International Journal of Forecasting 34 (2018) 551–565

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Residual value forecasting using asymmetric cost functions

Korbinian Dress [a], Stefan Lessmann [a,*], Hans-Jörg von Mettenheim [b]

[a] Humboldt-University of Berlin - School of Business and Economics, Berlin, Germany
[b] Leibniz University Hannover - Institute of Information Systems, Hannover, Germany

ARTICLE INFO

Keywords:
Asymmetric cost of error
Quantile regression
Neural networks
Ensemble learning
Automotive industry

ABSTRACT

Leasing is a popular channel for marketing new cars. However, the pricing of leases is complicated because the leasing rate must embody an expectation of the car's residual value after contract expiration. This paper develops resale price forecasting models in order to aid pricing decisions. One feature of the leasing business is that different forecast errors entail different costs. The primary objective of this paper is to identify effective ways of addressing cost asymmetry. Specifically, this paper contributes to the literature by (i) consolidating prior work in forecasting on asymmetric functions of the cost of errors; (ii) systematically evaluating previous approaches and comparing them to a new approach; and (iii) demonstrating that forecasting using asymmetric cost of error functions improves the quality of decision support in car leasing. For example, if the costs of overestimating resale prices are twice those of underestimating them, incorporating cost asymmetry into forecast model development reduces costs by about 8%.

1. Introduction

This paper concentrates on decision support in car leasing and the prediction of resale prices. In the leasing business, it is the lessor's obligation to take back and re-market returned cars following the expiration of the leasing contract. The (discounted) sum of payments implies an expectation of the car's residual value. When pricing leasing contracts, vendors require forecasts of the residual values of the cars after contract expiration. The difference between the car's original list price and its residual value determines the leasing rate. Forecasting is essential for supporting pricing decisions in the leasing business and assisting lessors in securing profits (Du, Xie, & Schroeder, 2009).

Forecasts are never fully accurate, but tend to over- or underestimate resale prices. Underestimating resale prices yields unexpected profits when selling the returned car in the second-hand market, but also implies that the lessor could have offered a lower leasing rate. This is important because high prices might have deterred some customers from signing contracts with the lessor. Thus, underestimating resale prices is associated with an opportunity costs of lost sales. On the other hand, overestimating resale prices implies that the lessor realizes lower profits from the overall contracts, or even faces losses. Given that the effects of different forecast errors are different, there is little reason to believe that equivalent costs arise from these errors.

Granger (1969) was the first to suggest that real-world forecasting tasks are rarely characterized by quadratic (i.e., symmetric) error costs. Since then, several others have echoed Granger's (1969) criticism, developed asymmetric cost of error functions (ACEF), and demonstrated their potential to improve the decision quality through empirical experimentation (e.g., Crone, 2010; Diebold & Mariano, 1995; Leitch & Tanner, 1991). Lessmann (2013) arrived at a similar conclusion for resale price modelling in the automotive industry.

The objectives of this paper are threefold: (i) to consolidate and integrate previous work in forecasting using ACEF; (ii) to provide a systematic comparison of various previously used and newly developed modelling

* Correspondence to: Humboldt-Universität zu Berlin - School of Business and Economics Unter den Linden 6 Berlin 10099, Germany.
E-mail address: stefan.lessmann@hu-berlin.de (S. Lessmann).

https://doi.org/10.1016/j.ijforecast.2018.01.008
0169-2070/

approaches; and (iii) to demonstrate the potential of such approaches for the improvement of pricing decision in the car leasing business. In pursuing these objectives, the paper makes the following contributions. First, it reintroduces several modelling strategies that can account for unequal costs of positive and negative deviations, including quantile regression, artificial neural networks (ANN), and other learning algorithms. Often, the techniques under study have been developed independently and have not been compared with each other. Thus, our benchmark provides insights into the relative merits of various existing modelling approaches in the focal application context.

Second, the paper develops a conceptual framework for identifying the stage in the forecasting process that best accounts for cost asymmetry. A comparison of alternative strategies shows that it is better to consider asymmetry during model estimation than to correct the predictions ex post. In response to this, we propose a novel ensemble modelling approach.

Third, the paper examines the interaction between the internal functioning of a forecast model and strategies for addressing the asymmetric costs of errors by systematically comparing linear and nonlinear forecasting models, as well as individual and ensemble models. Previous forecasting benchmarks have also analysed the observed results along these dimensions; however, such evidence is missing when forecasting using ACEF.[1]

Based on a broad set of empirical experiments and a subsequent sensitivity analysis, we identify a sizeable potential for improving the quality of decision support in the presence of asymmetric costs. For example, if the error of overestimating resale prices is weighted twice as heavily as the reverse error, the most efficient ACEF modelling strategy improves the decision quality by about 8% relative to the best conventional prediction method, while the improvement is even more substantial for higher degrees of asymmetry.

The remainder of the paper is organized as follows: Section 2 reviews related work. Section 3 introduces popular ACEFs, after which Section 4 discusses strategies for accounting for asymmetric error costs during forecast model development. Section 5 introduces the data used in the empirical comparison of forecasting models, and Section 6 elaborates on the corresponding results. Section 7 concludes the paper.

2. Related literature

Jerenz (2008) developed a decision support system for guiding price setting in the used car business, and showed the potential of the corresponding approaches to increase the revenue. Jerenz (2008) and others modelled resale prices by employing a hedonic framework using an ordinary linear least squares regression (e.g., Erdem & Sentürk, 2009; Prado, 2009). Linear regression provides an explanatory model that reveals the degree to which different car characteristics (i.e., independent variables) affect resale prices. However, while explanatory models are beneficial for providing insights into the formation of prices, a drawback is that they display a lower level of accuracy than data-driven prediction models such as neural networks or ensemble models (Shmueli & Koppius, 2011). When considering the provision of pricing decision support, the dependence of the leasing rate, and profits more generally, on resale price forecasts suggest that it is crucial to maximize the predictive accuracy (Du et al., 2009).

Lian, Zhao, and Cheng (2003) pioneered the use of advanced forecasting methods for predicting resale prices. They used a neural network and let an evolutionary algorithm select the most important variables and suitable meta-parameters for the forecasting model. Lessmann and Voß(2017) compared 19 state-of-the-art prediction models under different conditions in order to determine whether advanced methods improve the forecast accuracy. Their results show that random forest regression is an especially effective method for resale price forecasting.

Lessmann and Voß(2017) and Lian et al. (2003) assessed the forecast accuracy using symmetric cost functions such as the mean squared error (MSE). However, Granger (1969) criticizes a mismatch between symmetric cost functions and the cost functions in real-world business scenarios. He also shows that forecasts that are based on quadratic error functions do not lead to optimal estimators for general cost functions,[2] and therefore, the estimation and assessment of forecasts should focus on actual economic costs (Christoffersen & Diebold, 1997; Diebold & Mariano, 1995; Granger & Newbold, 1986; Leitch & Tanner, 1991). More generally, the assessment of a decision support system should focus on the system's ability to improve the decision quality and business performance (Lilien, Rangaswamy, Van Bruggen, & Starke, 2004). Again, this calls for the use of evaluation criteria such as profits and costs (Bharadwaj, 2000), and calls into question the use of symmetric cost functions.

In car leasing, the use of symmetric cost functions is inappropriate because different forecast errors carry different costs. Recall that an underestimation of resale prices may lead to opportunity costs due to lost sales, whereas an overestimation of resale prices decreases profits when remarketing the returned car, or may even lead to a loss. Prospect theory suggests that decision makers weight losses more heavily than gains even if the two have the same magnitudes and occur with equal probabilities (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Therefore, in this case, a manager will consider overestimation the more severe forecast error.

Some studies have employed asymmetric cost functions in real-world forecasting applications. In the car leasing context considered here, Lessmann (2013) reported preliminary findings, on which this study elaborates further. Examples in other domains include the study by Tian (2009), who considered a quadratic asymmetric loss function when forecasting Australian unemployment rates. She

[1] We acknowledge that the use of ACEF is popular in cost-sensitive learning. However, the corresponding studies consider discrete, often binary, response variables, which is different from the forecasting of a continuous response (e.g., resale prices) that we consider in this work.

[2] Granger uses the term *cost function* rather than *error* or *loss function*. We do not distinguish between these terms, and use them interchangeably throughout the paper.