Customer retention and price elasticity

Are motor insurance policies homogeneous with respect to loyalty?

Montserrat Guillén[1],
Ana M. Pérez Marín[1] and Leo Guelman[1,2]

[1]University of Barcelona and [2]Royal Bank of Canada

Motivation

- ✓ Insurance companies operate in a more competitive environment than they used to do in the past
- ✓ Customers easily switch from one insurer to another
- ✓ Nowadays, the central problem for insurance companies is not only to create and launch new products for the market, but additionally to achieve commercial success by retaining customers
- ✓ Policy cancellations and lapses strongly influence the position of the company in the market and its level of risk
- ✓ Understanding customer behavior is extremely valuable for insurers

Difficulties

- ✓ Many factors influence customer decisions, so it is difficult to predict the probability of a customer lapse and the impact of loosing a customer
- ✓ We should take into account the relationship between events affecting one particular contract and customer's decisions regarding other contracts held in the same company

- ✓ Which specific actions should a company design?
- ✓ Which groups of customers should be targeted in order to prevent profit losses due to policy lapses?

Background (I):

- ✓ The demand for insurance products: Hammond, Houston and Melander (1967); Ducker (1969); Mayers and Smith (1983); Doherty (1984); Babbel (1985); Showers and Shotick (1994); Ben-Arab, Brys and Schlesinger (1996); Gandolfi and Miners (1996)
- ✓ Customer satisfaction and loyalty: Crosby and Stephens (1987); Schlesinger and Schulenburg (1993); Wells and Stafford (1995); Stafford et al. (1998); Cooley (2002), Kuo, Tsai and Chen (2003); Bozzetto et al. (2004); Guillén, Nielsen and Pérez-Marín (2006)

Background (II):

- ✓ Customer Lifetime Value (CLV) in insurance: Jackson (1989); Berger and Nasr (1998); Verhoef and Donkers (2001), Ryals and Knox (2005), Donkers, Verhoef and Jong (2007), Haenlein, Kaplanb and Beeserc (2007), Guillén, Nielsen and Pérez-Marín (2008)
- ✓ Cross-selling and multiple contracts: Bonato and Zweifel (2002); Guillén, Gustafsson, Hansen and Nielsen (2008), Thuring, Nielsen, Guillén and Bolancé (2012).

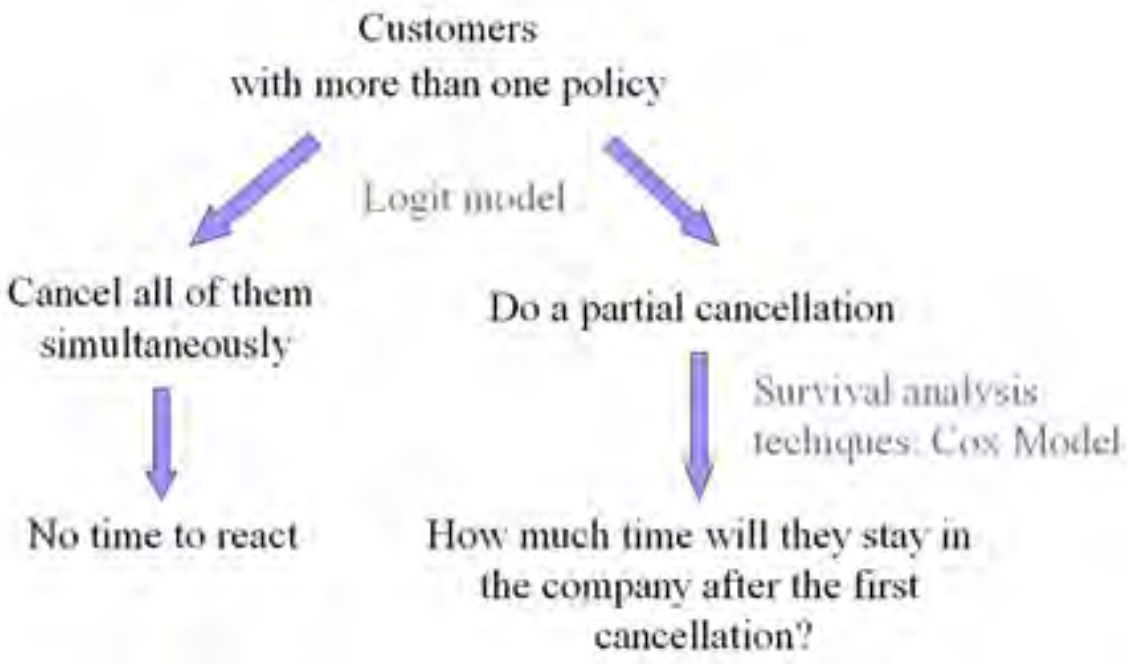
Households are customer units

- ✓ Brockett, P. L. et al. (2008) Survival Analysis of Household Insurance Policies: How Much Time Do You Have to Stop Total Customer Defection, Journal of Risk and Insurance 75, 3, 713-737.
- ✓ Guillen, M., Nielsen, J. P., Scheike, T. and Perez-Marin, A. M. (2011a) Time-varying effects in the analysis of customer loyalty: a case study in insurance, Expert Systems with Applications, 39, 3551-3558.

Households are customer units

- ✓ Using the household as the unit of analysis, we focus on the behavior of households having several policies (of three possible types: building, content and motor) in the same insurance company, and who cancel their first policy.
- ✓ How much time after the household's cancellation of the first policy does the insurer have to retain the customer and avoid customer defection on all policies to the competitor?
- ✓ What customer characteristics are associated with customer loyalty?

Methodology



Empirical study: customer information

- ✓ 151,290 Danish households with several insurance policies (of three possible types: building, content and motor), who cancel their first policy between 1997-2001.
- ✓ Customer information of all household policy holders : age, gender, tenure, occurrence of a claim, change of address, intervention of a external company, premium increase, policies in force, among other covariates.

Empirical study: results

Logistic Regression (Probability of Total Cancellation)

Parameter	Estimate	Standard Error	Odds Ratio	p-value
Constant	-2.201	0.112	-	<0.001
Change of address, less 2 m. ago	-0.596	0.060	0.551	<0.001
Change of address, 2 - 6 m. ago	-0.122	0.052	0.885	0.019
Change of address, 6 - 12 m. ago	-0.095	0.048	0.909	0.049
Change of address, 12 - 24 m. ago	0.229	0.041	1.258	<0.001
Change of address, more 24 m. ago	0.531	0.044	1.701	<0.001
Tenure	-0.011	0.001	0.989	<0.001
Claims, less 2 months ago	0.230	0.040	1.259	<0.001
Claims, 2 and 6 months ago	0.324	0.035	1.383	<0.001
Claims, 6 and 12 months ago	0.440	0.035	1.553	<0.001
Claims, 12 and 24 months ago	0.469	0.037	1.598	<0.001
Claims, more 2 years ago	0.546	0.054	1.727	<0.001
Contents0	0.277	0.087	1.319	0.001
Corecust	0.109	0.025	1.116	<0.001
Age	0.004	0.001	1.004	<0.001
External company A	2.548	0.041	12.779	<0.001
External company B	2.165	0.046	8.718	<0.001
External company C	1.893	0.048	6.637	<0.001
External company D	2.270	0.047	8.834	<0.001
Another known external company	1.686	0.035	9.680	<0.001
Gender (male)	0.099	0.028	1.104	<0.001
House0	-0.657	0.030	0.518	<0.001
Motor0	-1.253	0.033	0.286	<0.001
Newcontents	-0.113	0.043	0.893	0.008
Newhouse	0.073	0.060	1.076	0.225
Newmotor	-0.208	0.050	0.813	<0.001
Notice	-0.002	<0.001	0.998	<0.001
Pruning within past 12 months	-0.187	0.072	0.829	0.009
Pruning more than one year ago	0.086	0.111	1.089	0.442

Cox Regression (Customer Duration after First Policy Cancellation

Parameter	Estimate	Standard Error	Hazard Rate	p-value
Change of address less 2 m. ago	-0.245	0.019	0.783	<0.001
Change of address 2 - 6 m. ago	-0.083	0.020	0.920	<0.001
Change of address 6 - 12 m. ago	-0.023	0.020	0.978	0.252
Change of address 12 - 24 m. ago	0.044	0.019	1.045	0.021
Change of address more 24 m. ago	0.157	0.025	1.170	<0.001
Tenure	-0.003	0.001	0.997	<0.001
Claims, less 2 m. ago	0.096	0.016	1.100	<0.001
Claims, 2 - 6 m. ago	0.161	0.015	1.175	<0.001
Claims, 6 - 12 m. ago	0.185	0.015	1.203	<0.001
Claims, 12 - 24 m. ago	0.228	0.017	1.256	<0.001
Claims, more 24 m. ago	0.257	0.027	1.294	<0.001
Contents0	0.681	0.029	1.975	<0.001
Contents1	-0.869	0.016	0.419	<0.001
Corecust	-0.042	0.011	0.959	<0.001
Age	-0.003	<0.001	0.998	<0.001
External company A	1.727	0.018	5.625	<0.001
External company B	1.528	0.020	4.611	<0.001
External company C	1.652	0.019	5.217	<0.001
External company D	1.778	0.020	5.919	<0.001
Another known external company	1.643	0.012	5.170	<0.001
Gender (male)	0.103	0.012	1.109	<0.001
House0	0.222	0.017	1.249	<0.001
House1	-0.559	0.017	0.571	<0.001
Motor0	0.415	0.021	1.515	<0.001
Motor1	-0.529	0.018	0.589	<0.001
Newcontents	-0.059	0.016	0.942	<0.001
Newhouse	-0.109	0.024	0.897	<0.001
Newmotor	-0.076	0.017	0.927	<0.001
Notice	-0.001	<0.001	1.000	<0.001
Pruning within past 12 months	0.188	0.029	1.207	<0.001
Pruning more than 1 year ago	0.095	0.053	1.100	0.075

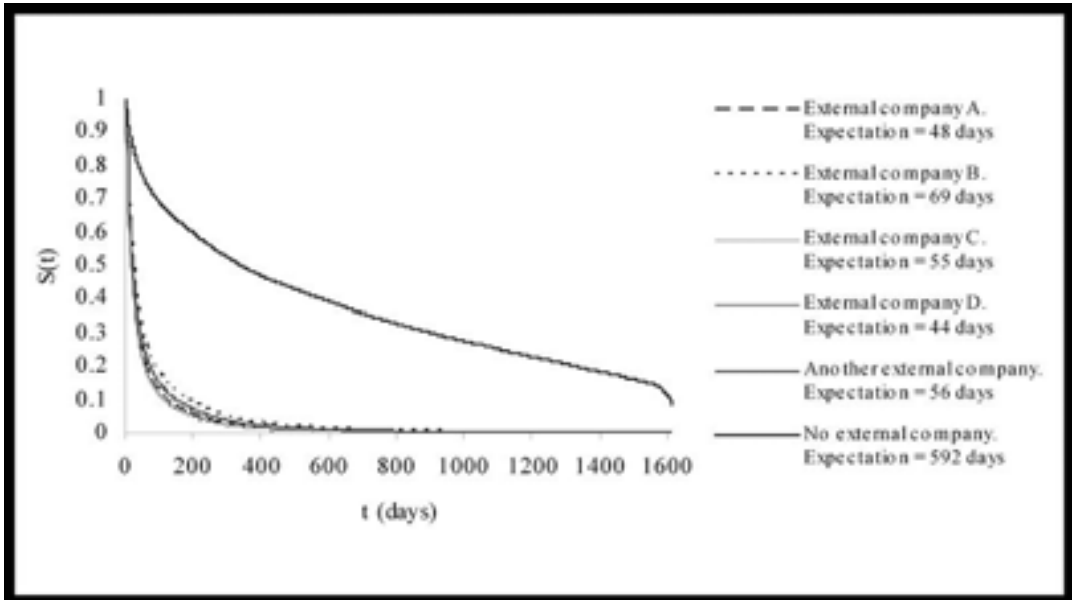
Conclusions from the empirical study

In Brockett et al. (2008) we identified relevant factors increasing the probability of total cancellation and also reducing customer duration:

- ✓ The most important factor that increases the probability of a total cancellation is the intervention of an external company.
- ✓ Claims, change of address (more than one year ago) and having a contents policy are factors that influence the probability of a total cancellation and also reduce customer duration after first policy cancellation.

Empirical study: duration after first cancellation

Example: survival function for a standard customer depending on the intervention of an external company.



Time-varying effect of covariates on loyalty

In Guillen et al. (2011a) we proved that some variables explaining customer loyalty have a time-varying effect:

- ✓ The kind of contracts held by the customer and the concurrence of an external competitor strongly influence customer loyalty right after the first cancellation, but the influence of those factors becomes much less significant some months later.
- ✓ Therefore, predictions of the probability of losing a customer can be readjusted over time to have a more precise estimation of customer duration.

Selling more policies to existing policyholders

- ✓ Guillen, M., Perez-Marin, A.M. and Alcañiz, M. (2011) A logistic regression approach to estimating customer profit loss due to lapses in insurance, Insurance Markets and Companies: Analyses and Actuarial Computations, 2, 2, 42-54.
- ✓ Thuring, F., Nielsen, J.P., Guillén, M. and Bolancé, C. (2012) Selecting prospects for cross-selling financial products using multivariate credibility, Expert Systems with Applications, 39, 10, 8809-8816.

Key Concepts

Guillen et al. (2011b) proposed a methodology for estimating profit loss due to policy cancellations when customers have more than one contract with the same insurer. Profits (based on the difference between premiums and cost of claims) are calculated by using three approaches:

- ✓ *Historical profit*: profit accumulated during a period of time.
- ✓ *Prospective profit*: profit that the company expects to receive if the customer keeps the policies in force. A logistic regression is proposed for estimating the probability of policy renewal.
- ✓ *Potential profit*: added profit when the customer underwrites new policies of a different type than those he currently has (cross-buying). A logistic regression is proposed for estimating the probability of selling new contracts.

Empirical application

- 79,599 customers with at least one policy in force (2005-2008).
- Two types of policies are considered: motor and diverse.
- Average profit loss due to policy lapse: 21.1 Euro per customer.

Example: Historical, Prospective and Potential Profits (2006-2007)

in Euros by type of policy in force

			Policy	
	Marital Status	Profit	Auto	Diverse
Men	Married	Historical	475,86	863,88
		Prospective	239,56	265,86
		Potential	7,34	78,39
	Single	Historical	547,29	732,23
		Prospective	239,16	265,86
		Potential	5,62	42,01
Women	Married	Historical	440,85	573,23
		Prospective	240,16	265,86
		Potential	5,43	29,89
	Single	Historical	522,21	488,79
		Prospective	239,90	265,86
		Potential	4,15	14,54

Targeting groups of customers

In Thuring et al. (2012) a method for selecting prospects for cross-selling insurance products was proposed:

- ✓ For many companies a possible way to expand its business is to sell more products to preferred customers in its portfolio.
- ✓ Historical data on customers' past behavior can be used to assess whether or not more products should be offered to a specific customer.
- ✓ In this paper we implement a method for using historical information of each individual customer, and the portfolio as a whole, to select a target group of customers to whom it would be interesting to offer more products.

Methodology

- ✓ We used multivariate credibility theory to estimate a customer specific latent risk profile and evaluate if a specific additional product, of a specific customer, is expected to contribute positively to the profit of the company, if that product is cross-sold to the customer.
- ✓ The profit is measured as the customer specific deviation between the a priori expected number of incidents (insurance claims) and the corresponding observed number.

Empirical study for cross-selling

- ✓ We analyzed a sample of 3,395 customers of an insurance company (between 1999 and 2004) who have owned all of the main insurance coverages (motor, building and content).
- ✓ We find that it is easier to identify the 20 percent of the data containing customers to avoid than the 20 percent of the data containing customers to target.
- ✓ By targeting all customers but the worst 20 percent the company could expect a subset of customer associated with less claims than a priori expected indifferent of which product is considered.
- ✓ The remaining 20 percent consist of customers with up to 64 percent more claims than a priori expected.

Retention of good customers: a new perspective

- ✓ Guelman, L., Guillen, M. and Perez-Marin, A. M. (2012) Random forest for uplift modeling: an insurance customer retention case, Lecture Notes in Business Information Processing, 115, 123-133.
- ✓ Guelman, L., Guillen, M. and Perez-Marin, A. M. (2013) Uplift random forests, Cybernetics and Systems: an International Journal, submitted.

Targeting the right customers

- An insurance company is interested in increasing the retention rate of its customers.
- The point is to decide which customers should be targeted by some retention action.
- Instead of targeting the most likely to leave customers, the authors advocate that the company should target those customers with a higher expected increase in the retention probability as a result of the marketing action by using uplift modeling.

If targeted by retention action	If NOT targeted by retention action	Remark
Churn	Churn	Unnecesary costs
Renew	Renew	Unnecesary costs
Churn	Renew	Negative effects
Renew	Churn	Best targets!

Methodology:

Notation:

- $X = \{X_1, ..., X_p\}$ a vector of predictor variables,
- $Y =$ binary response variable (1=renew, 0=lapse)
- t refers to the treatment $(t = 1)$ and control $(t = 0)$
- $L =$ a collection of observations $\{(y_\ell, x_\ell, t_\ell)\,; \ell = 1, ..., L\}$
- Uplift model
  $\widehat{f}^{uplitf}(x_\ell) = E\left(Y_\ell | x_\ell; t_\ell = 1\right) - E\left(Y_\ell | x_\ell; t_\ell = 0\right)$

Uplift model: indirect estimation

There are two general approaches: indirect and direct estimation

- Indirect uplift estimation:
  - Build two separate models, one using the treatment ($t = 1$) subset and another one using control data ($t = 0$).
    Predicted uplift is estimated by subtracting the class probabilities from the two models
    $P(Y = 1|x; t = 1) - P(Y = 1|x; t = 0)$
  - Alternatively, a single model can be obtained including an interaction term for every predictor in $X = \{X_1, ..., X_p\}$ and treatment t.
    This method does not work very well in practice, as the relevant predictors for the response are likely to be different from the relevant uplift predictors and the functional form of the predictors are likely to be different as well.

Uplift model: direct estimation

- Modeling uplift directly:
  - Requires modifying existing methods/algorithms or designing novel ones
  - Intuitively, tree-based algorithms are appropriate as they partition the input space into subgroups
  - Rzepakowski and Jaroszewicz (2011)and Radcliffe and Surry (2011) have proposed estimation algorithms
  - Our proposed algorithm: uplift Random Forests

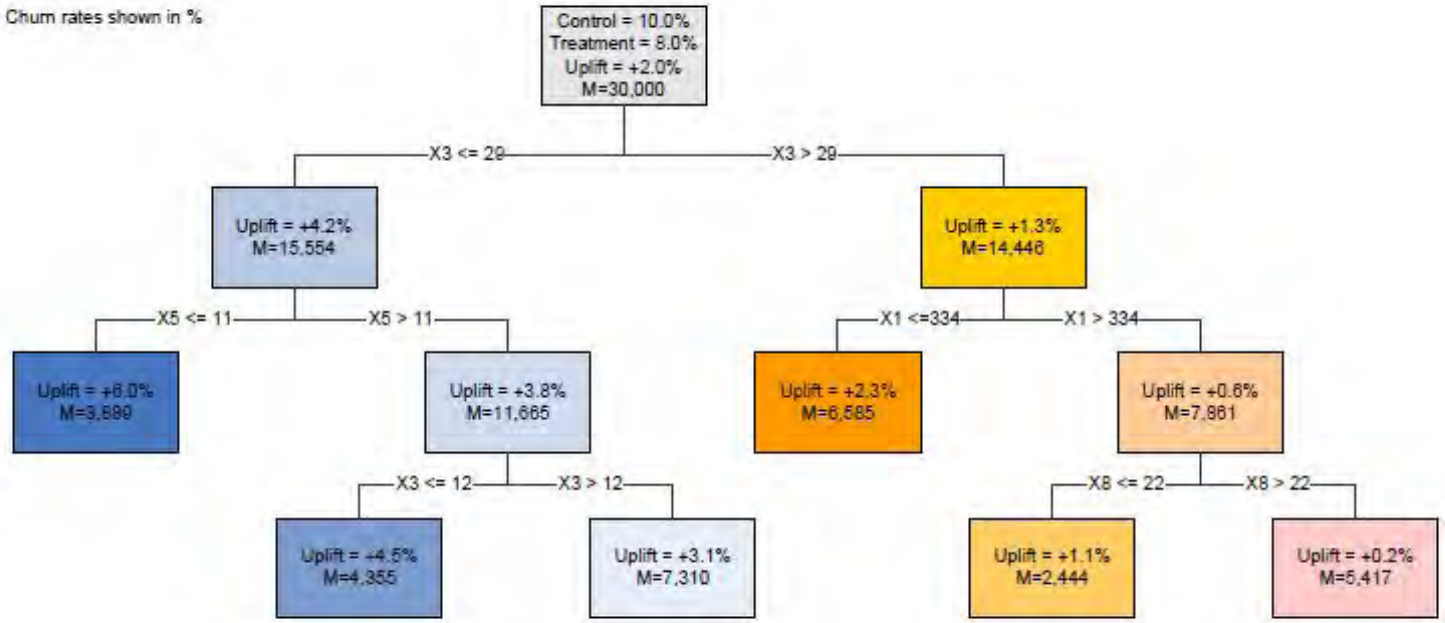
Methodology: uplift Random Forests

- In Guelman et al. (2012 and 2013) the proposed algorithm for modeling uplift directly is based on maximizing the distance in the class distributions between treatment and control groups
- Relative Entropy or *Kullback-Leibler distance KL* between two probability mass functions $P_t(Y)$ and $P_c(Y)$ is given by

$$KL\left(P_t(Y)||P_c(Y)\right) = \sum_{y \in Y} P_t(y) \log \frac{P_t(y)}{P_c(y)}$$

.

Methodology: illustration



Methodology

- Conditional on a given split Ω, KL becomes

$$KL\left(P_t(Y)||P_c(Y)|\Omega\right) = \sum_{\omega\in\Omega}\frac{M(\omega)}{M}KL\left(P_t(Y|\omega)||P_c(Y|\omega)\right)$$

  where $M = M_t + M_c$ (the sum of the number of training cases in treatment and control groups) and $M(\omega) = M_t(\omega) + M_c(\omega)$ (the sum of the number of training cases in which the outcome of the uplift Ω is ω in treatment and control groups).
- Define KL_{gain} as the increase in the KL divergence from a split Ω relative to the KL divergence in the parent node

$$KL_{gain}(\Omega) = KL(P_t(Y)||P_c(Y)|\Omega) - KL(P_t(Y)||P_c(Y))$$

Methodology

- Final split criterion is

$$KL_{ratio}(\Omega) = \frac{KL_{gain}(\Omega)}{KL_{norm}(\Omega)}$$

where KL_{norm} is a normalization factor that punishes:

  - splits with different treatment/control proportions on each branch
  - splits with unbalanced number of cases on each branch

Empirical study: targeting customers that react to campaigns

- Auto insurance portfolio from a large Canadian insurer
- A sample of approx. 12,000 customers coming up for renewal were randomly allocated into two groups:
  - Renewal letter+courtesy call: aim was to maximize customer retention
  - A control group: no retention efforts
  - Treatment is not effective if targets are selected randomly

	Attrition rates by group		
	Overall	Letter + Call	Control
Retained policies	10857	7492	3365
Cancelled policies	1111	757	354
Attrition rate	9.3%	9.2%	9.5%

Empirical study

We compare four uplift models:

- Uplift Random Forest Algorithm (upliftRF)
- The Two-Model Approach by using logistic regression (two-model)
- A Single Uplift Tree with Pruning (single-tree)
- and the approach based on explicitly adding an interaction term between each predictor and the treatment indicator by using logistic regression (int-model)

Top decile uplift

	Attrition rate (%)		
	Control	Treatment	Uplift
upliftRF	21.24	9.21	12.03
two-model	33.60	23.03	10.57
single-tree	13.98	5.21	8.77
int-model	27.41	20.60	6.81
random	9.50	9.20	0.30

Empirical study

Conclusions:

- None of the models dominates the others at all target volumes
- The *upliftRF* performs best in this application, specially for low target volumes: it is able to identify a 30 percent of customers for whom the retention program was highly effective (any additional targeted customer would result in a smaller reduction in attrition, as a result of negative effects of the campaign on the remaining customers)
- The *int-model* and *two-model* are able to identify the top 10 percent customers with highest attrition rate, but not those most impacted by the retention activity

Retention combined with price changes

Guelman, L. and Guillen, M. (2013) A causal inference approach to measure price elasticity in automobile insurance, Working Paper.

The role of price in customer retention

- Understanding price sensitivities at the individual policy holder level is extremely valuable for insurers.
- A rate increase has a direct impact on the premiums customers are paying, but there is also a causal effect on the customers decision to renew the policy term.
- It is difficult to measure price elasticity from most insurance datasets, as historical rate changes are reflective of a risk-based pricing exercise, therefore they are not assigned at random across the portfolio of policyholders.
- We propose a causal inference framework to measure price elasticity in the context of auto insurance.

Methodology

- L policyholders, $\ell = \{1, 2, \ldots, L\}$.
- vector of pre-treatment covariates $\mathbf{x}_\ell$.
- ordered treatment variable t (rate change levels), which takes values $t = \{1 < 2 < \ldots < T\}$ on a set $\Im$.
- $Z_{\ell t}$ set of T binary treatment indicators, $Z_{\ell t} = 1$ if subject ℓ received treatment t, and $Z_{\ell t} = 0$ otherwise.
- potential responses $r_{\ell t}$, renewal outcome that would be observed from policyholder ℓ if assigned to treatment t .
- observed response for subject ℓ is $R_\ell = \sum_{t \in \Im} Z_{\ell t} r_{\ell t}$.
- Our interest is to estimate price elasticity, defined as the renewal outcomes that result and are caused by the price change interventions.

Methodology

- We focus on the binary treatment case, $t = \{0, 1\} \in \Im$, $Z_\ell = 1$ if subject ℓ received the first treatment (the *treated* subjects), and $Z_\ell = 0$ if received the alternative treatment (the *control* subjects), without loss of generality.
- The average treatment effect (ATE) of treatment 1 relative to treatment 0 is:

$$\begin{aligned} ATE &= E[r_{\ell 1}|Z_\ell = 1] - E[r_{\ell 0}|Z_\ell = 0] \\ &= E[R_\ell|Z_\ell = 1] - E[R_\ell|Z_\ell = 0]. \end{aligned} \tag{1}$$

Methodology

Two basic assumptions:

- *unconfoundedness assumption*: conditional on $\mathbf{x}_\ell$, the outcomes $(r_{\ell 1}, r_{\ell 0})$ are independent of treatment Z_ℓ

$$(r_{\ell 1}, r_{\ell 0}) \perp Z_\ell | \mathbf{x}_\ell. \tag{2}$$

- *common support*: every unit in the population has a chance of receiving both treatments

$$0 < \pi(\mathbf{x}_\ell) \equiv P(Z_\ell = 1 | \mathbf{x}_\ell) < 1, \tag{3}$$

where $\pi(\mathbf{x})$ is known as the *propensity score*, conditional probability of assignment to treatment 1 given the pre-treatment covariates

Methodology

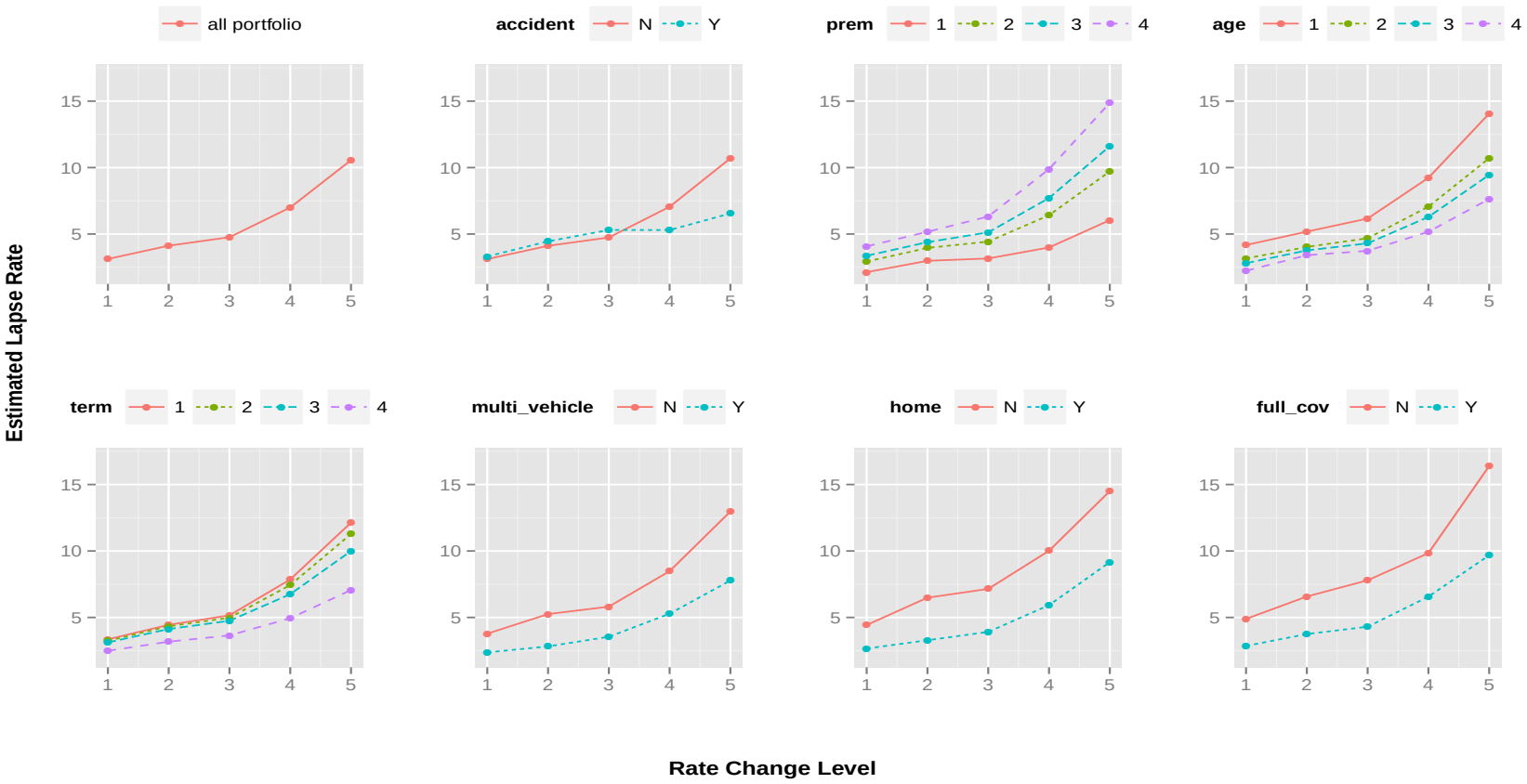
A matching algorithm is necessary:

- only one treatment is observed for each individual
- finding treated and control observations with similar values of the covariates will be impractical
- a matching algorithm based on the *propensity score* is implemented, which is sufficient to produce unbiased estimates of the average treatment effect

Empirical application: the data

- $L = 329,000$ auto insurance policies from a major Canadian insurer that have been given a renewal offer from June-2010 to May-2012 consisting on a new rate either lower, equal or higher than the current rate.
- more than 60 pre-treatment covariates (characteristics of the policy, the vehicle and driver).
- the treatment is the rate change: percentage change in premium from the current to the new rate, categorized into 5 ordered values $t = \{1 < 2 < \ldots < 5\}$.
- response variable: renewal outcome of the policy, measured 30 days after the effective date of the new policy term

Empirical application: estimated lapse rate



Empirical application: managerial implications

Which rate change should be applied to each policyholder to maximize the overall expected profit for the company subject to a fixed overall retention rate?

$$\underset{Z_{\ell t}\forall\ell\forall t}{\text{Max}} \sum_{\ell=1}^{L}\sum_{t=1}^{T} Z_{\ell t}\left[P_\ell(1+RC_t)(1-\hat{LR}_{\ell t})(1-\hat{r}_{\ell t})\right]$$

where P_ℓ is the current premium, RC_t is the actual rate change level associated with treatment t, $\hat{LR}_{\ell t}$ the predicted loss ratio (i.e., the ratio of the predicted insurance losses relative to premium), $\hat{r}_{\ell t}$ is the lapse probability of subject ℓ if exposed to rate change level t, and α the overall lapse rate of the portfolio.

Empirical application: managerial implications

The expected function to maximize is the expected profit of the portfolio

$$\underset{Z_{\ell t}\forall\ell\forall t}{\text{Max}} \sum_{\ell=1}^{L}\sum_{t=1}^{T} Z_{\ell t}\left[P_{\ell}(1+RC_t)(1-\hat{LR}_{\ell t})(1-\hat{r}_{\ell t})\right]$$

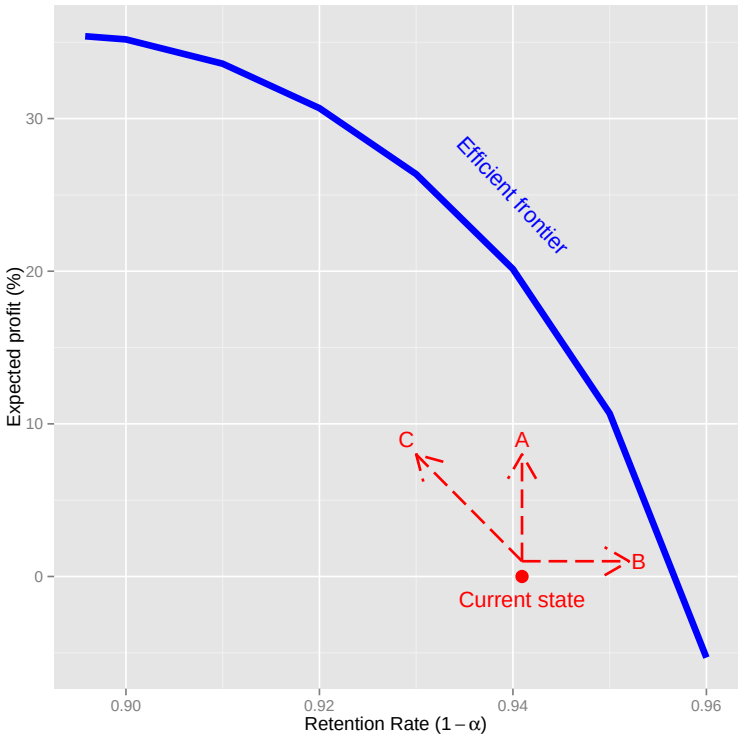
subject to the following constraints

$$\sum_{t=1}^{T} Z_{\ell t} = 1 \quad : \forall \ell$$

$$Z_{\ell t} \in \{0,1\}$$

$$\sum_{\ell=1}^{L}\sum_{t=1}^{T} Z_{\ell t}\hat{r}_{\ell t}/L \leq \alpha$$

Empirical application: managerial implications



Conclusions

- We have presented an approach to estimate price elasticity functions which allows for heterogeneous causal effects as a result of rate change interventions
- The model can assist managers in selecting an optimal rate change level for each policyholder for the purpose of maximizing the overall profits for the company
- Valuable insights can be gained by knowing the current company's position of growth and profitability relative to the optimal values given by the efficient frontier
- The managerial decision is to determine in which direction the company should move towards the frontier, as each decision point places a different weight on each of these objectives.

Sagrada Familia (Barcelona)

Work in progress...

Customer retention and price elasticity

Are motor insurance policies homogeneous with respect to loyalty?

Montserrat Guillén[1],
Ana M. Pérez Marín[1] and Leo Guelman[1,2]

[1]University of Barcelona and [2]Royal Bank of Canada