Introduction
000000

Transition to telematics
000000000

Data and results
0000000000000

Scenarios and conclusions
000

Going forward to optimal pricing
000

Is motor insurance ratemaking going to change with telematics and semi-autonomous vehicles?

Montserrat Guillén

University of Barcelona
mguillen@ub.edu
www.ub.edu/riskcenter

Riskcenter

UNIVERSITAT DE
BARCELONA

Risk Seminar, CDAR
Berkeley, August 28, 2018

1 Introduction
   - Telematics pricing
   - Some recent papers on telematics pricing
   - What we do

2 Transition to telematics motor insurance
   - Telematics information as complement/substitute of traditional risk factors
   - Models with an excess of zeros

3 Data and results
   - Information on the data set
   - Two step correction
   - Speed reduction
   - Visualizing results
   - Takeaways

4 Scenarios and conclusions

5 Going forward to optimal pricing

Riskcenter   UNIVERSITATde BARCELONA

1 Introduction

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing

Telematics pricing

- Usage-Based-Insurance (UBI). Telemetry provides the insurer with detailed information on the use of the vehicle and the **premium is calculated based on usage**.
  - Pay-As-You-Drive (PAYD) automobile insurance is a policy agreement linked to **vehicle driven distance**.
  - Pay-How-You-Drive (PHYD) considers **driving patterns**.

- **Distance** and **driving skills** of the drivers measure **speed**, **type of road** and **part of the day when the car is most frequently used**,... These new factors explain the risk of accident (Litman, 2005; Langford et al., 2008; Jun et al., 2007 and 2011; Ayuso, Guillen and Alcaniz, 2010; Verbelen et al., 2017; Henckaerts et al., 2017).

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing
○●○○○○○ | ○○○○○○○○○ | ○○○○○○○○○○○○○ | ○○○ | ○○○
Some recent papers on telematics pricing

1. The **relationship between the distance run by a vehicle and the risk of accident** has been discussed by many authors, most of them arguing that this relationship is **not proportional** (Litman, 2005 and 2011; Langford et al., 2008; Boucher et al., 2013).

2. There is **evidence of the relationship** between **speed, type of road, urban and nighttime driving** and the **risk of accident** (Rice et al., 2003; Laurie, 2011; Ellison et al, 2015; Verbelen et al. 2017).

3. **Telematics information can replace some traditional rating factors** and provide a pricing model with the same predictive performance (Verbelen et al. 2017; Ayuso et al., 2016b).

Gender: discrimination that turns out to be a proxy

Gender can be replaced by:

km/day (Barcelona approach) or
km/trip (Leuven approach)

Strong evidence exists

- Information on mileage and driving habits improves the prediction of the number of claims compared to traditional rating factors and coverage exclusively by time (usually one year).

- If the use of advanced driver assistance systems (ADAS) mitigates risk, because they transform driving patterns, then the transition towards semi-autonomous vehicles is expected to contribute to a lower frequency of motor accidents

Our question is:

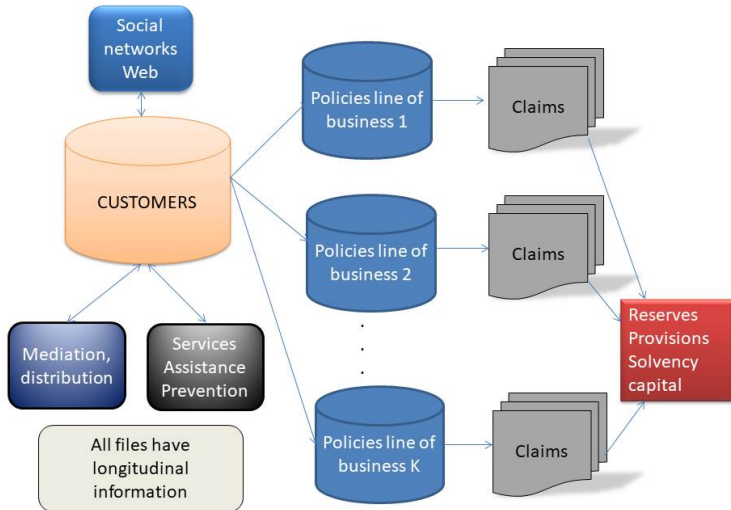What is the expected impact of telematics on motor insurance ratemaking?

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing

What we do

What we do:

- We explore the effect of distance driven (mileage, exposure to risk) and other telematics data on premium calculation.
- We show scenarios for a reduction of speed limit violations in semi-autonomous vehicles, a decrease in the expected number of accident claims and, ultimately, in insurance rates.

Our contribution:

- Propose a method to update premiums regularly with telematics data.
- Show that the price per mile depends on driving habits and is not proportional to distance driven. A zero claim is relatively more frequent for intensive users. Propose a predictive modeling approach for this purpose.
- Estimate the impact on prices and safety in a plausible scenario of assisted driving that would control speed limit violations.
- Derive some open-questions about risk measures to summarize telematics big data and optimal pricing when customers may lapse.

UNIVERSITAT de BARCELONA

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing

Why is insurance analytics a good example of big data in applied economics?

Source: Guillen, 2016

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing

How do telematics data look like?

Achieving Pay HOW You Drive

Sample Trip Summary Data – One Day

State Date	Start Time	Motorway Yards	Urban Yards	Other Yards	Motorway Seconds	Urban Seconds	Other Seconds	Total Speeding Yards	Total Speeding Seconds
3/3/2012	12:12:00	-	31	-	-	13,713	-	-	-
3/3/2012	14:17:11	-	3,355	-	-	7,934	-	-	-
3/3/2012	14:34:03	39,566	39,010	69,042	1,328	1,922	2,864	1,379	38
3/3/2012	15:47:59	-	11,346	907	-	858	60	-	-
3/3/2012	17:21:11	31,426	43,634	57,937	1,020	2,141	6,916	-	-
3/3/2012	19:36:07	-	4,501	5,401	-	2,912	330	-	-
3/3/2012	21:57:27	-	14,255	1,394	-	22,466	60	-	-
3/3/2012	22:24:43	-	-	-	-	386	-	-	-

2.6 MB per vehicle per week ✕ 52 weeks ✕ 50,000 vehicles = 6.8 TB per year

Source: Jim Janavich ideas.returnonintelligence.com

Riskcenter UNIVERSITAT DE BARCELONA

Introduction
Transition to telematics
Data and results
Scenarios and conclusions
Going forward to optimal pricing

2 Transition to telematics motor insurance

Introduction	Transition to telematics	Data and results	Scenarios and conclusions	Going forward to optimal pricing
000000	●00000000	0000000000000	000	000

Telematics information as complement/substitute of traditional risk factors

The classical ratemaking model is based on: a prediction of the number of claims (usually for one year) times the average claim cost plus some extra loadings.

- Subscript i denotes the ith policy holder in a portfolio of n insureds.
- Given $x_i = (x_{1i}, ... x_{ki})$ (vector of k covariates), the number of claims Y_i (dependent variable) follows a Poisson distribution with parameter λ_i, which is a function of the linear combination of parameters and regressors, $\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}$.

$$E(Y_i|x_i) = exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}) \tag{1}$$

The unkown parameters to be estimated are $(\beta_0, \ldots, \beta_k)$.

- Classical covariates are age, time since driver's license was issued, driving zone, type of car,...
- The **pure premium** equals the product of the expected number of claims times the average claim cost. Finally, the **premium** is obtained once additional margins and safety loadings are included.

Introduction	Transition to telematics	Data and results	Scenarios and conclusions	Going forward to optimal pricing
○○○○○○	○●○○○○○○○○	○○○○○○○○○○○○○	○○○	○○○

Telematics information as complement/substitute of traditional risk factors

CrossMark

Improving automobile insurance ratemaking using telematics: incorporating mileage and driver behaviour data

Mercedes Ayuso[1] · Montserrat Guillen[1] · Jens Perch Nielsen[2]

Abstract   We show how data collected from a GPS device can be incorporated in motor insurance ratemaking. The calculation of premium rates based upon driver behaviour represents an opportunity for the insurance sector. Our approach is based on count data regression models for frequency, where exposure is driven by the distance travelled and additional parameters that capture characteristics of automobile usage and which may affect claiming behaviour. We propose implementing a classical frequency model that is updated with telemetrics information. We illustrate the method using real data from usage-based insurance policies. Results show that not only the distance travelled by the driver, but also driver habits, significantly influence the expected number of accidents and, hence, the cost of insurance coverage. This paper provides a methodology including a transition pricing transferring knowledge and experience that the company already had before the telematics data arrived to the new world including telematics information.

Keywords   Tariff · Premium calculation · Pay-as-you-drive insurance · Count data models

Riskcenter   UNIVERSITAT DE BARCELONA

Introduction   **Transition to telematics**   Data and results   Scenarios and conclusions   Going forward to optimal pricing
○○○○○○   ○○●○○○○○○○   ○○○○○○○○○○○○○   ○○○   ○○○

Telematics information as complement/substitute of traditional risk factors

In **Transportation** (2018) we proposed a method for assessing the influence on the expected frequency of usage-based variables which can be viewed as a **correction of the classical ratemaking model**.
A two-step procedure:

- Step 1: Let $\hat{Y}_i$ be the frequency estimate obtained as a function of the classical explanatory covariates $x_i = (x_{i1}, \ldots, x_{ik})$.

- Step 2: Let $z_i = (z_{i1}, \ldots, z_{il})$ be the information collected periodically from a telematics GPS. Then, the prediction from usage-based insurance information is a correction such that:

$$E(Y_i^{UBI}|z_i, \hat{Y}_i) = \hat{Y}_i \, exp(\eta_0 + \eta_1 z_{i1} + \ldots + \eta_k z_{ik}), \qquad (2)$$

where the parameter estimates $(\eta_0, \ldots, \eta_l)$ can now be obtained using $\hat{Y}_i$ as an offset.

Note:

This approach is less efficient than a full information model, but it works well in practice. Telematics data are collected on a continuous basis and this correction can be implemented regularly (i.e. on a weekly basis)

UNIVERSITAT DE BARCELONA

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing

Models with an excess of zeros

Risk Analysis

The use of telematics devices to improve automobile insurance rates

Publication status

Guillen, M. et al (2018)

Most automobile insurance databases contain a large number of policyholders with zero claims. This high frequency of zeros may reflect the fact that some insureds make little use of their vehicle, or that they do not wish to make a claim for small accidents in order to avoid an increase in their premium, but it might also be because of good driving. We analyse information on exposure to risk and driving habits using telematics data from a Pay-as-you-Drive sample of insureds. We include distance travelled per year as part of an offset in a zero- inflated Poisson model to predict the excess of zeros. We show the existence of a learning effect for large values of distance travelled, so that longer driving should result in higher premium, but there should be a discount for drivers that accumulate longer distances over time due to the increased proportion of zero claims. We confirm that speed limit violations and driving in urban areas increase the expected number of accident claims. We discuss how telematics information can be used to design better insurance and to improve traffic safety.

Introduction   Transition to telematics   Data and results   Scenarios and conclusions   Going forward to optimal pricing
000000   000000●0000   000000000000   000   000
Models with an excess of zeros

In **Risk Analysis** (2018) we propose to include the distance travelled per year as an offset in a Zero Inflated Poisson model to predict the number of claims in *Pay as You Drive* insurance.

- *The Poisson model with exposure:* Let us call T_i the exposure factor for policy holder i, in our case $T_i = ln(D_i)$, where Di indicates distance travelled, then:

$$E(Y_i|x_i, T_i) = D_i \, exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}) = D_i \lambda_i \qquad (3)$$

Excess of zeros exists because:

- Some insureds do not use their car and so they do not have claims
- Some insured acquire exceptionally good driving skills and they do not have claims (*learning curve*).

Introduction   Transition to telematics   Data and results   Scenarios and conclusions   Going forward to optimal pricing
000000   000000●000   0000000000000   000   000

Models with an excess of zeros

- *The Zero-inflated Poisson (ZIP) model :* Now the probability of not suffering an accident is

$$P(Y_i = 0) = p_i + (1 - p_i)P(Y^* = 0) \tag{4}$$

where p_i is the probability of excess of zeros. Y_i^* follows a Poisson distribution with parameter $exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik})$, and p_i may depend on some covariates.

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing

Models with an excess of zeros

A *ZIP Poisson model with exposure*

We assume that p_i is the probability of an excess of zeros, and it is specified as a logistic regression model such that

$$p_i = \frac{exp(\alpha_0 + \alpha_1 ln(D_i))}{1 + exp(\alpha_0 + \alpha_1 ln(D_i))}. \tag{5}$$

The Poisson model for Y^* is specified as follows, with an exposure

$E(Y_i^*|x_i, T_i) = D_i exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}) = D_i \lambda_i = exp(ln(D_i))\lambda_i = exp(T_i)\lambda_i$, where $T_i = ln(D_i)$. The expectation of the Poisson part is:

$$(1 - p_i)E(Y_i^*|x_i, T_i) = \frac{1}{1 + exp(\alpha_0 + \alpha_1 ln(D_i))}D_i\lambda_i = D_i^*\lambda_i \tag{6}$$

where $D_i^* = \frac{D_i}{1+exp(\alpha_0+\alpha_1 ln(D_i))}$ is a transformation of the original measure of exposure (distance driven) D_i.

Introduction Transition to telematics Data and results Scenarios and conclusions Going forward to optimal pricing
000000   000000000●   0000000000000   000   000
Models with an excess of zeros

A *ZIP Poisson model with exposure*

So, when we include zero-inflation there is a transformation of the exposure in the Poisson part of the model.

- When D_i is big then $D_i^* = \frac{D_i}{1+exp(\alpha_0+\alpha_1 ln(D_i))}$ tends to zero if $\alpha_1 > 1$.

- When $\alpha_1 = 1$ then D_i^* tends to constant $\frac{1}{exp(\alpha_0)}$ when D_i increases.

- Assuming that $D_i \geq 1$, when $\alpha_1 > 1$ this is a concave transformation that scales exposure into the interval $\left[0, \frac{1}{1+exp(\alpha_0)}\right]$. So, the larger the exposure the smaller the value whereas the smaller the exposure the larger the value.

- Assuming that $D_i \geq 1$, when $\alpha_1 \leq 1$ then the transformation is a change of scale to the interval $\left[\frac{1}{1+exp(\alpha_0)}, +\infty\right)$.

Introduction | Transition to telematics | Data and results | Scenarios and conclusions | Going forward to optimal pricing
000000 | 00000000●0 | 0000000000000 | 000 | 000

Models with an excess of zeros

A *ZIP Poisson model with exposure*

If we look at the logistic regression part, we can also derive the following expression:

$$
p_i = \frac{exp(\alpha_0 + \alpha_1 ln(D_i))}{1 + exp(\alpha_0 + \alpha_1 ln(D_i))} = \frac{exp(\alpha_0 + \alpha_1 ln(D_i))}{1 + exp(\alpha_0 + \alpha_1 ln(D_i))} \frac{D_i}{D_i} =
$$
$$
exp(\alpha_0 + \alpha_1 ln(D_i)) \frac{D_i}{1 + exp(\alpha_0 + \alpha_1 ln(D_i))} \frac{1}{D_i} = exp(\alpha_0 + \alpha_1 ln(D_i)) \frac{D_i^*}{D_i}
$$
$$
\tag{7}
$$

So, the probability of zero excess (p_i) can be understood as a rescaling of the relative transformed exposure.

Interestingly, when $\alpha_1 < 0$ then note that p_i tends to zero when D_i increases, whereas when $\alpha_1 > 0$ then p_i tends to one when D_i increases. In the empirical part we find $\alpha_1 > 0$, which means that there is a learning effect and the excess of zeros is more important than the Poisson part when distance driven increases.

Riskcenter · UNIVERSITAT DE BARCELONA

3 Data and results

Introduction   Transition to telematics   **Data and results**   Scenarios and conclusions   Going forward to optimal pricing
○○○○○○   ○○○○○○○○○   ●○○○○○○○○○○○○   ○○○   ○○○
Information on the data set

Zero-inflation for the Number of Claims

Empirical application based on 25,014 insureds with car insurance coverage throughout 2011, that is, individuals exposed to the risk for a **full year**.

Table I. Frequency of claims per driver (n=25,014)
in the Spanish insurance dataset (all claims, at fault, and not at fault)

| Number of claims | Absolute frequency per driver | | |
	All claims	Claims at fault	Claims not at fault
0	20,608	22,837	22,432
1	3,310	1,750	2,111
2	889	385	424
3	165	37	40
4	34	4	6
5	7	1	1
6	1	0	0

One insured driver had 6 claims, 2 were at fault and 4 where not at fault.

Table II. Descriptive statistics for the risk exposure indicator
(total kilometres travelled per year in 000s)

	All Sample n = 25,014	Drivers with no claims n = 20,608 (82.4%)	Drivers with claims n = 4,406 (17.6%)
Mean	7.16	6.99	7.96
1st Quartile	4.14	4.00	4.87
Median	6.46	6.28	7.22
3rd Quartile	9.40	9.22	10.30
Standard Deviation	4.19	4.14	4.35

Riskcenter   UNIVERSITAT DE BARCELONA

21 / 46

Introduction ○○○○○○ | Transition to telematics ○○○○○○○○○ | **Data and results** ○●○○○○○○○○○○○○ | Scenarios and conclusions ○○○ | Going forward to optimal pricing ○○○

Information on the data set

Table 2 Descriptive statistics by claims (quantitative variables)

	All sample N=25,014		Drivers with no claims N=20,608 (82.4%)		Drivers with claims N=4406 (17.6%)	
	Mean	SD	Mean	SD	Mean	SD
Age	27.57	3.09	27.65	3.09	27.18	3.10
Age driving licence	7.17	3.05	7.27	3.07	6.73	2.94
Vehicle age	8.75	4.17	8.76	4.19	8.69	4.11
Power	97.22	27.77	96.98	27.83	98.36	27.46
Km per year (000s)	7.16	4.19	6.99	4.14	7.96	4.35
Km per year at night (%)	6.91	6.35	6.85	6.32	7.16	6.49
Km per year over speed limit (%)	6.33	6.83	6.28	6.87	6.60	6.59
Urban km per year (%)	25.87	14.36	25.51	14.31	27.56	14.47

Table 3 Descriptive statistics by claims (categorical variables)

	All sample N=25,014		Drivers with no claims N=20,608 (82.4%)		Drivers with claims N=4406 (17.6%)	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Gender						
Men	12,235	48.91	10,018	48.61	2217	50.32
Women	12,779	51.09	10,590	51.39	2189	49.68
Parking						
Yes	19,356	77.38	15,912	77.21	3444	78.17
No	5658	22.62	4696	22.79	962	21.83

Riskcenter · UNIVERSITAT DE BARCELONA

| Introduction | Transition to telematics | **Data and results** | Scenarios and conclusions | Going forward to optimal pricing |
| ○○○○○○ | ○○○○○○○○○ | ○○●○○○○○○○○○○○ | ○○○ | ○○○ |

Two step correction

Poisson model results. All types of claims.

	All variables		Non-telematics		Telematics		Telematics with offsets	
	Coefficient	(p-value)	Coefficient	(p-value)	Coefficient	(p-value)	Coefficient	(p-value)
Intercept	-1.503	0.122	0.135	0.888	-3.427	<.0001	-1.807	<.0001
Age	-0.132	0.064	-0.101	0.153				
Age^2	0.002	0.066	0.002	0.208				
Male	-0.040	0.155	0.084	0.003				
Age Driving License	-0.061	<.0001	-0.061	<.0001				
Vehicle Age	0.010	0.003	0.002	0.549				
Power	0.003	<.0001	0.003	<.0001				
Parking	0.031	0.347	0.037	0.252				
Log of Km per year (thousands)	0.644	<.0001			0.645	<.0001	0.620	<.0001
Km per year at night (%)	-0.004	0.295			-0.001	0.761	-0.007	0.067
Km per year at night (%)^2	0.0002	0.140			0.0001	0.413	0.0002	0.041
Km per year over speed Limit (%)	0.026	<.0001			0.026	<.0001	0.022	<.0001
Km per year over speed Limit (%)^2	-0.001	<.0001			-0.001	<.0001	-0.001	<.0001
Urban km per year (%)	0.023	<.0001			0.024	<.0001	0.022	<.0001
AIC	29,464.858		30,315.914		29,640.186		29,483.041	
BIC	29,578.638		30,380.931		29,697.076		29,539.931	

Riskcenter · UNIVERSITAT DE BARCELONA

Introduction | Transition to telematics | **Data and results** | Scenarios and conclusions | Going forward to optimal pricing
000000 | 000000000 | 0000●00000000 | 000 | 000
Two step correction

Poisson model results. All types of claims.

Table 6. Poisson model results with offset km per year. All claim types (n=25,014)

	All variables		Non-telematics		Telematics		Telematics with offsets (Log of prediction of Non-telematics model - Column 2)	
	Coefficient	(p-value)	Coefficient	(p-value)	Coefficient	(p-value)	Coefficient	(p-value)
Intercept	-2.193	0.024	-0.472	0.625	-4.219	<.0001	-0.731	<.0001
Age	-0.145	0.043	-0.200	0.005				
Age^2	0.003	0.040	0.004	0.005				
Male	-0.086	0.002	-0.049	0.076				
Age Driving License	-0.061	<.0001	-0.076	<.0001				
Vehicle Age	0.015	<.0001	0.022	<.0001				
Power	0.003	<.0001	0.001	0.063				
Parking	0.034	0.292	0.034	0.299				
Log of km per year (000s)	1.000	--	1.000	--	1.000	--	1.000	--
Km per year at night (%)	-0.008	0.051			-0.005	0.161	-0.009	0.017
Km per year at night (%)^2	0.0002	0.062			0.0001	0.193	0.0002	0.033
Km per year over speed Limit (%)	0.015	0.004			0.014	0.006	0.019	<.001
Km per year over speed Limit (%)^2	-0.001	0.001			-0.001	0.003	-0.001	<.001
Urban km per year (%)	0.029	<.0001			0.031	<.0001	0.028	<.0001
AIC	29,631.281		30,624.100		29,809.179		29,658.447	
BIC	29,736.934		30,689.117		29,857.942		29,707.210	
LogL	-13,742.650		-14,244.060		-13,838.600		-13,763.230	
Chi-2	1,357.220	<0.001	354.400	<0.001	1,165.320	<0.001	1,316.060	<0.001

Riskcenter | UNIVERSITAT de BARCELONA

Introduction   Transition to telematics   **Data and results**   Scenarios and conclusions   Going forward to optimal pricing
○○○○○○   ○○○○○○○○○   ○○○○●○○○○○○○○○   ○○○   ○○○
Two step correction

Concordant predictions of all models (in percentages).

	All variables	Non-telematics	Telematics	Telematics with offsets
Poisson model results. All types of claims	62.28	55.91	61.34	62.10
Poisson model results with offsets (Log of Km per year in thousands). All types of claims	62.15	58.60	61.18	62.05
Poisson model results. Claims where the policyholder is guilty	62.70	57.72	61.13	62.65
Poisson model results with offsets (Log of Km per year in thousands). Claims where the policyholder is guilty	62.38	58.96	60.89	62.43

Riskcenter   UNIVERSITAT DE BARCELONA

Introduction	Transition to telematics	Data and results	Scenarios and conclusions	Going forward to optimal pricing
oooooo	ooooooooo	oooooo●oooooooo	ooo	ooo

Two step correction

Prediction with telematics and offset

Prediction with all variables

Introduction	Transition to telematics	**Data and results**	Scenarios and conclusions	Going forward to optimal pricing
○○○○○○	○○○○○○○○○	○○○○○○●○○○○○○	○○○	○○○

Two step correction

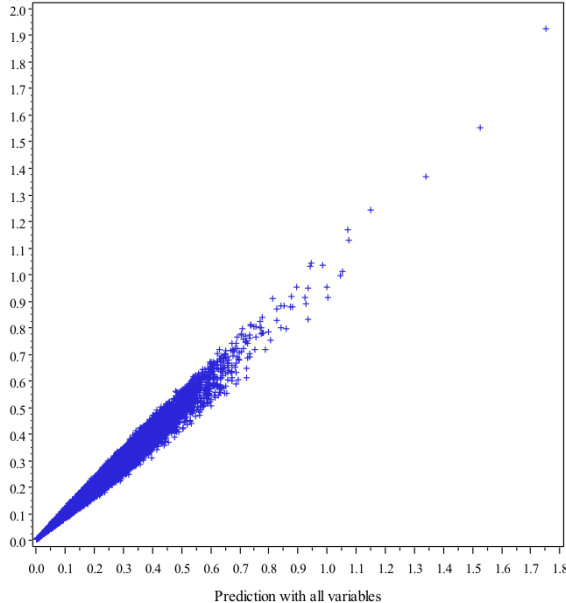Table IV. Zero-inflated Poisson model with offsets (Log of km per year in 000s). All types of claims.

	All variables		(Only significant)		Non-telematics		Telematics	
	Coefficient	(p-value)	Coefficient	(p-value)	Coefficient	(p-value)	Coefficient	(p-value)
Poisson part								
Intercept	-2.148	0.045	-3.396	<.001	-0.829	0.440	-3.461	<.001
Age	-0.094	0.232			-0.123	0.121		
Age^2	0.002	0.221			0.002	0.131		
Male	-0.068	0.029	-0.074	0.017	-0.011	0.719		
Age Driving Licence	-0.059	<.001	-0.056	<.001	-0.067	<.001		
Vehicle Age	0.014	<.001	0.014	<.001	0.017	<.001		
Power	0.003	<.001	0.003	<.001	0.001	0.017		
Parking	0.029	0.420			0.032	0.381		
Log of km per year (thousands) - offset	1.000	--	1.000	--	1.000	--	1.000	--
Km per year at night (%)	-0.004	0.312					-0.001	0.771
Km per year at night (%)^2	0.0001	0.467					0.000	0.931
Km per year over speed limit (%)	0.019	0.001	0.019	0.001			0.018	0.001
Km per year over speed limit (%)^2	-0.001	0.001	-0.001	0.001			-0.001	0.003
Urban km per year (%)	0.026	<.001	0.026	<.001			0.027	<.001
Zero-inflation part								
Intercept (Logit)	-0.847	<.001	-0.857	<.001	-1.639	<.001	-0.795	<.001
Log of km per year (thousands) (Logit)	0.404	<.001	0.410	<.001	0.824	<.001	0.406	<.001
AIC	28,877.112		28,870.556		29,427.423		29,005.172	
BIC	28,999.019		28,951.828		29,508.694		29,070.189	

Riskcenter — UNIVERSITAT de BARCELONA

Introduction   Transition to telematics   Data and results   Scenarios and conclusions   Going forward to optimal pricing
○○○○○○   ○○○○○○○○○   ○○○○○○○●○○○○○   ○○○   ○○○

Two step correction

Concordant predictions of all models (in percentages).

	All variables	Non-telematics	Telematics	Telematics with offsets
Zero Poisson model results with offsets (Log of Km per year in thousands). All types of claims	62.36	59.10	61.39	62.20
Poisson model results with offsets (Log ok Km per year in thousands). All types of claims	62.15	58.60	61.18	62.05
Zero Poisson model results with offsets (Log of Km per year in thousands). Claims where the policyholder is at fault	62.71	59.85	61.17	62.77
Poisson model results with offsets (Log ok Km per year in thousands). Claims where the policyholder is at fault	62.38	58.96	60.89	62.43

Introduction Transition to telematics **Data and results** Scenarios and conclusions Going forward to optimal pricing
○○○○○○   ○○○○○○○○○   ○○○○○○○○○●○○○○   ○○○   ○○○

Speed reduction

Changing driving habbits

We study a sample of 9,614 young drivers with a pay-how-you-drive (PHYD) policy in force during the entire year 2010 in a Spanish insurer.

| Variable	Definition	Mean	Std. Dev.
km	Distance traveled during the year measured in kilometers	13,063.71	7,715.80
speed	% of kilometers traveled at speeds above the limit	9.14	8.76
urban	% of kilometers traveled on urban roads	26.29	14.18
age	Age of the driver	24.78	2.82
claims	Number of "at fault" accident claims during the year	0.10	0.32

Riskcenter UNIVERSITAT DE BARCELONA

One of the basic modeling approaches is the classical Poisson model Even
under the presence of overdispersion, the Poisson model provides
consistent parameter estimates of the linear predictor.

Poisson model results:

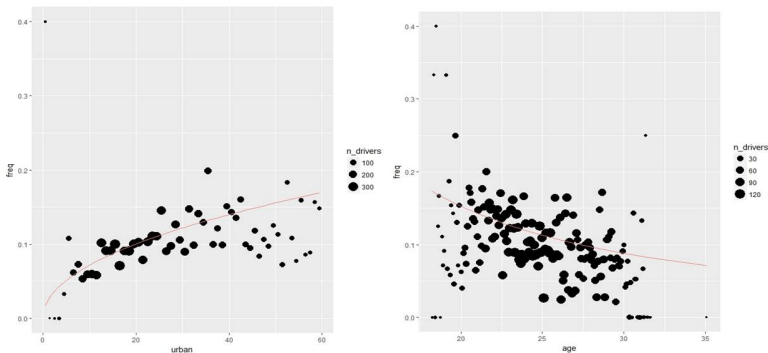| Parameter	Estimate	Std. error
Intercept	-3.2465	1.1662
ln(km)	0.3931	0.0593
ln(speed)	0.0653	0.0345
ln(urban)	0.4794	0.0692
ln(age)	-1.3580	0.2827

Guillen, Pérez-Marín (2018)

Figure: Expected number of claims vs. Distance (Km) and Excess Speed (%)
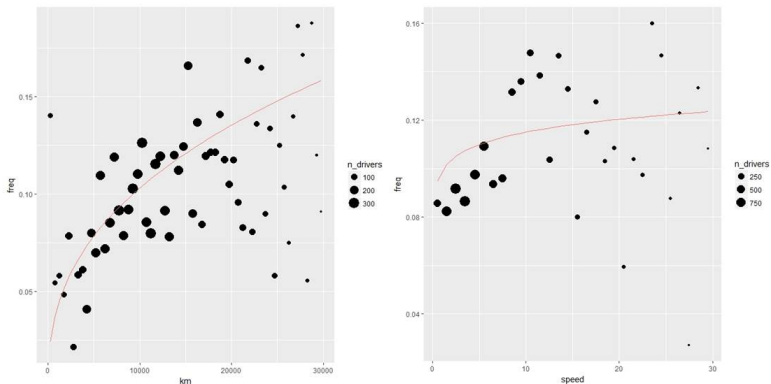
Guillen, Pérez-Marín (2018)

Figure: Expected number of claims vs. Urban (%) and Age (years)

Guillen, Pérez-Marín (2018)

Introduction   Transition to telematics   **Data and results**   Scenarios and conclusions   Going forward to optimal pricing
○○○○○○   ○○○○○○○○○   ○○○○○○○○○○○○○●   ○○○   ○○○

Takeaways

- The expected number of claims at fault increases with the distance travelled, but the increase is not proportional.

- Driving above the speed limits and extensive urban driving also increase the expected number of claims.

- The expected number of claims decreases with age.

4 Scenarios

Speed	Expected number of claims per 1,000 drivers
0%	**63.04**
1%	98.99
2%	103.57
5%	109.95
7%	112.39
9%	**114.25**
10%	115.04
12%	116.42
15%	118.13
17%	119.10
20%	**120.37**

All values in the matrix are in percentage. Cource: Guillen, Pérez-Marín (2018)

At the average level of speed limit violation (9%) the expected number of claims at fault per 1,000 drivers is 114. Elimination of speed limit violations (0%) leads to 63 claims per 1,000 drivers
Therefore, the initial level would be reduced by approximately one half.

Introduction   Transition to telematics   Data and results   **Scenarios and conclusions**   Going forward to optimal pricing
○○○○○○   ○○○○○○○○○   ○○○○○○○○○○○○○   ●○○   ○○○

Impact of changes in speed control

Figure: Relative change in expected number of claims per 1,000 drivers due to a change in the level of speed violation

Before	After										
	0%	1%	2%	5%	7%	9%	10%	12%	15%	17%	20%
0%	0	35.94	40.52	46.90	49.35	51.21	51.99	53.37	55.08	56.05	57.32
1%	-35.94	0	4.58	10.96	13.40	15.26	16.05	17.43	19.14	20.11	21.38
2%	-40.52	-4.58	0	6.38	8.82	10.68	11.47	12.85	14.56	15.53	16.80
5%	-46.90	-10.96	-6.38	0	2.44	4.30	5.09	6.47	8.18	9.14	10.42
7%	-49.35	-13.41	-8.82	-2.44	0	1.86	2.65	4.02	5.73	6.70	7.97
9%	-51.21	-15.26	-10.68	-4.30	-1.86	0	0.79	2.17	3.87	4.84	6.11
10%	-51.99	-16.05	-11.47	-5.09	-2.65	-0.79	0	1.38	3.09	4.05	5.33
12%	-53.37	-17.43	-12.85	-6.47	-4.02	-2.17	-1.38	0	1.71	2.68	3.95
15%	-55.08	-19.14	-14.56	-8.18	-5.73	-3.87	-3.09	-1.71	0	0.97	2.24
17%	-56.05	-20.11	-15.53	-9.14	-6.70	-4.84	-4.05	-2.68	-0.97	0	1.27
20%	-57.32	-21.38	-16.80	-10.42	-7.97	-6.11	-5.33	-3.95	-2.24	-1.27	0

Guillen, Pérez-Marín (2018)

Introduction
○○○○○○

Transition to telematics
○○○○○○○○○

Data and results
○○○○○○○○○○○○○

Scenarios and conclusions
○●○

Going forward to optimal pricing
○○○

Impact on motor insurance

Figure: Percentage of variation in the price of insurance due to a change in the level of speed violation, with 20% premium loading

	After										
Before	0%	1%	2%	5%	7%	9%	10%	12%	15%	17%	20%
0%	0%	45.6%	51.4%	59.5%	62.6%	65.0%	66.0%	67.7%	69.9%	71.1%	72.7%
1%	-29.1%	0%	3.7%	8.9%	10.8%	12.3%	13.0%	14.1%	15.5%	16.2%	17.2%
2%	-31.3%	-3.5%	0%	4.9%	6.8%	8.2%	8.9%	9.9%	11.2%	12.0%	13.0%
5%	-34.1%	-8.0%	-4.6%	0%	1.8%	3.1%	3.7%	4.7%	5.9%	6.6%	7.6%
7%	-35.1%	-9.5%	-6.3%	-1.7%	0%	1.3%	1.9%	2.9%	4.1%	4.8%	5.7%
9%	-35.9%	-10.7%	-7.5%	-3.0%	-1.3%	0%	0.5%	1.5%	2.7%	3.4%	4.3%
10%	-36.1%	-11.2%	-8.0%	-3.5%	-1.8%	-0.5%	0%	1.0%	2.1%	2.8%	3.7%
12%	-36.7%	-12.0%	-8.8%	-4.4%	-2.8%	-1.5%	-0.9%	0%	1.2%	1.8%	2.7%
15%	-37.3%	-13.0%	-9.9%	-5.5%	-3.9%	-2.6%	-2.1%	-1.1%	0%	0.7%	1.5%
17%	-37.6%	-13.5%	-10.4%	-6.1%	-4.5%	-3.2%	-2.7%	-1.8%	-0.6%	0%	0.8%
20%	-38.1%	-14.2%	-11.2%	-6.9%	-5.3%	-4.1%	-3.5%	-2.6%	-1.5%	-0.8%	0%

Guillen, Pérez-Marín (2018)

Riskcenter   UNIVERSITAT DE BARCELONA

Assume that all vehicles were equipped with automated speed control devices (excess speed 0%):

- the total number of fatal victims would be reduced by 28.6 deaths per one million drivers
- the number of accidents with victims (bodily injuries and/or death) would be reduced by 1.66 accidents per 1,000 drivers.

Spain has 26.5 milion drivers (DGT, 2017)

- 1,200 fatalities - 28.6*26.5= 1,200-742 = 458 fatalities (62% reduction)
- 4,837 severely injured using the same proportion-¿ 4,937-2,991=1,846 severely injured

6 Going forward to optimal pricing

Introduction
oooooo

Transition to telematics
ooooooooo

Data and results
ooooooooooooo

Scenarios and conclusions
ooo

Going forward to optimal pricing
ooo

Recap:

- Proposed a method to update premiums regularly with telematics data.

- Showed that the price per mile depends on driving habits and is not proportional to distance driven. A zero claim is relatively more frequent for intensive users.

- Estimated the impact on prices and safety in a plausible scenario of assisted driving that would control speed limit violations.

- Still want derive some **open-questions** about risk measures to summarize telematics big data and optimal pricing when customers may lapse.

Introduction
oooooo

Transition to telematics
ooooooooo

Data and results
oooooooooooooo

Scenarios and conclusions
ooo

Going forward to optimal pricing
ooo

Summary

```
┌─────────┐   ┌─────────┐   ┌─────────┐
│  Trip   ▶   │   Day   ▶   │  Week   ▶
└─────────┘   └─────────┘   └─────────┘
```

• Linear models	• Longitudinal and panel data models	• Bayesian regression models
• Regression with categorical dependent variables	• Linear mixed models	• Generalized additive models and nonparametric regression
• Regression with count-dependent variables	• Credibility and regression modeling	• Non-linear mixed models
• Generalized linear models	• Fat-tailed regression models	• Claims triangles/loss reserves
• Frequency and severity models	• Spatial modeling	• Survival models
	• Unsupervised learning	• Transition modeling

… and then correct premium

Pricing and Personalization

In dependent modelling claims, lapse and usage are all interconnected

Pricing and Personalization

Usage

Competition
between insurers

Road safety,
responsible driving

Risk

Lapse

Claims

Company profit
Insurance market

Introduction	Transition to telematics	Data and results	Scenarios and conclusions	Going forward to optimal pricing
oooooo	ooooooooo	ooooooooooooo	ooo	●oo

Further research

The future of telematics in insurance is usage based pricing and personalized services.

- The past is modelling number of claims and severity and the future is scoring the driver based also on **detailed information** (breaks, abrupt accelerations, traffic environment and driving style).

- Telematics makes **real time dynamic pricing** possible.

- With our proposal, daily estimates of usage correct the insurance premium and work effectively as a **modern version of experience rating**.

Award

An extended version of the semi-autonomous vehicle insurance article has been awarded the **Best Paper in the Non-Life Section of the International Congress of Actuaries** (Berlin, June 2018). The full version can be found at: www.ub.edu/riskcenter

UNIVERSITAT DE BARCELONA

Ayuso, M., Guillen, M. and Pérez-Marín, A.M. (2014) "Time and distance to first accident and driving patterns of young drivers with pay-as-you-drive insurance", **Accident Analysis and Prevention**, 73, 125-131.

Ayuso, M., Guillen, M., Pérez-Marín, A.M. (2016a) "Telematics and gender discrimination: some usage-based evidence on whether men's risk of accident differs from women's", **Risks**, 2016, 4, 10.

Ayuso, M., Guillen, M., and Pérez-Marín, A.M. (2016b) "Using GPS data to analyse the distance travelled to the first accident at fault in pay-as-you drive insurance", **Transportation Research Part C** 68, 160-167.

Boucher, J.P., Coté, S. and Guillen, M. (2018). "Exposure as duration and distance in telematics motor insurance using generalized additive models", **Risks**, 5(4), 54.

Ayuso, M., Guillen, M. and Nielsen, J.P. (2018). "Improving automobile insurance ratemaking using telematics: incorporating mileage and driver behaviour data", **Transportation**, accepted, in press.

Guillen, M., Nielsen, J.P., Ayuso, M. and Pérez-Marín, A.M. (2018). "The use of telematics devices to improve automobile insurance rates", **Risk Analysis**, accepted, in press.

Pérez-Marín, A.M. and Guillen, M. (2018). "The transition towards semi-autonomous vehicle insurance: the contribution of usage-based data ". Under revision in **Accident Analysis and Prevention**.

Is motor insurance ratemaking going to change with telematics and semi-autonomous vehicles?

Montserrat Guillén

University of Barcelona
mguillen@ub.edu
www.ub.edu/riskcenter

Riskcenter

UNIVERSITAT DE BARCELONA

Risk Seminar, CDAR
Berkeley, August 28, 2018