



Appl. Statist. (2018)
67, Part 5, pp. 1275–1304

Unravelling the predictive power of telematics data in car insurance pricing

Roel Verbelen,
KU Leuven, Belgium

Katrien Antonio
KU Leuven, Belgium, and University of Amsterdam, The Netherlands

and Gerda Claeskens
KU Leuven, Belgium

[Received January 2017. Revised March 2018]

Summary. A data set from a Belgian telematics product aimed at young drivers is used to identify how car insurance premiums can be designed based on the telematics data collected by a black box installed in the vehicle. In traditional pricing models for car insurance, the premium depends on self-reported rating variables (e.g. age and postal code) which capture characteristics of the policy(holder) and the insured vehicle and are often only indirectly related to the accident risk. Using telematics technology enables tailor-made car insurance pricing based on the driving behaviour of the policyholder. We develop a statistical modelling approach using generalized additive models and compositional predictors to quantify and interpret the effect of telematics variables on the expected claim frequency. We find that such variables increase the predictive power and render the use of gender as a rating variable redundant.

Keywords: Compositional predictors; Generalized additive models; Pay as you drive insurance; Risk classification; Structural 0s; Usage-based insurance

1. Introduction

For a unique Belgian portfolio of young drivers in the period between 2010 and 2014, telematics data on how many kilometres are driven, during which time slots and on which type of roads were collected using black box devices installed in the insured drivers' cars. Our aim is to incorporate this information in statistical rating models, where we focus on predicting the number of claims, to set premium levels adequately on the basis of individual policyholder's driving habits.

Determining a fair and correct price for an insurance product (also called *rate making*, *pricing* or *tarification*) is crucial for both insured drivers and insurance companies. Car insurance is traditionally priced on the basis of self-reported information from the insured person, most importantly age, licence age, postal code, engine power, use of the vehicle and claims history. However, these observable risk factors are only proxy variables, not reflecting current driving habits and driving style. Telematics technology—the integrated use of telecommunication and informatics—may fundamentally change the car insurance industry. The use of this technology

Address for correspondence: Roel Verbelen, ORSTAT, KU Leuven, Naamrestraat 69, Bus 3555, 3000 Leuven, Belgium.
E-mail: roel.verbelen@kuleuven.be

© 2018 The Authors Journal of the Royal Statistical Society: Series C (Applied Statistics) 0035–9254/18/671275
Published by John Wiley & Sons Ltd on behalf of the Royal Statistical Society.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.