Asymmetric Information in Automobile Insurance: Evidence from Telematic Data

Daniela Kremslehner and Alexander Muermann

Department of Finance, Accounting and Statistics

WU Vienna

Effect of private information on contract choice and risk

This paper is based on a telematic data set of insured cars which

- provides detailed data about driving behavior
- is not accessible to the corresponding insurance company

Controlling for the insurer's risk classification we find

- number of car rides (controlling for distance driven) and relative distance driven on weekends are risk factors
- average speeding and number of car rides are both negatively to the level of third-party liability coverage
- number of car rides and relative distance at night are positively related to the level of first-party coverage

Results suggest overlapping effects of private information based on risk preferences and driving behavior

Residual correlation test

- Overlapping effects of asymmetric information pose challenge to empirical tests based on residual correlation between level of insurance coverage and risk
  - hidden degree of risk preferences (de Meza and Webb, 2001, Chiappori et al., 2006)
- Applying the residual correlation test to our insurance data spuriously alludes to different interpretations
  - no correlation between first-party coverage and risk
  - positive correlation between liability coverage and risk

Alternative empirical tests

- include survey data
  - Life: Cawley and Philipson (1999) - self-assessed risk not relevant
  - Long-term care: Finkelstein and McGarry (2006) - self-assessed prob. of nursing home use is positively related to both subsequent nursing home use and insurance coverage, evidence for risk-based selection
  - Medigap: Fang et al. (2008) - evidence for selection based on cognitive ability
- include "unused observables"
  - Annuity: Finkelstein and Poterba (2014) - postcode information (socio-economic characteristics) related to survival prob. and insurance coverage
  - Auto: Saito (2006) - postcode information (accident characteristics) not related to insurance coverage

Our contribution

We provide direct evidence of existence and effect of private information by combining insurance data with telematic data which

- contains detailed private information about driving behavior provided by independent, unbiased third party
- is contractually not accessible to the insurance company
  - "unused observables" might be used by insurance company through channels different than pricing, e.g., set of contracts offered, claims settlement process, cancellation policy
- is based on real decisions and behavior that is of direct interest to insurance company
  - survey responses might be biased, in particular if related to self-reported probabilities (Hurd et al., 1998; Hurd, 1999; Gan et al., 2005)

Background

- in Nov 2007, an insurance company introduced a pay-as-you-drive insurance contract with a telematic device using GPS
- independent telematic company develops hardware and collects and manages telematic data
- features of the contract
  - premium depends on aggregate distance driven (country roads and motorways are scaled by 0.8)
  - emergency device
  - crash sensor
  - fixed installation fee and monthly fee for safety services
  - discount on premium for full comprehensive coverage

Telematic data set

- 2340 cars
- 3.7 Mio data points
  - between Feb. 1st and April 30th 2009
  - each data point includes date, time, GPS-coordinates, direction, speed, distance, ignition status and road type
  - data point is recorded when engine is started, after every two km (1.24 mi) driven, and when engine is switched off
- data cleaning
  - only completed car rides
  - exclude car rides with speed above 200km/h (124 mph)
  - exclude car rides with distances between data points above 99.9% quantile

Summary statistics telematic data

	urban	country road	motorway	total
Number of cars				2,340
Number of car rides				537,181
Number of data points	1,717,049	686,042	744,542	3,147,633
Average speed in km/h	48	74	113	78
(mph)	(30)	(46)	(70)	(48)
Distance driven in km	2,041,466	1,195,018	1,567,140	4,803,624
(mi)	(1,268,508)	(742,550)	(973,776)	(2,984,834)
Avg. distance per ride in km				8.94
(mi)				(5.56)

Access to telematic data

- insurance company has only access to the telematic data necessary for the pricing of the contract
  - aggregate distance driven per road type
- insurance company contractually refrains from accessing any other telematic data
- telematic data provides us with detailed private information on driving behavior which is inaccessible to the insurance company

Insurance data set

- 2000 insurance contracts (includes all privately insured cars in the telematic data set)
  - all the information used for pricing the contract
  - Feb. 1st, 2009 and Feb. 1st, 2010
- data cleaning
  - car insured for entire period
  - only cars with more than 4 kW (5.4 hp)
- 1849 insurance contracts with corresponding telematic data

Insurance data set

- **Car:** year of construction, brand, engine power, value of car
- **Policyholder:** age, gender, postal code, bonus-malus class
- **Risk:** downgrade in Bonus-Malus class = submission of at least one liability claim during the year
- **Coverage:**
  - first-party: none, comprehensive (e.g., vandalism, theft, weather), full comprehensive (in addition at-fault collision)
  - liability: € 10 million ($ 12.8m), € 15 million ($ 19.2m)

Summary statistics insurance data

	Mean				
	total	none/compr.	full compr.	liab. 10m	liab. 15m
age in years	3.5	6.7	1.9	3.7	3.0
kW	87	84	89	87	88
(hp)	(117)	(112)	(119)	(116)	(118)
value of car in €	26,709	26,204	27,023	26,656	26,835
($)	(34,206)	(33,559)	(34,608)	(34,138)	(34,368)
age in years	48.7	48.2	48.9	48.13	49.9
male	0.61	0.61	0.61	0.61	0.62
urban	0.44	0.42	0.45	0.45	0.42
Bonus-Malus class	0.52	0.55	0.51	0.52	0.51
downgrade BM in %	7.6	9.7	6.6	7.9	7.0
N	1849	595	1254	1293	556

Driving behavior

We use the following indices of driving behavior:

- **AvgSpeeding:** average speed driven above the legal speed limits

$$\textbf{AvgSpeeding} = \frac{\sum_j \sum_{i \in \Delta_n} (v_{ij} - u_j)}{n}$$

  - j road type, u_j legal speed limit for road type j: 50 (31), 100 (62), 130 (81) km/h (mph)
  - $i = 1, ..., n$ is data point, v_{ij} is the speed of the car at data point i on road type j
  - $\Delta_n = \{i = 1, \ldots, n | v_{ij} > u_j\}$ set of data points where the speed of the car is above the legal speed limit
- **#Rides:** number of car rides
- **DistWE/Dist:** distance driven on weekends relative to total distance (data points between Fri and Sun midnight)
- **DistNight/Dist:** distance driven at night relative to total distance (data points between sunset and sunrise)

Summary statistics driving indices

		first-party		liability		$\triangle BM$ class	
	total	0	1	0	1	0	1
Mean AvgSpeeding	3.15	3.05	3.19	3.2	3.02	3.15	3.16
Mean #Rides	242	217	255	246	236	238	293
Mean DistWE/Dist	0.268	0.286	0.26	0.268	0.269	0.266	0.288
Mean DistNight/Dist	0.083	0.083	0.083	0.084	0.081	0.085	0.062
N	1849	595	1254	1293	556	1708	141

Econometric model: residual correlation test

- bivariate probit model by Chiappori and Salanié (2000)

$$Coverage = 1(X\beta_1 + \varepsilon > 0)$$
$$Risk = 1(X\gamma_1 + \eta > 0)$$

  - X vector of all risk classifying variables used by insurance company
  - test for the correlation ρ of the error terms ε and η
  - positive correlation $\rho > 0$ is interpreted as existence of relevant private information
- BUT with hidden degree of risk aversion and multidimensional hidden information rejecting (not rejecting) the hypothesis of non-dependence between insurance coverage and risk may not be indicative of existence (absence) of private information (Chiappori et al., 2006; Finkelstein and Poterba, 2006; Koufopoulos, 2007)

Econometric model: our model

- Our model: trivariate probit model based on Finkelstein and Poterba (2014) and Finkelstein and McGarry (2006)

$$
\begin{aligned}
CovLiab &= 1(X\beta_1 + Y\beta_2 + \varepsilon_1 > 0) \\
CovFP &= 1(X\gamma_1 + Y\gamma_2 + \varepsilon_2 > 0) \\
\Delta BM &= 1(X\delta_1 + Y\delta_2 + \varepsilon_3 > 0)
\end{aligned}
$$

  - Y vector of private information $(AvgSpeeding, \#Rides, DistWE/Dist, DistNight/Dist)$
  - $CovLiab$, $CovFP$ liability and first-party coverage
  - ΔBM downgrade in Bonus Malus class
  - coefficients of interest β_2, γ_2, and δ_2

Coefficients of driving behavior

	CovLiab	*CovFP*	$\triangle BM$
AvgSpeeding	-.0111*	.003	.0026
	(.0059)	(.0069)	(.0031)
#Rides	-.0002*	.0003**	.0002***
	(.0001)	(.0001)	(.0001)
DistWE/Dist	.0307	-.1335	.0934*
	(.0951)	(.1119)	(.0495)
DistNight/Dist	.1295	.3345**	-.0463
	(.1365)	(.1673)	(.0802)
Pseudo-R^2	.0160	.3532	.0508
N	1849	1849	1849

Other significant coefficients

	CovLiab	*CovFP*	$\triangle BM$
kW	0.0015**	-0.0005	0.0003
	(0.0008)	(0.0009)	(0.0005)
age of car	-0.0116***	-0.1003***	0.0073***
	(0.0031)	(0.0049)	(0.0014)
urban	-0.0509**	0.0733**	0.0231**
	(0.0224)	(0.0266)	(0.0126)
Bonus-Malus class	-0.0342	-0.3012***	0.043*
	(0.0766)	(0.0884)	(0.0384)
age of policyholder	0.0014*	-0.0003	0.0005
	(0.0008)	(0.001)	(0.0004)
total distance driven	-2.23e-08	7.49e-08***	1.82e-08
	(2.08e-08)	(2.67e-08)	(1.08e-08)

Number of car rides

# car rides / day	quantile	predicted probability of ΔBM	
		adj. for distance driven	not adj.
1.22	25%	4.27%	4.90%
2	42.5%	5.58%	5.95%
2.73	50%	7.09%	7.09%
3.78	75%	9.78%	9.02%
4	78.3%	10.44%	9.48%

Discussion

- risk factors
  - the number of car rides (controlling for distance driven): start and end of drive particularly exposed to accident risk
  - weekend drivers
- liability coverage
  - negative relations of average speeding and number of car rides to liability coverage indicate a selection and incentive effect based on hidden risk preferences (risk aversion, overconfidence)
- first-party coverage
  - positive relation of number of car rides to first-party coverage and to downgrade in BM class suggests selection and/or incentive effect based on driving behavior

Correlations of residual error terms

	without private information	with private information
$\rho_{Liab,FP}$	0.113**	0.128***
	(0.0106)	(0.0041)
$\rho_{Liab,\triangle BM}$	0.061*	0.06*
	(0.0822)	(0.0944)
$\rho_{FP,\triangle BM}$	-0.015	0.000
	(0.7311)	(0.9967)
N	1849	1849

Discussion

- consistent with literature, no statistically significant correlation of error terms between first-party coverage and downgrade in BM class **BUT** our direct test shows that private information does matter for contract choice and downgrade in BM class
- positive correlation of error terms between liability coverage and downgrade in BM class **BUT** our direct test suggest opposite selection based on risk preferences
- highly statistically significant, positive correlation of error terms between first-party and liability coverage
- these results suggest overlapping and counteracting effects of asymmetric information based on risk preferences and driving behavior

Selection bias

- since pay-as-you-drive insurance contract is offered for choice, there might be a selection bias
- we apply three bivariate probit models with selection as a Heckman correction method
- selection probit equation includes insurance data of a randomly selected sample of 2000 policyholders who chose not to switch to the pay-as-you-drive contract

$$selection = 1(X\gamma_1 + \varepsilon_1 > 0)$$

- outcome probit equation for *CovLiab*, *CovFP*, and ΔBM

$$outcome = 1(X\beta_1 + Y\beta_2 + \varepsilon_2 > 0)$$

- *male* as instrumental variable: gender of person who purchases insurance but not necessarily of the car driver

Results selection equation

	coefficients
kW	0.0014**
value of car in €	4.04e-06**
urban	0.2522***
male	-0.0622***
Bonus-Malus class	1.7238***
age of policyholder	-0.0055***
CovFP	0.4082***
Pseudo-R^2	0.2429
N	3985

Results outcome equation

	CovLiab	*CovFP*	$\triangle BM$
AvgSpeeding	-.011*	.0022	.0027
	(.0058)	(.0048)	(.0033)
#Rides	-.0002*	.0002*	.0002***
	(.0001)	(.0001)	(.0001)
DistWE/Dist	.0297	-.09	.0991
	(.0943)	(.0741)	(.0619)
DistNight/Dist	.13	.2263**	-.0539
	(.1353)	(.109)	(.0814)
N	1849	1849	1849

Disposable income

- average annual purchasing power per resident on post-code level for 2009
- mean € 17,572

	CovLiab	*CovFP*	$\triangle BM$
AvgSpeeding	-0.0116*	0.0037	0.0024
	(0.0059)	(0.0072)	(0.0031)
#Rides	-0.0002*	0.0003**	0.0002***
	(0.0001)	(0.0001)	(0.0001)
%DistWE	0.0300	-0.1335	0.0935*
	(0.0954)	(0.1119)	(0.0582)
%DistNight	0.1295	0.3345**	-0.0459
	(0.1365)	(0.1673)	(0.0770)

Conclusion

- we provide direct evidence that private information about driving behavior is related to contract choice and risk
  - number of car rides positively related to level of **first-party coverage** and downgrade in BM class
  - number of car rides and average speeding negatively related to **liability coverage**
- results suggest interaction of risk-based and preference-based selection effects
- residual correlation test can be misleading
  - no statistically significant correlation between **first-party coverage** and downgrade in BM class
  - positive correlation between **liability coverage** and downgrade in BM class