Actuarial Pricing in a Competitive Insurance Market

Arthur Charpentier

UQAM QUANTACT

AMSE Marseille, December 2020
Insurance & Actuarial Science

“Insurance is the contribution of the many to the misfortune of the few”

What would be a “fair contribution”? see O’Neill (1997)

- pure actuarial fairness contributions for individual policyholders should perfectly reflect their predicted risk levels → predictive modeling

- choice-sensitive fairness contributions should take into account only risks that result from choices - luck-egalitarianism (Cohen (1989) or Arneson (2011))
Insurance Pricing

“Insurance is the contribution of the many to the misfortune of the few”

mutualization (of risk): spread of risk among several parties see pooling, sharing \( \pi = \mathbb{E}_P[S_1] \)

(market) segmentation: division of a market into identifiable groups see differentiation, customization \( \pi(\omega) = \mathbb{E}_P[S_1|\Omega = \omega] \)
for some (unobservable) risk factor \( \Omega \)

Use of features (covariates) \( x \) as a proxy
\[
\pi(x) = \mathbb{E}_P[S_1|X = x] = \mathbb{E}_{P_x}[S_1]
\]
Risk Transfert without Segmentation

<table>
<thead>
<tr>
<th></th>
<th>Insured</th>
<th>Insurer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>$\mathbb{E}[S]$</td>
<td>$S - \mathbb{E}[S]$</td>
</tr>
<tr>
<td>Average Loss</td>
<td>$\mathbb{E}[S]$</td>
<td>0</td>
</tr>
<tr>
<td>Variance</td>
<td>0</td>
<td>$\text{Var}[S]$</td>
</tr>
</tbody>
</table>

All the risk - $\text{Var}[S]$ - is kept by the insurance company. **Remark:** interpretations are discussed in Denuit & Charpentier (2004).
Risk Transfer with Segmentation and Perfect Information

Assume that information $\Omega$ is observable,

\[
\begin{array}{|c|c|c|}
\hline
 & \text{Insured} & \text{Insurer} \\
\hline
\text{Loss} & \mathbb{E}[S|\Omega] & S - \mathbb{E}[S|\Omega] \\
\text{Average Loss} & \mathbb{E}[S] & 0 \\
\text{Variance} & \text{Var}\left[\mathbb{E}[S|\Omega]\right] & \text{Var}\left[S - \mathbb{E}[S|\Omega]\right] \\
\hline
\end{array}
\]

Observe that $\text{Var}\left[S - \mathbb{E}[S|\Omega]\right] = \mathbb{E}\left[\text{Var}[S|\Omega]\right]$, so that

\[
\text{Var}[S] = \mathbb{E}\left[\text{Var}[S|\Omega]\right] + \text{Var}\left[\mathbb{E}[S|\Omega]\right].
\]

→ insurer  \quad \rightarrow \quad \rightarrow \quad \text{insured}
Segmentation and Imperfect Information

Assume that $X \subset \Omega$ is observable

<table>
<thead>
<tr>
<th>Risk Factor (X)</th>
<th>Insured</th>
<th>Insurer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>$E[S</td>
<td>X]$</td>
</tr>
<tr>
<td>Average Loss</td>
<td>$E[S]$</td>
<td>0</td>
</tr>
<tr>
<td>Variance</td>
<td>$\text{Var}[E[S</td>
<td>X]</td>
</tr>
</tbody>
</table>

Now

$$E[\text{Var}[S|X]] = E[E[\text{Var}[S|\Omega]|X]] + E[\text{Var}[E[S|\Omega]|X]]$$

$$= E[\text{Var}[S|\Omega]] + E \left\{ \text{Var}[E[S|\Omega]|X] \right\}.$$

spiral of segmentation...

@freakonometrics  🦊 freakonometrics  🌐 freakonometrics.hypotheses.org-  Arthur Charpentier, UQAM, Dec. 2020
Assume that $X \subset \Omega$ is observable

<table>
<thead>
<tr>
<th></th>
<th>Insured</th>
<th>Insurer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss</td>
<td>$\mathbb{E}[S</td>
<td>X]$</td>
</tr>
<tr>
<td>Average Loss</td>
<td>$\mathbb{E}[S]$</td>
<td>0</td>
</tr>
<tr>
<td>Variance</td>
<td>$\text{Var}[\mathbb{E}[S</td>
<td>X]]$</td>
</tr>
</tbody>
</table>

Now

$$\mathbb{E}\left[\text{Var}[S|X]\right] = \mathbb{E}\left[\mathbb{E}\left[\text{Var}[S|\Omega]|X\right]\right] + \mathbb{E}\left[\text{Var}\left[\mathbb{E}[S|\Omega]|X\right]\right]$$

$$= \mathbb{E}\left[\text{Var}[S|\Omega]\right] + \mathbb{E}\left\{\text{Var}\left[\mathbb{E}[S|\Omega]|X\right]\right\}.$$ 

**spiral of segmentation...**
Consider a simple model: number of claims $N \in \{0, 1\}$ and fixed cost, say $1,000$. Simple classification problem.

Assume that $N \sim B(\theta)$ and we have some covariate $x$.

Two important features:

- the dispersion of the heterogeneity, i.e. the variance of $\theta$
  e.g. assume that $\theta$ has a Beta distribution on $[0, 1]$
- the dependence between heterogeneity $\theta$ and covariate $x$
the dispersion of the heterogeneity, $d = 10\%$ or $40\%$

the dependence between heterogeneity $\theta$ and covariate $x$
(correlation of the underlying copula function)

very difficult to reach high AUC (area under the ROC curve)

More complicated for insurance premiums...

Frees et al. (2014) defined a ROC-type curve, inspired by Lorenz curve: given observed losses $s_i$ and premiums $\hat{\pi}(x_i)$, policyholders ordered by premiums, $\hat{\pi}(x_1) \geq \hat{\pi}(x_2) \geq \cdots \geq \hat{\pi}(x_n)$,

plot $\{F_i, L_i\}$ with 
\[
F_i = \frac{i}{n} \quad \text{proportion of insured} \\
L_i = \frac{\sum_{j=1}^{i} s_j}{\sum_{j=1}^{n} s_j} \quad \text{proportion of losses}
\]
Actuarial Pricing Model

Premium is \( \mathbb{E}[S|X = x] = \mathbb{E} \left[ \sum_{i=1}^{N} Y_i \bigg| X = x \right] = \mathbb{E}[N|X = x] \cdot \mathbb{E}[Y|X = x] \)

Statistical and modeling issues to approximate based on some training datasets, with claims frequency \( \{n_i, x_i\} \) and individual losses \( \{y_i, x_i\} \).

Use GLM to approximate \( \mathbb{E}[N|X = x] \) and \( \mathbb{E}[Y|X = x] \)

“Most firms . . . rely on traditional generalised linear models (GLMs) . . . .

A small number of firms use non-linear methods (e.g. decision trees) as input to GLMs” FCA (2016)

(see also Yiao (2013))
From Econometric to ‘Machine Learning’

In a competitive market, insurers can use different sets of variables and different models, e.g. GLMs, $N_t|\mathbf{X} \sim \mathcal{P}(\lambda_\mathbf{X} \cdot t)$ and $Y|\mathbf{X} \sim \mathcal{G}(\mu_\mathbf{X}, \varphi)$

$$\hat{\pi}_j(\mathbf{x}) = \hat{\mathbb{E}}[N_1|\mathbf{X} = \mathbf{x}] \cdot \hat{\mathbb{E}}[Y|\mathbf{X} = \mathbf{x}] = \exp(\hat{\alpha}^T \mathbf{x}) \cdot \exp(\hat{\beta}^T \mathbf{x})$$

that can be extended to GAMs,

$$\hat{\pi}_j(\mathbf{x}) = \exp \left( \sum_{k=1}^{d} \hat{s}_k(x_k) \right) \cdot \exp \left( \sum_{k=1}^{d} \hat{t}_k(x_k) \right)$$

or Tweedie models on $S_t$ (compound Poisson, see Tweedie (1984)) conditional on $\mathbf{X}$
From Econometric to ‘Machine Learning’

(see Charpentier & Denuit (2005) or Kaas et al. (2008)) or any other statistical model

\[ \hat{\pi}_j(x) \text{ where } \hat{\pi}_j \in \arg\min_{m \in F_j : X_j \rightarrow \mathbb{R}} \left\{ \sum_{i=1}^{n} \ell(s_i, m(x_i)) + \lambda \cdot \text{penalty}(m) \right\} \]

For some loss function \( \ell : \mathbb{R}^2 \rightarrow \mathbb{R}_+ \) (usually an \( L_2 \) based loss, \( \ell(s, y) = (s - y)^2 \) since \( \arg\min \{ \mathbb{E}[\ell(S, m)], \ m \in \mathbb{R} \} \) is \( \mathbb{E}(S) \), interpreted as the pure premium).

For instance, consider regression trees, forests, neural networks, or boosting based techniques to approximate \( \pi(x) \), and various techniques for variable selection, such as LASSO (see Hastie et al. (2009) or Charpentier et al. (2017) for a description and a discussion).
Competitive Insurance Markets

\[ \hat{\pi}_1(\cdot) \]
Competitive Insurance Markets

model 1 \( \hat{\pi}_1(\cdot) \)

model 2 \( \hat{\pi}_2(\cdot) \)
Competitive Insurance Markets

model 1 \( \hat{\pi}_1(\cdot) \)

model 2 \( \hat{\pi}_2(\cdot) \)
Before discussing the use of those models in insurance, note that the same issues exist in credit, see Hardt, Price & Srebro (2017).

“the shift from traditional to machine learning lending models may have important distributional effects for consumers […] machine learning would offer lower rates to racial groups who already were at an advantage under the traditional model, but it would also benefit disadvantaged groups by enabling them to obtain a mortgage in the first place” Fuster, Goldsmith-Pinkham, Ramadorai & Walther (2017)
Field experiment: actuarial pricing games

Actuarial pricing is data based, and model based.

To understand how model influence pricing we ran some actuarial pricing games.

With $d$ competitors, each insured $i$ has to choose among $d$ premiums,

$$\pi_i = (\hat{\pi}_1(x_i), \cdots, \hat{\pi}_d(x_i)) \in \mathbb{R}_+^d$$
Insurance Ratemaking Before Competition

(impact of various categorical variables)
Premiums Tail Correlations (strong)

Strong tail dependence between (say) \( \pi_1 \) and \( \pi_2 \)

\[
\lambda(u) = \begin{cases} 
\mathbb{P}[X_1 \leq F_1^{-1}(u) \mid X_2 \leq F_2^{-1}(u)] & \text{if } u \in (0, 1/2) \\
\mathbb{P}[X_1 > F_1^{-1}(u) \mid X_2 > F_2^{-1}(u)] & \text{if } u \in (1/2, 1),
\end{cases}
\]

e.g.

\[
\lambda(u) = \frac{\mathbb{P}[X_1 \leq F_1^{-1}(u), X_2 \leq F_2^{-1}(u)]}{\mathbb{P}[X_2 \leq F_2^{-1}(u)]} \quad \text{if } u \in (0, 1/2)
\]

estimated by (with \( U_{1,i} = \hat{F}_1(\pi_1(x_i)) \) and \( U_{2,i} = \hat{F}_1(\pi_2(x_i)) \))

\[
\hat{\lambda}(u) = \begin{cases} 
\frac{1}{nu} \sum_{i=1}^{n} 1[U_{1,i} \leq u, U_{2,i} \leq u] & \text{if } u \in (0, 1/2) \\
\frac{1}{n(1-u)} \sum_{i=1}^{n} 1[U_{1,i} > u, U_{2,i} > u] & \text{if } u \in (1/2, 1),
\end{cases}
\]

from Joe (1990), see also Charpentier (2012).
Premiums Tail Correlations (weak)

Weak tail dependence between (say) \( \pi_1 \) and \( \pi_2 \)

\[
\chi(u) = \begin{cases} 
\frac{\log \Pr[X_2 \leq F_2^{-1}(u)]}{\log \Pr[X_1 \leq F_1^{-1}(u), X_2 \leq F_2^{-1}(u)]} & \text{if } u \in (0, 1/2) \\
\frac{\log \Pr[X_2 > qF_2^{-1}(u)]}{\log \Pr[X_1 > F_1^{-1}(u), X_2 > F_2^{-1}(u)]} & \text{if } u \in (1/2, 1), 
\end{cases}
\]

estimated by

\[
\hat{\chi}(u)^{-1} = \begin{cases} 
\frac{1}{u} \log \left( \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}[U_{1,i} \leq u, U_{2,i} \leq u] \right) & \text{if } u \in (0, 1/2) \\
\frac{1}{1-u} \log \left( \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}[U_{1,i} > u, U_{2,i} > u] \right) & \text{if } u \in (1/2, 1), 
\end{cases}
\]

from Ledford & Tawn (1996), see also Charpentier (2012).
Premiums Correlations
Premiums Correlations

Pearson correlation = 0.829
Spearman correlation = 0.868
Pearson correlation = 0.821
Spearman correlation = 0.875
Premiums Correlations
We need a **Decision Rule** to select premium chosen by insured $i$

<table>
<thead>
<tr>
<th></th>
<th>Ins1</th>
<th>Ins2</th>
<th>Ins3</th>
<th>Ins4</th>
<th>Ins5</th>
<th>Ins6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>787.93</td>
<td>706.97</td>
<td>1032.62</td>
<td>907.64</td>
<td>822.58</td>
<td>603.83</td>
</tr>
<tr>
<td></td>
<td>170.04</td>
<td>197.81</td>
<td>285.99</td>
<td>212.71</td>
<td>177.87</td>
<td>265.13</td>
</tr>
<tr>
<td></td>
<td>473.15</td>
<td>447.58</td>
<td>343.64</td>
<td>410.76</td>
<td>414.23</td>
<td>425.23</td>
</tr>
<tr>
<td></td>
<td>337.98</td>
<td>336.20</td>
<td>468.45</td>
<td>339.33</td>
<td>383.55</td>
<td>672.91</td>
</tr>
</tbody>
</table>
Insurance Ratemaking Competition

Basic ‘rational rule’ \( \pi_i = \min\{\hat{\pi}_1(x_i), \ldots, \hat{\pi}_d(x_i)\} = \hat{\pi}_{1:d}(x_i) \)

<table>
<thead>
<tr>
<th>Ins1</th>
<th>Ins2</th>
<th>Ins3</th>
<th>Ins4</th>
<th>Ins5</th>
<th>Ins6</th>
</tr>
</thead>
<tbody>
<tr>
<td>787.93</td>
<td>706.97</td>
<td>1032.62</td>
<td>907.64</td>
<td>822.58</td>
<td>603.83</td>
</tr>
<tr>
<td>170.04</td>
<td>197.81</td>
<td>285.99</td>
<td>212.71</td>
<td>177.87</td>
<td>265.13</td>
</tr>
<tr>
<td>473.15</td>
<td>447.58</td>
<td>343.64</td>
<td>410.76</td>
<td>414.23</td>
<td>425.23</td>
</tr>
<tr>
<td>337.98</td>
<td>336.20</td>
<td>468.45</td>
<td>339.33</td>
<td>383.55</td>
<td>672.91</td>
</tr>
</tbody>
</table>
## Insurance Ratemaking Competition

A more **realistic rule** $\pi_i \in \{\hat{\pi}_1(d(x_i)), \hat{\pi}_2(d(x_i)), \hat{\pi}_3(d(x_i))\}$

<table>
<thead>
<tr>
<th></th>
<th>Ins1</th>
<th>Ins2</th>
<th>Ins3</th>
<th>Ins4</th>
<th>Ins5</th>
<th>Ins6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ins1</td>
<td>787.93</td>
<td>706.97</td>
<td>1032.62</td>
<td>907.64</td>
<td>822.58</td>
<td>603.83</td>
</tr>
<tr>
<td>Ins2</td>
<td>170.04</td>
<td>197.81</td>
<td>285.99</td>
<td>212.71</td>
<td>177.87</td>
<td>265.13</td>
</tr>
<tr>
<td>Ins3</td>
<td>473.15</td>
<td>447.58</td>
<td>343.64</td>
<td>410.76</td>
<td>414.23</td>
<td>425.23</td>
</tr>
<tr>
<td>Ins4</td>
<td>337.98</td>
<td>336.20</td>
<td>468.45</td>
<td>339.33</td>
<td>383.55</td>
<td>672.91</td>
</tr>
</tbody>
</table>
Insurer 4
GLM for frequency and standard cost (large claims were removed, above 15k), Interaction Age and Gender
Actuary working for a *mutuelle* company

Insurer 11
Use of two XGBoost models (bodily injury and material), with correction for negative premiums
Actuary working for a private insurance company
Insurer 6 (market 3)
Team of two actuaries (degrees in Engineering and Physics), in Vancouver, Canada. Used GLMs (Tweedie), no territorial classification, no use of information about other competitors, 

“Segments with high market share and low loss ratios were also given some premium increase”
Insurer 7 (market 1)
Actuary in France, used random forest for variable selection, and GLMs
Insurer 15 (market 2)

Actuary, working as a consultant, Margin Method with iterations, MS Access & MS Excel
Actuarial Pricing Game, 2017

Insurer 21 (market 1)

Actuary, working as a consultant, used GLMs, with variable selection using LARS and LASSO

Iterative learning algorithm (codes available on github)
Insurer 4 (market 2)
Actuary, working as a consultant, used XGBOOST, used GLMs for year 3.
Insurer 8 (market 3)
Mathematician, working on Solvency II software in Austria
Generalized Additive Models with spatial variable
Segmentation + Overall Level

\[
\hat{\pi}(x) = \exp(\hat{\alpha}^T x) \cdot \exp(\hat{\beta}^T x) = e^\gamma_0 \cdot e^{\gamma_1 x_1} e^{\gamma_2 x_2} \cdots e^{\gamma_k x_k}
\]

Poisson \( P(\lambda_x) \) \hspace{1cm} Gamma \( G(\mu_x, \varphi) \)

Why not try to change \( e^{\gamma_0} \)?
What could we do when we observe competitors’ prices?
What could we do when we observe competitors’ prices?
In some cases, it seems possible to increase the market share and decrease the loss ratio...
Ajust to Competitors’ Prices

- If \( \pi_* (x_i) = \pi_1: d(x_i) \) then \( \pi_* (x_i) = \pi_2: d(x_i) - \epsilon \)
- If \( \pi_* (x_i) = \pi_2: d(x_i) \) and if \( \pi_* (x_i) - \pi_1: d(x_i) \leq \alpha \) then \( \pi_* (x_i) = \pi_1: d(x_i) - \epsilon \)
Using simulations, we created a few models, some *linear* and some *non-linear*. We created markets of 10 companies.

The correlation between the total earned premium (of the market) and the proportion of *non-linear* models is strongly negative.

The more nonlinear models are used, the lower the total earned premium (the worst the market loss ratio...)

ongoing research with Ali Farzanehfar, Florimond Houssiau and Yves-Alexandre de-Montjoye (Imperial College)
On-going research (see https://pricing-game.com/)

- hard to derive theoretical properties of competition market
- use field studies (but hard to get players...)
- use simulated models and markets