using approach #2 (fit first the Nelson–Siegel model). We observe positive correlations between τ_1 and τ_2, as well as negative correlations between β_2 and β_3. These observations are regardless of which approach for initial values is applied, however most obvious when using approach #1, #2, #3 and #4.[10] This is in line with De Pooter (2007) who reported a correlation of -0.47 between the values of β_2 and β_3 derived from the fitted Svensson model over the period 1984-2003.[11] The correlations observed here are even stronger. For example, for all trading days from February 2012 to May 2013 there is a correlation of 0.99 between τ_1 and τ_2. Furthermore, the correlation between β_2 and β_3 is -1.00 for all trading days between 2012 and 2013, as well as

9 Results are available upon request.

10 Results are available upon request.

11 See table 5 in De Pooter (2007). The author did not report correlation values involving τ_1 and τ_2.



(a) τ_2 subtracted from τ_1, both derived from the Svensson model, fitted by using approach #2 for initial values.

(b) Sum of β_2 and β_3 derived from the Svensson model, fitted by using approach #2 for initial values, subtracted from β_2 derived from the Nelson-Siegel model, fitted by using approach #1 for initial values.

Fig. 8 Parameter values for all trading days between 2000 and 2019 in ascending order, when yield curve models are fitted by minimizing yield errors. Each plot in the figure has an unique color representing the trading day, which goes from blue for 1st of January 2000 to red for 31st of December 2019

− 0.96 throughout all trading days between 2000 and 2019. In summary, these findings indicate difficulties in forming expectations about each curvature factor of the Svensson model, since they have innovations that are dependent on the other, as suggested by Lengwiler & Lenz (2010). Furthermore, this interconnection indicates confounding effects between the two curvature factors, implying that one of them is superfluous.

Figures 8a and b show parameter values for all trading days between 2000 and 2019 in ascending order derived from different models. Figure 8a shows that the values of τ_1 and τ_2, derived from the fitted Svensson model, are very similar and often the difference is zero. This means that the locations of the hump or trough of the curvature factors coincide, and the loadings of the third and fourth term in equation (5) become equal. As a consequence, the parameters β_2 and β_3 cannot be identified separately, and only their sum can be interpreted. Thus, the extra flexibility by introducing the additional curvature term in the Svensson model is most of the time not exploited. This is confirmed by Figure 8b, which shows the difference between the magnitude of the single curvature factor of the Nelson-Siegel model (β_2) and the sum of the two magnitudes of the curvature factors of the Svensson model (β_2 and β_3). Most of the time, differences are close to zero, and the Svensson model does not provide a better fit than the less flexible Nelson-Siegel model. In summary, these findings are another evidence of the confounding effects in the curvature factors of the Svensson model.

To assess if and when the additional curvature factor of the Svensson model is beneficial compared to the Nelson-Siegel and Bliss models, we evaluate the goodness of fit for each individual yield curve over the whole sample period. Let Λ_j^{mod} be the average of the absolute values of all the yield errors $\epsilon_i^{mod}(\gamma) =$

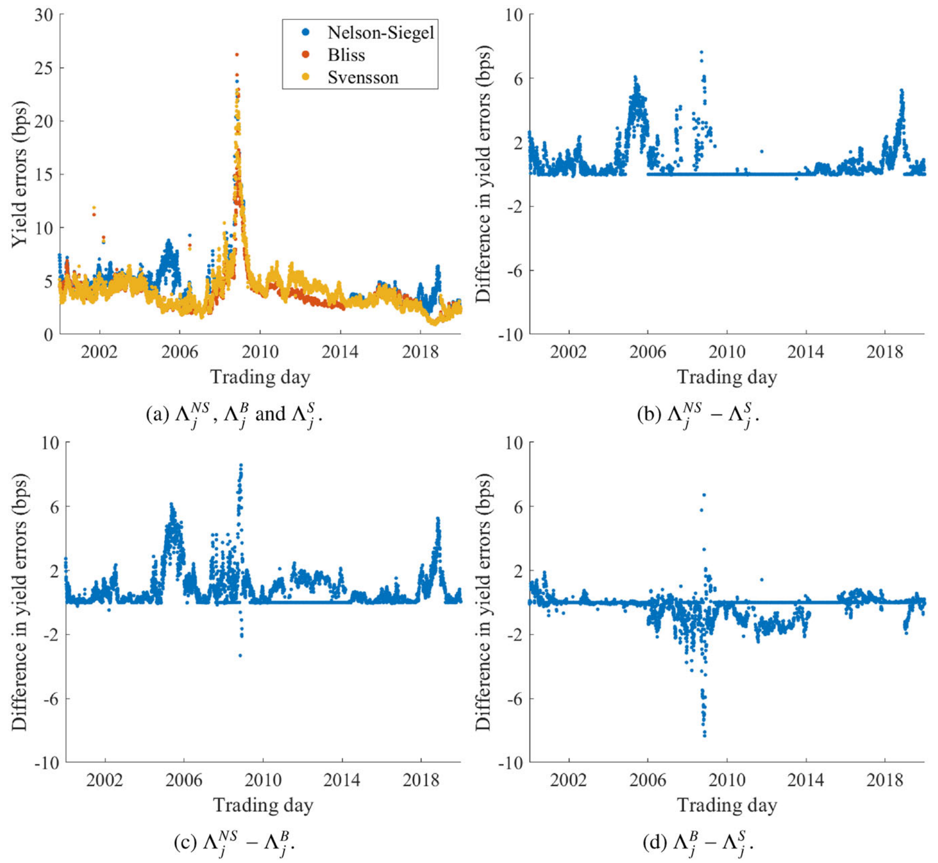


Fig. 9 Evolution of the averages of absolute yield errors in basis points (bps) on each trading day j between 2000 and 2019, when yield curves are fitted with the Nelson-Siegel (Λ_j^{NS}), Bliss (Λ_j^{B}) and Svensson (Λ_j^{S}) models by minimizing yield errors and using approach #1 for initial values

$y_i - \hat{y}_i^{mod}(\gamma)$ of all the instruments $i = \{1, \ldots, N\}$ given in bps for trading day j, defined as

$$\Lambda_j^{mod} = \frac{1}{N} \sum_{i=1}^{N} |\epsilon_i^{mod}(\gamma)|$$

where *mod* has the value *NS*, *B* or *S* indicating if the yield curve is fitted with the Nelson-Siegel, Bliss or Svensson model, respectively. Figure 9a shows Λ_j^{NS}, Λ_j^{B} and Λ_j^{S} obtained when the yield curve models are fitted by minimizing yield errors and using approach #1 for initial values. As before, we observe a worse goodness of fit from late 2007 to mid 2009 for all models, which corresponds to the Financial Crisis of 2007–2008. Again, no significant change in goodness of fit can be found during the shocks on the global markets after the 9/11 terror attacks in 2001. Furthermore, from the comparison with Fig. 2 we observe that times of normal, flat and inverted yield curves, as well as rising and falling interest rates, are not indicators for the choice of a specific model. We observe a better goodness of fit when using the Svensson model compared to the Nelson-Siegel model, as illustrated by the difference $\Lambda_j^{NS} - \Lambda_j^{S}$ in Fig. 9b. In addition, we observe a better goodness of fit when using the Bliss model compared to the Nelson-Siegel model, as illustrated by the difference $\Lambda_j^{NS} - \Lambda_j^{B}$ in Fig. 9c. This better goodness of fit when using the Svensson and Bliss models, compared to the Nelson–Siegel model, can be attributed to their extra flexibility. We also observe a better goodness of fit when using the Bliss model compared to using the Svensson model, even if the latter is more flexible, as illustrated by the difference $\Lambda_j^{B} - \Lambda_j^{S}$ in Fig. 9d. This stems from the fact that the optimization algorithm gets often trapped in a sub-optimal local minimum. Due to the higher dimensionality of the parameter space, the Svensson model is more sensitive to the choice of initial values when the non-convex data fitting problem is solved. Nevertheless, these differences in goodness of fit in Fig. 9b, c and d are so small that we do not consider them relevant when using the yield curve for monetary policy analysis. The difference is often close to zero, and the averages of the data shown in Fig. 9b,c and d are 0.57 bps, 0.76 bps and − 0.19 bps, respectively. In summary, we find that the extra flexibility of the Svensson model does not bring a significant contribution to the goodness of fit. It may even lead to a poorer goodness of fit compared to the less flexible Bliss model due to the challenge of identifying a “good” local optimum for the non-convex data fitting problem.[12]

To sum up, our findings confirm the statement of Söderlind & Svensson (1997) that the less flexible Nelson-Siegel model gives a satisfactory fit in many cases, as well as the conclusion of Dahlquist & Svensson (1996) that it is well above what is needed for monetary policy analysis. In particular, our findings are consistent with those of Diebold et al. (2006) and De Pooter (2007) that the Nelson-Siegel model gives a satisfactory fit compared to more flexible models, and illustrate that a lower number of factors in the yield curve model is typically adequate (Diebold & Rudebusch, 2013).

Furthermore, we observe that the model choice has an impact on the variation of parameter values, as also found by De Pooter (2007). This becomes evident in Fig. 10, which displays the evolution of the estimated values of β_0 and β_1 when yield curves are fitted by minimizing yield errors with approach #1 for initial values. In particular, we observe most variation in parameter values for the Svensson model, as shown in the top panels of Fig. 10. However, this variation is reduced with the Bliss model (middle panels of Fig. 10). The parameter values variate least when fitting the Nelson–Siegel model (bottom panels). Moreover, we observe that the variation of parameter values is not dependent on financial crises, times of different yield curve shapes or regimes of rising or falling interest rates. A similar pattern of variation in parameter values does also apply for the other parameters, but we have omitted their presentation for reasons of space.[13] Table 4 summarizes for

[12] These findings persist when models were fitted by minimizing duration-weighted price errors instead of yield errors. Results are available upon request.

[13] Results are available upon request.

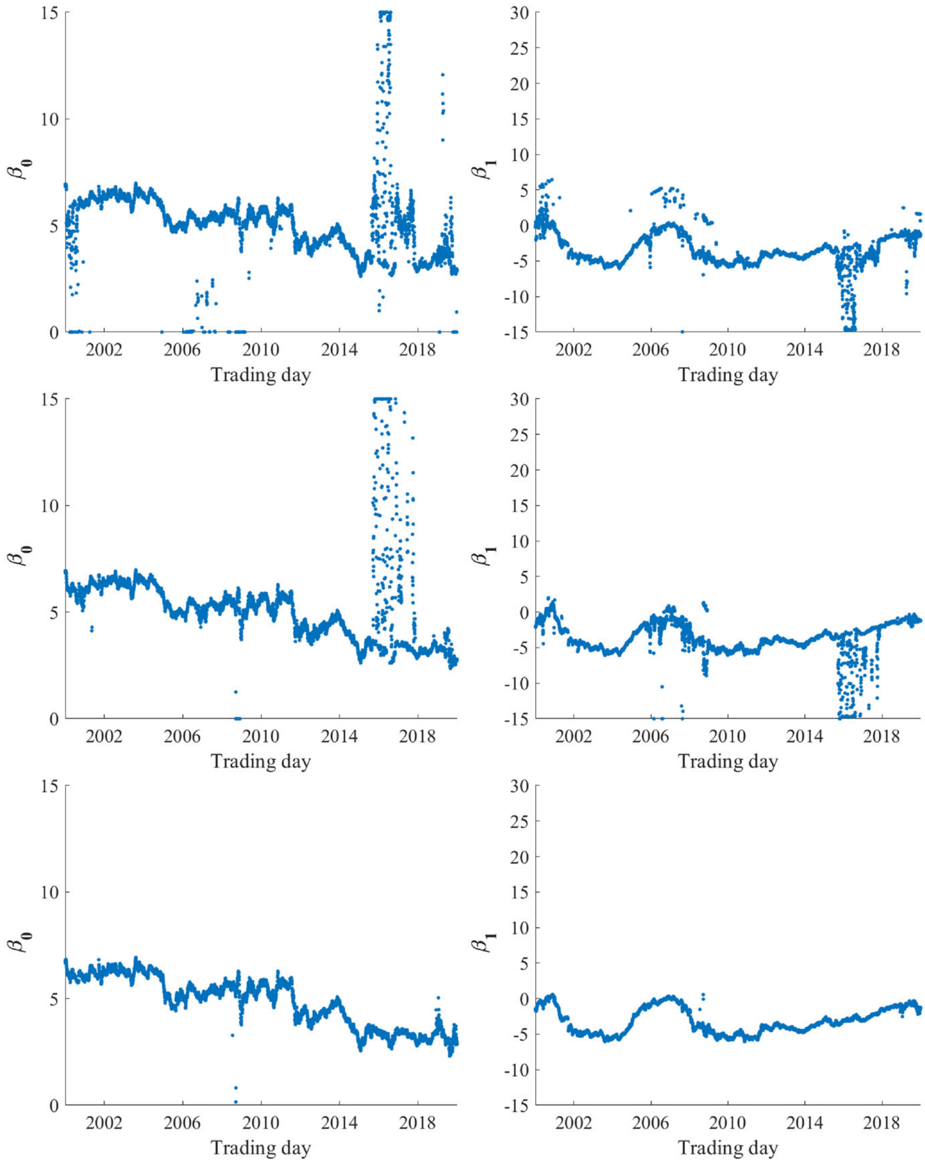


Fig. 10 Estimated values of β_0 and β_1 across trading days when yield curves are fitted by minimizing yield errors and using approach #1 for initial values. Top panels show values when using the Svensson model. Middle panels display values when using the Bliss model. Bottom panels present values when using the Nelson-Siegel model

all three models the standard deviations of the complete set of estimated parameters.[14]

[14] Similar results were obtained for the minimization of duration-weighted price errors and are available upon request.

Table 4 Standard deviation across all trading days between 2000 and 2019 of estimated parameter values derived from the Svensson, Bliss and Nelson–Siegel models, respectively, fitted by minimizing yield errors and using approach #1 for initial values

Model	β_0	β_1	β_2	β_3	τ_1	τ_2
Svensson	1.76	2.48	8.09	9.33	3.57	4.38
Bliss	1.78	2.37	4.62		5.95	4.31
Nelson-Siegel	1.18	1.80	2.40		1.80	

Overall, if the focus is on employing the estimated parameters for monetary policy decisions, we conclude that the Nelson-Siegel model is a better choice than the Bliss and Svensson models.

4.3 Preimposing Restrictions on the Distance Between τ_1 and τ_2

If the motivation for the second curvature factor in the Svensson model is a better fit for the long end of the yield curve, we would expect $\tau_2 > \tau_1$. However, in our results above, where we preimpose no restrictions on the distance between τ_1 and τ_2 like in Gürkaynak et al. (2007), this is most often not the case, as illustrated in Fig. 8a. Furthermore, using approach #5 for initial values results in solutions with $\tau_2 < \tau_1$ for all trading days. In addition, regardless of the approach for initial values, we observe less outliers and more stability in all estimated parameter values for trading days when $\tau_2 < \tau_1$, compared to trading days when $\tau_2 > \tau_1$.[15]

These counter-intuitive insights, and the observation that confounding effects are partly due to correlations between τ_1 and τ_2, are the motivation for testing the implications on our findings when preimposing restrictions on the distance between τ_1 and τ_2. First, we regenerate results when making sure that τ_2 is larger than τ_1, like in Ferstl & Hayden (2010) and Sasongko et al. (2019). Second, we regenerate results when making sure that τ_1 is larger than τ_2, like in De Pooter (2007). In particular, we investigate the implications on our findings by refitting the yield curve with the Svensson model by minimizing yield errors, using approach #1 for initial values and adding the constraints $\tau_2 \geq \tau_1 + 0.5$ and $\tau_1 \geq \tau_2 + 0.5$, respectively.[16]

Figure 11 shows yield errors when preimposing no restriction, when preimposing $\tau_2 \geq \tau_1 + 0.5$ and preimposing $\tau_1 \geq \tau_2 + 0.5$, respectively. We observe that in most cases the restrictions have an insignificant effect on the goodness of fit. Furthermore, we still observe positive correlations between τ_1 and τ_2 and negative correlations between β_2 and β_3, which indicates that confounding effects in the curvature factors of the Svensson model persist.[17] However, we observe that preimposing restrictions on the distance between τ_1 and τ_2 has a significant effect on the variation in parameter values across trading days. Indeed, the variation of

[15] Results are available upon request.

[16] The initial values in Table 1 were adjusted correspondingly.

[17] Results are available upon request.

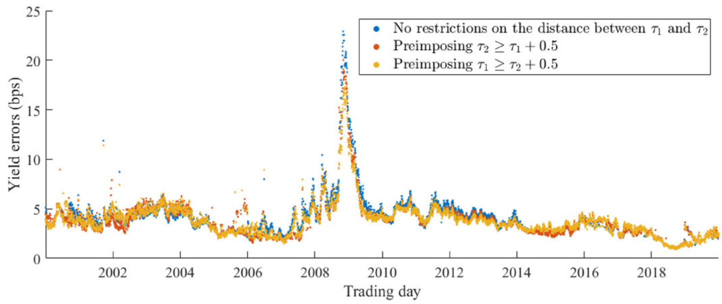


Fig. 11 Evolution of average absolute yield errors $\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i(\gamma))|$ in basis points (bps) of the N instruments taken into account on each trading day between 2000 and 2019, when the yield curve is fitted with the Svensson model by minimizing yield errors, using approach #1 for initial values and preimposing different restrictions on the distance between τ_1 and τ_2

estimated values increases for all parameters. This is displayed in Fig. 12, in which we again restrict ourselves to the presentation of β_0 and β_1. The increasing variation can also be seen in Table 5, which exhibits the standard deviations of the complete parameter set for the entire sample period. Based on these results, we recommend not to preimpose restrictions on the distance between τ_1 and τ_2 when using the yield curve for monetary policy decisions.

4.4 Robustness Checks

In this section, we present case studies where we use subsets of the total data set to regenerate results for checking the robustness of our findings. Our focus is on confounding effects in the curvature factors of the Svensson model, parameter stability and goodness of fit. Initial values for the fitting procedure are derived from approaches #1 and #2, respectively. For reasons of space we show only results for the former.[18] The various case studies are *(i)* excluding certain instruments that behave differently than others, namely Treasury bills, and *(ii)* controlling for the observed clustering of instruments across time to maturity by restricting the maturity segments with different concentration of available instruments. The effects on goodness of fit in both cases are presented in Figure 13, which compares yield errors when using the different subsets of data.

In the first case study, we investigate the effects of excluding Treasury bills from the data. This was suggested by Gürkaynak et al. (2007), who motivated it with the observation that bills are priced measurably differently from notes and bonds with less than one year to maturity due to liquidity, taxes and other effects. They referred here to Duffee (1996), who found that movements in bill yields are often disconnected from yields of notes and bonds. However, we find that excluding Treasury bills from the data has an insignificant effect on the goodness of fit, as shown in Fig. 13. In addition, the effect on the evolution of parameters is marginal, which can be seen in the middle panels of Fig. 14 for the example of β_0 and β_1, but the findings prevail for the other parameters as well. This can be seen also in

[18] Results for approach #2 are available upon request.

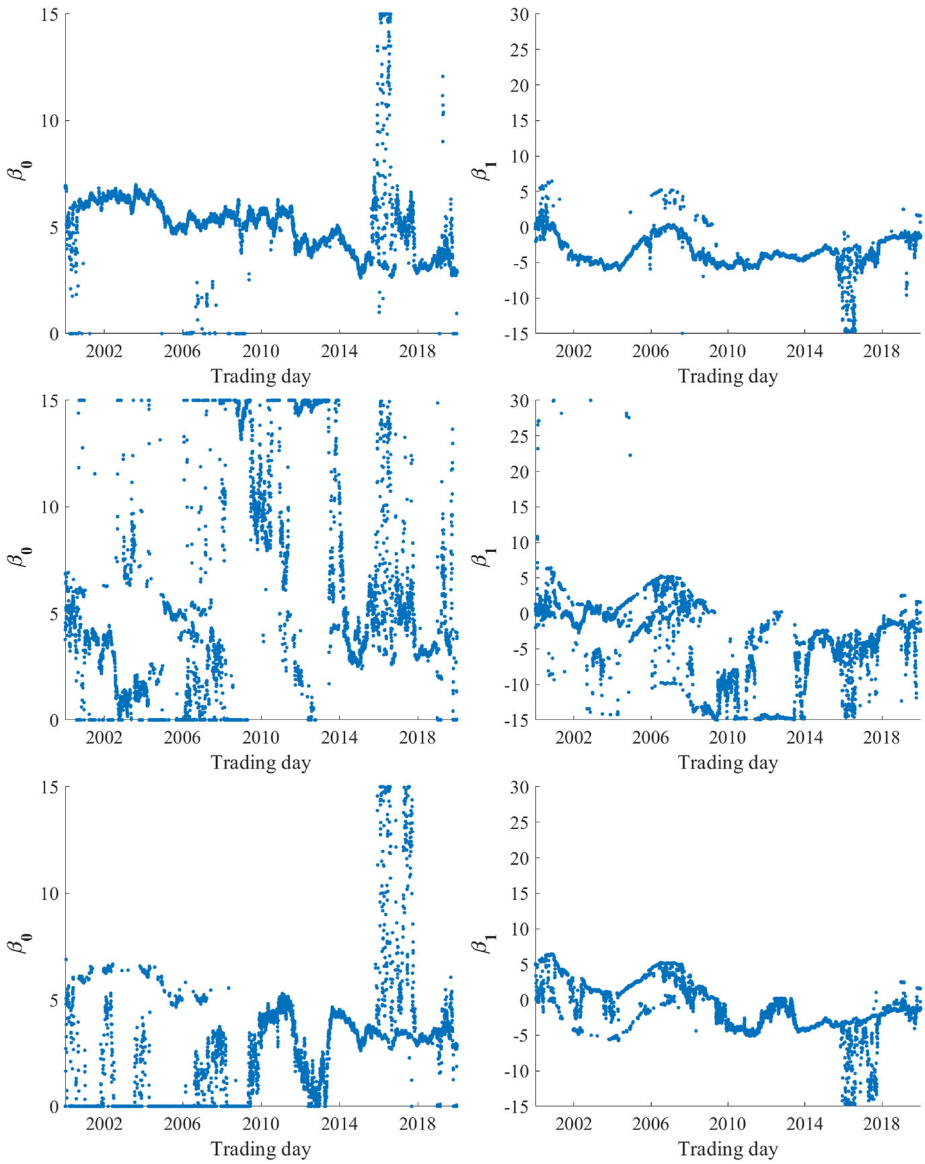


Fig. 12 Estimated values of β_0 and β_1 across trading days when yield curves are fitted to the Svensson model by minimizing yield errors and using approach #1 for initial values. Top panels show values when preimposing no restrictions on the distance between τ_1 and τ_2. Middle panels display values when preimposing $\tau_2 \geq \tau_1 + 0.5$. Bottom panels present values when preimposing $\tau_1 \geq \tau_2 + 0.5$

Table 6, which shows again the standard deviations of estimated parameters across all trading days between 2000 and 2019 when different subsets of data are used. Insignificant effects on the goodness of fit and parameter stability are also observed when fitting the Nelson–Siegel model. We still observe positive correlations

Table 5 Standard deviation across all trading days between 2000 and 2019 of estimated parameter values derived from the Svensson model fitted by minimizing yield errors, using approach #1 for initial values and preimposing different restrictions on the distance between τ_1 and τ_2

Model	β_0	β_1	β_2	β_3	τ_1	τ_2
Preimposing no restrictions	1.76	2.48	8.09	9.33	3.57	4.38
Preimposing $\tau_2 \geq \tau_1 + 0.5$	4.95	6.09	8.10	17.48	3.17	9.88
Preimposing $\tau_1 \geq \tau_2 + 0.5$	2.75	3.66	14.16	11.73	7.07	3.88

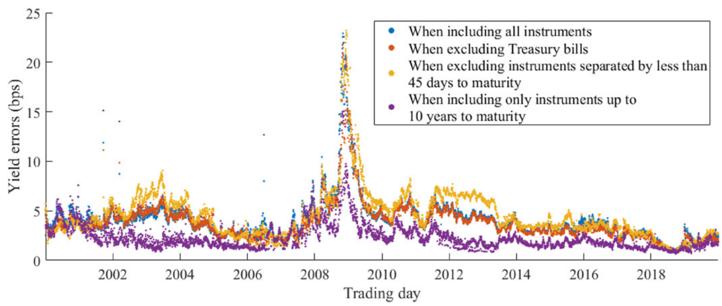


Fig. 13 Evolution of average absolute yield errors $\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i(\gamma))|$ in basis points (bps) of the N instruments taken into account on each trading day between 2000 and 2019, when the yield curve is fitted with the Svensson model by minimizing yield errors, using approach # 1 for initial values and for different subsets of data

between τ_1 and τ_2 and negative correlations between β_2 and β_3, which indicate confounding effects in the curvature factors of the Svensson model.[19]

As a consequence of the Treasury's issuing policy, certain maturity segments contain a larger number of instruments than others. This clustering is illustrated in Fig. 15a, which shows the number of instruments in the original data set per trading day within different intervals of years to maturity. Since parts of the yield curve with higher concentration of data points have a higher contribution to the error measure, the goodness of fit in maturity segments with less observations may degrade. Therefore, we investigate in a second case study whether a clustering of instruments has any impact on our findings. First, we exclude instruments separated by less than 45 days to maturity. In particular, if any two instruments at any specific trading day are separated by less than 45 days to maturity, the instrument with the smallest outstanding amount is excluded. The number of instruments per trading day within different intervals of years to maturity after this exclusion is shown in Fig. 15b. Second, since various authors restrict their data sets to instruments with maturities up to 10 years only, we investigate if excluding the very long end of the yield curve affects our findings. We observe that confounding effects in the curvature factors of the Svensson model persist. The smaller number of instruments in the data leads to a higher variation in parameter values for both procedures. This

[19] Results are available upon request.

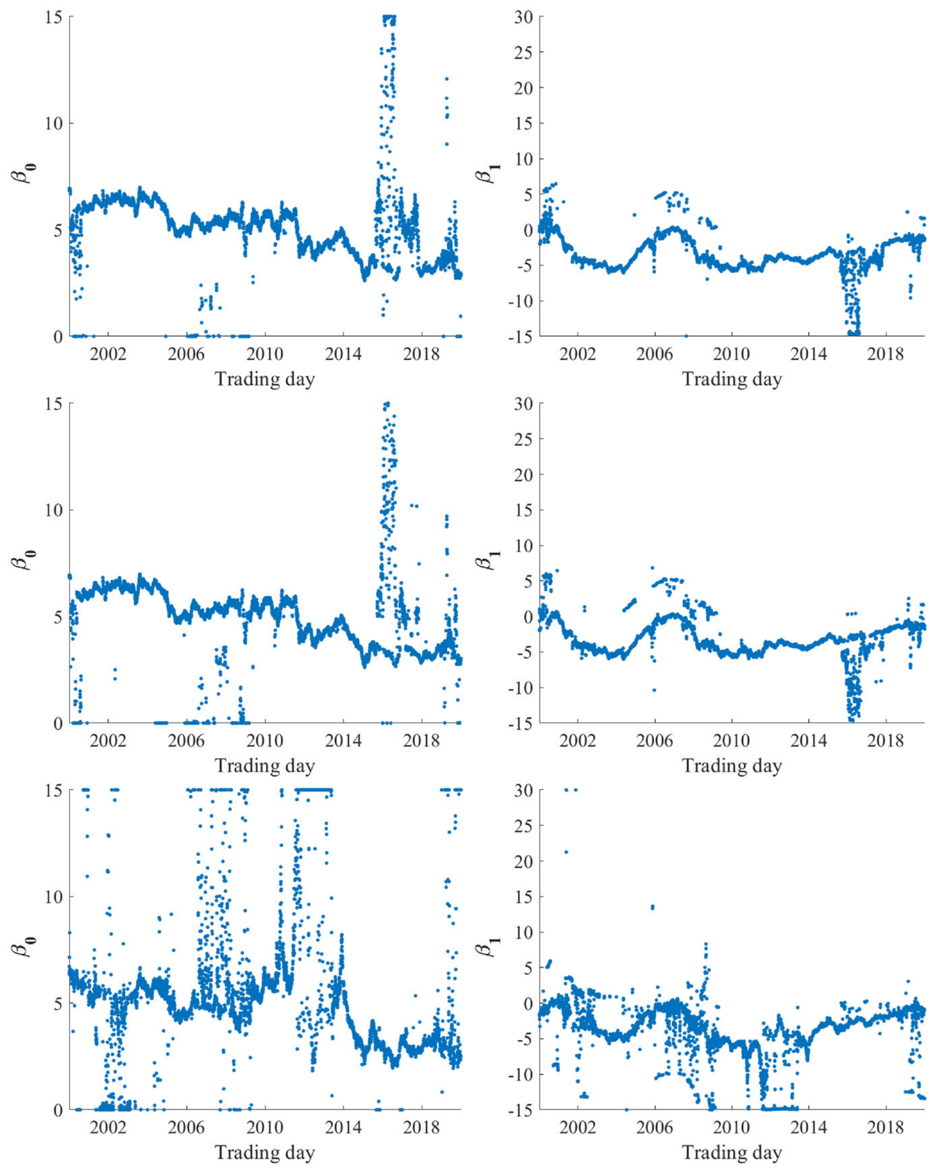
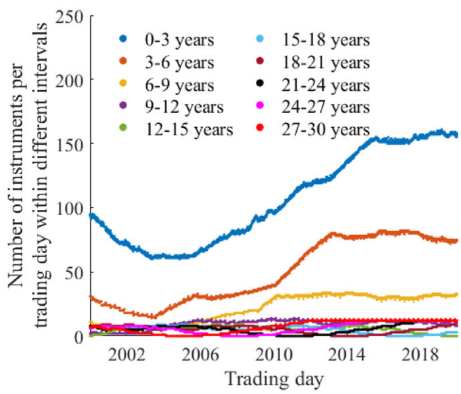


Fig. 14 Estimated values of β_0 and β_1 across trading days when yield curves are fitted to the Svensson model by minimizing yield errors and using approach #1 for initial values. Top panels show values when including all instruments in the data. Middle panels display values when excluding Treasury bills. Bottom panels present values when including only instruments up to 10 years to maturity

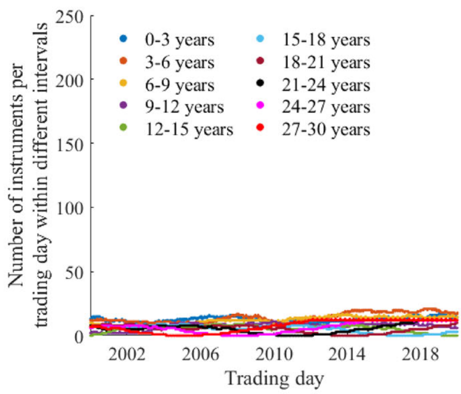
is evident in the standard deviations across all trading days between 2000 and 2019 shown in Table 6, as well as in the bottom panels of Fig. 14 that show the evolution of β_0 and β_1 when including only instruments up to 10 years to maturity. Findings prevail when considering the evolution of parameters after excluding instruments

Table 6 Standard deviation across all trading days between 2000 and 2019 of estimated parameter values derived from the Svensson model fitted by minimizing yield errors, using approach #1 for initial values and using different subsets of data

	β_0	β_1	β_2	β_3	τ_1	τ_2
Including all instruments	1.76	2.48	8.09	9.33	3.57	4.38
Excluding Treasury bills	1.80	2.54	8.15	9.47	4.25	3.90
Excluding instruments separated by less than 45 days to maturity	1.84	2.60	11.98	12.66	3.98	2.93
Including only instruments up to 10 years to maturity	3.70	4.17	12.99	13.39	3.18	4.24



(a) When including all instruments.



(b) When excluding instruments separated by less than 45 days to maturity.

Fig. 15 Number of instruments in the data per trading day within different intervals of years to maturity

separated by less than 45 days to maturity, also with respect to β_2, β_3, τ_1 and τ_2.[20] In particular, the higher variation in the values of β_0 in the case of including only instruments up to 10 years to maturity means that including instruments with maturities up to 30 years leads to a better approximation of the long end of the yield curve.

In conclusion, we observe that goodness of fit and confounding effects in the curvature factors hold for all cases. However, for the sake of the parameter stability, we recommend not to reduce the clustering of instruments across time to maturity.[21]

[20] Results are available upon request.

[21] We also found an insignificant effect on the goodness of fit, a persistence of confounding effects in the curvature factors and a reduction in parameter stability when excluding instruments separated by less than other than 45 days to maturity, as well as when fitting yield curve models only to instruments up to 3 and 5 years to maturity, respectively.

5 Conclusions

We assess and make recommendations concerning modelling and estimation choices relevant for central banks when using parsimonious yield curve models for monetary policy decisions. In this context, we illustrate that winning the objective function race is not a relevant criterion since different choices result in negligible differences in the goodness of fit, rather the stability of model parameters becomes relevant as they have a specific financial interpretation. For every trading day between 2000 and 2019, we fit the Nelson–Siegel, Svensson and Bliss models to observed US Treasury securities with maturities up to 30 years. Following the practice of central banks, we do not fix any model parameters. Consequently, parameters are estimated by solving a non-linear optimization problem, which requires a predefinition of initial parameter values. Our study is the first in the literature that compares the stability of estimated model parameters *(i)* among different parsimonious models and *(ii)* for different approaches for predefining initial parameter values. Furthermore, it investigates the impact of *(iii)* constraints on the parameters that define the location of humps and troughs as well as *(iv)* filter criteria for the selection of instruments considered in the estimation on parameter stability, confounding effects and goodness of fit.

To obtain the most stable and intuitive parameter estimates over time, we recommend that central banks employ the Nelson-Siegel model by taking initial parameter values derived from the observed yields. Our findings are consistent with previous studies (Diebold & Rudebusch, 2013) and confirm that the Nelson–Siegel model gives a satisfactory fit compared to more flexible models (Diebold et al., 2006; De Pooter, 2007) and is also well above what is needed for monetary policy analysis (Söderlind & Svensson, 1997; Dahlquist & Svensson, 1996). The recommendation of using the Nelson-Siegel model is further supported by the concluding result that the Svensson model is often superfluous due to confounding effects between the curvature factors. In general, our findings hold regardless of whether parameters are estimated by minimizing yield errors or duration-weighted price errors. We observe that neither regimes of normal, flat or inverted yield curve shapes, financial crises, rising/falling interest rates are indicators for the choice of a specific model.

The observed confounding effects in the Svensson model are partly due to correlations between the parameters controlling the location of the humps or troughs of the yield curve. Consequently, we study the implications of constraining them as suggested by De Pooter (2007), Ferstl & Hayden (2010) and Sasongko et al. (2019). Indeed, to our knowledge, we are the first to investigate the implications of such constraints on the stability of estimated parameters and the goodness of fit. Our findings suggest not to use such constraints as they result in reduced parameter stability, while the impacts on confounding effects and goodness of fit are insignificant.

Since there is evidence that yields of Treasury bills are often disconnected from yields of notes and bonds (Duffee, 1996; Gürkaynak et al., 2007), we investigate the impact of excluding them from the data. Our finding is that an exclusion of bills has insignificant impact on the goodness of fit, parameter stability and confounding effects in the Svensson model. Furthermore, as the maturity dates of observed bonds are not uniformly distributed along the curve, we assess the impact of a concentration of instruments in certain maturity segments on our results. An elimination of instruments in segments with higher concentration neither improves the goodness of fit nor eliminates confounding effects. In particular, we observe that the exclusion of instruments with maturities above ten years, which is often done in empirical studies, leads to higher parameter instability. Therefore, including also the available long-term instruments provides a better approximation for the long end of the yield curve.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital).

Appendix A. Numerical Solution

The fitting procedures of all parsimonious models are implemented in Matlab. The spot rates shown in Fig. 2 were derived from prices of Treasury instruments with the function "bootstrap" from the Financial Instrument Toolbox. While the latter contains also standard functions to fit the Nelson–Siegel and the Svensson model, we have implemented our own estimation routines for all three models that allow also to take into account constraints like on the distance between parameters. We use the interior point solver "fmincon" from the Optimization Toolbox to solve the non-linear optimization problem (8) with optional constraints and analytical gradients when appropriate. For the minimization of yield errors without additional constraints, we used the solver "lsqnonlin", which implements a trust-region reflective least-squares algorithm (Moré & Sorensen, 1983; Sorensen, 1997), with numerical gradients. The parameter for the termination tolerance on the first-order optimality was set to 10^{-12}. The implementation of the Nelder-Mead method used for approach #4 is taken from the NLopt library (Johnson, 2017).

For the computationally more demanding yield curve error minimization, the solution of the most complex approach #6 that solves the non-linear optimization problem 105 times with different starting values takes about 45 minutes on a PC with Intel i7 processor at 1.9 GHz. This is numerically feasible in the daily practice of a financial institution. However, we used the NTNU IDUN computing cluster (Själander et al., 2019) to carry out the various case studies for each of the 4996 trading days. Each study was performed twice: For the minimization of errors in yields and errors in duration-weighted prices. The cluster has more than 70 nodes and 90 GPGPUs. Each node contains two Intel Xeon cores, at least 128 GB of main memory, and is connected to an Infiniband network. Half of the nodes are equipped with two or more Nvidia Tesla P100 or V100 GPGPUs. Idun's storage is provided by two storage arrays and a Lustre parallel distributed file system.

References

Ang, A., & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics, 50,* 745–787.

Ang, A., Piazzesi, M., & Wei, M. (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics, 131,* 359–403.

Audrino, F. (2012). What drives short rate dynamics? A functional gradient descent approach. *Computational Economics, 39,* 315–335.

Bernanke, B. S., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics, 120,* 387–422.

Bianchi, C., & Cleur, E. M. (1996). Indirect estimation of stochastic differential equation models: Some computational experiments. *Computational Economics, 9,* 257–274.

Bianchi, D., Büchner, M., Hoogteijling, T., & Tamoni, A. (2020a). Corrigendum: Bond risk premiums with machine learning. *The Review of Financial Studies*. (forthcoming).

Bianchi, D., Büchner, M., & Tamoni, A. (2020b). Bond risk premiums with machine learning. *The Review of Financial Studies, 34,* 1046.

BIS (2005). Zero-coupon yield curves: Technical documentation. BIS Papers No. 25. Bank of International Settlements.

Björk, T., & Christensen, B. J. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance, 9,* 323–348.

Bliss, R. R. (1997). Testing term structure estimation methods. *Advances in Futures and Options Research, 9,* 197–232.

Bolder, D., & Stréliski, D. (1999). *Yield curve modelling at the Bank of Canada*. Bank of Canada: Technical Report.

Box, M. J. (1965). A new method of constrained optimization and a comparison with other methods. *The Computer Journal, 8,* 42–52.

Caldeira, J. F., Moura, G. V., & Santos, A. A. (2015). Measuring risk in fixed income portfolios using yield curve models. *Computational Economics, 46,* 65–82.

Carriero, A. (2011). Forecasting the yield curve using priors from no-arbitrage affine term structure models. *International Economic Review, 52,* 425–459.

Carriero, A., Kapetanios, G., & Marcellino, M. (2012). Forecasting government bond yields with large Bayesian vector autoregressions. *Journal of Banking & Finance, 36,* 2026–2047.

Christensen, J. H. E., Diebold, F. X. & Rudebusch, G. D. (2007). The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models. NBER Working Paper No. 13611. National Bureau of Economic Research.

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2009). An arbitrage-free generalized Nelson-Siegel term structure model. *The Econometrics Journal, 12,* C33–C64.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica, 53,* 385–407.

Cœuré, B. (2017). Speech at the annual meeting of the ECB's Bond Market Contact Group, by Benoît Cœuré, Member of the Executive Board of the ECB. https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp170516.en.html. Online; accessed 6th of January 2021.

Dahlquist, M., & Svensson, L. E. O. (1996). Estimating the term structure of interest rates for monetary policy analysis. *The Scandinavian Journal of Economics, 98,* 163–183.

De Pooter, M. (2007). Examining the Nelson-Siegel class of term structure models: In-sample fit versus out-of-sample forecasting performance. Working Paper. Tinbergen Institute.

De Rossi, G. (2010). Maximum likelihood estimation of the Cox-Ingersoll-Ross model using particle filters. *Computational Economics, 36,* 1–16.

Diaz, D., Theodoulidis, B., & Dupouy, C. (2016). Modelling and forecasting interest rates during stages of the economic cycle: A knowledge-discovery approach. *Expert Systems with Applications, 44,* 245–264.

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics, 130,* 337–364.

Diebold, F. X., Piazzesi, M., & Rudebusch, G. D. (2005). Modeling bond yields in finance and macroeconomics. *The American Economic Review, 95,* 415–420.

Diebold, F. X., & Rudebusch, G. D. (2013). *Yield curve modeling and forecasting: The dynamic Nelson-Siegel approach*. Princeton: Princeton University Press.

Diebold, F. X., Rudebusch, G. D., & Boragan Aruoba, S. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics, 131,* 309–338.

Doshi, H., Jacobs, K., & Liu, R. (2020). Information in the term structure: A forecasting perspective. *Management Science*. (forthcoming).

Duffee, G. (2011). Forecasting with the term structure: The role of no-arbitrage restrictions. IDEAS Working Paper Series from RePEc .

Duffee, G. (2013). Chapter 7 – forecasting interest rates. Handbook of Economic Forecasting 2, 385–426.

Duffee, G. R. (1996). Idiosyncratic variation of Treasury bill yields. *Journal of Finance, 51,* 527–551.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance, 6,* 379–406.

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. *Review of Economics and Statistics, 85,* 629–644.

Estrella, A., & Trubin, M. R. (2006). The yield curve as a leading indicator: Some practical issues. *Current Issues in Economics and Finance, 12,* 1–7.

Eydeland, A. (1996). A spectral algorithm for pricing interest rate options. *Computational Economics, 9,* 19–36.

Fama, E. F., & Bliss, R. R. (1987). The information in long-maturity forward rates. *American Economic Review, 77,* 680–692.

Ferstl, R., & Hayden, J. (2010). Zero-coupon yield curve estimation with the package termstrc. *Journal of Statistical Software, 36,* 1–34.

Gilli, M., Große, S. & Schumann, E. (2010). Calibrating the Nelson–Siegel–Svensson model. COMISEF Working Paper 31. University of Geneva.

Gilli, M. & Schumann, E. (2010). A note on 'good starting values' in numerical optimisation. COMISEF Working Paper 44. University of Geneva.

Gimeno, R., & Nave, J. M. (2009). A genetic algorithm estimation of the term structure of interest rates. *Computational Statistics & Data Analysis, 53,* 2236–2250.

Gogas, P., Papadimitriou, T., Matthaiou, M., & Chrysanthidou, E. (2015). Yield curve and recession forecasting in a machine learning framework. *Computational Economics, 45,* 635–645.

Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The U.S. treasury yield curve: 1961 to the present. *Journal of Monetary Economics, 54,* 2291–2304.

Hagan, P. S., & West, G. (2006). Interpolation methods for curve construction. *Applied Mathematical Finance, 13,* 89–129.

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.

Ho, T. S. Y., & Lee, S. B. (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance, 41,* 1011–1029.

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,* 1–8.

Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies, 3,* 573–592.

James, J., & Webber, N. (2000). *Interest rate modelling*. Chichester: Wiley (Wiley Series in Financial Engineering).

Johnson, S. G. (2017). The NLopt nonlinear-optimization package, http://github.com/stevengj/nlopt.

Ju, Y. J., Kim, C. E., & Shim, J. C. (1997). Genetic-based fuzzy models: Interest rate forecasting problem. *Computers & Industrial Engineering, 33,* 561–564.

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,* 85–95.

Koopman, S. J., Mallee, M. I. P., Wel, M. V., & d. . (2010). Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters. *Journal of Business & Economic Statistics, 28,* 329–343.

Lengwiler, Y., & Lenz, C. (2010). Intelligible factors for the yield curve. *Journal of Econometrics, 157,* 481–491.

Longstaff, F. A., & Schwartz, E. S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance, 47,* 1259–1282.

Manousopoulos, P., & Michalopoulos, M. (2009). Comparison of non-linear optimization algorithms for yield curve estimation. *European Journal of Operational Research, 192,* 594–602.

Mishkin, F. S. (1990a). The information in the longer-maturity term structure about future inflation. *Quarterly Journal of Economics, 105,* 815–828.

Mishkin, F. S. (1990b). What does the term structure tell us about future inflation? *Journal of Monetary Economics, 25,* 77–95.

Moench, E. (2008). Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. *Journal of Econometrics, 146,* 26–43.

Moré, J. J., & Sorensen, D. C. (1983). Computing a trust region step. *SIAM Journal on Scientific Statistical Computing, 4,* 553–572.

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives, 31,* 87–106.

Müller, R. (2002). Zur Berechnung der Obligationenrenditen im statistischen Monatsheft der SNB. *Swiss National Bank, Quartarly Report, 2,* 64–73.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal, 7,* 308–313.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *The Journal of Business, 60,* 473–489.

Nymand-Andersen, P. (2018). Yield curve modelling and a conceptual framework for estimating yield curves: evidence from the European Central Bank's yield curves. Report 9289933569. ECB Statistics Paper.

Oh, K. J., & Han, I. (2000). Using change-point detection to support artificial neural networks for interest rates forecasting. *Expert Systems with Applications, 19,* 105–115.

Rudebusch, G. D., & Williams, J. C. (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics, 27,* 492–503.

Rudebusch, G. D., & Wu, T. (2008). A macro-finance model of the term structure, monetary policy and the economy. *The Economic Journal, 118,* 906–926.

Sasongko, A., Utama, C. A., Wibowo, B., & Husodo, Z. A. (2019). Modifying hybrid optimisation algorithms to construct spot term structure of interest rates and proposing a standardised assessment. *Computational Economics, 54,* 957–1003.

Shiller, R. & Campbell, J. (1991). Yield spreads and interest rate movements: A bird's eye view. Review of Economic Studies 58.

Själander, M., Jahre, M., Tufte, G. & Reissmann, N. (2019). EPIC: An Energy-Efficient, High-Performance GPGPU Computing Research Infrastructure. Technical Report. Norwegian University of Science and Technology (NTNU).

Sorensen, D. C. (1997). Minimization of a large-scale quadratic function subject to a spherical constraint. *SIAM Journal on Optimization, 7,* 141–161.

Svensson, L. E. O. (1994). Estimating and Interpreting Forward Interest Rates: Sweden 1992–1994. NBER Working Paper Series. National Bureau of Economic Research.

Svensson, L. E. O. (1995). *Estimating forward interest rates with the Extended Nelson-Siegel method.* Sveriges Riksbank: Quarterly Review.

Söderlind, P., & Svensson, L. (1997). New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics, 40,* 383–429.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics, 5,* 177–188.

Authors and Affiliations

Ranik Raaen Wahlstrøm[1] · **Florentina Paraschiv**[1,2] · **Michael Schürle**[2]

✉ Ranik Raaen Wahlstrøm
ranik.raaen.wahlstrom@ntnu.no

[1] NTNU Business School, Norwegian University of Science and Technology, 7491 Trondheim, Norway

[2] Institute for Operations Research and Computational Finance, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland

Article 2

Title:
Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies

Authors:
Akarsh Kainth
Ranik Raaen Wahlstrøm

Reference:
Kainth, A., Wahlstrøm, R.R., 2021. Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies. Journal of Risk and Financial Management 14, 123. doi:10.3390/jrfm14030123.

Journal of
Risk and Financial Management

Article

Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies

Akarsh Kainth [1,*] and Ranik Raaen Wahlstrøm [2]

[1] Department of International Business, Norwegian University of Science and Technology, 6009 Ålesund, Norway
[2] NTNU Business School, Norwegian University of Science and Technology, 7491 Trondheim, Norway; ranik.raaen.wahlstrom@ntnu.no
* Correspondence: akarsh.kainth@ntnu.no; Tel.: +47-70161407

Abstract: The purpose of our paper is to investigate whether any differences between International Financial Reporting Standards (IFRS) and local Generally Accepted Accounting Principles (GAAP) impact the transparency of financial reporting of non-listed companies through bankruptcy prediction. This contributes to extant research that has focused on the effects of IFRS adoption in the context of listed companies. For our investigation, we used logistic regression, well-established accounting-based predictors, and a sample of financial statements from privately held Swedish companies using IFRS, and Norwegian companies using Norwegian GAAP. The results indicate that financial statements made under IFRS may be better suited for bankruptcy prediction than those made under Norwegian GAAP. Our findings suggest that the use of IFRS could aid in increasing the informativeness of financial reports by promoting transparency and prevent managers of firms facing insolvency from engaging in creative accounting practices. Our results should, however, be applied with caution, as they may be due to the differences in characteristics across firms that are not captured by our research design. We leave this issue open to future research.

Keywords: IFRS; accounting standards and principles; bankruptcy prediction; transparency; privately held companies; Norwegian GAAP; logistic regression; accounting-based predictors

check for updates

Citation: Kainth, Akarsh, and Ranik Raaen Wahlstrøm. 2021. Do IFRS Promote Transparency? Evidence from the Bankruptcy Prediction of Privately Held Swedish and Norwegian Companies. *Journal of Risk and Financial Management* 14: 123. https://doi.org/10.3390/jrfm14030123

Academic Editor: S. Abraham Ravid

Received: 16 February 2021
Accepted: 11 March 2021
Published: 15 March 2021

1. Introduction

Predicting company bankruptcy is at the core of credit risk management and thus important for academics, regulators and practitioners (Bărbută-Misu and Madaleno 2020). Since the input variables of models used for bankruptcy predicting often are derived from financial statements, it is important that these are transparent. International Financial Reporting Standards (IFRS) are widely used for financial reporting and promote cross-country comparability and more transparency than local Generally Accepted Accounting Principles (GAAP) through the use of fair values and more disclosure requirements (Diamond and Verrecchia 1991; Levitt 1998; Botosan and Plumlee 2002; International Standards Accounting Board 2002; Lambert et al. 2007; George et al. 2016; Fossung et al. 2020). Thus, the use of IFRS can prevent managers from engaging in creative accounting practices in order to mask the credit risk of their companies (Bhat et al. 2014; Bodle et al. 2016). All of this should make financial statements based on IFRS more relevant to stakeholders than those based on local GAAP. In this paper, we evaluate this by investigating whether the use of IFRS relative to local GAAP improves the transparency of financial reporting through bankruptcy prediction.

A financial report presents the financial position and performance of a company. When preparing a financial report, the choice of accounting regulations is of great importance, since different regulations yield different accounting figures, resulting in varied perceptions of a company. For example, the 110 companies listed on the Oslo Stock Exchange in 2005 experienced a 17% increase in net income on average, after restating their 2004 financial statements from the Norwegian GAAP (NGAAP) to IFRS, which Gjerde et al. (2008) argue is mainly due to differences in accounting for goodwill and intangible assets under the two sets of regulations. First, development expenditures are recognized as intangible assets under IFRS, while NGAAP provide a widely used option to expense them immediately. Second, goodwill is subject to amortization under NGAAP, while IFRS require that it be tested annually for impairment. Third, expenditure on brands is recognized as an intangible asset under NGAAP but not under IFRS (Norwegian Accounting Standards Board 2012; Picker et al. 2016; Bodle et al. 2016; IFRS Foundation 2021). Given these differences, we expect that the use of IFRS will lead to a change in transparency and thus the assessment of bankruptcy prediction, especially when using accounting-based variables.

We used a comprehensive sample of 2,290,551 annual financial statements from privately held Swedish and Norwegian companies based on IFRS and NGAAP, respectively, spanning the time period of 2006–2018. Furthermore, we predicted company bankruptcy using logistic regression (LR) and the input variables from Altman (1968) and the SEBRA model developed by the central bank of Norway. Our findings suggest that financial statements based on IFRS yield better bankruptcy prediction models, compared to those based on NGAAP, both in terms of in-sample fit and out-of-sample performance.

We contribute to the literature in the following ways. First, our study focuses on the role of accounting standards in bankruptcy prediction—an area which to the best of our knowledge has not been researched extensively. Second, our study is among the few that focus on the benefits of IFRS for creditors while the majority of the existing literature has investigated the effects of IFRS adoption on equity markets, cost of capital, cross country comparability and corporate investment efficiency (George et al. 2016). Third, to the best of our knowledge, we are the first to investigate whether the alleged benefits of IFRS also apply to bankruptcy prediction in the Scandinavian market. Our choice of market thus differs from most studies on bankruptcy prediction, which have used data from the United States of America (Appiah et al. 2015; Bodle et al. 2016). A study related to ours is that of Bodle et al. (2016), which found that financial reporting under IFRS yields better capabilities in terms of bankruptcy prediction models, compared to financial reporting under Australian GAAP. However, the authors used only listed companies, which is, indeed, the convention in the bankruptcy prediction literature (Appiah et al. 2015). By contrast, our analysis of medium- and large-sized privately held Swedish companies is particularly relevant, as such companies can choose to prepare their consolidated financial statements under IFRS. Consequently, IFRS have considerable legitimacy in Sweden (IFRS Foundation 2016; Marton 2017).

The rest of this paper is structured as follows. Section 2 reviews the literature on IFRS and their impact on value relevance, forecasting accuracy, credit ratings, and bankruptcy prediction. Section 3 describes the data and sampling choices, and Section 4 describes the research method. The results are presented and discussed in Section 5. Finally, the conclusions are given in Section 6.

2. Literature Review

We review the literature on the impact of IFRS on (i) value relevance, (ii) forecasting accuracy, and (iii) credit ratings and bankruptcy prediction. All of these areas can be regarded as providing evidence on the transparency of financial reporting (Singleton-Green 2015).

2.1. IFRS and Value Relevance

Financial reports are value relevant if their accounting numbers are correlated with stock market prices. Thus, if the economic reality is reflected in market prices, then value relevant financial reports are transparent, as their accounting numbers will reflect the economic reality. This also results in other benefits, such as an increased comparability of financial reports and improved efficiency of capital markets (George et al. 2016).

Studies have reported that using IFRS rather than local GAAP leads to an increased value relevance of financial reports (Bartov et al. 2005; Singleton-Green 2015). For instance, Barth et al. (2008) found that adopting International Accounting Standards (IAS) yields more value relevant financial reports in a sample of 327 adopters and non-adopters across 21 countries in the time period of 1992–2009.[1] Moreover, Horten and Serafeim (2010) suggested that financial reports under IFRS promote more value relevance through the credible communication of bad news, compared to financial reports under UK GAAP.

On the contrary, several studies found a weak relationship or no relationship between IFRS and value relevance. For instance, Hung and Subramanyam (2007) studied a sample of 80 German companies that adopted IAS during the time period of 1998–2002 and found that accounting standards did not have a major impact on value relevance. They found only weak evidence of a higher timeliness of IAS income, compared to local GAAP income, and that IAS adjustments were value relevant for the book value of equity, but not for net income. Furthermore, Oliveira et al. (2010) studied 32 Portuguese companies over the time period of 1998–2008 and found that using IFRS, instead of local GAAP, yields a lower value relevance of earnings, no change in the value relevance of the book value of equity and intangibles, and a higher value relevance of goodwill. In a similar vein, Christiansen et al. (2015) and Günther et al. (2009) found no change in value relevance when using IFRS, instead of German GAAP. Moreover, Ates (2021) used a sample of listed companies from 11 European Union countries and found that the use of IFRS led to increased value relevance of earnings per share and no significant impact on the value relevance of book value per share.

Some studies have found that the value relevance of intangibles is lower when using IFRS. For instance, a study by Cordazzo and Rossi (2020) based on a sample of non-financial listed Italian firms from 2000 to 2015 found that intangibles as a whole were not value relevant under IFRS, except for goodwill and research and development expenditures. However, when they divided the sample into intangible-intensive or non-intangible-intensive firms, the value relevance of research and development expenditures fell after the IFRS adoption. In a similar vein, a study by Paolone et al. (2020) based on a sample of Italian listed firms in the period 2010–2018 found that intangibles such as goodwill and research and development expenditures were positively related to stock prices. By contrast, Güleç (2021) claimed no change in the value relevance of research and development expenditures under IFRS.

In summary, there is a lack of consensus on whether using IFRS, instead of local GAAP, affects the value relevance of financial reports. Moreover, the different findings in the above-mentioned studies reflect the differences in the markets and time periods covered. The differences in findings could also be due to any changes in the accounting standards and principles over time.

2.2. IFRS and Forecasting Accuracy

Tan et al. (2011) and Choi et al. (2013) found that financial analysts' earnings forecasts were more accurate when based on financial reports under IFRS, compared to financial reports under UK GAAP for mandatory UK IFRS adopters in the period of 2003–2007. Furthermore, Wang et al. (2008) examined the effects of mandatory IFRS adoption in 2005 for a sample of 1438 firms in 17 European countries during the period of 2005–2006. They found significantly more accurate financial analysts' earnings forecasts for the post-IFRS adoption period than for the pre-IFRS adoption period. However, this finding was not so obvious when the authors divided the countries into legal origin groups. In particular, they found more accurate financial analysts' earnings forecasts for the French legal origin group but no significant change in accuracy for the German legal origin group. Byard et al. (2010) found that financial analysts' earnings forecasts were more accurate and less dispersed when using accounting numbers based on IFRS for 1168 mandatory adopters in 20 European countries for the time period of 2005–2006. However, this applied

[1] IAS are related to IFRS to a high degree. The IAS were published by the International Accounting Standards Committee (IASC) between 1973 and 2000. In 2000, IASC restructured itself into the International Accounting Standards Board (IASB), adopted all the IAS standards, and named the future standards IFRS (IFRS Foundation 2020).

only to IFRS adopters domiciled in countries with both strong enforcement regimes and significantly different reporting practices under local GAAP, compared to IFRS. Further, Kwon et al. (2019) found that the use of IFRS led to more accurate earnings forecasts for a sample of firms listed on the Korean Stock Exchange. In a similar vein, Masoud (2017) investigated 66 companies listed on the Amman Stock Exchange and found that earnings forecasts were more accurate under IFRS. Hence, it appears that using IFRS results in better earnings forecast accuracy.

2.3. IFRS, Credit Ratings, and Bankruptcy Prediction

Bodle et al. (2016) used a sample of 46 listed Australian companies that went bankrupt in the period of 1991–2004 and found that the accounting numbers based on IFRS predicted bankruptcy better than those based on Australian GAAP due to the increased transparency and conservative accounting rules for intangibles under IFRS. This is consistent with Florou and Kosi (2015), who found lower bond yield spreads for companies using IFRS, and Florou et al. (2017), who found that the accounting numbers of listed companies explained credit ratings better after the introduction of mandatory IFRS reporting in 2005. Furthermore, Charitou et al. (2015) found that IFRS were beneficial to the market, as companies with a higher default risk exhibited deteriorating characteristics after they started using IFRS. In addition, Wu and Zhang (2014) documented a significant increase in the sensitivity of credit ratings with the adoption of IFRS. On the other hand, Kraft and Landsman (2020) found no clear evidence of the credit relevance of accounting numbers after mandatorily switching to IFRS. Similarly, Bhat et al. (2014) found that adopting IFRS yields no change in the ability of earnings, the book value of equity, and the leverage to explain credit risk prices.

3. Data

It is difficult to obtain enough financial reports for model estimation and evaluation from the same companies and same accounting years based on IFRS and local GAAP separately as most companies make financial statements for an accounting year under a single set of accounting standards. Consequently, we consider two similar Scandinavian countries by including annual financial statements of (i) Swedish companies made under IFRS, retrieved through the Orbis database, and (ii) Norwegian companies made under NGAAP, provided by the Norwegian governmental agency, Brønnøysund Register Centre.[2] By using accounting data from the two countries, we obtain enough financial statements to compare the results of using local GAAP and IFRS, within the same accounting years. Thus, we eliminate any time period effects. However, as national cultures can impact accounting measurements and financial reporting practices, our results could potentially be affected by cross-country differences (Kanagaretnam et al. 2014). Guermazi and Halioui (2020) found that individualism and uncertainty avoidance are two important dimensions of national culture that influence behavior in terms of the implementation of accounting standards. Norway scores 69 and 50, whereas Sweden scores 71 and 29 for these dimensions, respectively. The scores for individualism are almost similar, meaning that both nations are characterized by an individualistic culture. The scores differ, however, in terms of uncertainty avoidance. The score of 50 for Norway does not indicate any preference for avoiding uncertainty, while the score of 29 for Sweden indicates a very low preference (Hofstede Insights 2020). Overall, it appears that both countries are similar in terms of cultural dimensions that could impact the accounting practice. Hence, we assume that any effects due to cross-country differences in our data can be deemed negligible. Moreover, the local GAAP of Norway and Sweden are also very similar to each other in practice (Kristoffersen 2020).

We include financial statements of privately held limited liability companies, spanning over the time period of 2006–2018. Furthermore, we include only financial statements of

[2] Restrictions apply to the availability of these data. The web pages for the data providers are www.orbis.bvdinfo.com and www.brreg.no, respectively.

medium- or large-sized companies, which we define in accordance with the Orbis database as those with a turnover above EUR 1 million or total assets above EUR 2 million.[3] Following the common convention in the literature, we exclude all financial statements operating in the banking, real estate, and public utility sectors (Mansi et al. 2012). Further, we exclude financial statements that have missing values for any of the accounting indicators used for deriving our input variables.[4] In accordance with the central bank of Norway, we categorize financial statements as bankrupt if they are the last of their company to which they belong, and the company has filed for bankruptcy (Bernhardsen and Larsen 2007). All other financial statements are categorized as non-bankrupt. Our final data sample consists of 1,892,294 financial statements using NGAAP, of which 1.8% are categorized as bankrupt, and 347,159 financial statements using IFRS, of which 1.5% are categorized as bankrupt.

In keeping with the recent bankruptcy prediction literature (e.g., Tian et al. 2015), we chose to use the real population of observations and to refrain from performing any matching in order to achieve a balanced dataset with an equal number of bankrupt and non-bankrupt financial statements. This is in line with Zmijewski (1984), who argued that the capability of a bankruptcy prediction model is distorted if it is estimated using a constructed dataset with a ratio of bankrupt to non-bankrupt observations that deviates from the real population.

The Input Variables

Since we are studying the effects of accounting standards, we use accounting-based input variables. Our initial set of variables is taken from the bankruptcy prediction model of Altman (1968) and the SEBRA model for bankruptcy prediction developed by the central bank of Norway. The former has been proven to perform well across different country settings (Altman et al. 2017) and is widely used by practitioners and academics (Begley et al. 1996; Grice and Ingram 2001; Mansi et al. 2012; Appiah et al. 2015; Tian et al. 2015; Bodle et al. 2016; Tian and Yu 2017). The latter is developed for Norwegian companies and is used by the Financial Supervisory Authority of Norway (Bernhardsen and Larsen 2007). Paraschiv et al. 2021) proved empirically that the variables of the SEBRA model yield good predictions when used with recent financial statements from privately held companies. Our initial set of variables is shown in Table 1 and measures liquidity, profitability, leverage, solvency, and company size. We do not consider the variable measuring activity which is present in the model of Altman (1968) as it has been found to be insignificant and industry sensitive (Altman 1968, 1993). Furthermore, this is also consistent with Vo et al. (2019) and Ntoung et al. (2020) who also predicted bankruptcy with the accounting-based variables of Altman (1968) but without the variable measuring activity.

Altman (1968) used the market value of equity in the numerator of the variable, BVEQ/TL. Instead, we follow the revised model of Altman (1993), using the book value of equity. This is in accordance with the claim that book–debt ratios are better than market–debt ratios, as debt issued against the latter can distort future investment decisions, which is due to the fact that market values incorporate present values of future growth opportunities (Moyer 1977; Shyam-Sunder and Myers 1999). Further, as the book value of equity is not directly available in the Orbis database, we calculate it by subtracting total liabilities from total assets. Moreover, as retained earnings are not commonly reported by privately held companies in the Orbis database, we use "Other shareholder funds" as a proxy. This item also includes profits for the fiscal year, treasury reserves, voluntary provisions, and other minority interests (Orbis 2007). However, this is deemed acceptable, as all of these items reflect the company's savings (Fan and Kalemli-Ozcan 2015).

[3] The Orbis database also uses the number of employees for defining size. We, however, do not rely on this, as it is not available for all our data.
[4] This constitutes 0.5% of the financial statements in our remaining data set.

Table 1. The initial set of input variables used in this paper. The first four are taken from Altman (1968) with the book value of equity in the variable BVEQ/TL, as suggested by Altman (1993). The remaining are taken from the SEBRA model developed by the central bank of Norway (Bernhardsen and Larsen 2007).

Variable	Category	Description
WC/TA	Liquidity	Working capital to total assets
RE/TA	Leverage	Retained earnings to total assets
EBIT/TA	Profitability	Earnings before interest and taxes to total assets
BVEQ/TL	Solvency	Book value of equity to total liabilities
BVEQ/TA	Leverage	Book value of equity to total assets
dEQ	Solvency	Dummy: one if book value of equity is less than paid in capital
LIQ/REV	Liquidity	Cash and cash equivalents less current liabilities to operating revenue
logTA	Size	The natural logarithm om total assets in EUR
PA/TA	Liquidity	Trade payables to total assets

Following the existing literature,[5] we restrict the values of the non-dummy input variables between the 5th and 95th percentiles across the financial statements based on IFRS and NGAAP, respectively, for each accounting year. If the denominator of a ratio variable is zero and its numerator is positive (negative) then the variable value is set to the maximum (minimum), i.e., the 95th (5th) percentile. If both the numerator and denominator are zero, the variable value is set to zero.

To avoid multicollinearity, we exclude several input variables from our initial set in Table 1. The exclusions are based on the Pearson correlation coefficient between the pairs of input variables and the variance inflation factor for each input variable i calculated as:

$$VIF_i = \frac{1}{1 - R^2_{OLS}} \tag{1}$$

where R^2_{OLS} is the coefficient of determination of an ordinary least squares (OLS) model with variable i as the regressand and the remaining variables as regressors. VIF_i can take any value above one, where the lower the VIF_i the lower the multicollinearity (Gareth et al. 2017). We recalculate the VIF_i values each time a variable is excluded from our variable set.

We find that the input variables RE/TA, BVEQ/TL, BVEQ/TA and LIQ/REV are highly correlated with each other, especially for the financial statements based on IFRS, resulting in very high VIF_i values of above 100. To ensure that leverage is measured by the final variable set, we select only BVEQ/TA from these four variables. Next, we exclude logTA as it has a high VIF_i value even though it is not highly correlated with any other single variable. In the remaining variable set, we have two variables measuring liquidity. Among these, we exclude WC/TA as it has the highest VIF_i value and is highly correlated with BVEQ/TA while the other liquidity variable, PA/TA, is only weakly correlated with any other variable.

Tables 2 and 3 show the VIF_i values for each variable in the final variable set when using the financial statements based on IFRS and NGAAP, respectively, as well as the correlations between the pairs of variables. We observe no evidence of unacceptable multicollinearity for the final variable set.

[5] e.g., Campbell et al. (2008) and Gupta et al. (2018).

Table 2. The variance inflation factor (VIF_i) values for each variable in the final set and the correlations between the pairs of variables when using the financial statements based on International Financial Reporting Standards (IFRS). The description of the variables is provided in Table 1.

Variable	VIF_i	EBIT/TA	BVEQ/TA	dEQ
EBIT/TA	1.66			
BVEQ/TA	1.77	0.23		
dEQ	1.07	−0.18	−0.25	
PA/TA	1.44	−0.05	−0.39	0.13

Table 3. The variance inflation factor (VIF_i) values for each variable in the final set and the correlations between any pairs of variables when using the financial statements based on Norwegian Generally Accepted Accounting Principles (NGAAP). The description of the variables is provided in Table 1.

Variable	VIF_i	EBIT/TA	BVEQ/TA	dEQ
EBIT/TA	1.27			
BVEQ/TA	1.22	0.32		
dEQ	1.37	−0.45	−0.57	
PA/TA	1.32	−0.20	−0.41	0.23

4. Methodology

Earlier bankruptcy prediction studies used a linear discriminant analysis to derive their models.[6] However, there are several issues with using this method in economics and finance, including its assumption of a multivariate normal distribution of the input variables and equal variance–covariance matrices across the groups of classes (Joy and Tollefson 1975; Deakin 1976; Eisenbeis 1977). Consequently, LR was introduced for bankruptcy prediction by Ohlson (1980). The benefits of using LR are that it requires less restrictive assumptions and gives more intuitive outputs.[7]

Let the vector $\hat{\boldsymbol{y}} = \{\hat{y}_n\}_{n=1\ldots N} \in [0,1]^N$ determine the predicted probabilities of bankruptcy given by:

$$\hat{\boldsymbol{y}} = \boldsymbol{\iota} \oslash (\boldsymbol{\iota} + \exp(-\boldsymbol{X\beta} - \boldsymbol{\iota}\beta_0)) \tag{2}$$

where $\boldsymbol{X} = \left\{x_{(n,i)}\right\}_{n=1,\ldots,N,i=1,\ldots,I}$ is a matrix of values for the input variables and i, derived from the financial statements, n, $\boldsymbol{\beta} = \{\beta_i\}_{i=1,\ldots,I}$ and β_0 are the model coefficients, $\oslash$ denotes the Hadamard (elementwise) division, and $\boldsymbol{\iota}$ is an $N \times 1$ vector of ones.

The coefficients are estimated by maximizing the likelihood function given by:

$$\prod_{\boldsymbol{n=1}}^{\boldsymbol{N}} (\hat{\boldsymbol{y}}_{\boldsymbol{n}})^{\boldsymbol{y_n}} (\boldsymbol{1} - \hat{\boldsymbol{y}}_{\boldsymbol{n}})^{\boldsymbol{1-y_n}} \tag{3}$$

where $\boldsymbol{y} = \{y_n\}_{n=1\ldots N} \in \{0,1\}^N$ is the vector of actual classifications of bankrupt (1) or non-bankrupt (0) for the financial statements n. In practice, instead of maximizing the likelihood function, we minimize the negative of the log likelihood function given by:

$$\ell(\boldsymbol{\beta}, \beta_0) = \sum_{n=1}^{N} [\boldsymbol{y} \odot (\boldsymbol{X\beta} + \boldsymbol{\iota}\beta_0) - \log(\boldsymbol{\iota} + \exp(\boldsymbol{X\beta} + \boldsymbol{\iota}\beta_0))] \tag{4}$$

6 e.g., Altman (1968); Meyer and Pifer (1970); Deakin (1972); Wilcox (1973); Blum (1974); Libby (1975); Altman and Loris (1976); Ketz (1978); and Pettway and Sinkey (1980).

7 When predicting bankruptcy using LR, where bankrupcy is labeled 1, the frequency of bankruptcies in the training data, i.e., the data used for estimating the coefficients, will always correspond to the average of the outputs from the trained LR model across all observations in the training data. Consequently, the output of the LR bankruptcy prediction model for any specific observation can be interpreted as the probability of banktuptcy.

where $\odot$ denotes the Hadamard (element-wise) product. The minimization is conducted by following an iterative optimization algorithm.[8]

We predict bankruptcy in a one-year horizon, which corresponds with the practice of most practitioners and academics (Hillegeist et al. 2004; Tian et al. 2015; Tian and Yu 2017). Indeed, Appiah et al. (2015) found that one-year data were most often considered among all the bankruptcy prediction studies they reviewed and that such studies achieved remarkable results. Further, several studies on bankruptcy prediction have shown that the best prediction was made when the forecasting horizon was one year or shorter.[9] By using a one-year time horizon, we comply with the Basel III regulatory framework, which states that the probability of default for bank and corporate exposures is the prediction of a one-year-ahead probability of default (Bank of International Settlements 2017).

We use an eight-fold cross-validation procedure with forward validation and a rolling window to divide the sample into eight subsamples, as illustrated in Table 4 (Kaastra and Boyd 1996; Keles et al. 2016).[10] For each of the eight subsamples, we evaluate the out-of-sample performance using test data consisting of all financial statements from one of the accounting years during the period of 2011–2018. We name the subsamples in accordance with the year used for measuring out-of-sample performance. Furthermore, we train the model, i.e., estimate the coefficients $\boldsymbol{\beta}$ and β_0, and evaluate the in-sample fit separately for each subsample using training data consisting of all financial statements from the five previous years. The procedure is carried out separately for all financial statements based on IFRS and local GAAP.

Table 4. For each of the eight subsamples, we separately train the model and evaluate the in-sample fit using training data consisting of all financial statements from five subsequent accounting years, which are given in green. We further evaluate the out-of-sample performance using test data consisting of all financial statements from the following accounting year, which are given in blue.

Subsample	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2011													
2012													
2013													
2014													
2015													
2016													
2017													
2018													

Table 5 shows the number of financial statements based on IFRS and local GAAP within the training and test data for each of the eight subsamples. It also shows the fractions of financial statements categorized as bankrupt. We restrict the values of the input variables between the 5th and 95th percentiles as explained in Section 3 across financial statements based on IFRS and NGAAP, respectively, for each accounting year.

We evaluate the model performance using the area under the receiver operating characteristic curve (AUC), which is a widely used metric for evaluating bankruptcy prediction models.[11] The receiver operating characteristic curve is a plot of the false positive rate against the true positive rate at different thresholds for defining the predicted class that an observation belongs to (Fawcett 2006; Hosmer et al. 2013). The AUC can, in practice, have any value between 0.5 and 1, where a higher value indicates a higher

8 For minimization, we use the L-BFGS-B algorithm (Byrd et al. 1995; Zhu et al. 1997). Further, we use zero as the initial value for all coefficients $\boldsymbol{\beta}$ and β_0 for the algorithm.

9 e.g., Altman (1968); Blum (1974); Altman et al. (1977); Moyer (1977); Ohlson (1980); Aziz et al. (1988); Altman et al. (1994, 1995); Dimitras et al. (1999); Tian et al. (2015); and Tian and Yu (2017).

10 Our results are robust to using an expanding window for enabling the utilization of all previous data when training models. Results are available upon request.

11 e.g., Duffie et al. (2007); Altman et al. (2010); Tian et al. (2015); Tian and Yu (2017); and Gupta et al. (2018).

explanatory power.[12] As a rule of thumb, an AUC between 0.7 and 0.8 is considered acceptable, while a value above 0.8 is considered excellent (Hosmer et al. 2013). When comparing different prediction models, the AUC has been found to be superior to other statistics, as it takes into account both error costs and class skewness within the data (Huang and Ling 2005).

Table 5. Number of financial statements within the training and test data for each subsample, based on International Financial Reporting Standards (IFRS) and Norwegian Generally Accepted Accounting Principles (NGAAP), respectively. The fractions of financial statements categorized as bankrupt are shown below each number.

Subsample	2011	2012	2013	2014	2015	2016	2017	2018
IFRS								
Training data	35,772	69,046	104,801	141,986	180,936	200,151	215,406	231,768
Bankrupt	3.3%	2.7%	2.4%	2.0%	1.6%	1.4%	1.2%	1.1%
Test data	34,803	37,727	39,564	41,932	46,125	50,058	54,089	58,187
Bankrupt	1.9%	1.8%	1.4%	0.9%	1.0%	1.1%	1.1%	1.1%
GAAP								
Training data	620,395	634,431	650,756	672,262	702,234	738,792	778,695	818,318
Bankrupt	2.0%	2.0%	1.8%	1.8%	1.7%	1.7%	1.7%	1.7%
Test data	128,715	136,584	145,556	153,781	161,706	167,824	175,798	181,244
Bankrupt	1.7%	1.8%	1.8%	1.7%	1.6%	1.8%	1.6%	1.3%

We also measure in-sample fit using the pseudo-R squared (R^2) of McFadden (1974), which is given as:

$$R^2 = 1 - \frac{\ell(\boldsymbol{\beta}, \beta_0)}{\ell(\beta_0)} \in [0, 1] \tag{5}$$

where $\ell(\beta_0)$ is the log likelihood of the null model, which does not contain any independent variables, but only the intercept coefficient β_0.

To determine the significance of the estimated coefficients, we use Wald statistics to assess the z-score of any coefficient of any input variable.[13] This is given for input variable i by:

$$z_i = \frac{\beta_i}{s_{\beta_i}} \tag{6}$$

where the denominator is the standard deviation of the numerator, which is given as $s_{\beta_i} = \sqrt{C_{i,i}}$, where $\boldsymbol{C} = \left\{C_{j,k}\right\}_{j=1,\ldots,N,k=1,\ldots,N}$ is the variance covariance matrix, given as $(\boldsymbol{X}'\boldsymbol{D}\boldsymbol{X})^{-1}$, and $\boldsymbol{D} = \left\{d_{j,k}\right\}_{j=1,\ldots,N,k=1,\ldots,N}$ is a diagonal matrix with $d_{l,l} = \hat{y}_l(1 - \hat{y}_l)$.

5. Results and Discussions

Tables 6 and 7 show the estimated coefficient values and model evaluations across the eight subsamples for the accounting years of 2011–2018, when considering the financial statements based on IFRS and NGAAP, respectively. The values in parentheses are the z-scores. The in-sample fit is evaluated by R^2 and AUC, while the out-of-sample performance is evaluated using AUC.

12 In theory, AUC can have a value below 0.5 which represents an unrealistic model.

13 The reader is referred to page 330 in Ryan (2018) and page 40 in Hosmer et al. (2013) for details on Wald statistics.

Table 6. Results across the eight subsamples for the accounting years of 2011–2018, when considering the financial statements based on International Financial Reporting Standards (IFRS). We show the estimated coefficient values of the logistic regression (LR) model, with the z-scores in parentheses. The input variables are detailed in Table 1. We report the in-sample fit using R^2 and area under the receiver operating characteristic curve (AUC) and the out-of-sample performance using AUC.

Variable/Metric	2011	2012	2013	2014	2015	2016	2017	2018
constant	−2.75 (−89.98)	−2.89 (−117.53)	−3.05 (−144.18)	−3.18 (−161.81)	−3.51 (−179.37)	−3.63 (−183.18)	−3.72 (−184.4)	−3.81 (−186.99)
EBIT/TA	−2.56 (−11.51)	−2.75 (−14.46)	−2.67 (−15.87)	−2.58 (−15.83)	−2.73 (−15.92)	−2.52 (−14.71)	−2.35 (−13.77)	−2.17 (−12.96)
BVEQ/TA	−3.84 (−27.29)	−4.17 (−37.02)	−4.32 (−44.30)	−4.78 (−50.18)	−4.54 (−49.01)	−4.60 (−49.13)	−4.51 (−48.38)	−4.32 (−47.48)
dEQ	0.53 (8.58)	0.74 (14.94)	0.87 (20.62)	0.97 (25.22)	1.05 (27.39)	1.11 (29.00)	1.12 (27.95)	1.12 (26.96)
PA/TA	1.39 (13.40)	1.80 (20.70)	2.29 (30.32)	2.60 (36.87)	2.94 (41.77)	3.00 (41.79)	2.91 (39.01)	2.78 (36.25)
R^2	0.19	0.22	0.23	0.24	0.22	0.22	0.22	0.20
In-sample AUC	0.80	0.82	0.83	0.84	0.84	0.84	0.83	0.82
Out-of-sample AUC	0.84	0.84	0.85	0.83	0.83	0.81	0.80	0.82

Table 7. Results across the eight subsamples for the accounting years of 2011–2018, when considering the financial statements based on Norwegian Generally Accepted Accounting Principles (NGAAP). We show the estimated coefficient values of the logistic regression (LR) model, with the z-scores in parentheses. The input variables are detailed in Table 1. We report the in-sample fit using R^2 and area under the receiver operating characteristic curve (AUC) and the out-of-sample performance using AUC.

Variable/Metric	2011	2012	2013	2014	2015	2016	2017	2018
Constant	−4.89 (−522.56)	−4.93 (−528.47)	−4.98 (−522.25)	−4.95 (−518.98)	−4.91 (−531.88)	−4.93 (−543.10)	−4.88 (−552.90)	−4.84 (−559.44)
EBIT/TA	−1.21 (−40.20)	−1.16 (−38.68)	−1.11 (−36.64)	−1.07 (−34.21)	−1.03 (−34.08)	−0.88 (−30.33)	−0.86 (−31.68)	−0.80 (−30.80)
BVEQ/TA	−0.89 (−38.44)	−0.85 (−37.41)	−0.89 (−39.53)	−0.94 (−42.57)	−0.92 (−44.17)	−0.83 (−41.86)	−0.81 (−44.24)	−0.84 (−49.19)
dEQ	1.11 (102.57)	1.12 (103.42)	1.07 (96.73)	0.98 (87.60)	0.93 (85.33)	0.93 (86.65)	0.87 (82.60)	0.81 (77.60)
PA/TA	2.61 (93.90)	2.76 (98.63)	2.79 (97.55)	2.83 (99.07)	2.95 (107.14)	3.00 (109.81)	2.95 (110.33)	2.86 (108.58)
R^2	0.22	0.23	0.22	0.21	0.20	0.19	0.18	0.17
In-sample AUC	0.83	0.83	0.83	0.82	0.81	0.81	0.80	0.80
Out-of-sample AUC	0.83	0.82	0.81	0.78	0.79	0.80	0.80	0.82

We observe that the values of R^2, in-sample AUC, and out-of-sample AUC are mostly higher for the financial statements based on IFRS, compared to those based on NGAAP.[14] This could be attributed to the increased transparency under IFRS, which prevents managers from using creative accounting practices to manipulate accounting reports in order to hide their true situation. Our findings also seem consistent with the literature, claiming improvements in financial reporting quality through IFRS adoption. Moreover, the AUC is close to or above 0.8 in all cases, which indicates that the accounting-based input variables in our study can accurately predict bankruptcy. Further, we observe stable coefficient estimates across the years, with high z-scores, indicating that all are significant.

The coefficient estimates for EBIT/TA have negative signs in all cases. This is expected, as higher values of these variables translate to relatively higher earnings and savings and thus a lower probability of going bankrupt. Furthermore, the magnitudes of these coefficients are greater under IFRS than NGAAP. The reason for this may be that NGAAP allow for the amortization of goodwill, while IFRS demand an impairment test, which yields more transparency. Furthermore, IFRS require the classification of development expenditures as intangible assets, whereas NGAAP allow them to be recognized as expenses. This

14 In this regard it should be noted that models resulting in only slight improvement in AUC scores have been shown to be superior at predicting bankruptcies resulting in potentially huge profit gains for creditors who use such models for credit decisions (Agarwal and Taffler 2008; Paraschiv et al. 2021).

can lead to an understatement of net income under NGAAP, which indicates that IFRS could be better suited to predicting company bankruptcy (Franzen et al. 2007).

The coefficient estimates for BVEQ/TA are negative under both IFRS and NGAAP. This is also in accordance with our expectations as it suggests that a higher rate of equity compared to debt lowers the probability of going bankrupt. However, the magnitudes of the coefficients are greater under IFRS than NGAAP. This may be due to IFRS being more conservative when it comes to accounting for intangibles. For instance, brands cannot be recognized as intangible assets under IFRS (IFRS Foundation 2021) while under NGAAP, they can. Another reason may be that IFRS require the recognition of more liabilities, such as long- and short-term employee benefits, termination benefits, and pension obligations, all of which increase liabilities and salary expenses. This lowers the value of BVEQ as higher expenses decrease retained earnings.

The coefficient estimates for dEQ are positive in all cases, which is logical as it suggests an increase in the probability of going bankrupt if the book value of equity falls below the paid-in equity. The coefficient values are higher under NGAAP than IFRS. This may be due to the lower retained earnings under IFRS as discussed above, which results in companies being worse off when dEQ is equal to one under NGAAP rather than IFRS.

The magnitudes of the coefficients for PA/TA are relatively similar under both NGAAP and IFRS. In all cases the signs are positive, which is as expected since it means that more trade payables increase the probability of going bankrupt.

Limitations and Suggestions for Future Research

Our study has some limitations. First, while we argue that cross country differences can be deemed negligible in our study, our research design does not capture any differences in firm characteristics. Second, it is difficult to obtain enough financial reports for model estimation and evaluation from the same companies and same accounting years reported based on both IFRS and local GAAP, respectively, as most companies make financial statements for any accounting year under a single set of accounting standards. Moreover, very few privately held companies report under IFRS in Norway.

While we analyze the role of IFRS through bankruptcy prediction only with Swedish and Norwegian data, we recommend future research to explore this issue in other markets with local GAAP of other countries. Furthermore, if possible, we recommend investigating the differences between IFRS and local GAAP by considering the same financial reports based on both IFRS and local GAAP separately.

6. Conclusions and Implications

This paper examined the impact of using IFRS on the transparency of financial reporting through bankruptcy prediction models using accounting-based input variables. We started with a set of variables taken from Altman (1968) and the SEBRA model developed by the central bank of Norway. We then excluded variables such that our final variable set showed no evidence of unacceptable multicollinearity. By using logistic regression and a comprehensive sample of privately held Norwegian and Swedish companies, our results indicate that financial reports using IFRS may yield better bankruptcy prediction models compared to financial reports using NGAAP. While our results show that the use of IFRS can help in providing a better picture of a company's financial health, our research design does not capture all differences in firm characteristics. This is due to the difficulties in obtaining the same financial statements of any particular company, derived under both IFRS and local GAAP separately, as most companies, especially private ones, make financial statements for any given accounting year under a single set of accounting standards. Hence, we urge caution while interpreting our results.

Our findings have implications for several stakeholders, as well as for the development and application of accounting. The increased performance of bankruptcy prediction models under IFRS could mean that the strict accounting regulations under IFRS improve transparency, which prevents managers of firms facing insolvency from hiding their company's true situation by engaging in creative accounting practices or window dressing of the accounts. For example, IAS 38 constrains managers from capitalizing on certain intangible assets such as brands (IFRS Foundation 2021), thereby limiting the opportunity for the overstatement of total assets. Overall, improved transparency under IFRS should aid in providing a clearer picture to investors and creditors who can then make a sound decision on investing or lending funds to companies. For standard setters, our results provide empirical evidence of the benefits of aligning accounting standards towards IFRS, and how abandoning strict accounting practices could impact bankruptcy prediction. While there are several benefits of using IFRS, it may, however, generate extra costs for companies' accounting departments.

Author Contributions: Conceptualization, A.K. and R.R.W.; methodology, A.K. and R.R.W.; software, R.R.W.; validation, A.K. and R.R.W.; formal analysis, R.R.W.; investigation, A.K. and R.R.W.; resources, A.K. and R.R.W.; data curation, A.K. and R.R.W.; writing—original draft preparation, A.K.; writing—review and editing, A.K. and R.R.W.; visualization, A.K. and R.R.W. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We appreciate the many insightful comments from the editor, three anonymous reviewers, Trond Kristoffersen and Per Bjarte Solibakke. We are also grateful for the helpful comments from the conference participants at the NTNU Business School Conference 2018.

Conflicts of Interest: The authors declare no conflict of interest.

References

Agarwal, Vineet, and Richard Taffler. 2008. Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance* 32: 1541–51. [CrossRef]

Altman, Edward I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23: 558–609. [CrossRef]

Altman, Edward I. 1993. *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting & Avoiding Distress and Profiting from Bankruptcy*, 2nd ed. New York: Wiley.

Altman, Edward I., and Bettina Loris. 1976. A financial early warning system for over-the-counter broker-dealers. *The Journal of Finance* 31: 1201–17. [CrossRef]

Altman, Edward I., Robert G. Haldeman, and Pual Narayanan. 1977. ZETA analysis a new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance* 1: 29–54. [CrossRef]

Altman, Edward I., Giancarlo Marco, and Franco Varetto. 1994. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neutral networks (the Italian experience). *Journal of Banking and Finance* 18: 505–29. [CrossRef]

Altman, Edward I., Young Ho Eom, and Dong Won Kim. 1995. Failure prediction: Evidence from Korea. *Journal of International Financial Management and Accounting* 6: 230–49. [CrossRef]

Altman, Edward I., Gabriele Sabato, and Nicholas Wilson. 2010. The value of non-financial information in small and medium-sized enterprise risk management. *The Journal of Credit Risk* 6: 95–127. [CrossRef]

Altman, Edward I., Malgorzata Iwanicz-Drozdowska, Erkki K. Laitinen, and Arto Suvas. 2017. Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management and Accounting* 28: 131–71. [CrossRef]

Appiah, Kingsley Opoku, Amon Chizema, and Joseph Arthur. 2015. Predicting corporate faliure: A systematic literature review of methodological issues. *International Journal of Law and Management* 57: 461–85. [CrossRef]

Ates, Sinem. 2021. Accounting Convergence, Mandatory Adoption of IFRS and Accounting Quality. In *Handbook of Research on Institutional, Economic, and Social impacts of Globalization and Liberalization*. Edited by Yilmaz Bayar. Hershey: IGI Global, pp. 605–27.

Aziz, Abdul, David C. Emanuel, and Gerald H. Lawson. 1988. Bankruptcy prediction—An investigation of cash flow based models. *Journal of Management Studies* 25: 419–37. [CrossRef]

Bank of International Settlements. 2017. *Basel III: Finalising Post-Crisis Reforms*. Report. Basel: Bank of International Settlements.

Bărbută-Misu, Nicoleta, and Mara Madaleno. 2020. Assessment of Bankruptcy Risk of Largen Companies: European Countries Evolution Analysis. *Journal of Risk and Financial Management* 13: 58. [CrossRef]

Barth, Mary E., Wayne R. Landsman, and Mark H. Lang. 2008. International accounting standards and accounting quality. *Journal of Accounting Research* 46: 467–98. [CrossRef]

Bartov, Eli, Stephen R. Goldberg, and Myungsun Kim. 2005. Comparative value relevance among German, U.S., and International Accounting Standards: A German stock market perspective. *Journal of Accounting, Auditing & Finance* 20: 95–119.

Begley, Joy, Jin Ming, and Susan Watts. 1996. Bankruptcy classification errors in the 1980′s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies* 1: 267–84. [CrossRef]

Bernhardsen, Eivind, and Kai Larsen. 2007. Modelling credit risk in the enterprise sector – further development of the SEBRA model. *Economic Bulletin* 78: 102–8.

Bhat, Gauri, Jeffrey L. Gallen, and Dan Segal. 2014. Credit risk and IFRS: The case of Credit Default Swaps. *Journal of Accounting, Auditing & Finance* 29: 129–62. [CrossRef]

Blum, Marc. 1974. Failing company discriminant analysis. *Journal of Accounting Research* 12: 1–25. [CrossRef]

Bodle, Kerry Anne, Patti J. Cybinski, and Reza Monem. 2016. Effect of IFRS adoption on financial reporting quality: Evidence from bankruptcy prediction. *Accounting Research Journal* 29: 292–312. [CrossRef]

Botosan, Christine A., and Marlene A. Plumlee. 2002. A Re-examination of Disclosure Level and the Expected Cost of Equity Capital. *Journal of Accounting Research* 40: 21–40. [CrossRef]

Byard, Donal, Ying Li, and Yong Yu. 2010. The effect of mandatory IFRS adoption on financial analysts' information environment. *Journal of Accounting Research* 49: 69–96. [CrossRef]

Byrd, Richard H., Peihuang LU, Jorge Nocedal, and Ciyou Zhu. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16: 1190–208. [CrossRef]

Campbell, John Y., Jens Hilscher, and Jan Szilagyi. 2008. In search of distress risk. *The Journal of Finance* 63: 2899–939. [CrossRef]

Charitou, Andreas, Irene Karamanou, and Neophytos Lambertides. 2015. Who are the losers of IFRS adoption in Europe? An empirical examination of the cash flow effect of increased disclosure. *Journal of Accounting, Auditing & Finance* 30: 150–80. [CrossRef]

Choi, Young-Soo, Ken Peasnell, and Joao Toniato. 2013. Has the IASB been successful in making accounting earnings more useful for prediction and valuation? UK evidence. *Journal of Business Finance & Accounting* 40: 741–68. [CrossRef]

Christiansen, Hans B, Edward Lee, Martin Walker, and Cheng Zheng. 2015. Incentives or standards: What determines accounting quality changes around IFRS adoption? *European Accounting Review* 24: 31–61. [CrossRef]

Cordazzo, Michela, and Paola Rossi. 2020. The influence of IFRS mandatory adoption on value relevance of intangible assets in Italy. *Journal of Applied Accounting Research* 21: 415–36. [CrossRef]

Deakin, Edward B. 1972. A discriminant analysis of predictors of business failure. *Journal of Accounting Research* 10: 167–79. [CrossRef]

Deakin, Edward B. 1976. Distributions of fnancial accounting ratios: Some empirical evidence. *The Accounting Review* 51: 90–96.

Diamond, Douglas W., and Robert E. Verrecchia. 1991. Disclosure, Liquidity and the Cost of Capital. *The Journal of Finance* 46: 1325–59. [CrossRef]

Dimitras, Augustinos I., Roman Slowinski, Robert Susmaga, and Constantin Zopounidis. 1999. Business faliure prediction using rough sets. *European Journal of Operational Research* 114: 263–80. [CrossRef]

Duffie, Darrell, Leandro Saita, and Ke Wang. 2007. Multi-period corporate default prediction withstochastic covariates. *Journal of Financial Economics* 83: 635–65. [CrossRef]

Eisenbeis, Robert A. 1977. Pitfalls in the application of discriminant analysis in Business, Finance and Economics. *The Journal of Finance* 32: 875–900. [CrossRef]

Fan, Jingting, and Sebnem Kalemli-Ozcan. 2015. Emergence of Asia: Reforms, corporate savings, and global imbalances. *IMF Economic Review* 64: 239–67. [CrossRef]

Fawcett, Tom. 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27: 861–74. [CrossRef]

Florou, Annita, and Urska Kosi. 2015. Does mandatory IFRS adoption facilitate debt financing? *Review of Accounting Studies* 20: 1407–56. [CrossRef]

Florou, Annita, Urska Kosi, and Peter F. Pope. 2017. Are international accounting standards more credit relevant than domestic standards? *Accounting and Business Research* 47: 1–29. [CrossRef]

Fossung, Micheal Frozeh, Lious Agbor Tabot Ntuong, Helena Maria Santos de Oliveira, Cláudia Maria Ferreira Pereira, Susana Adelina Moreira Carvalho Bastos, and Liliana Marques Pimentel. 2020. Transition to the Revised OHADA Law on Accounting and Financial Reporting: Corporate Perceptions of Costs and Benefits. *Journal of Risk and Financial Management* 13: 172. [CrossRef]

Franzen, Laurel A., Kimberly J. Rodgers, and Timothy T. Simin. 2007. Measuring Distress Risk: The Effectof R&D Intensity. *The Journal of Finance* 62: 2391–967. [CrossRef]

Gareth, James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2017. *An Introduction to Statistical Learning: With Applications in R*, 7th ed. New York: Springer.

George, Emmanuel T. De, Xi Li, and Lakshmanan Shivakumar. 2016. A review of the IFRS adoption literature. *Review of Accoutning Studies* 21: 898–1004. [CrossRef]

Gjerde, Øystein, Kjell Knivsflå, and Frode Sættem. 2008. The value-relevance of adopting IFRS: Evidence from 145 NGAAP restatements. *Journal of International Accounting, Auditing and Taxation* 17: 97–112. [CrossRef]

Grice, John Stephen, and Robert W. Ingram. 2001. Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research* 54: 53–61. [CrossRef]

Guermazi, Walid, and Khamoussi Halioui. 2020. Do differences in national cultures affect cross-country conditionalconservatism behavior under IFRS? *Research in International Business and Finance* 52: 1–13. [CrossRef]

Güleç, Ömer Faruk. 2021. Value Relevance of Intangibles: A Literature Review. In *Ethics and Sustainability in Accounting and Finance*. Edited by Kıymet Tunca Çalıyurt. Singapore: Springer Nature Singapore Pte Ltd.

Günther, Nina, Bernhard Gegenfurtner, Christoph Kaserer, and Ann-Kristin Achleitner. 2009. *International Financial Reporting Standards and Earnings Quality: The Myth of Voluntary vs. Mandatory Adoption*. Working Paper. München: Center for Entrepreneurial and Financial Studies (CEFS), Technische Universität München, Available online: https://www.econstor.eu/handle/10419/48427 (accessed on 15 February 2021).

Gupta, Jairaj, Andros Gregoriou, and Tahera Ebrahimi. 2018. Empirical comparison of hazard models in predicting SMEs failure. *Quantitative Finance* 18: 437–66. [CrossRef]

Hillegeist, Stephen A., Elizabeth K. Keating, Donald P. Cram, and Kyle G. Lundstedt. 2004. Assessing the probability of bankruptcy. *Review of Accounting Studies* 9: 5–34. [CrossRef]

Hofstede Insights. 2020. Country Comparison. Available online: https://www.hofstede-insights.com/country-comparison/norway,sweden/ (accessed on 1 January 2021).

Horten, Joanne, and George Serafeim. 2010. Market reaction to and valuation of IFRS reconciliation adjustments: First evidence from the UK. *Review of Accounting Studies* 15: 725–51. [CrossRef]

Hosmer, David W., Stanley Lameshow, and Rodney X. Sturdivant. 2013. *Applied Logistic Regression*. New York: John Wiley & Sons.

Huang, Jin, and Charles X. Ling. 2005. Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering* 17: 299–310. [CrossRef]

Hung, Mingyi, and K. R. Subramanyam. 2007. Financial statement effects of adopting international accounting standards: The case of Germany. *Review of Accounting Studies* 12: 623–57. [CrossRef]

IFRS Foundation. 2016. IFRS Application around the World. Available online: https://www.ifrs.org/-/media/feature/around-the-world/jurisdiction-profiles/sweden-ifrs-profile.pdf (accessed on 15 February 2021).

IFRS Foundation. 2020. IFRS-Who We Are. Available online: https://www.ifrs.org/about-us/who-we-are/#history (accessed on 15 February 2021).

IFRS Foundation. 2021. *IFRS Standards*. London: IFRS Foundation.

International Standards Accounting Board. 2002. *International Accounting Standards Comittee Foundation*. London: International Standards Accounting Board.

Joy, O. Maurice, and John O. Tollefson. 1975. On the financial applications of discriminant analysis. *The Journal of Financial and Quantitative Analysis* 10: 723–39. [CrossRef]

Kaastra, Iebeling, and Milton Boyd. 1996. Designing a neural network for forecasting financial and economic time series. *Neurocomputing* 10: 215–36. [CrossRef]

Kanagaretnam, Kiridaran, Chee Yeow Lim, and Gerald J. Lobo. 2014. Influence of national culture on accounting conservatism and risk-taking in the banking industry. *The Accounting Review* 89: 1115–49. [CrossRef]

Keles, Dogan, Jonathan Scelle, Florentina Paraschiv, and Wolf Fichtner. 2016. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Applied Energy* 162: 218–30. [CrossRef]

Ketz, J. Edward. 1978. The effect of general price-level adjustments on the predictive ability of financial ratios. *Journal of Accounting Research* 16: 273–84. [CrossRef]

Kraft, Pepa, and Wayne R. Landsman. 2020. Effect of mandatory IFRS adoption on accounting-based prediction models for CDS spreads. *European Accounting Review*. [CrossRef]

Kristoffersen, Trond. 2020. Opphev bestemmelsene om grunnleggende regnskapsprinsipper. *Revisjon og Regnskap* 4: 41–47.

Kwon, Soo Young, Kyunga Na, and Jongil Park. 2019. The economic effects of IFRS adoption in Korea. *Asia-pacific Journal of Accounting and Economics* 26: 321–61. [CrossRef]

Lambert, Richard, Christian Leuz, and Robert E. Verrecchia. 2007. Accounting information, disclosure and the cost of capital. *Journal of Accounting Research* 45: 385–420. [CrossRef]

Levitt, Arthur. 1998. The importance of high quality accounting standards. *Accounting Horizons* 12: 79–82.

Libby, Robert. 1975. Accounting ratios and the prediction of failure: Some behavioral evidence. *Journal of Accounting Research* 13: 150–61. [CrossRef]

Mansi, Sattar A., William F. Maxwell, and Andrew Zhang. 2012. Bankruptcy prediction models and the cost of debt. *The Journal of Fixed Income* 21: 25–42. [CrossRef]

Marton, Jan. 2017. The Role and Current Status of IFRS in the Completion of National Accounting Rules—Evidence from Sweden. *Accounting in Europe* 14: 207–16. [CrossRef]

Masoud, Najeb. 2017. The effects of mandatory IFRS adoption on financial analysts' forecast: Evidence from Jordan. *Cogent Business and Management* 4: 1290331. [CrossRef]

McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behaviour. In *Frontier in Econometrics*. Edited by Paul Zarembka. New York: New York Academic Press, pp. 105–42.

Meyer, Paul A., and Howard W. Pifer. 1970. Prediction of bank failures. *The Journal of Finance* 25: 853–68. [CrossRef]

Moyer, R. Charles. 1977. Forecasting financial failure: A re-examination. *Financial Management* (pre-1986) 6: 11–17. [CrossRef]

Ntoung, Lious Agbor Tabot, Helena Maria Santos de Oliveira, Benjamim Manuel Ferreira de Sousa, Liliana Marques Pimentel, and Susana Adelina Moreira Carvalho Bastos. 2020. Are Family Firms Financially Healthier Than Non-Family Firm? *Journal of Risk and Financial Management* 13: 5. [CrossRef]

Ohlson, James A. 1980. Financial ratios and the probablistic prediction of bankruptcy. *Journal of Accounting Research* 18: 109–31. [CrossRef]

Oliveira, Lidia, Lucia Lima Rodriguez, and Russel Craig. 2010. Intangible assets and value relevance: Evidence from the Portuguese stock exchange. *The British Accounting Review* 42: 241–52. [CrossRef]

Orbis. 2007. *ORBIS Internet User Guide*. Brussels: Bureau van Dijk Electronic Publishing.

Norwegian Accounting Standards Board. 2012. NRS 19 Immaterielle Eiendeler. Available online: https://www.regnskapsstiftelsen.no/regnskap/regnskapsstandarder/nrs-19-immaterielle-eiendeler (accessed on 8 February 2021).

Paolone, Francesco, Ricardo Tiscini, and Laura Martiniello. 2020. Accounting and Value Relevance of Intangibles: A longitudinal Study in Italian Stock Exchange 2010–2018. *Journal of Modern Accounting and Auditing* 16: 201–10. [CrossRef]

Paraschiv, Florentina, Markus Schmid, and Ranik Raaen Wahlstrøm. 2021. *Bankruptcy Privately Held SME's Using Feature Selection Methods*. Working Paper. Trondheim: NTNU Business School, Norwegian University of Science and Technology, St. Gallen: University of St. Gallen.

Pettway, Richard H., and Joseph F. Sinkey. 1980. Establishing on-site bank examination priorities: An early warning system using accounting and market information. *The Journal of Finance* 35: 137–50. [CrossRef]

Picker, Ruth, Kerry Clark, John Dunn, David Kolitz, Gilad Livne, Janice Loftus, and Leo van der Tas. 2016. *Applying IFRS Standards*, 4th ed. Great Britain: Wiley.

Ryan, Thomas P. 2018. *Modern Regression Methods*, 2nd ed. Hoboken: John Wiley & Sons.

Shyam-Sunder, Lakshmi, and Stewart C. Myers. 1999. Testing static tradeoff against pecking order models of capital structure. *Journal of Financial Economics* 51: 219–44. [CrossRef]

Singleton-Green, Brian. 2015. *The Effects of Mandatory IFRS Adoption in the EU: A Review of Empirical Research. Information for Better Markets*. ICAEW Financial Reporting Quality. London: ICAEW.

Tan, Hongping, Shiheng Wang, and Michael Welker. 2011. Analyst following and forecast accuracy after mandated IFRS adoptions. *Journal of Accounting Research* 49: 1307–57. [CrossRef]

Tian, Shaonan, and Yan Yu. 2017. Financial ratios and bankruptcy prediction: An international Evidence. *International Review of Economics and Finance* 51: 510–26. [CrossRef]

Tian, Shaonan, Yan Yu, and Hui Guo. 2015. Variable selection and corporate bankruptcy forecasts. *Journal of Banking and Finance* 52: 89–100. [CrossRef]

Vo, Duc Hong, Binh Ninh Vo Pham, Chi Minh Ho, and Michael McAleer. 2019. Corporate Financial Distress of Industry Level Listings in Vietnam. *Journal of Risk and Financial Management* 12: 155. [CrossRef]

Wang, Xin, Danqing Young, and Zili Zhuang. 2008. *The Effects of Mandatory Adoption of International Financial Reporting Standards on Information Environments*. Working Paper. Hong Kong: The Chinese University of Hong Kong.

Wilcox, Jarrod W. 1973. A prediction of business failure using accounting data. *Journal of Accounting Research* 11: 163–79. [CrossRef]

Wu, Joanna Shuang, and Ivy Xiying Zhang. 2014. The adoption of Internationally Recognized Accounting Standards: Implications for the credit markets. *Journal of Accounting, Auditing and Finance* 29: 95–128. [CrossRef]

Zhu, Ciyou, Richard H. Byrd, Peihuang Lu, and Jorge Nocedal. 1997. Algorithm 7788: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software* 23: 550–60. [CrossRef]

Zmijewski, Mark E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22: 59–82. [CrossRef]

Article 3

Title:
Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods

Authors:
Florentina Paraschiv
Markus Schmid
Ranik Raaen Wahlstrøm

Reference:
Paraschiv, F., Schmid, M., Wahlstrøm, R.R., 2021. Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods. Working Paper, Norwegian University of Science and Technology and University of St. Gallen, to be submitted to the Review of Finance.

ISBN 978-82-326-5700-1 (printed ver.)
ISBN 978-82-326-6261-6 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology