A Telematics Case Study

CAS Research Paper Series on Race and Insurance Pricing

Mallika Bender, FCAS, MAAA - Staff Actuary

Casualty Actuarial Society

7/29/25 – NAIC CASTF Book Club

CAS Research Paper Series on Race and Insurance Pricing

Phase 1

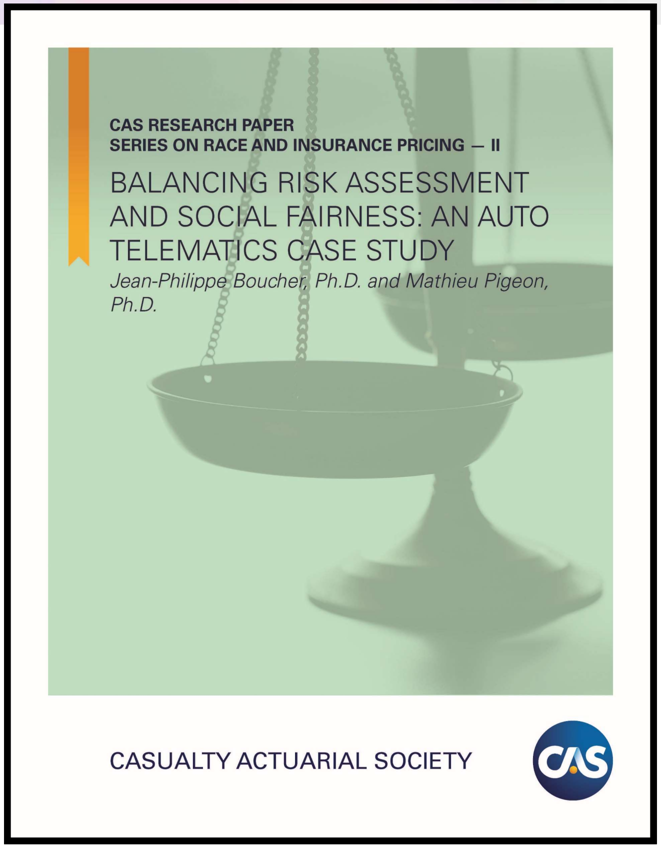
CAS Research Paper Series on Race and Insurance Pricing Phase 2

All Papers Now Available!

Scan code above or visit casact.org/raceandinsuranceresearch

Balancing Risk Assessment and Social Fairness

A Telematics Case Study



Authors: Jean-Philippe Boucher, Ph.D. & Mathieu Pigeon, Ph.D.

Purpose: To explore the potential for telematics or usage-based insurance rating variables to reduce insurers' reliance on protected or sensitive information (e.g. sex, age, marital status, territory, credit)

Telematics Technology & Data

Collected via onboard diagnostics device or phone app

Informs Usage-Based Insurance products

<u>Benefits</u>

- Pricing Accuracy
- Personalization
- Encourage/incentivize safe driving

<u>Challenges</u>

- Implementation cost
- Consumer privacy
- Barriers to take-up:
  - lack of smart phone
  - older vehicles
  - trust

Existing Telematics Research

- Pricing Models perform better when one or more telematics variables are included.
  - Distance driven and driving habits significantly impact claims experience

- Use of one or more telematics elements can replace some sensitive variables such as sex or age.

Telematics: Approach

Data Used: synthetic database generated from real insurance/telematic data

- generated from major Canadian insurer data

Models: claims frequency & claims severity,

- using both GLM and GBM/”Black Box” approaches

Analysis: compare the model residuals by sensitive variables between:

1. Model with traditional non-sensitive + sensitive variables
2. Model with traditional non-sensitive + Telematics variables

Are the sensitive variables still predictive in model 2?

Validation: compare the outcomes from the synthetic data to the outcomes using the original insurer data

Sensitive Covariates of Interest

- Age
- Sex

Sometimes considered “Protected Information”

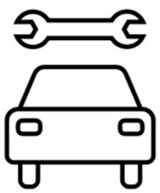
- Marital Status
- Credit Score
- Territory of Residence

Can be correlated to protected info like race/ethnicity

Traditional Non-Sensitive Covariates

- Policy Duration
- Car Age
- Years without Claim

- Region
- Car Use
- Annual miles Driven



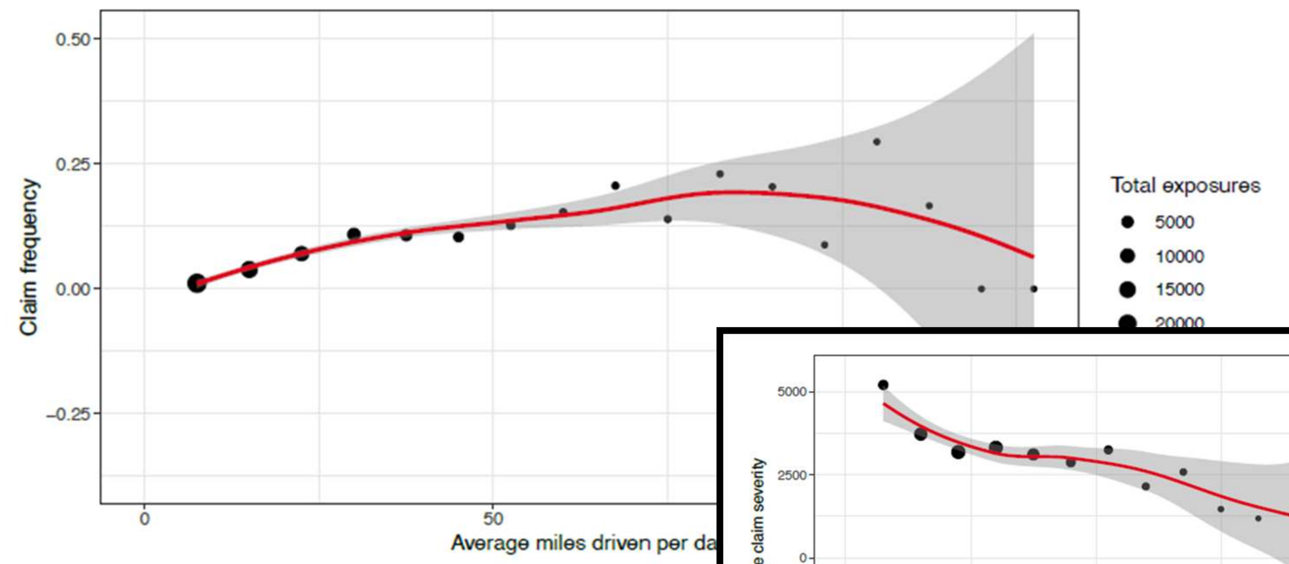
Telematics Data Elements

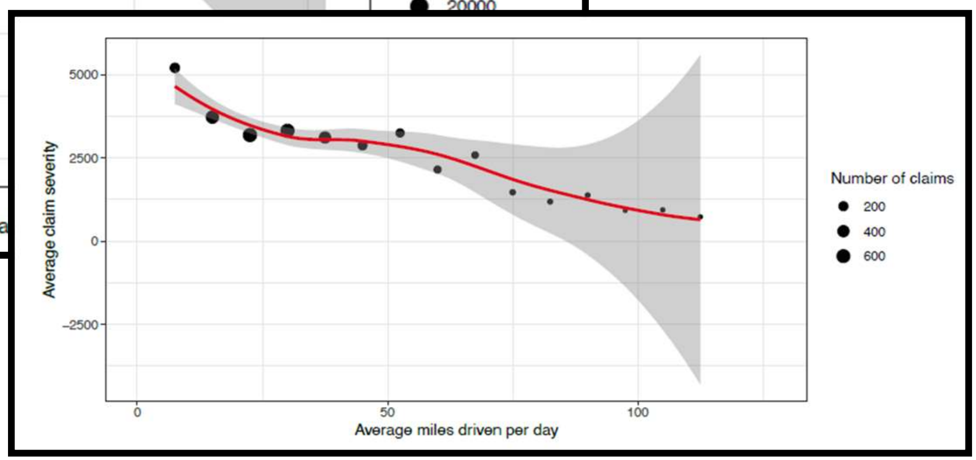
- Distance Driven → Avg Miles Driven Per Day
- Number of Days / Days of Week / Weekend
- Hard Acceleration
- Hard Braking
- Left/Right Turn Intensity
- Long vs Short trips
- Rush Hour Driving

*Can also include detailed GPS location (not used in this study)

Telematics – Avg Miles Per Day

Figure 3.12. Average Miles Driven per Day





Telematics – Days of the Week

Figure 3.13. Average Number of Days per Week the Car is Used

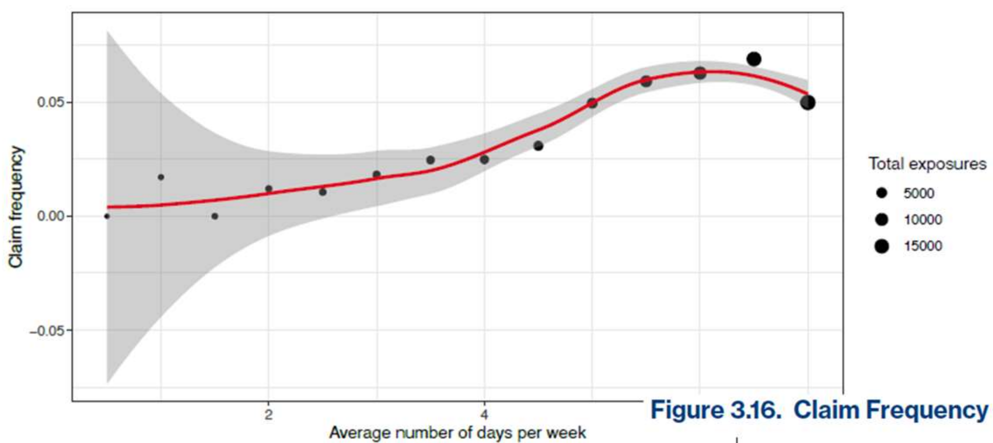
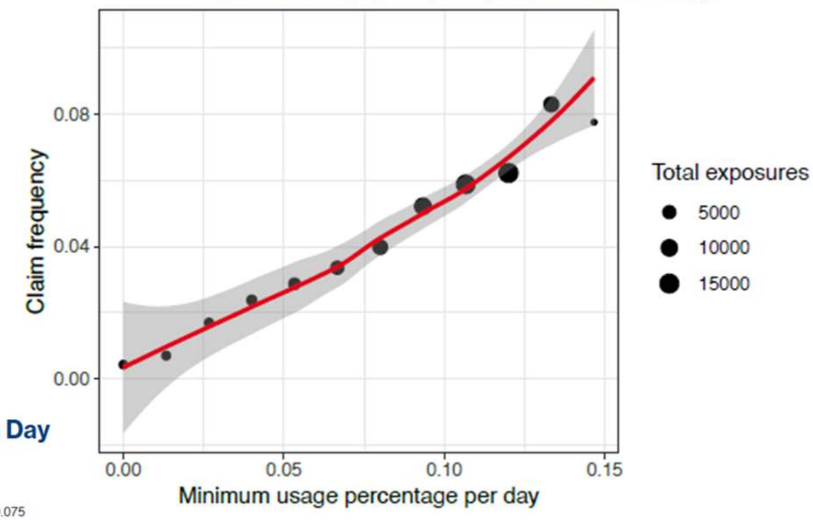
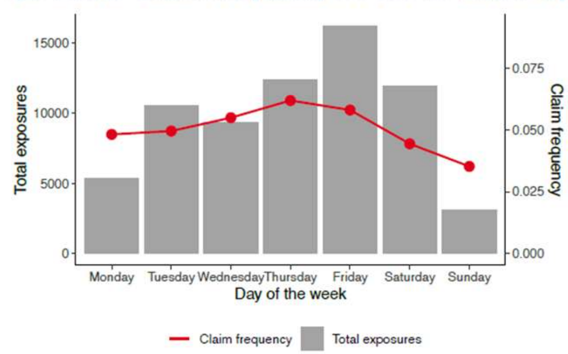


Figure 3.15. Claim Frequency vs. Use for Each Day



(b) Minimum Use

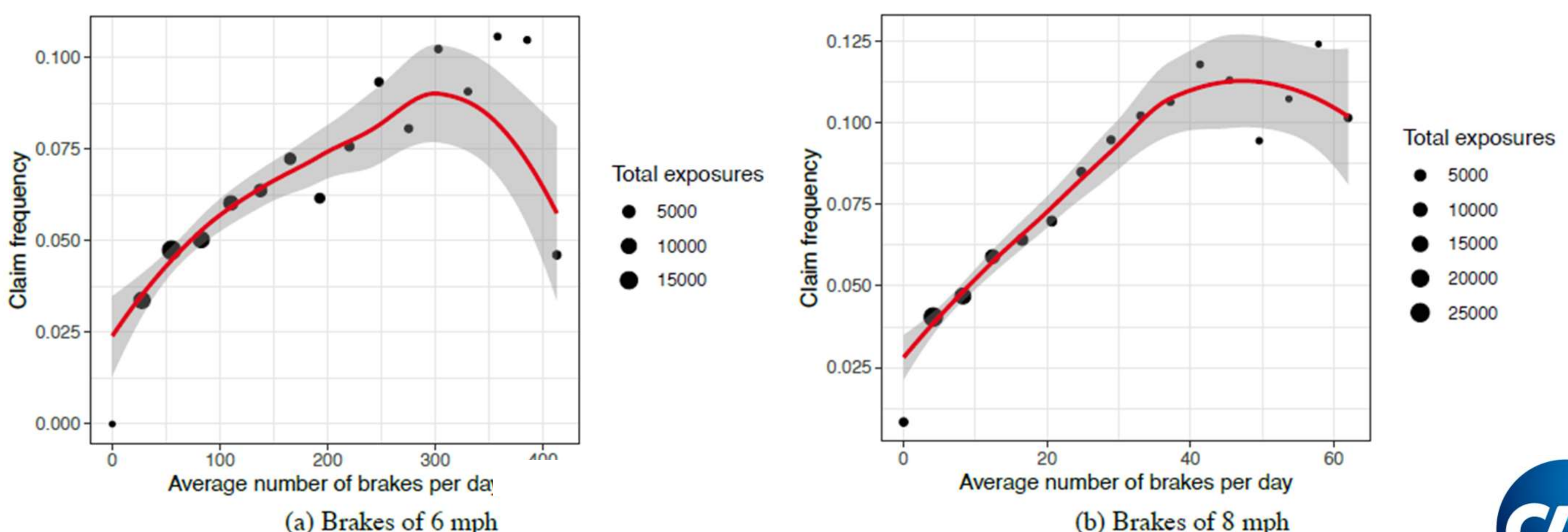
Figure 3.16. Claim Frequency vs. Use for Each Day



(a) Day with the maximum use

Telematics – Hard Braking

Figure 3.20. Claim Frequency vs. Average Number of Brakes



Similar trends seen in frequency for hard accelerations, fast left turns, and fast right turns.

Telematics: Impacts on Sensitive Vars

- Insured Age – lost most significance
- Marital Status – lost most significance
- Sex – reduced significance

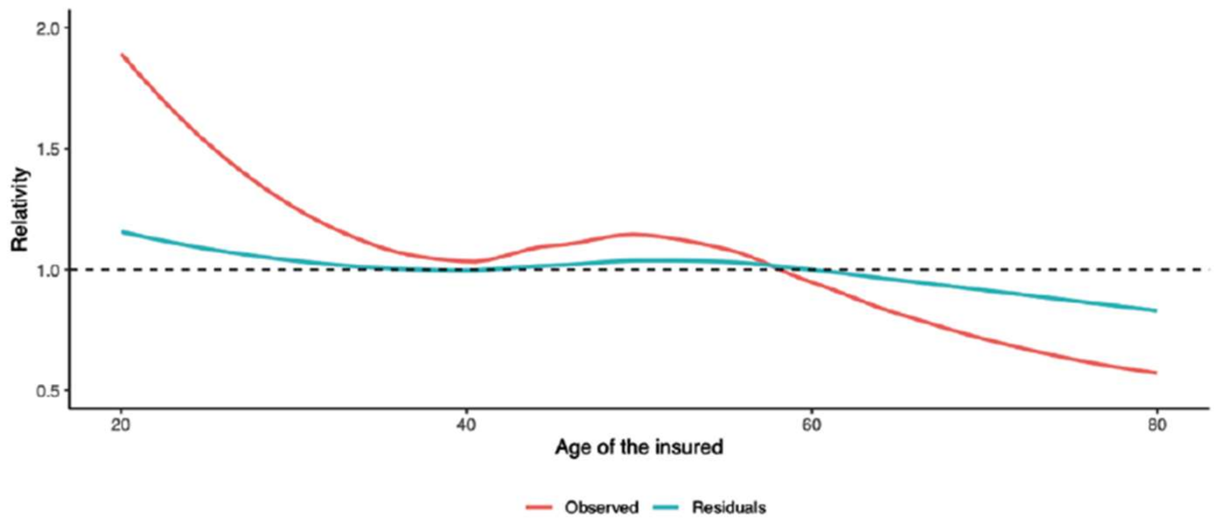
Insurer data validated results

- Territory – lost most significance
- Credit Score – reduced significance somewhat

Insurer data did not validate results

Results – Insured Age

Figure 2.2. Impact of Insured's Age Before (Red Line) and After (Blue Line) Adding Telematics Variables in an XGBoost Model

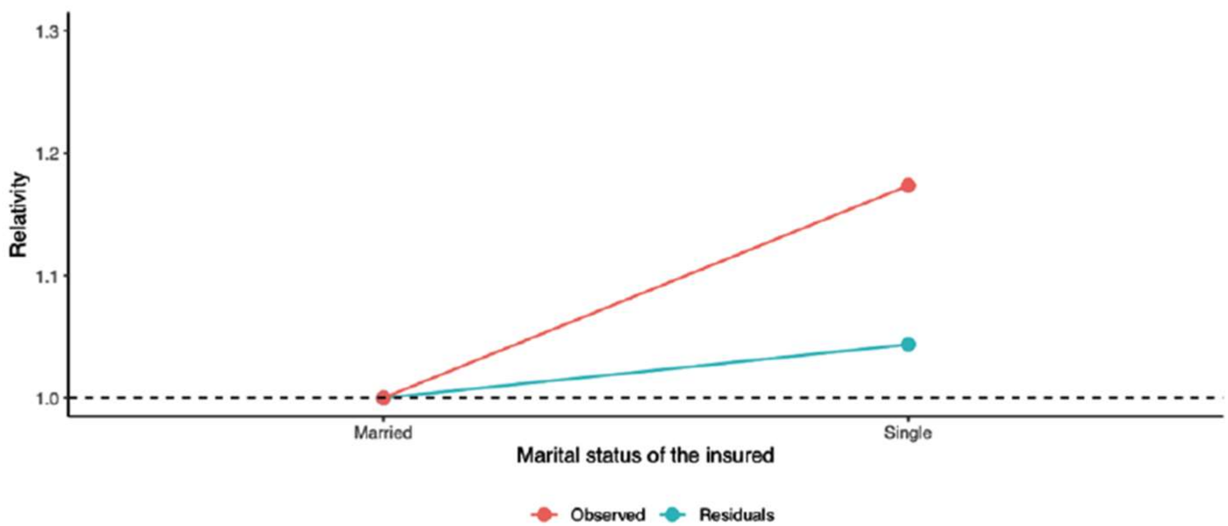


(a) Frequency

Similar results when tested on original insurer data

Results – Marital Status

Figure 2.4. Impact of Marital Status Before (Red Line) and After (Blue Line) Adding Telematics Variables in an XGBoost Model

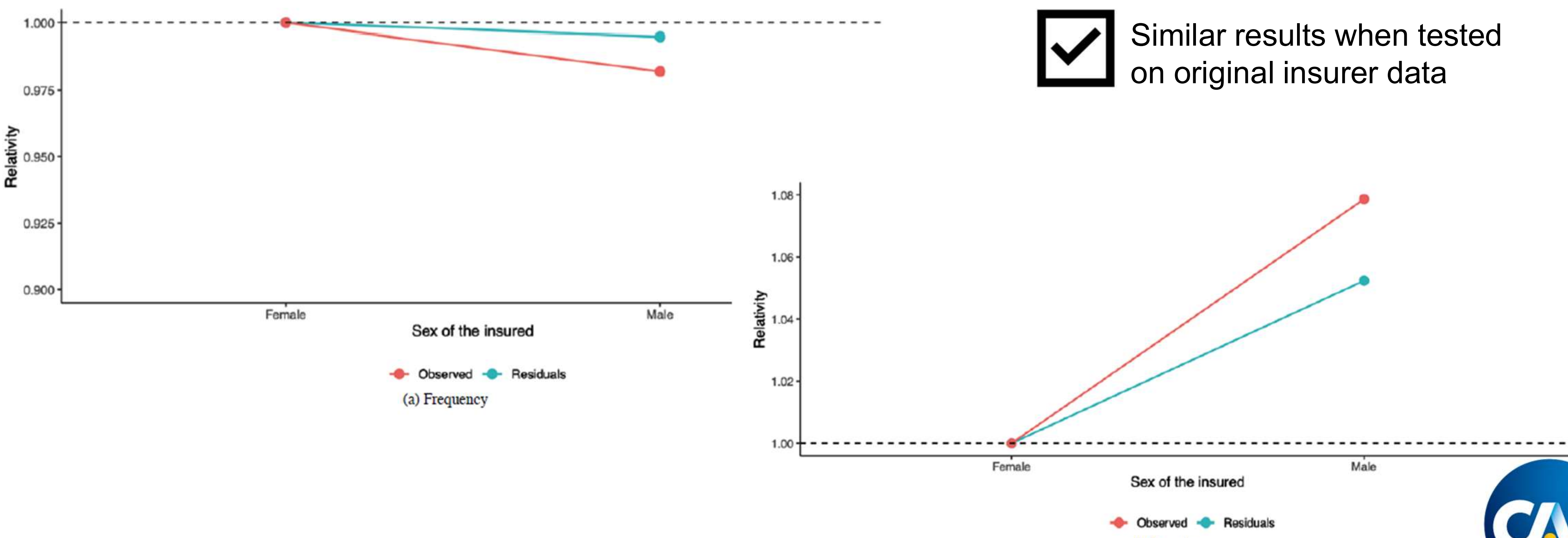


(a) Frequency

Similar results when tested on original insurer data

Results – Insured Sex

Figure 2.5. Impact of Insured's Sex Before (Red Line) and After (Blue Line) Adding Telematics Variables in an XGBoost Model

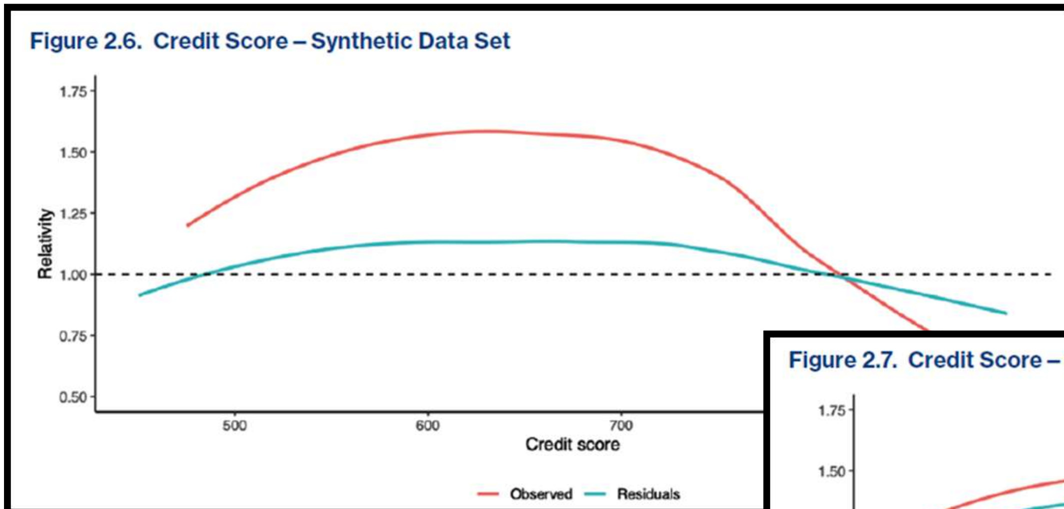


(a) Frequency

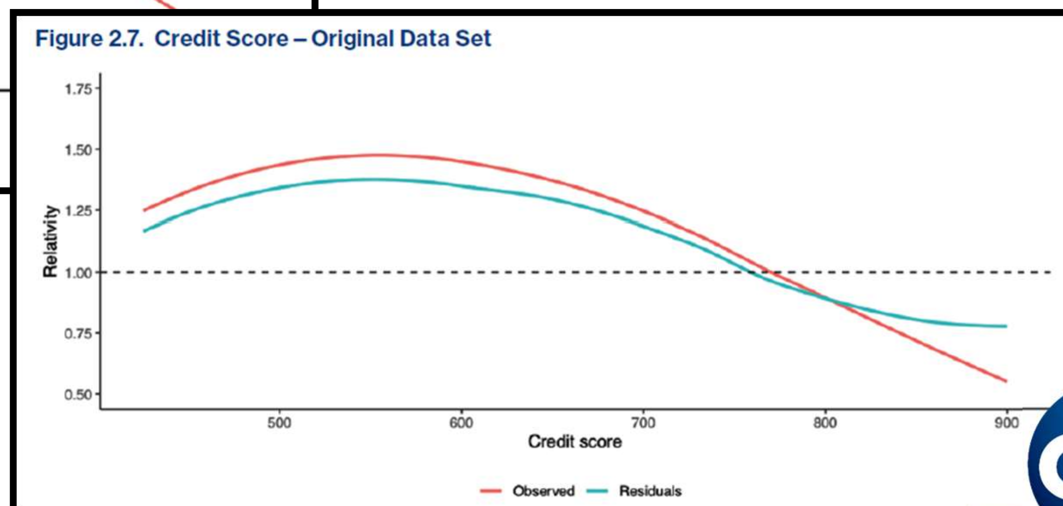
(b) Severity

Similar results when tested on original insurer data

Results vs Validation – Credit Score



Insurer data
<u>did not</u>
validate
results



Results vs Validation – Territory

Figure 2.8. Territory – Synthetic Data Set

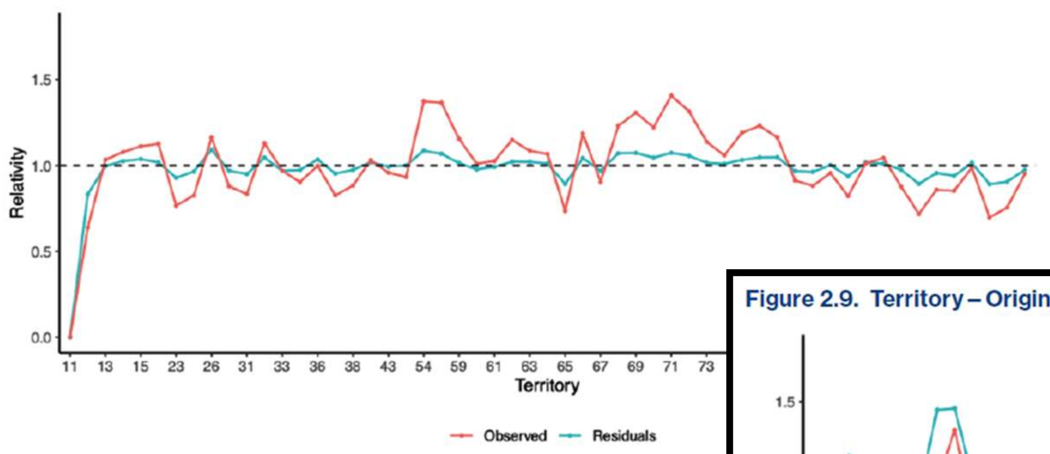
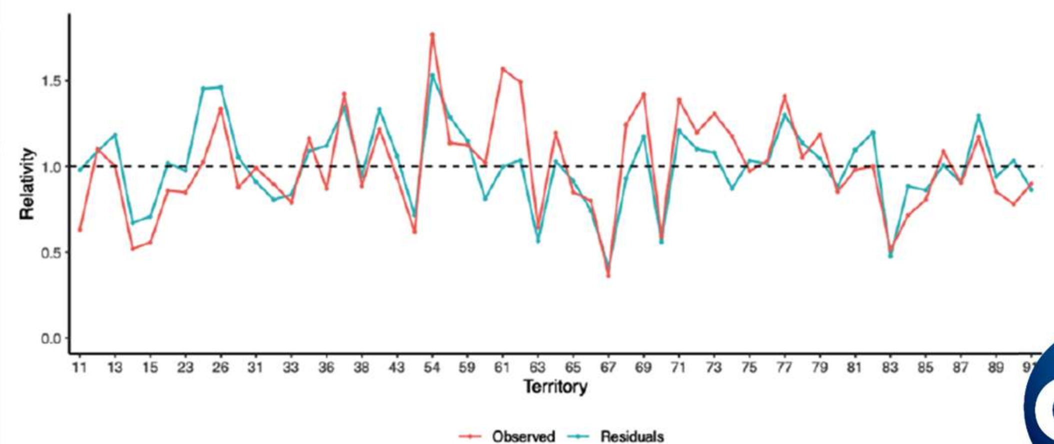


Figure 2.9. Territory – Original Data Set



Insurer data <u>did not</u> validate results

Note on Model Selection

- Black Box Models (GBM) may unlock greater potential from telematics data

+ Flexibility, parameters, interaction effects

+ Greater reduction in reliance on sensitive information

✕ Less transparent, difficult to implement / explain

More Testing is Warranted

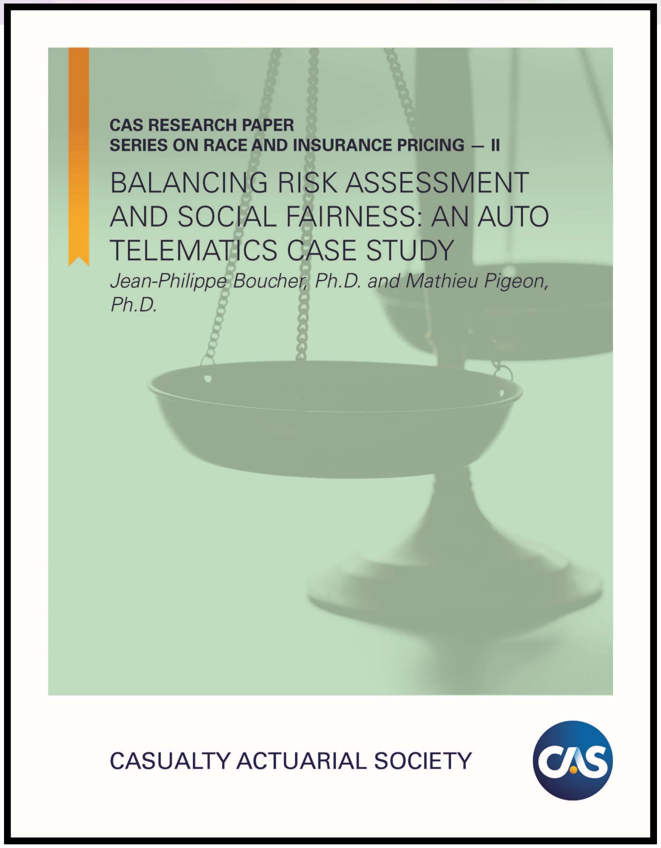
- Individual insurers may arrive at different conclusions

Unique policyholder mix

Varying definition of traditional rating factors (ex: credit score in CAN vs CBIS in USA) and

Varying approaches to use of telematics data

Non-representative sample (which policyholders opt in to telematics)



CAS RESEARCH PAPER
SERIES ON RACE AND INSURANCE PRICING — II

BALANCING RISK ASSESSMENT AND SOCIAL FAIRNESS: AN AUTO TELEMATICS CASE STUDY

Jean-Philippe Boucher, Ph.D. and Mathieu Pigeon, Ph.D.

CASUALTY ACTUARIAL SOCIETY

Questions?

www.casact.org/RaceAndInsuranceResearch

Email: mbender@casact.org