Finance Research Letters xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Approximating risk-free curves in sparse data environments☆

C.J. van der Merwe[*,a,b], D. Heyman[b], T. de Wet[a]

[a] Department of Statistics and Actuarial Science, Stellenbosch University, South Africa
[b] Department of Financial Economics, Ghent University, Belgium

ARTICLE INFO

Keywords:
Sparse data
Fair value
Simulation
Risk-free curves

JEL classification:
C530
C580
G120
M410
M420

ABSTRACT

Accounting standards require one to minimize the use of unobservable inputs when calculating fair values of financial assets and liabilities. In emerging markets and less developed countries, zero curves are not as readily observable over the longer term, as data are often more sparse than in developed countries. A proxy for the extended zero curve, calculated from other observable inputs, is found through a simulation approach by incorporating two new techniques, namely permuted integer multiple linear regression and aggregate standardized model scoring. A Nelson Siegel fit, with a mixture of average forward rates as proxies for the long term zero point, and some discarding of initial data points, was found to perform relatively well in the training and testing data sets.

1. Introduction

The International Financial Reporting Standards (IFRS) 13 states that when a fair value is calculated, the use of observable inputs should be maximized. These inputs are classified into three levels, with the highest, or first level, being broadly directly observable inputs, the second level inputs that are derived via models from other observable inputs, and the third, or lowest level, unobservable inputs.

Some emerging markets and less developed countries, however, lack these observable inputs. While there exists a vast amount of research on parameterizing the yield curve, extrapolating it, and forecasting it as well, very little research has been and is being done on extending it in sparse data environments. The approach followed in this research was to simulate such sparse environments from data rich environments and find methods that perform well in extrapolating curves under these conditions.

This paper focuses specifically on zero risk-free curves, and finds a proxy that can be used to obtain the extended curve through other observable inputs and specified models. The approximations are found through a phased simulation approach which incorporates two new techniques, namely permuted integer multiple linear regression (PIMLR) and aggregate standardized model scoring (ASMS).

Kumarasiri and Fisher (2011) summarized Pacter's (2007) concerns regarding the application of fair value measurement in developing countries. The first being that inactive markets cause unreliable fair value estimates due to infrequent transactions, large bid-ask spreads, and market prices only being influenced by a few market participants or transactions. Secondly, there exists a trade-off between cost and benefit of implementing sophisticated fair valuation techniques. Furthermore, there are significant skill shortages in these countries - not only in-house, but externally as well. Lack of valuation standards and guidance on how to determine fair value also raises additional concerns.

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Conflicts of interest: none.
* Corresponding author at: Department of Statistics and Actuarial Science, Stellenbosch University, South Africa.
E-mail addresses: cjvdmerwe@sun.ac.za (C.J. van der Merwe), Dries.Heyman@UGent.be (D. Heyman), tdewet@sun.ac.za (T. de Wet).

https://doi.org/10.1016/j.frl.2017.12.016
Received 7 July 2017; Accepted 17 December 2017
1544-6123/

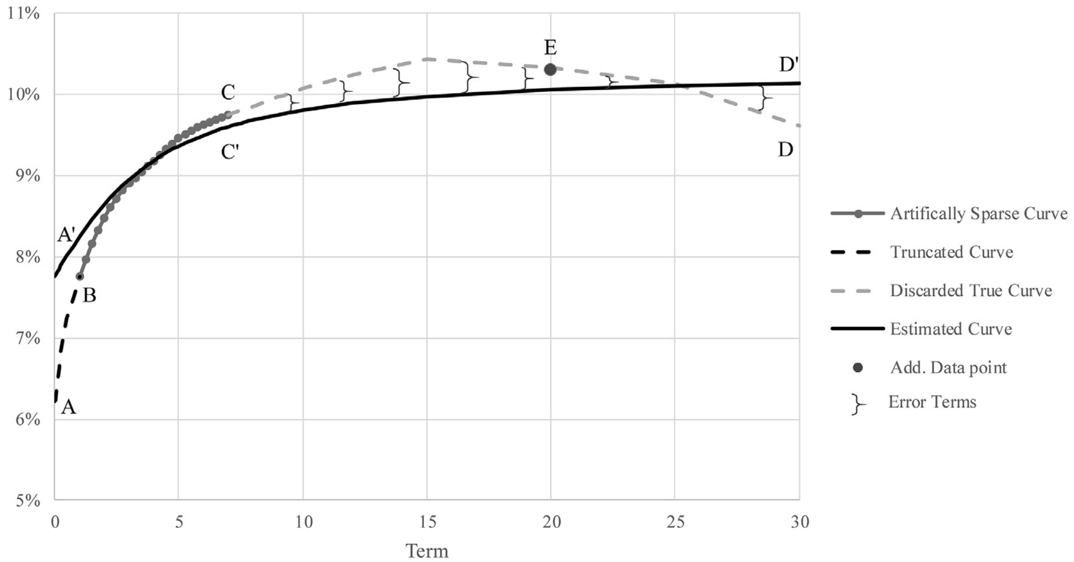

Fig. 1. Example of the various input data used in Phase II, where the Additional Data point is an output from Phase I.

They identified various areas for future research, one of which includes the extension of their study to other developing countries due to the concerns raised regarding the credibility of financial statements prepared on the basis of fair value accounting in developing markets. They further stated that there is a perceived lack of technical guidance for preparers and auditors regarding fair valuation in developing countries, and that further research should consider the optimal nature, form, and source of such guidance.

Palea and Maino (2013) investigated whether the application of IFRS 13 for private equity valuation actually does contribute to the enhancing of transparency and comparability in financial statements. Benston (2008) found that fair values other than those that are directly observable in the market, could be manipulated easily and often are difficult to verify. Laux and Leuz (2009) discussed the different views regarding fair value accounting, and pointed to further research. Barth (2004) investigated the impact of the volatility of estimates on financial statements due to fair value, Penman (2007) discussed the benefits and disadvantages of fair value over historical cost estimates for various assets, while Ryan (2008) criticized the definition and measurement of fair value during a financial crisis.

The paper is organized as follows: Section 2 discusses the research methodology, Section 3 describes the modeling framework, Section 4 provides a description of the data used in the study, Section 5 presents the main results obtained, and Section 6 concludes.[1]

2. Research methodology

A simulation study was designed and performed in order to find the overall average optimal proxy for the risk-free rates where there are no data in the longer end of the curve. This is done through fitting various different models to an artificially created sparse environment and comparing it to the actual observed rates. The models were then scored to determine the average best performing model across various data sets. Out-of-sample data sets were used to test for consistency of the overall results.

In Fig. 1 assume that the data represented by the line *AD* are known, and an artificially sparse environment is created by discarding *CD*. This allows for various curves (*A′D′*) to be fitted to the remaining *AC* data points. In order to improve a fit over the area of interest *CD*, some initial data (*AB*) is truncated and a data point *E* is added. The goodness of fit of the newly fitted curve *C′D′*, based on *BC* and *E*, is then measured by considering the squared differences between *CD* and *C′D′*. The single additional data point was chosen in order to 'pull' the longer term estimated curve towards the true values, something which would be difficult to accomplish by only considering the sparse data.

The simulation study was split into two phases, the estimation methods for the additional data point forms the basis of Phase I, while simulating the different variations of the fitted curve was done in Phase II.

The framework for the additional data point followed an approach whereby a training data set was used to obtain a number of additional data point models from observable data through permuted integer multiple linear regression (PIMLR). These models were then scored using the aggregate standardized model scoring (ASMS) technique in order to obtain the better performing regression output across the training set.

Once these models were obtained, they were incorporated into the simulation design. The simulation fitted a number of different Nelson Siegel and Svensson parameterizations to artificially created sparse data environments, and found the better performing

[1] This paper is a shortened version of the authors' working paper, which can be consulted for a more detailed description of methodologies and data used, as well as results. It is available from the authors upon request.