Milliman

Usage-Based Insurance: A European Case Study using Machine Learning

Peggy Brinkmann, FCAS, MAAA

CAS Ratemaking and Product Management Seminar

Concurrent Session 5

March 16, 2016

Antitrust Notice

- The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.
- Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.
- It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

Outline

- UBI a la Europa
- Machine Learning Case Study

UBI Value propositions vary

UK: 35M cars, 1% UBI
Young driver mileage accounts, Mileage bonus for safe driving

CH: 4.3M cars, >1% UBI
Crash recorder (Young driver)
Theft protection (Luxury cars)
Emergency call (Safety)

France: 32M cars, <1% UBI
Attractive offerings for safe drivers
Roadside assistance

Spain: 22M cars, <1% UBI
Fleets
Track & Trace

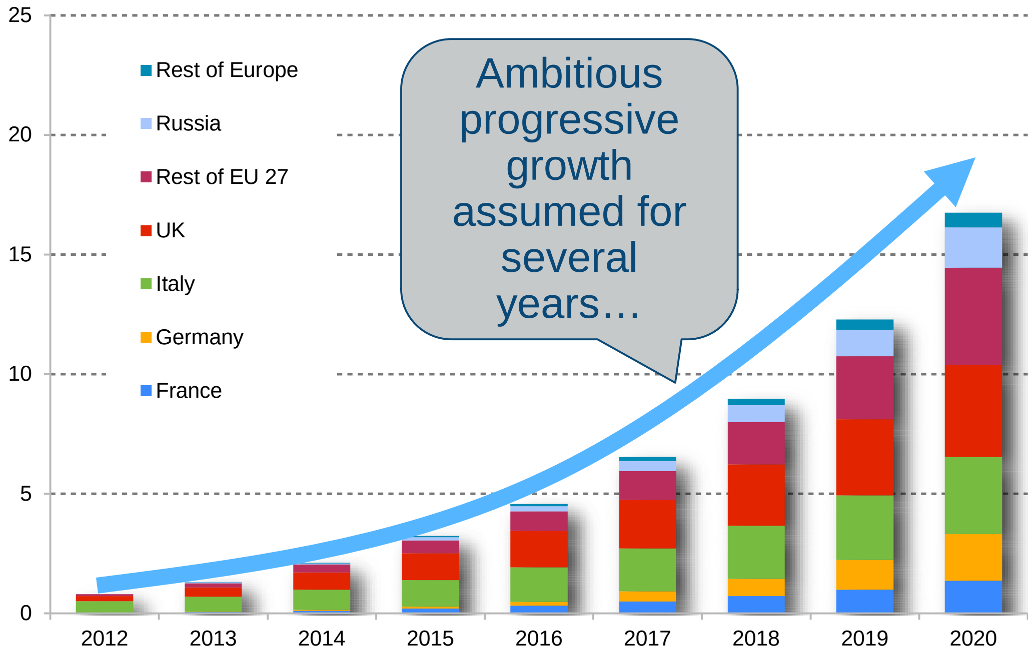
Germany: 43M cars, 0.005%
Strong privacy concerns

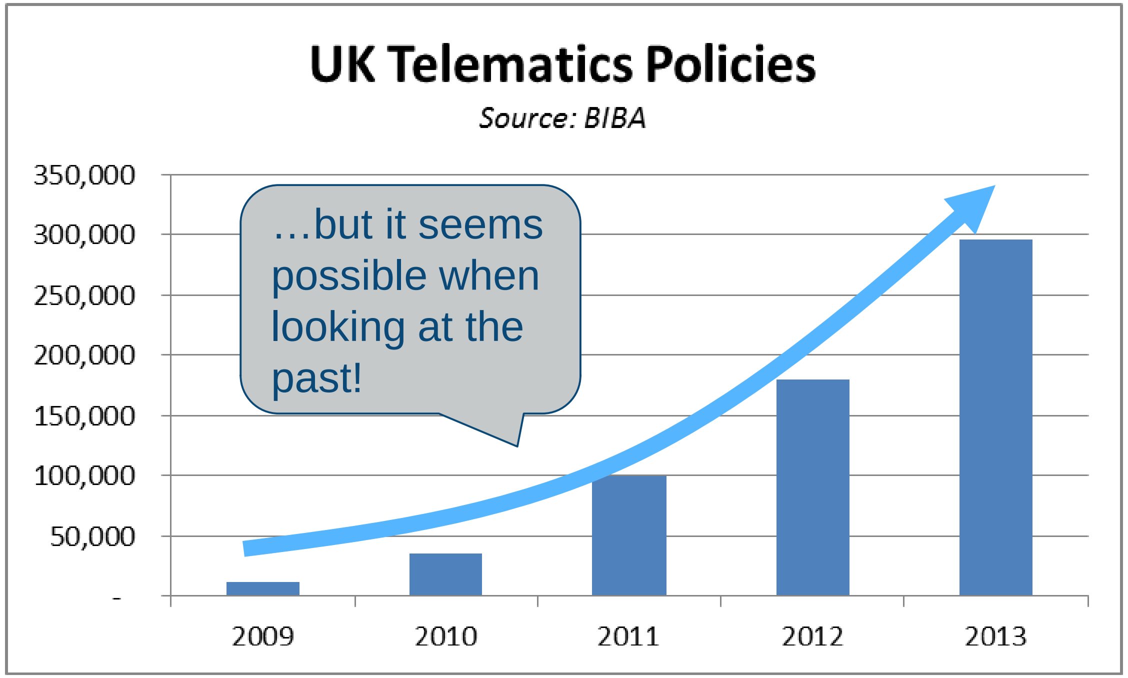
Austria: 4.5M cars, >1% UBI
Technology-friendly country, Safety first

Italy: 37M cars, 4% UBI
Low mileage discounts

Sources: 1) PRNewswire, Insight Report: Technology in Action - A Roadmap for Insurance Telematics
2) United Nations Economic Commission for Europe

Expected UBI growth





The Insurance Telematics (or UBI) will represent more than 35 million policies in 2020 or around 15% of the European personal lines market.

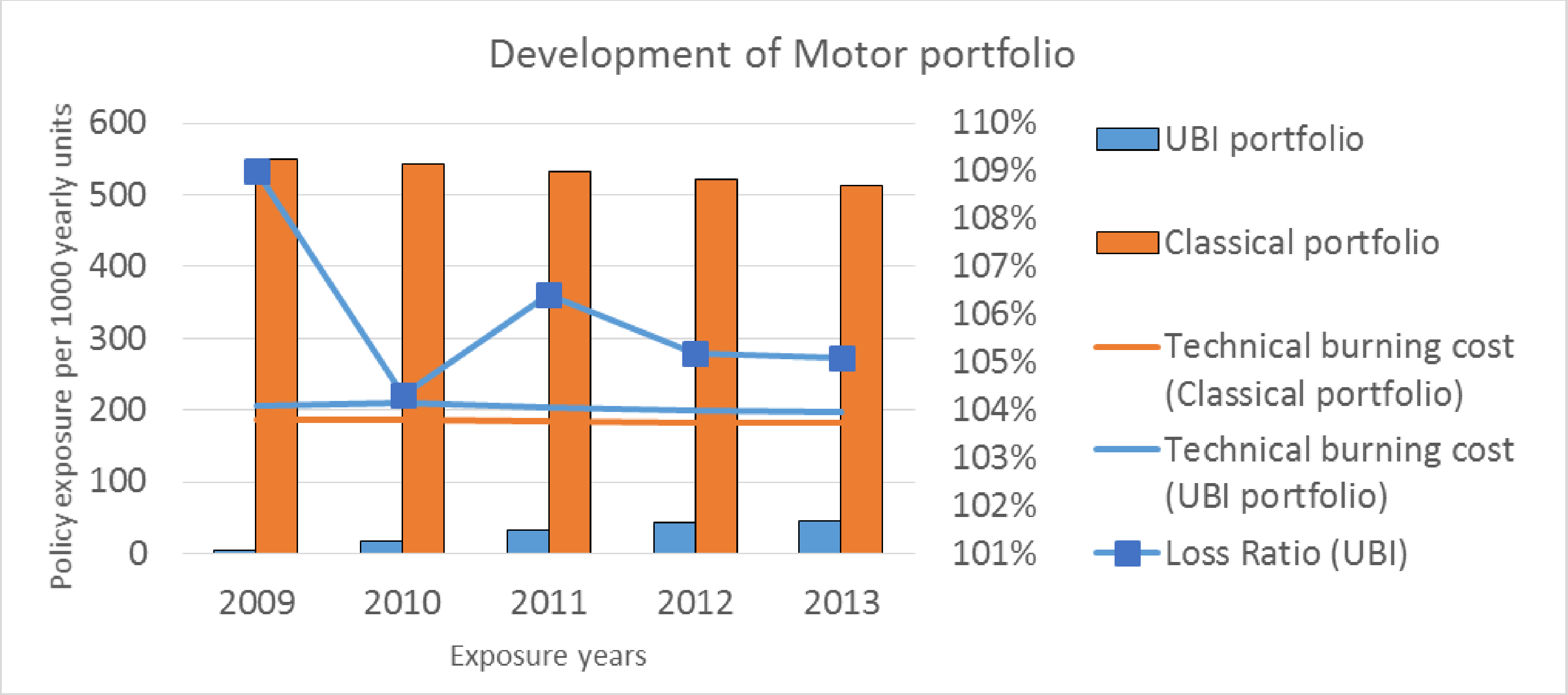
Source: 1) Ptolemus USAGE-BASED INSURANCE Global Study 2013
2) British Insurance Brokers' Association (BIBA)

Insurers weighing benefits and barriers

Benefits	Barriers
–Attractive pricing	–Technology choices/cost
–Claims handling	–IT infrastructure/costs
–Enhance product	–Privacy/data ownership

Case study: Background

- European client with private UBI Motor business
- UBI portfolio loss ratio 5% higher than standard portfolio
- UBI pricing model built using GLM



Case Study: two questions, two models

1. How well does the current GLM price the UBI business?
   Target = Loss / Current GLM-based premium
   Predictors = Classical and UBI rating variables
   Method = Generalized Boosted Model (GBM)

2. What are the most profitable/unprofitable UBI segments?
   Target = Loss / GBM prediction excluding UBI variables
   Predictors = Classical and UBI rating variables
   Method = Regression Tree

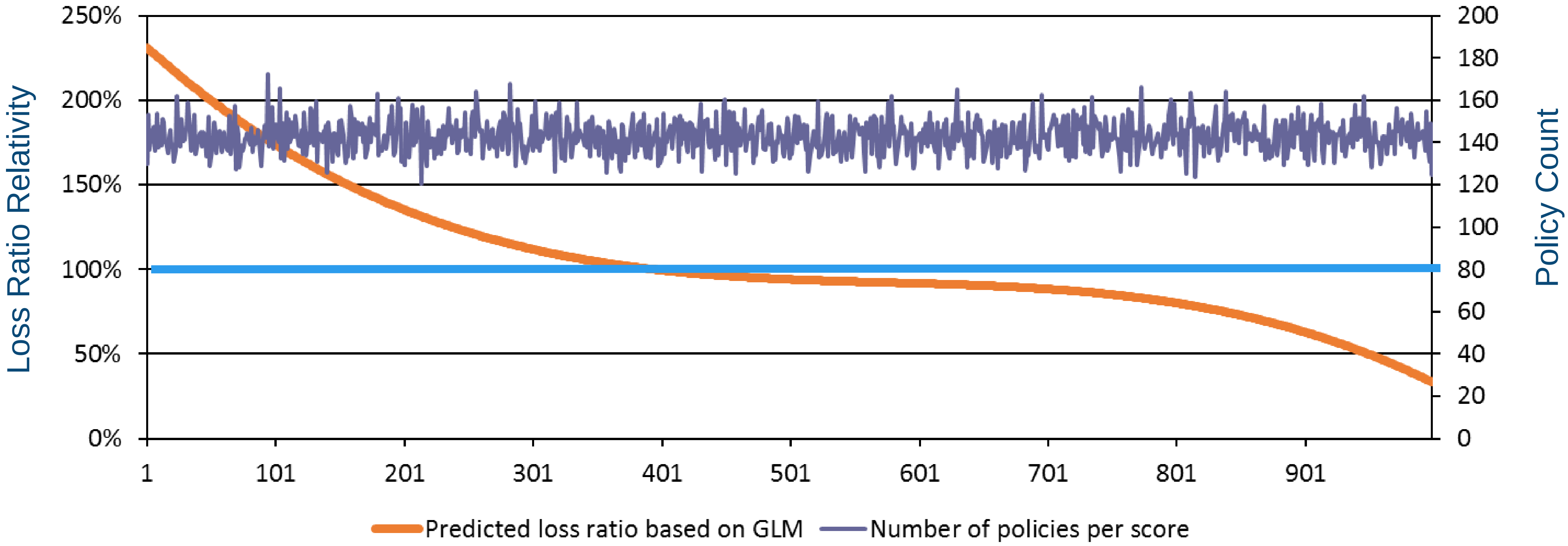
Why use machine learning?

- UBI devices capture a lot of data - automated approaches for variable selection can be useful
- Machine learning techniques are useful for detecting interaction, and UBI variables interact with each other and traditional rating characteristics
- For example: drivers with a speeding violation are on average worse risks but
  - Some speed on highways; some on rural road
  - Some speed constantly and got caught once; some just had a bad day
  - Some speed during the day; some speed at night
- In other words, the importance of this indicator (having a speeding violation) will be different for different drivers, and the dependencies become ever more important with additional data

Why GBMs and trees?

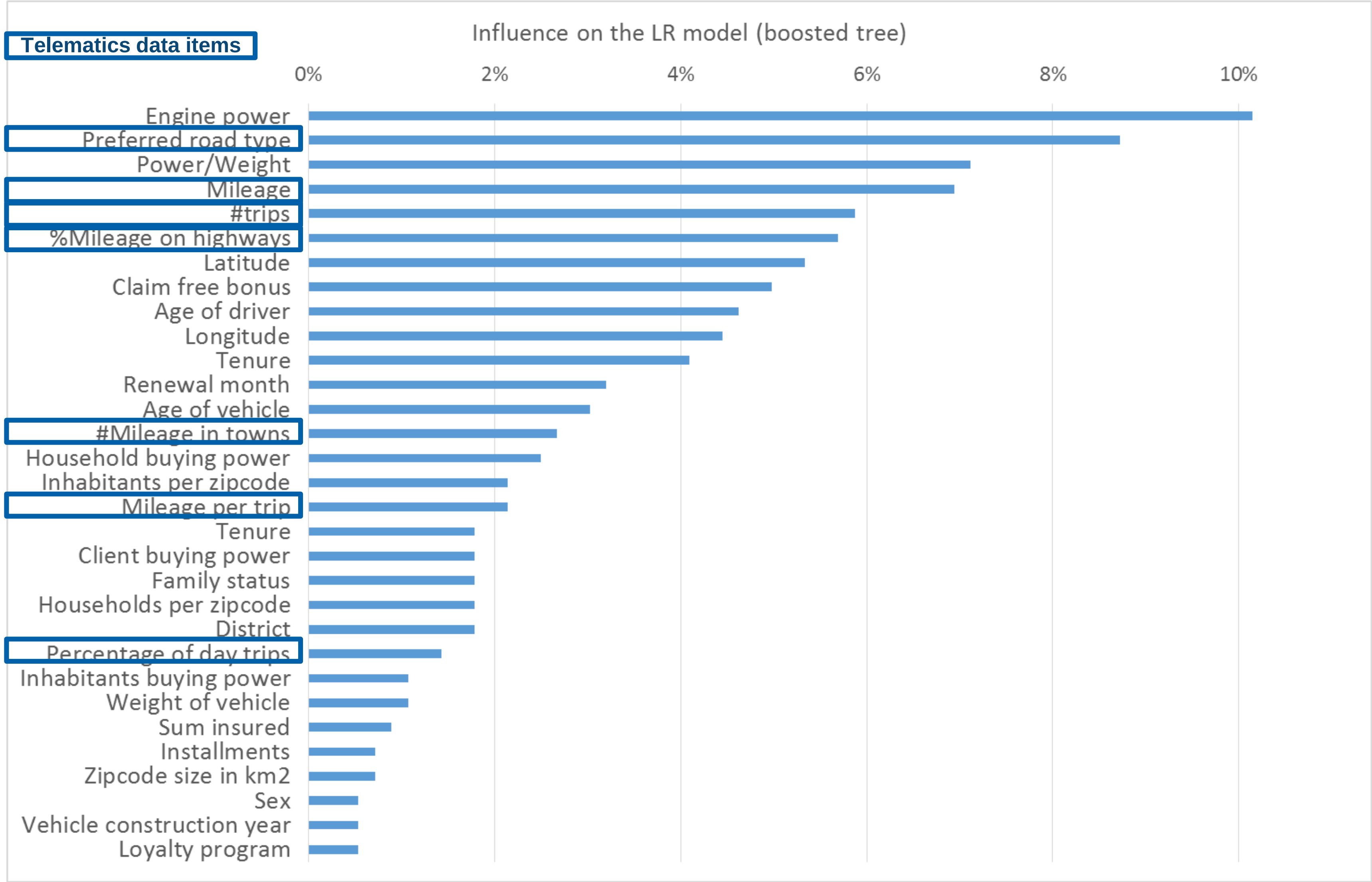
- Machine learning has many, many approaches
- Trees are useful because:
  - Trees are all about local interactions.
  - Single trees can be simple and transparent. Relationships are there to see.
- GBMs (boosted trees) are
  - Smooth and powerful, the results stable
  - Transparent, even if they are complex.
- Remember that all automated routines run an extra risk of overfitting the data. You *must* validate these models.

Model 1 results



The score on the x-axis represents the score from the GBM, from highest expected loss ratio to lowest expected loss ratio.

Model 1 variable importance



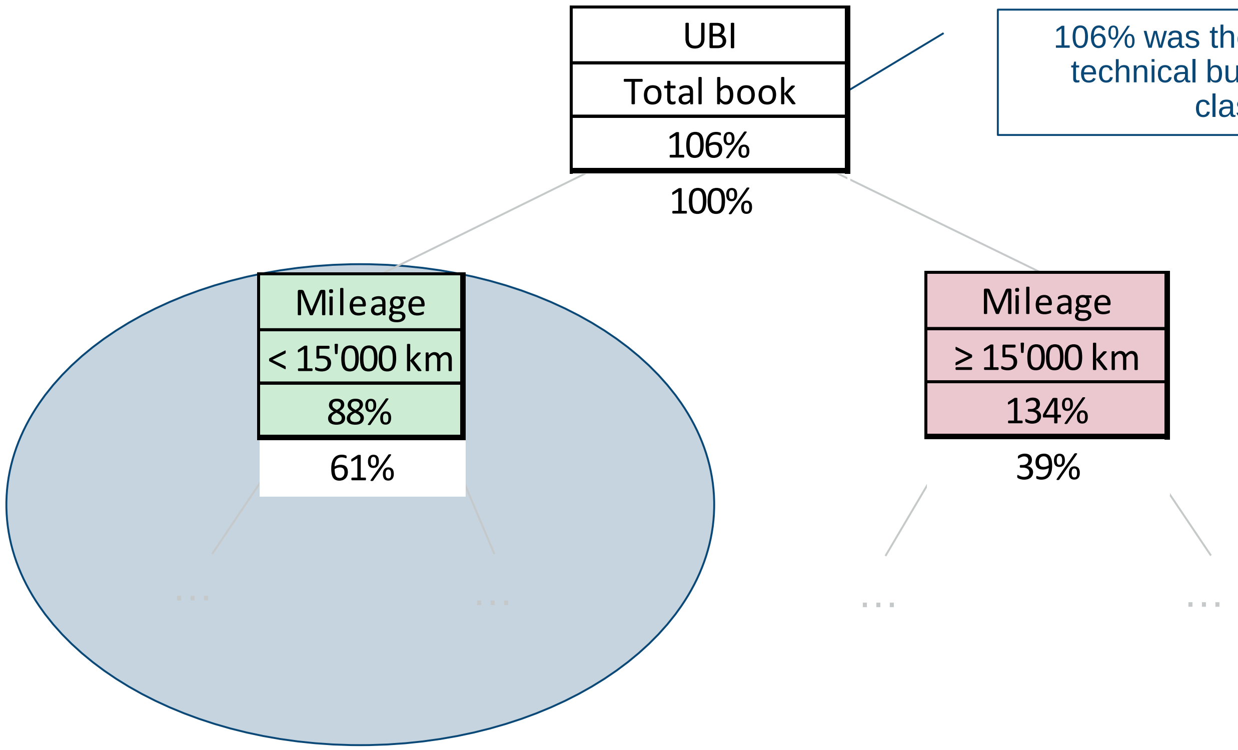
GBM found local interactions among the risk factors even though mileage was already included in the GLM as the strongest Telematics factor!

Model 2 results

Segment	Volume	Loss Ratio Relativity*
Under 15,000k • Mostly highway driving	12%	63%
4,000-8,000k • Mostly country lanes	13%	77%
<4,000k or 8,000-15,000k • Country lanes • Non-Metro • High Power/Weight	15%	82%
15,000-20,000k • Mostly highway/country lanes	11%	97%
<4,000k or 8,000-15,000k • Country lanes • Non-Metro • Low Power/Weight	12%	108%
<4,000k or 8,000-15,000k • Country lanes • Metro	8%	129%
15,000-20,000k • City driving	7%	139%
Over 20,000k • High trips/year	14%	139%
Over 20,000k • Low trips/year	5%	184%

* Premium = GBM prediction excluding UBI variables

Model 2: As a tree (Root)



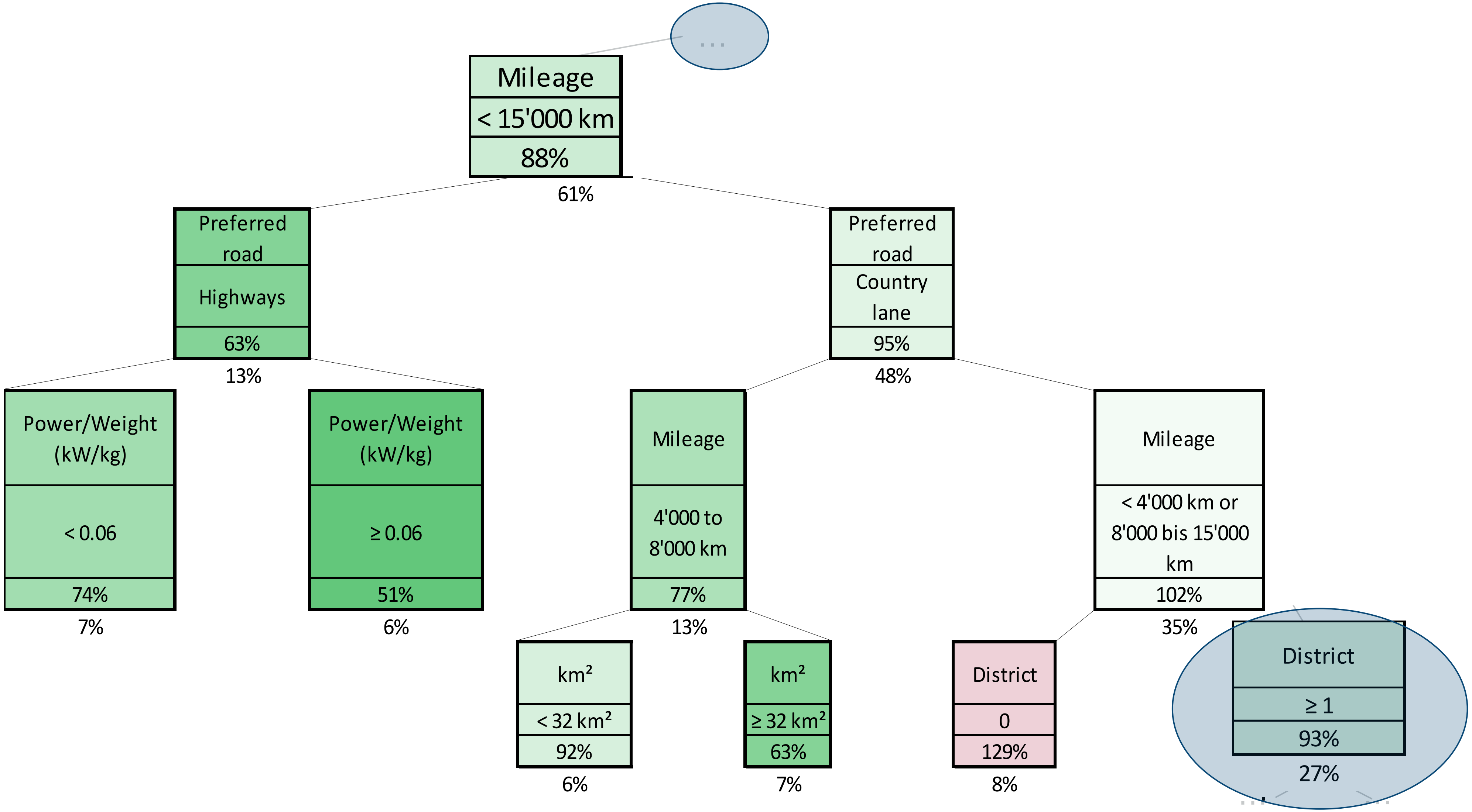
106% was the loss ratio relative to the technical burning cost model for the classical business

Legend:

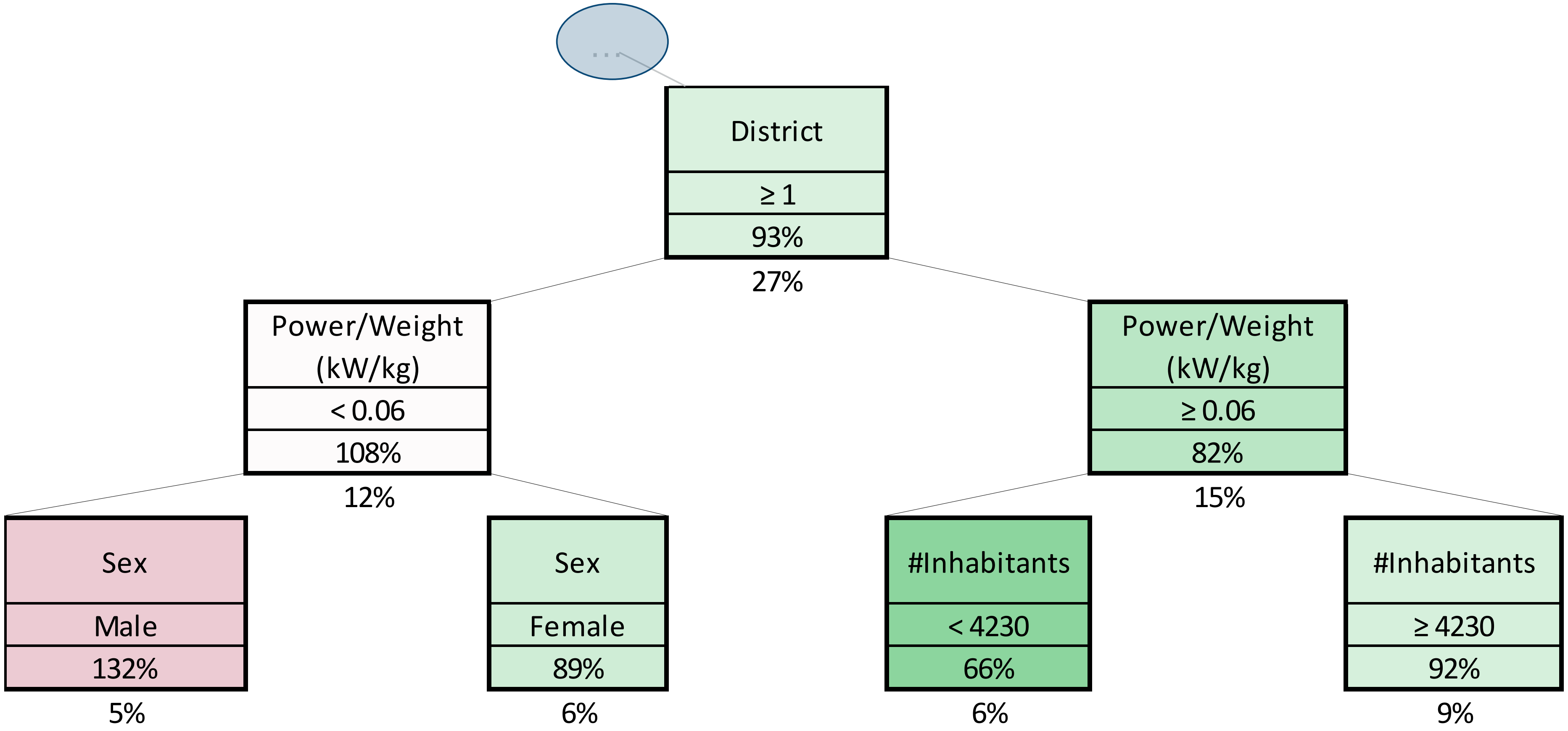
Variable
Category
LR in %

% of book

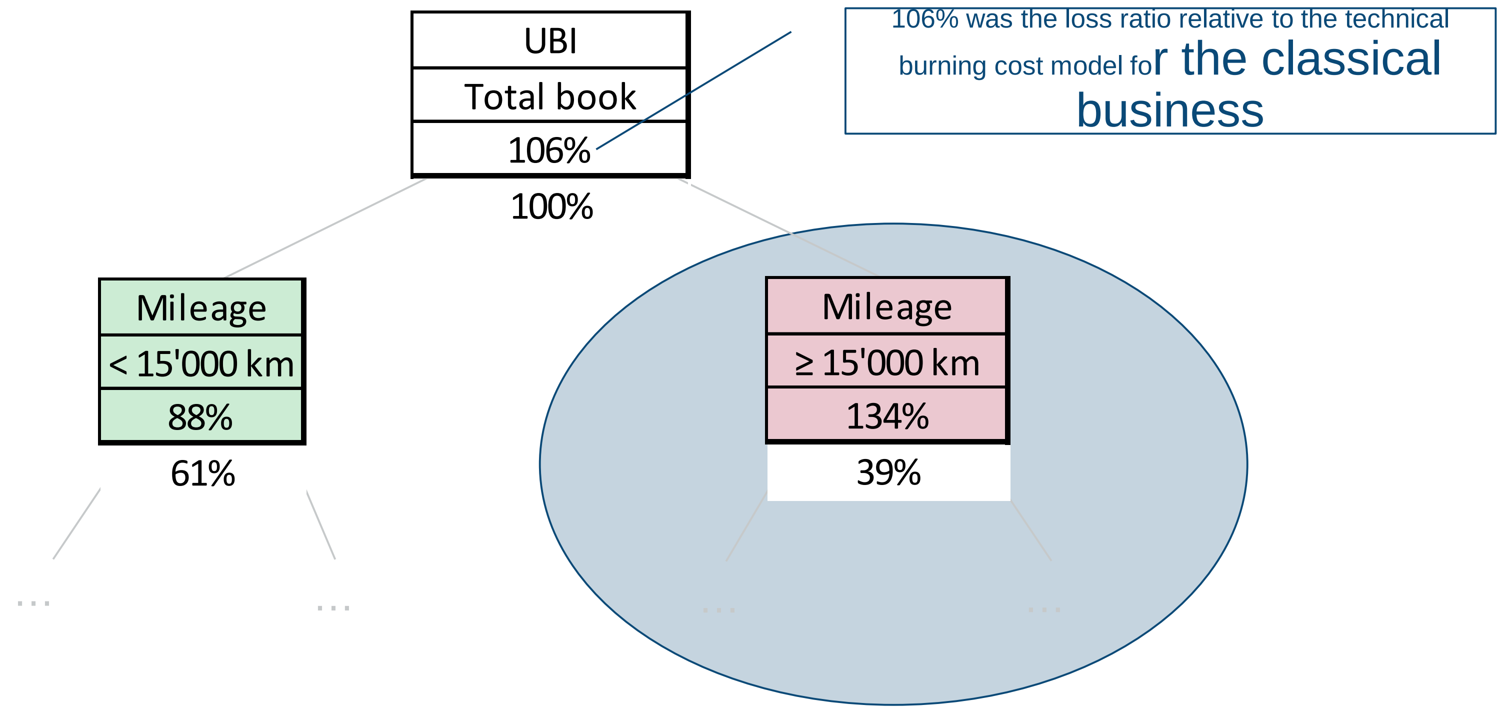
Model 2: The good guys (left branch)



Model 2: Still some good guys (left branch)



Model 2: The poor guys (right branch)

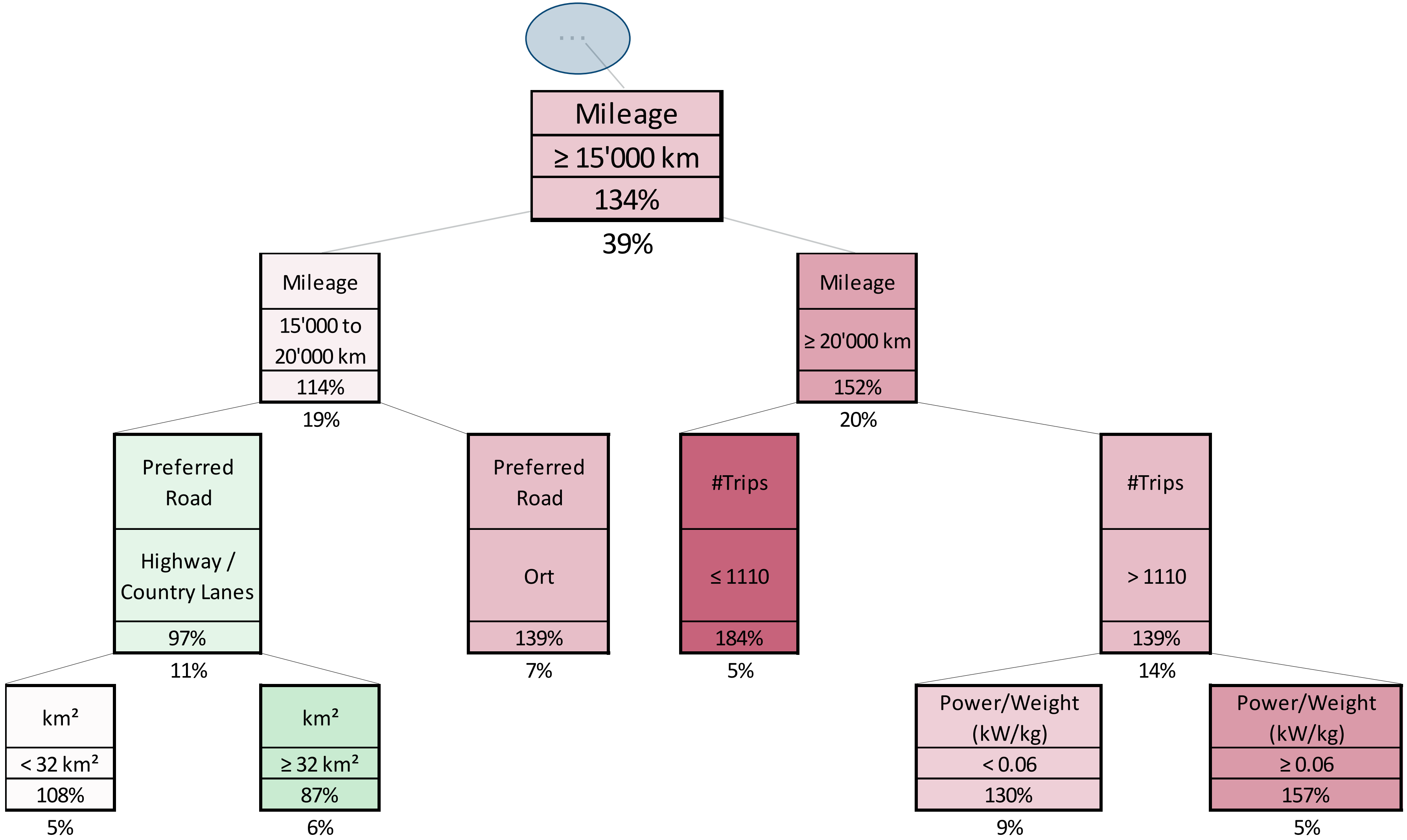


Legend:

Variable
Category
LR in %

% of book

Model 2: The poor guys (right branch)



Conclusions

- UBI business is still a child (not fully grown up at least in Europe)
- UBI comes with big data and this mine is barely tapped
- There are complex interactions among UBI variables and traditional rating factors that are difficult to fit with GLMs distributions
- Moving beyond GLM and introducing machine learning techniques with telematics data may enable insurers to leverage key competitive advantages

Thank you

Peggy Brinkmann, FCAS, MAAA
Peggy.brinkmann@Milliman.com