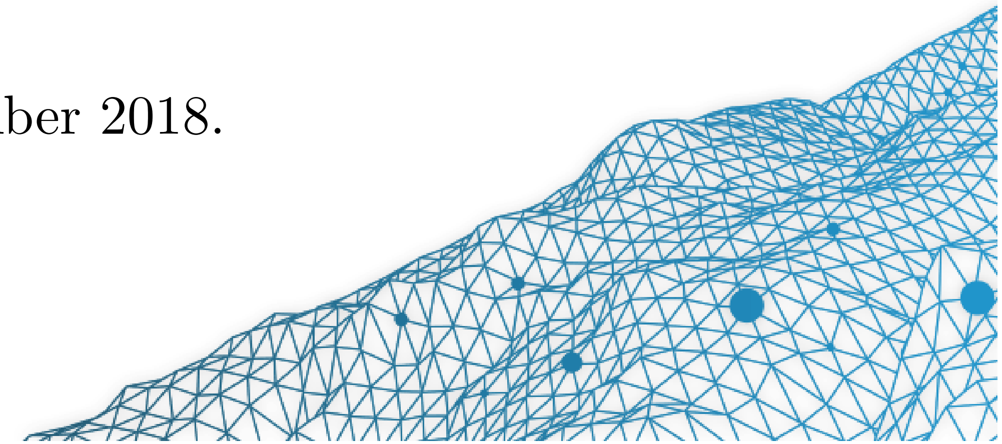
Segmentation et mutualisation, les deux faces d'une même pièce ?

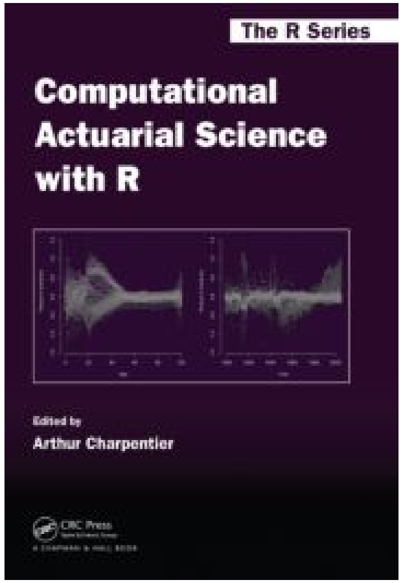
A. Charpentier (Université du Québec à Montréal)

Rencontres Mutualistes, Beaune, November 2018.



A. Charpentier (Université du Québec à Montréal)

Professor Mathematics Department, Université du Québec à Montréal
previously Econ. Dept, Université de Rennes & ENSAE Paristech
actuary in Hong Kong, IT & Stats FFA
director Data Science for Actuaries Program, Institute of Actuaries
PhD in Statistics (KU Leuven), Fellow of the Institute of Actuaries
MSc in Financial Mathematics (Paris Dauphine) & ENSAE
Research Chair :
ACTINFO (valorisation et nouveaux usages actuariels de l'information)
Editor of the freakonometrics.hypotheses.org's blog
Editor of Computational Actuarial Science, CRC
Author of Mathématiques de l'Assurance Non-Vie (2 vol.), Economica



Insurance, "segmentation" & "mutualisation"

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

- what is actuarial pricing ?
- why a "spirale de la segmentation" ?
- how to compare actuarial models ?
- field experiment : pricing games

SEGMENTATION ET MUTUALISATION
LES DEUX FACES D'UNE MÊME PIÈCE ?

Arthur Charpentier
Professeur à l'Université du Québec, Montréal

Michel Denuit
Professeur à l'Université catholique de Louvain

Romuald Elie
Professeur à l'Université de Marne-la-Vallée

L'assurance repose fondamentalement sur l'idée que la mutualisation des risques entre des assurés est possible. Cette mutualisation, qui peut être vue comme une relecture actuarielle de la loi des grands nombres, n'a de sens qu'au sein d'une population de risques « homogènes » [Charpentier, 2011]. Cette condition (actuarielle) impose aux assureurs de segmenter, ce que confirment plusieurs travaux économiques (1). Avec l'explosion du nombre de données, et donc de variables tarifaires possibles, certains assureurs évoquent l'idée d'un tarif individuel, semblant remettre en cause l'idée même de mutualisation des risques. Entre cette force qui pousse à segmenter et la force de rappel qui tend (pour des raisons sociales mais aussi actuarielles, ou au moins de robustesse statistique (2)) à imposer une solidarité minimale entre les assurés, quel équilibre va en résulter dans un contexte de forte concurrence entre les sociétés d'assurance ?

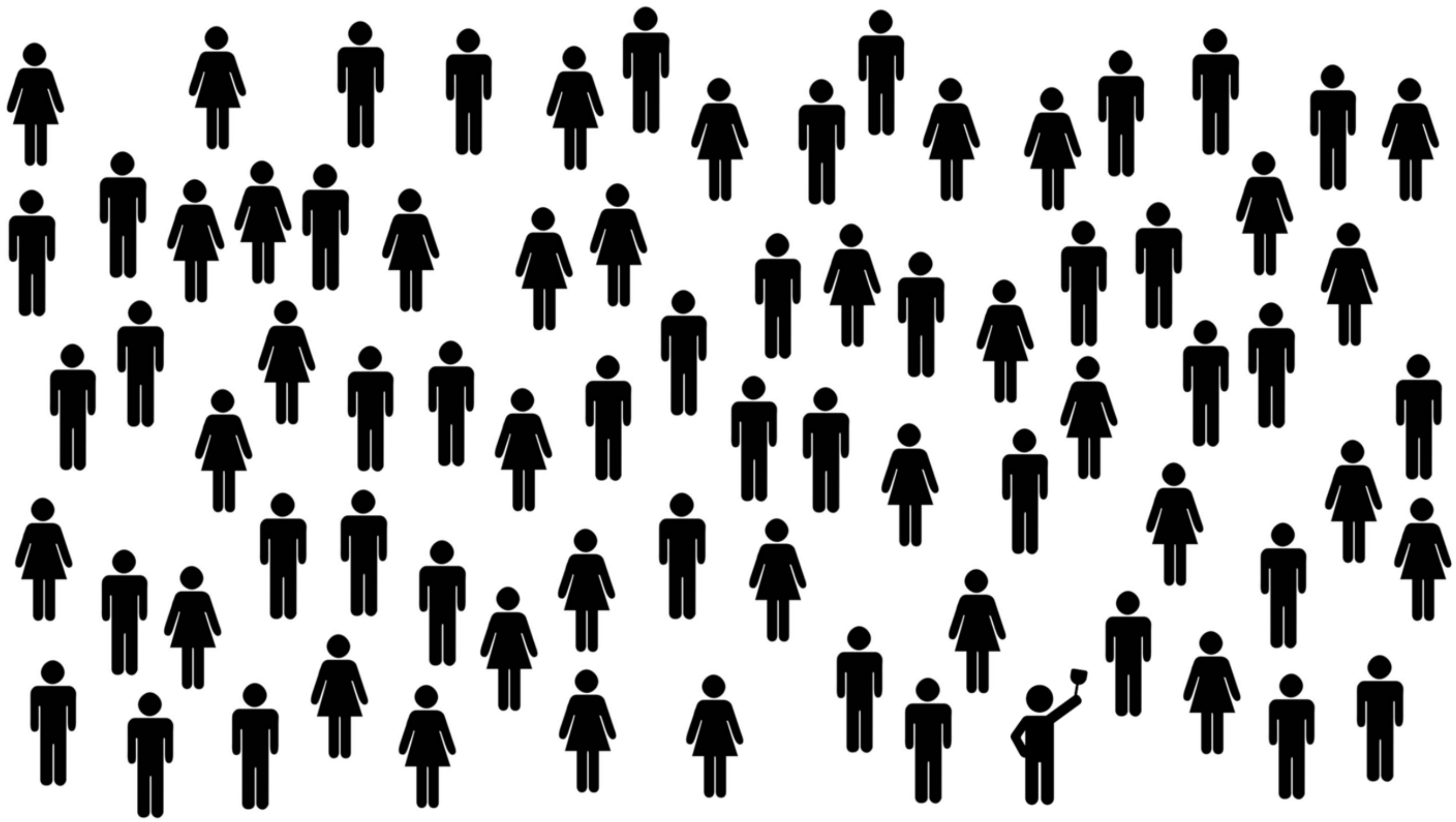
Tarification sans segmentation

Sans segmentation, le « prix juste » d'un risque est l'espérance mathématique de la charge annuelle. C'est l'idée du théorème fondamental de la valorisation actuarielle : en moyenne, la somme des primes doit permettre d'indemniser l'intégralité des sinistres survenus dans l'année. Afin d'illustrer les différents aspects de la construction du tarif et ses conséquences, on va utiliser les données présentées dans le tableau 1 (voir p. xx), qui indique la fréquence annuelle de sinistres.

Les facteurs de risque sont ici le lieu d'habitation et l'âge de l'assuré, et on observe la fréquence de sinistre par classe. Le coût unitaire, supposé fixe, équivaut à 1 000 euros. La prime pure est alors E [S] = 1 000 x E [N]. Dans cet exemple, la prime pure sans segmentation sera de 82,30 euros.

Risques n° 103 19

Actuarial Pricing in a Nutshell

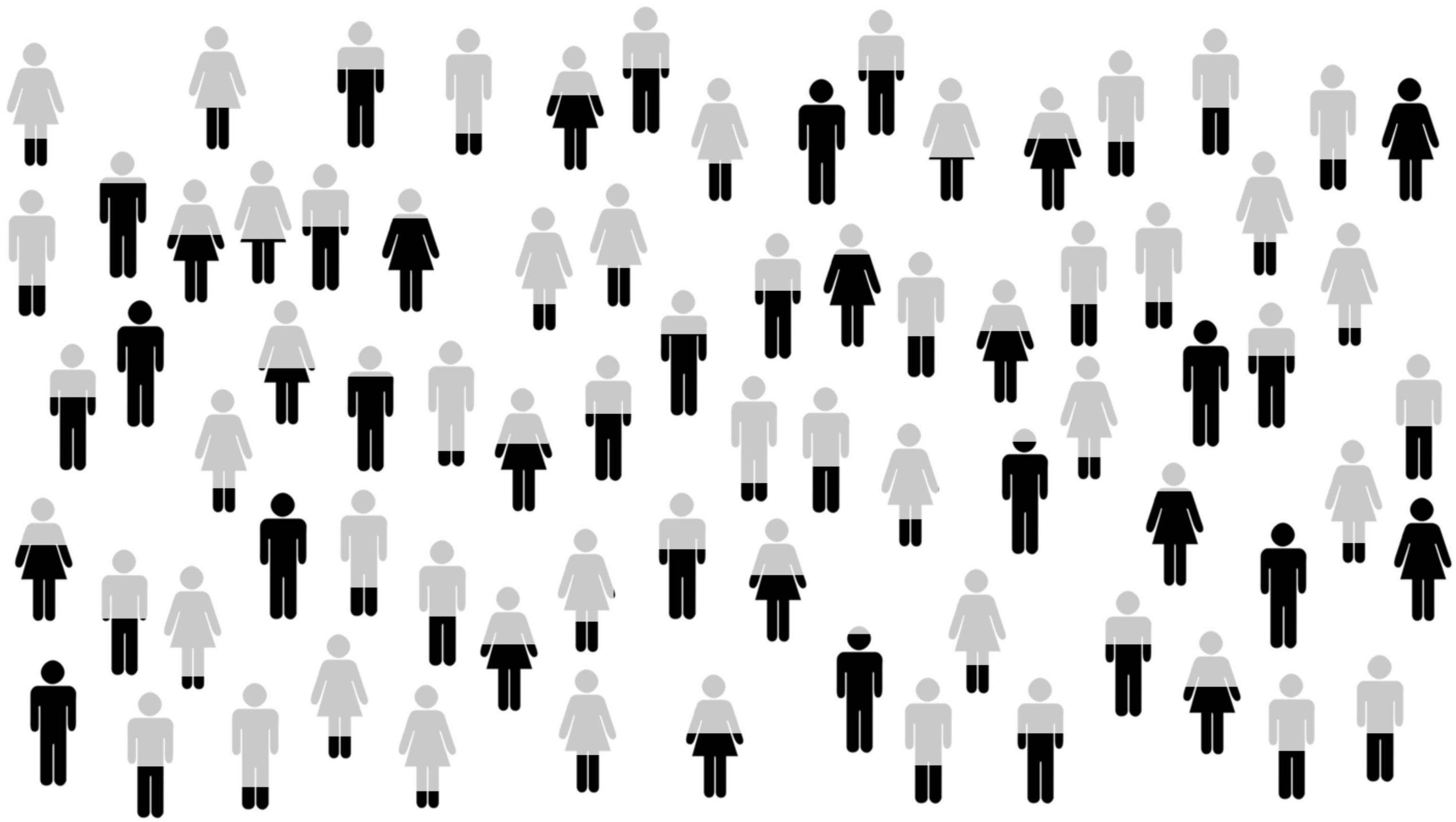


Consider some portfolio with n insured

Actuarial Pricing in a Nutshell

Consider some portfolio with n insured

Actuarial Pricing in a Nutshell

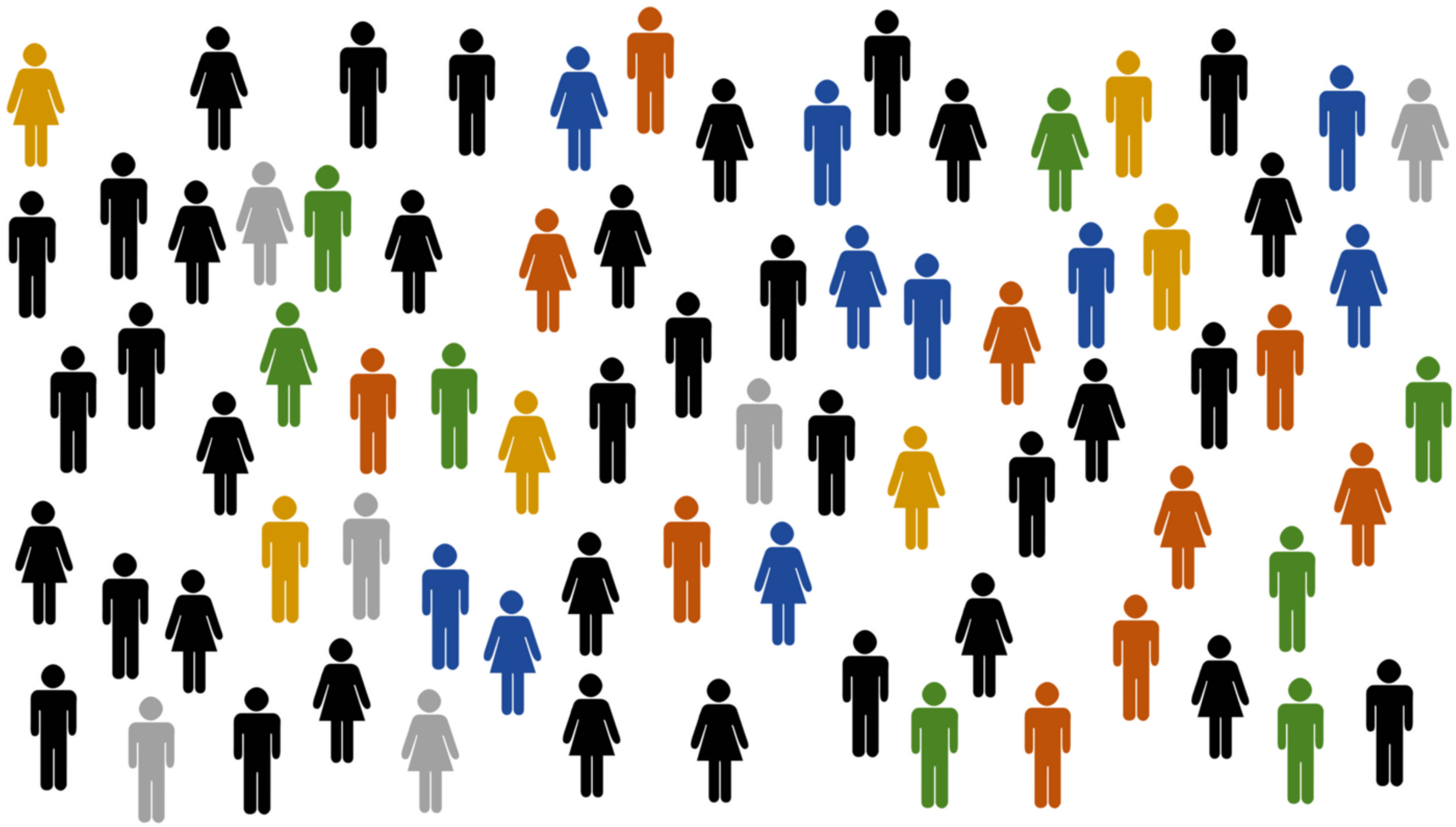


The n insured have heterogeneous **unobservable** risk factors $\omega_1, \cdots, \omega_n$

Actuarial Pricing in a Nutshell

Nevertheless, they have characteristics $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ (that can approximate ω_i's)

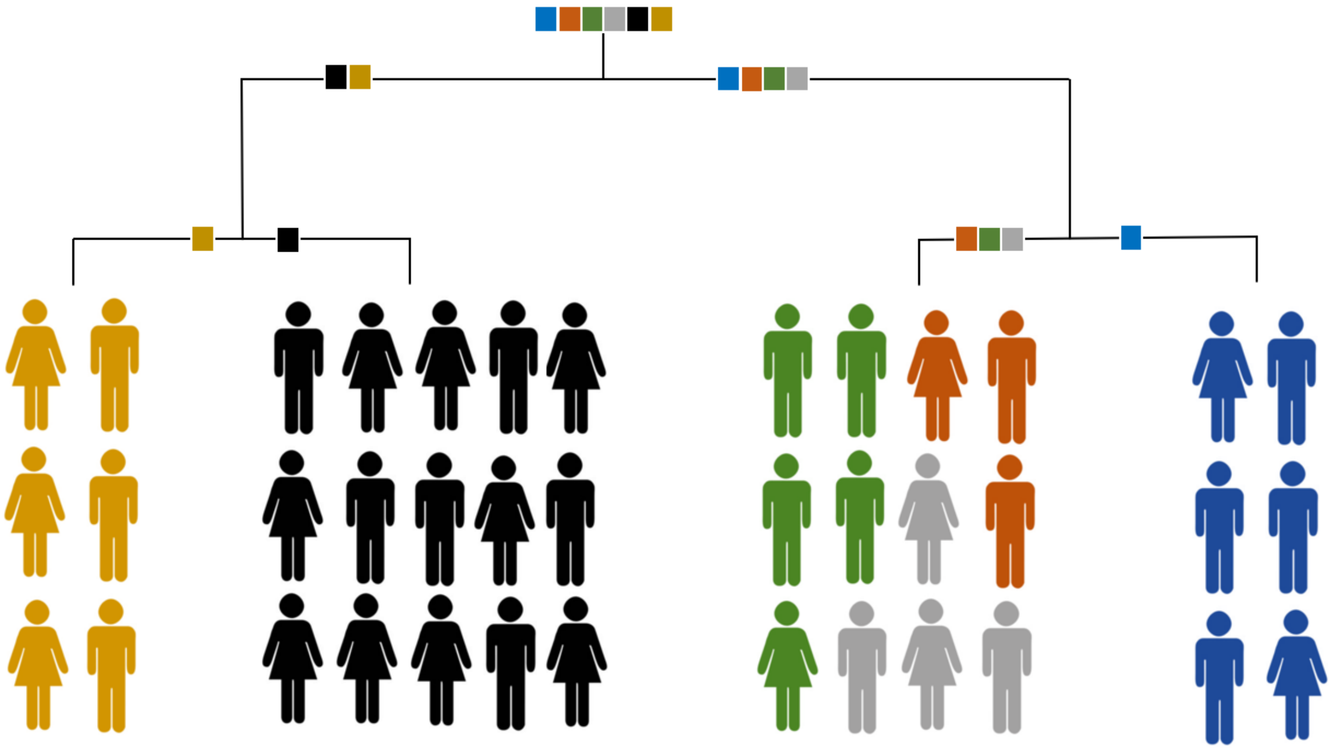
Actuarial Pricing in a Nutshell



e.g. x_1 can be the color of the car, x_2 can the gender, etc.

Those characteristics are **observable**, a priori.

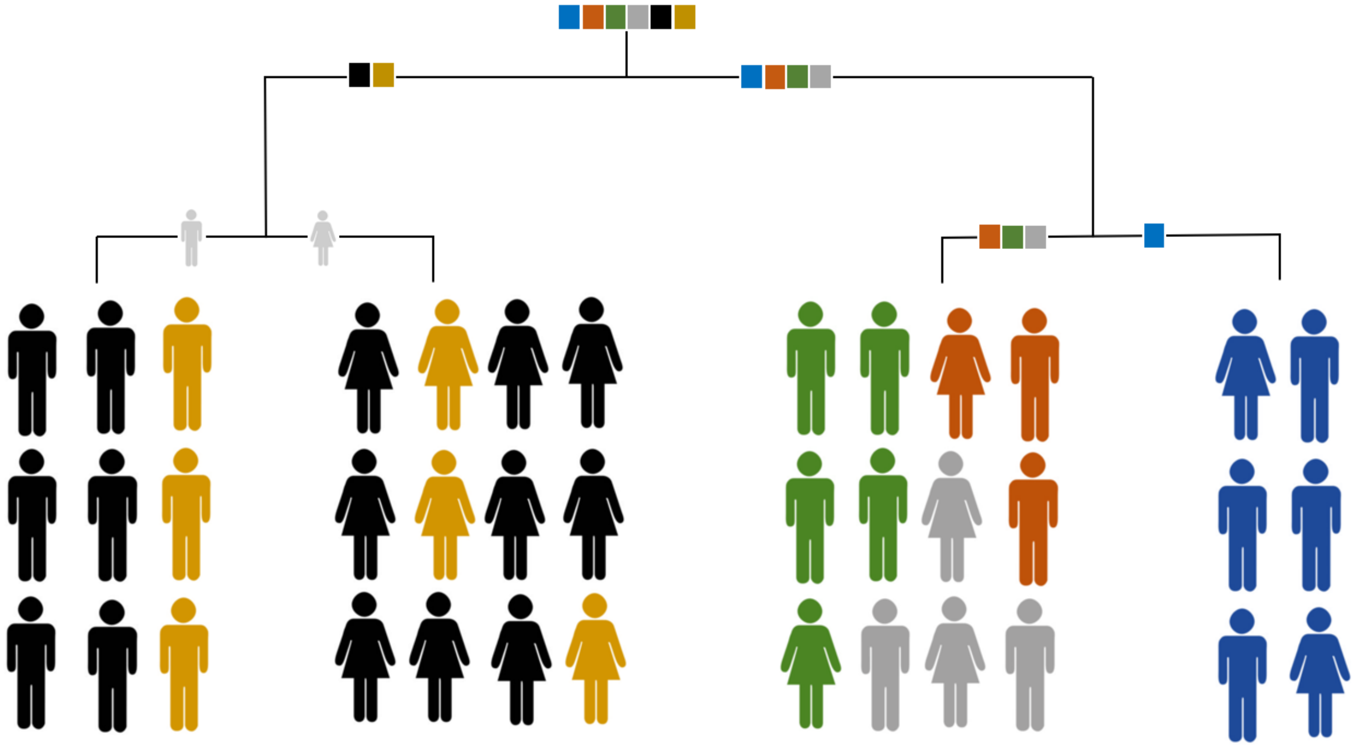
Actuarial Pricing in a Nutshell



Actuaries can use a regression tree to create homogeneous rate classes (or GLMs)

Here classes are based on x_1 (only)

Actuarial Pricing in a Nutshell



Actuaries can use a regression tree to create homogeneous rate classes (or GLMs)

But classes can be based on x_1 and x_2 – see Bailey (1963)

Risk Transfert without Segmentation

	Insured	Insurer
Loss	$\mathbb{E}[S]$	$S - \mathbb{E}[S]$
Average Loss	$\mathbb{E}[S]$	0
Variance	0	$\text{Var}[S]$

All the risk - $\text{Var}[S]$ - is kept by the insurance company.

Remark: all those interpretation are discussed in Denuit & Charpentier (2004).

Risk Transfert with Segmentation and Perfect Information

Assume that information $\boldsymbol{\Omega}$ is observable,

	Insured	Insurer
Loss	$\mathbb{E}[S\vert\boldsymbol{\Omega}]$	$S-\mathbb{E}[S\vert\boldsymbol{\Omega}]$
Average Loss	$\mathbb{E}[S]$	0
Variance	$\text{Var}\big[\mathbb{E}[S\vert\boldsymbol{\Omega}]\big]$	$\text{Var}\big[S-\mathbb{E}[S\vert\boldsymbol{\Omega}]\big]$

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

$$\text{Var}[S]=\underbrace{\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big]}_{\rightarrow\text{ insurer}}+\underbrace{\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big]}_{\rightarrow\text{ insured}}.$$

Risk Transfert with Segmentation and Imperfect Information

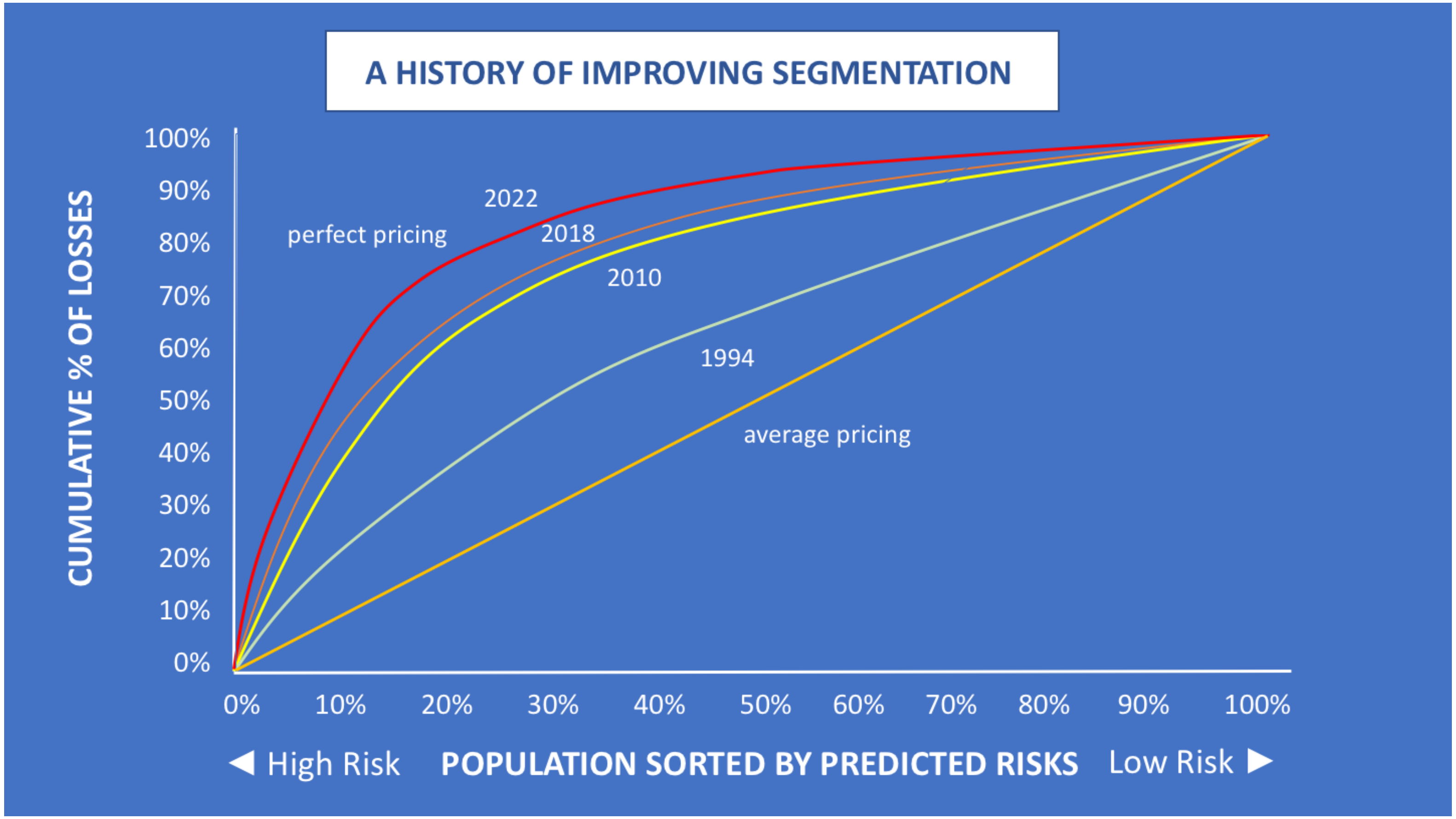
Assume that $\boldsymbol{X} \subset \boldsymbol{\Omega}$ is observable

	Insured	Insurer
Loss	$\mathbb{E}[S\vert\boldsymbol{X}]$	$S - \mathbb{E}[S\vert\boldsymbol{X}]$
Average Loss	$\mathbb{E}[S]$	0
Variance	$\text{Var}\big[\mathbb{E}[S\vert\boldsymbol{X}]\big]$	$\mathbb{E}\big[\text{Var}[S\vert\boldsymbol{X}]\big]$

Now

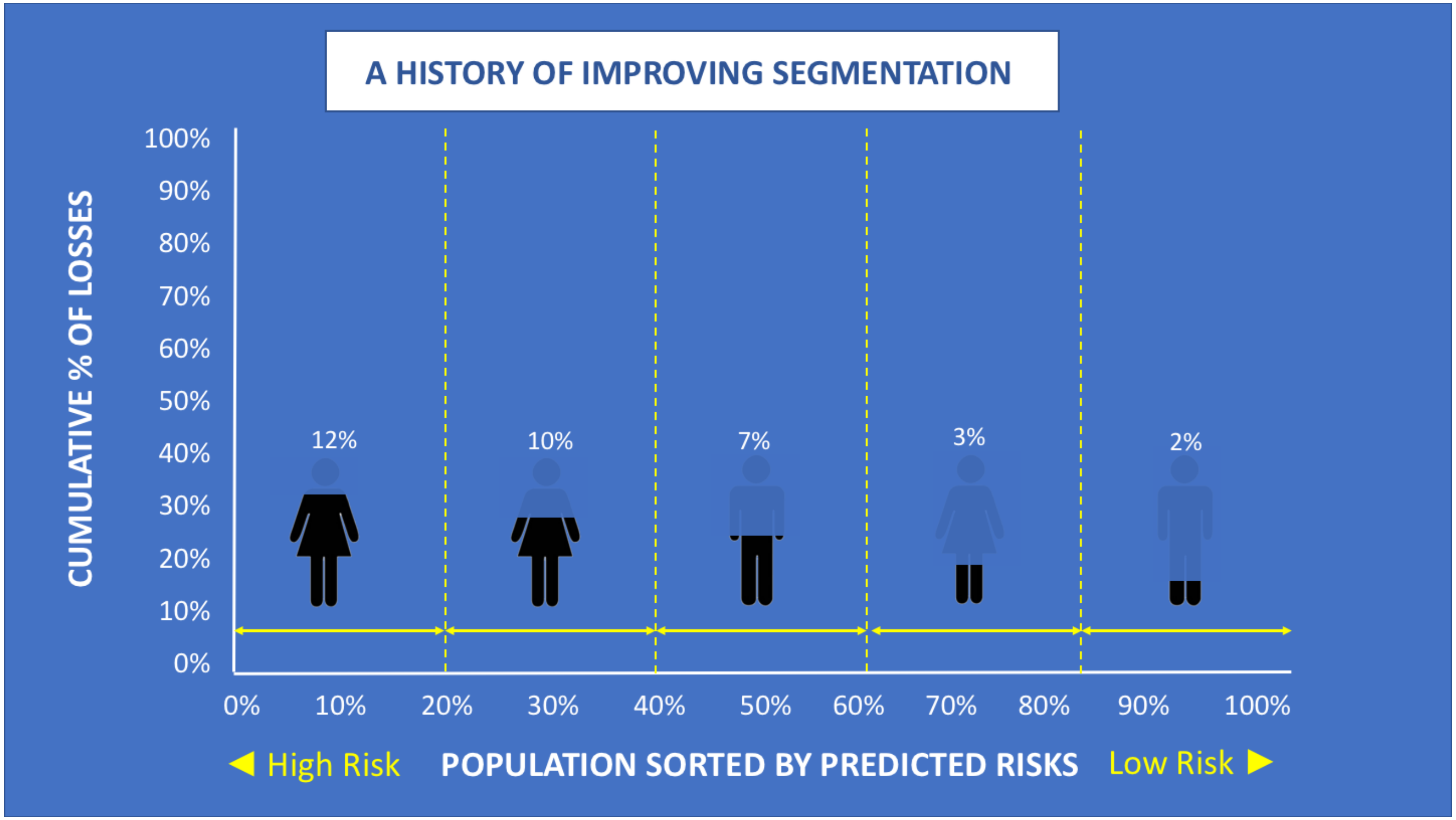
$$
\begin{aligned}
\mathbb{E}\Big[\text{Var}[S|\boldsymbol{X}]\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big] + \mathbb{E}\Big[\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big] \\
&= \underbrace{\mathbb{E}\Big[\text{Var}[S|\boldsymbol{\Omega}]\Big]}_{\text{perfect pricing}} + \underbrace{\mathbb{E}\Big\{\text{Var}\Big[\mathbb{E}[S|\boldsymbol{\Omega}]\Big|\boldsymbol{X}\Big]\Big\}}_{\text{misfit}}.
\end{aligned}
$$

How can we visualize the goodness of a model ?



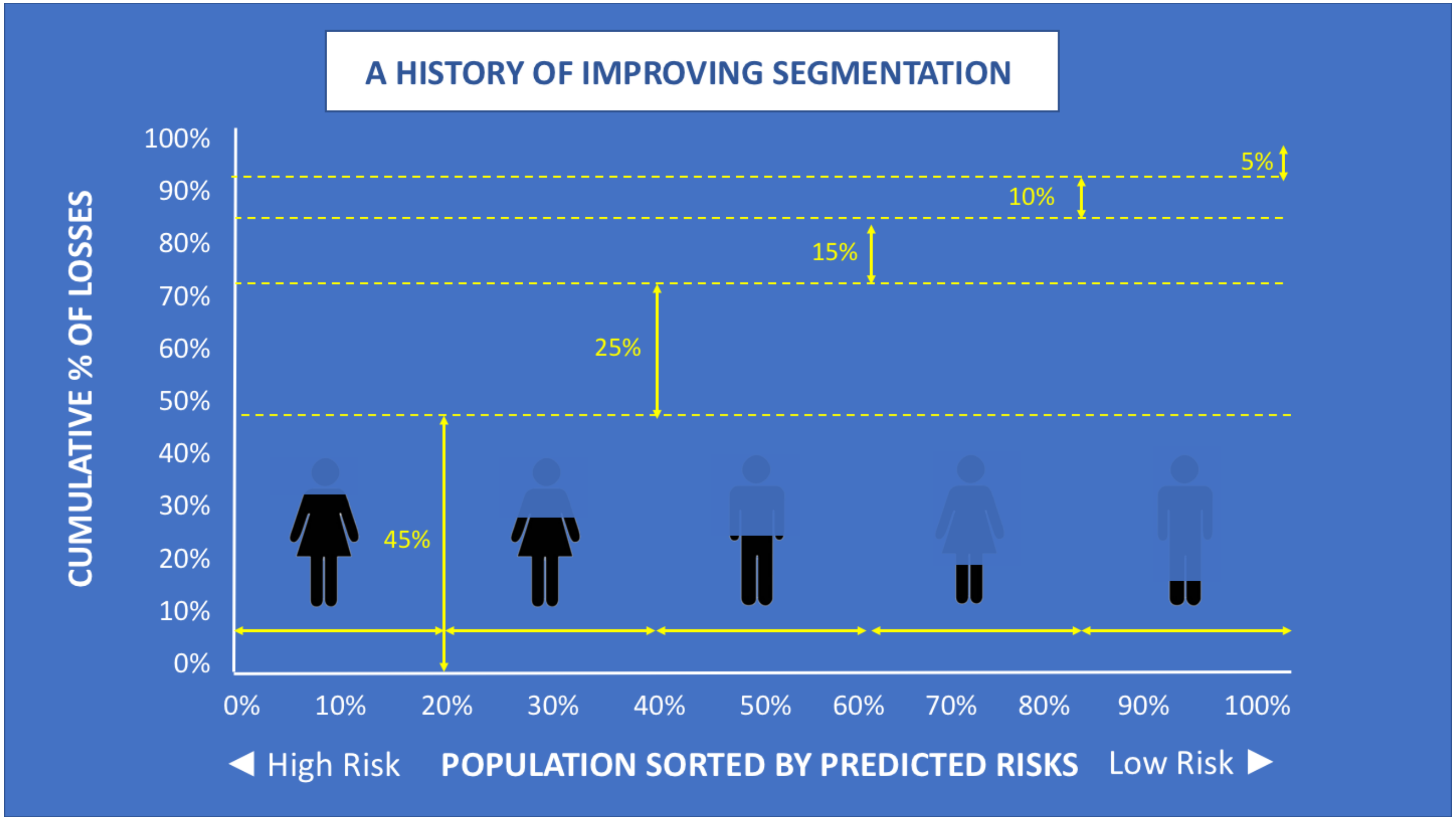
Source : https://www.progressive.com/jobs/analyst-program/

Constructing the (*pseudo*)-Lorenz curve



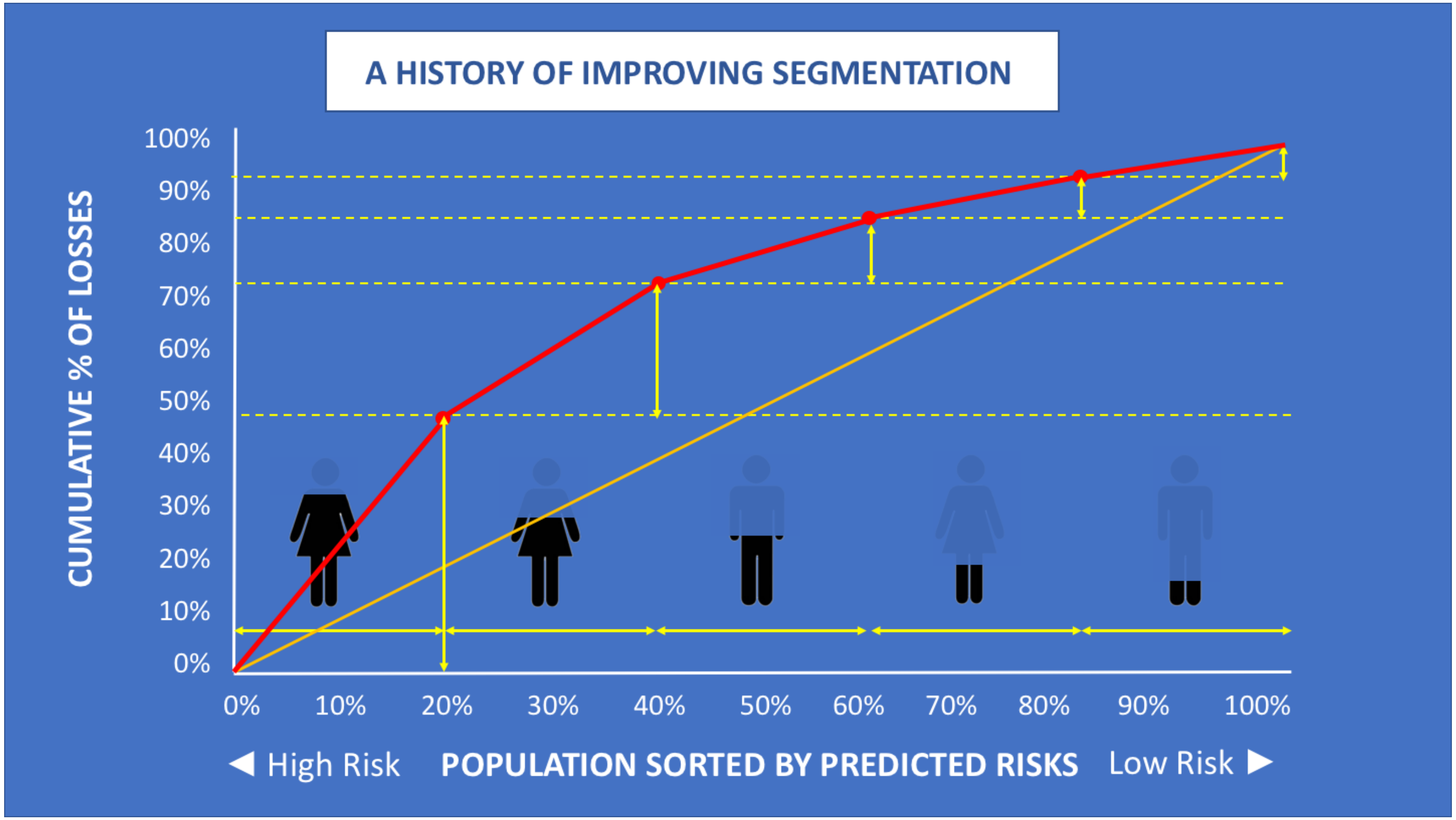
Sort the n risks according to the model $m(\boldsymbol{x}_1) \geq m(\boldsymbol{x}_2) \geq \cdots m(\boldsymbol{x}_n)$

Constructing the (*pseudo*)-Lorenz curve



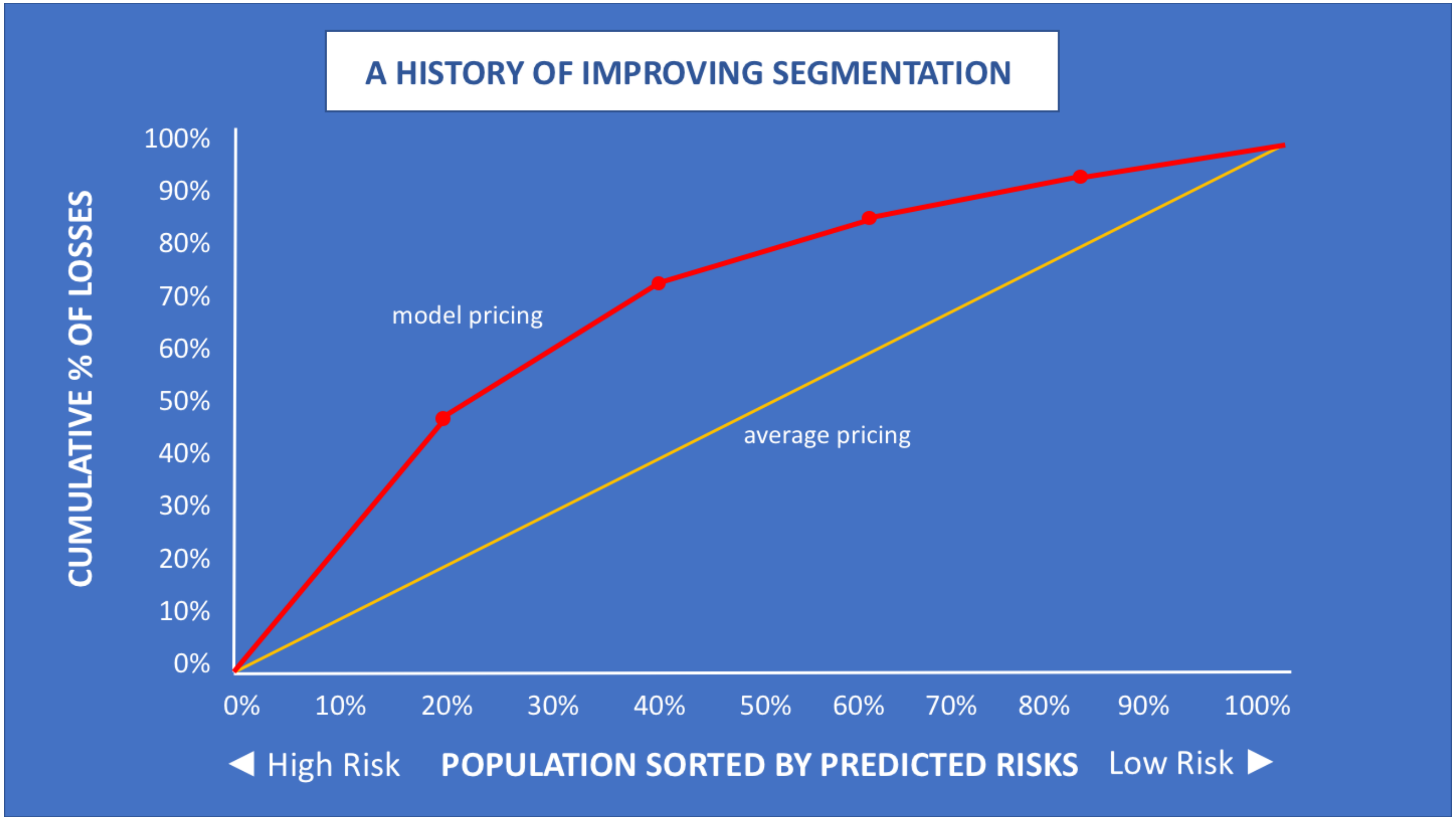
On the x-axis, $x_i = i/n$, on the y-axis, $y_i = \sum_{j=1}^{i} y_j / \sum_{j=1}^{n} y_j$

Constructing the (*pseudo*)-Lorenz curve



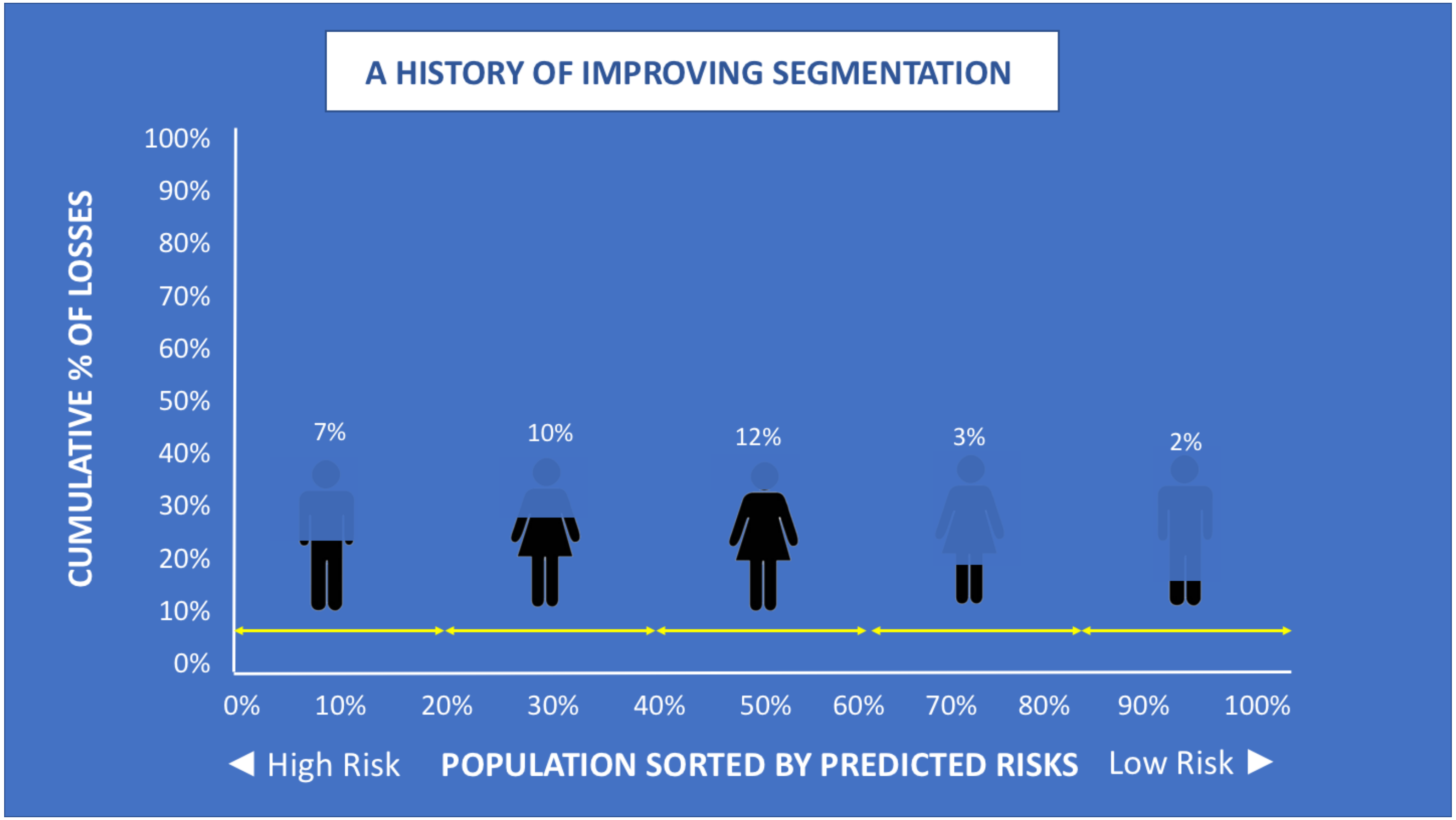
Connect points (x_i, y_i)

Constructing the (*pseudo*)-Lorenz curve



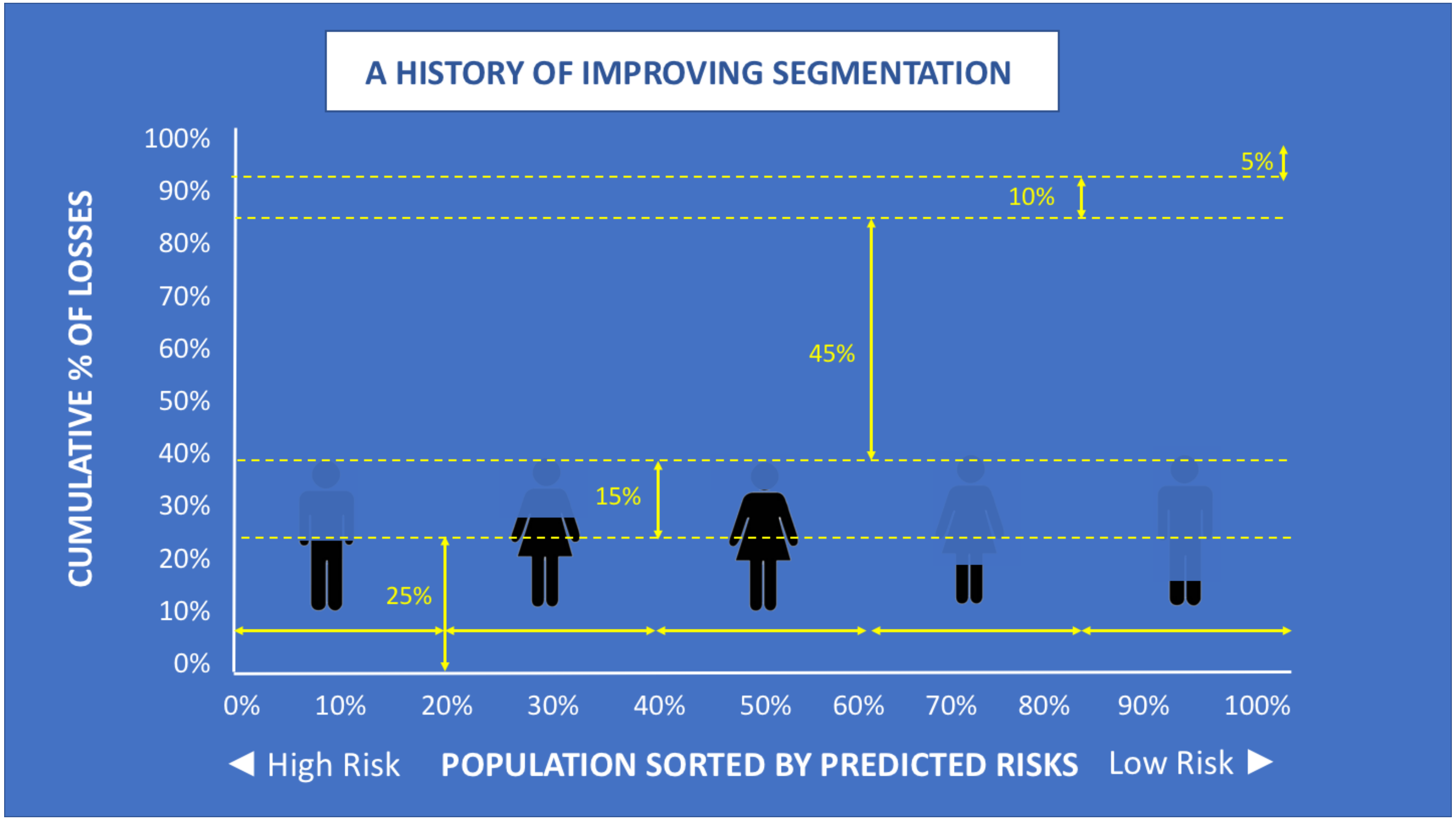
see Frees, Meyers & Cummins (2014).

Practice of (*pseudo*)-Lorenz curves



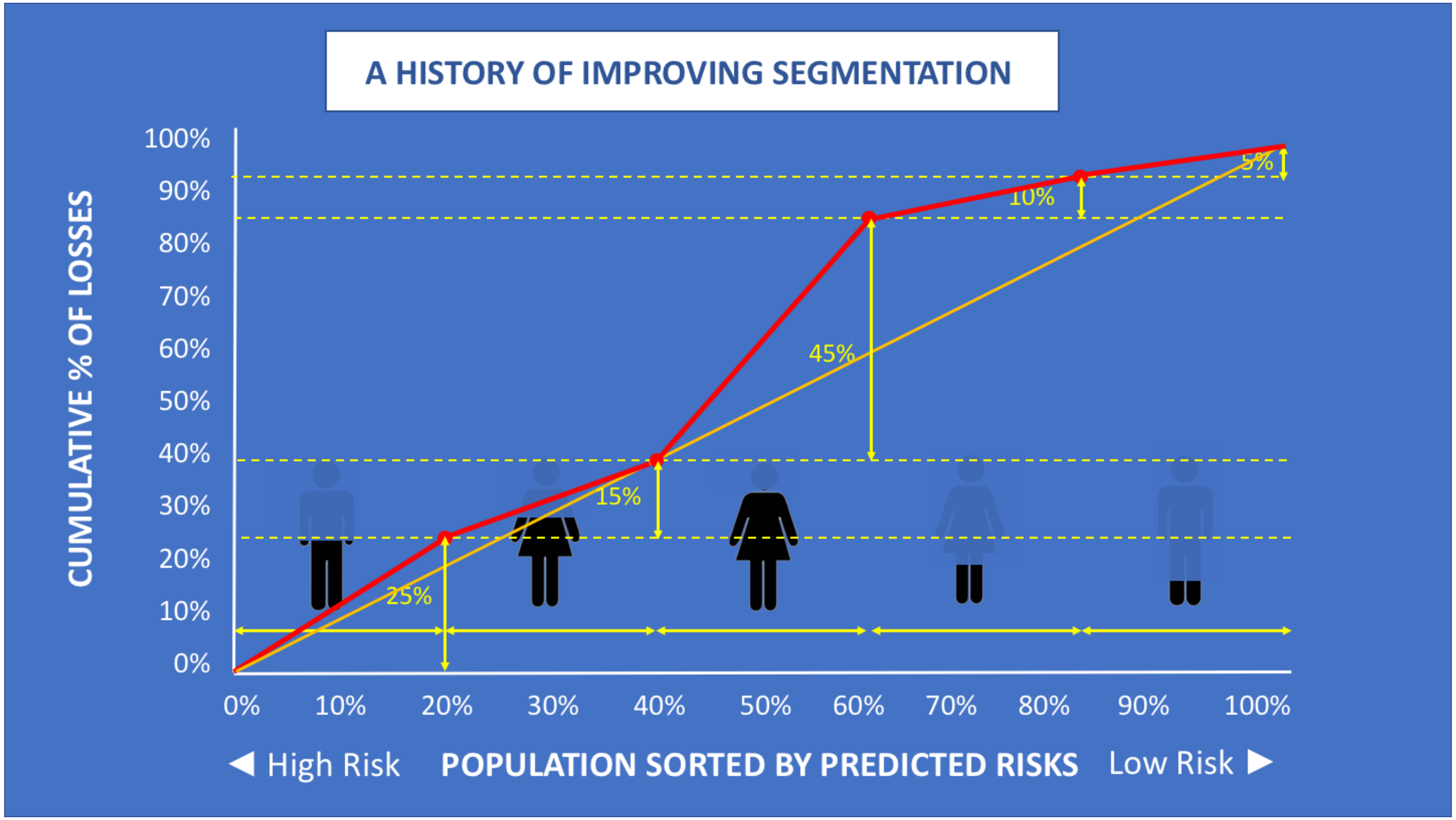
What if $\widehat{m}$ and m are not perfectly correctly correlated...?

Practice of (*pseudo*)-Lorenz curves



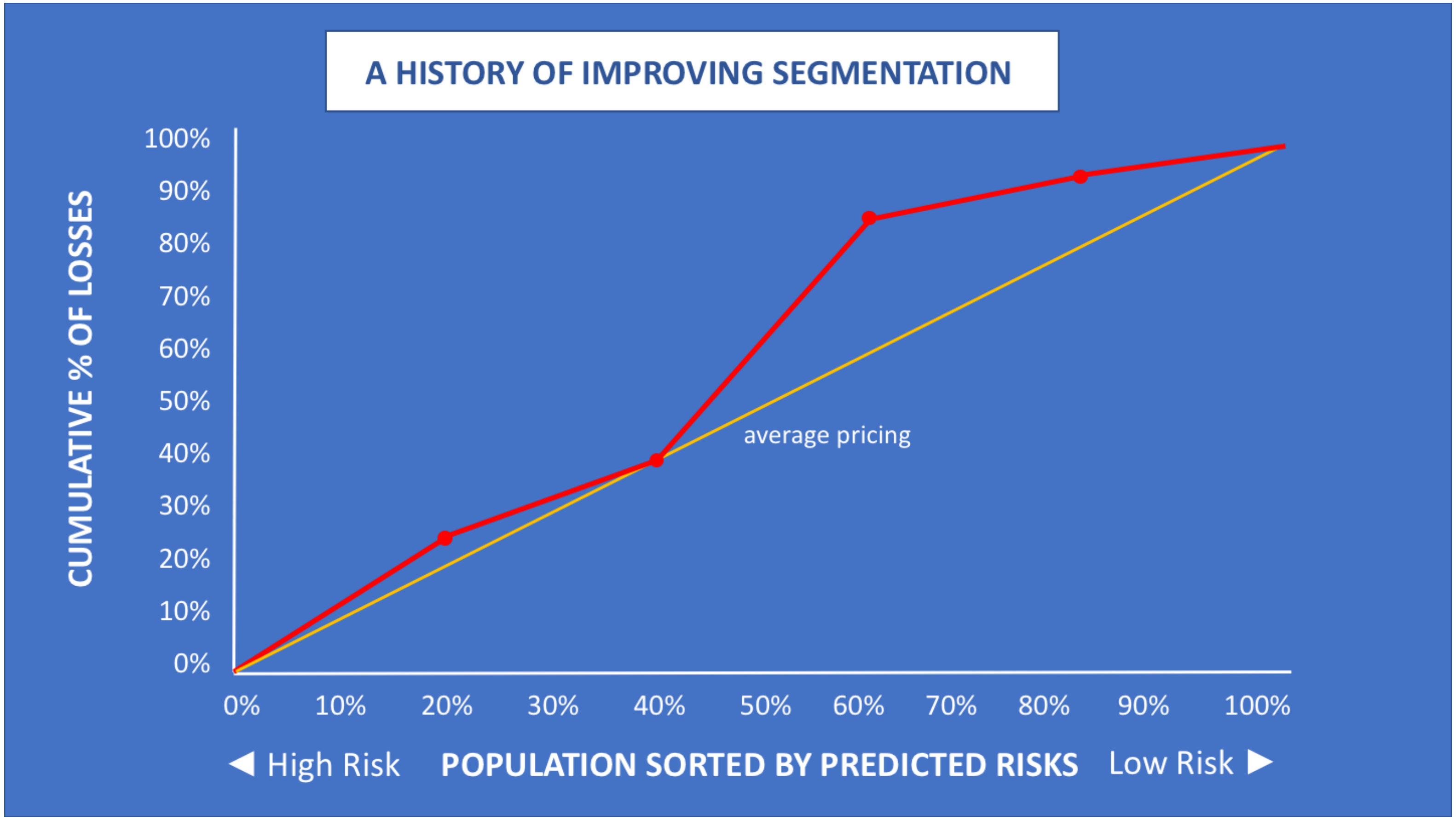
What if $\widehat{m}$ and m are not perfectly correctly correlated...?

Practice of (*pseudo*)-Lorenz curves



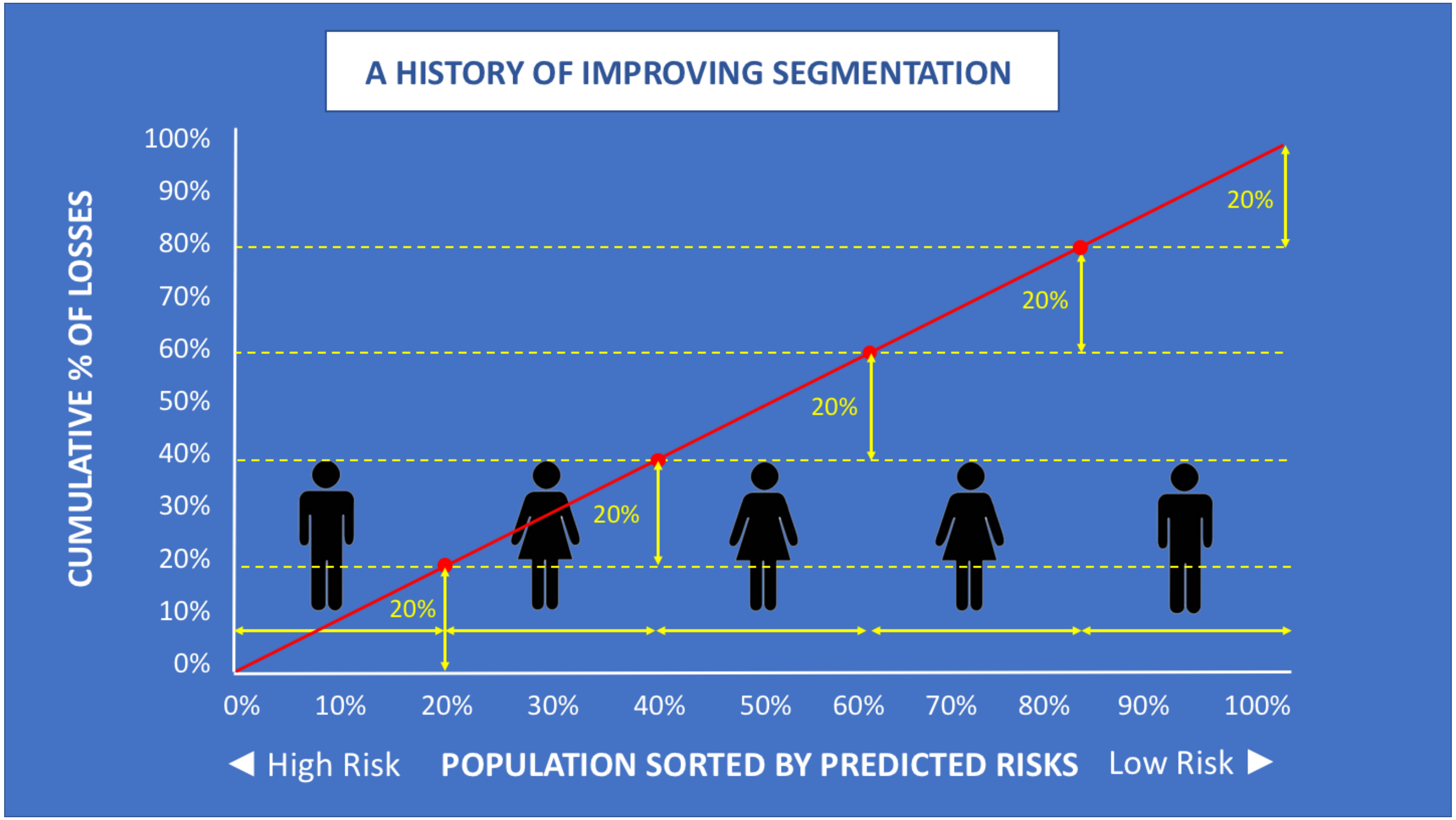
What if $\widehat{m}$ and m are not perfectly correctly correlated...?

Practice of (*pseudo*)-Lorenz curves



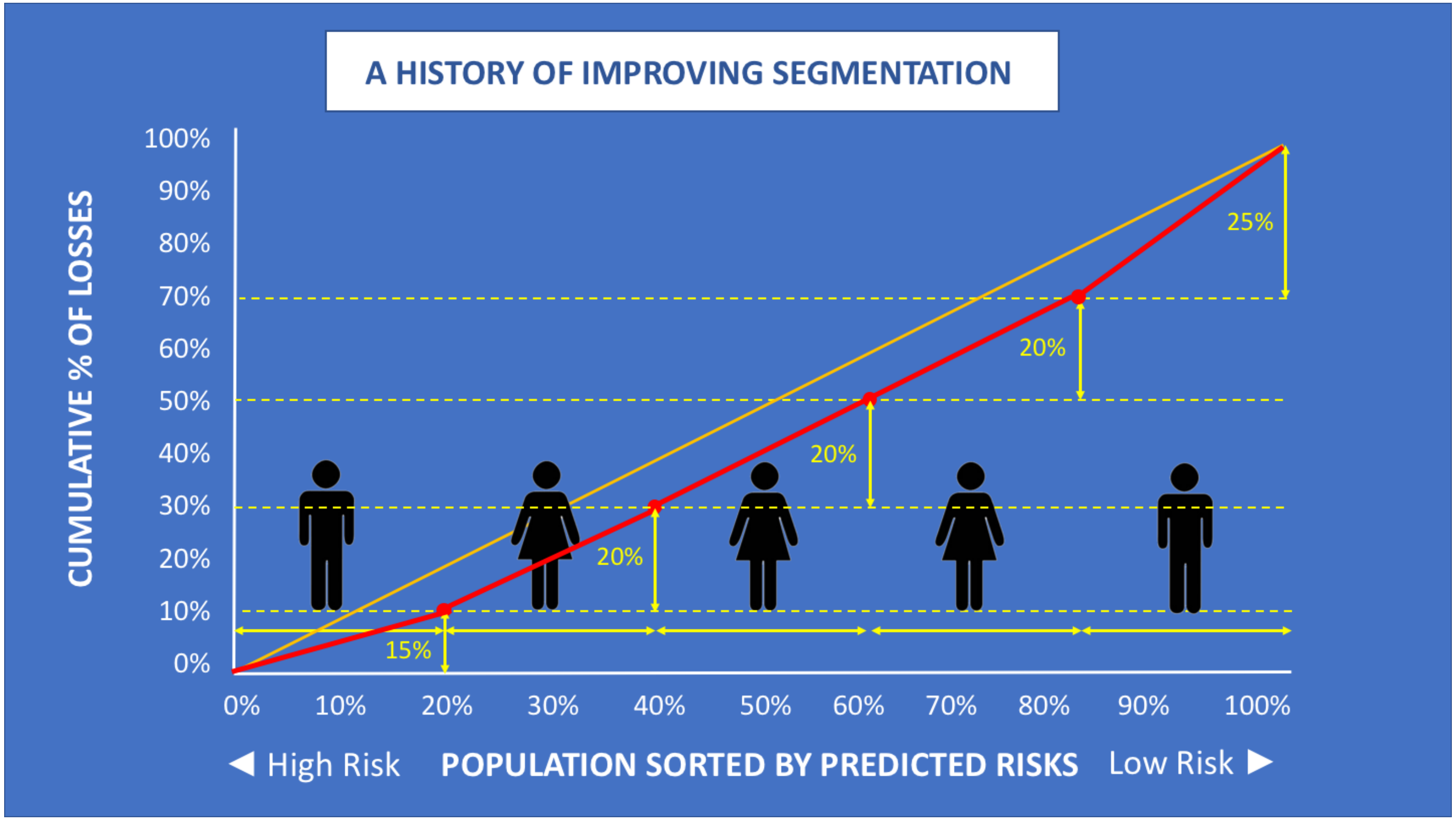
What if $\widehat{m}$ and m are not perfectly correctly correlated...?

What is the “average” model ?



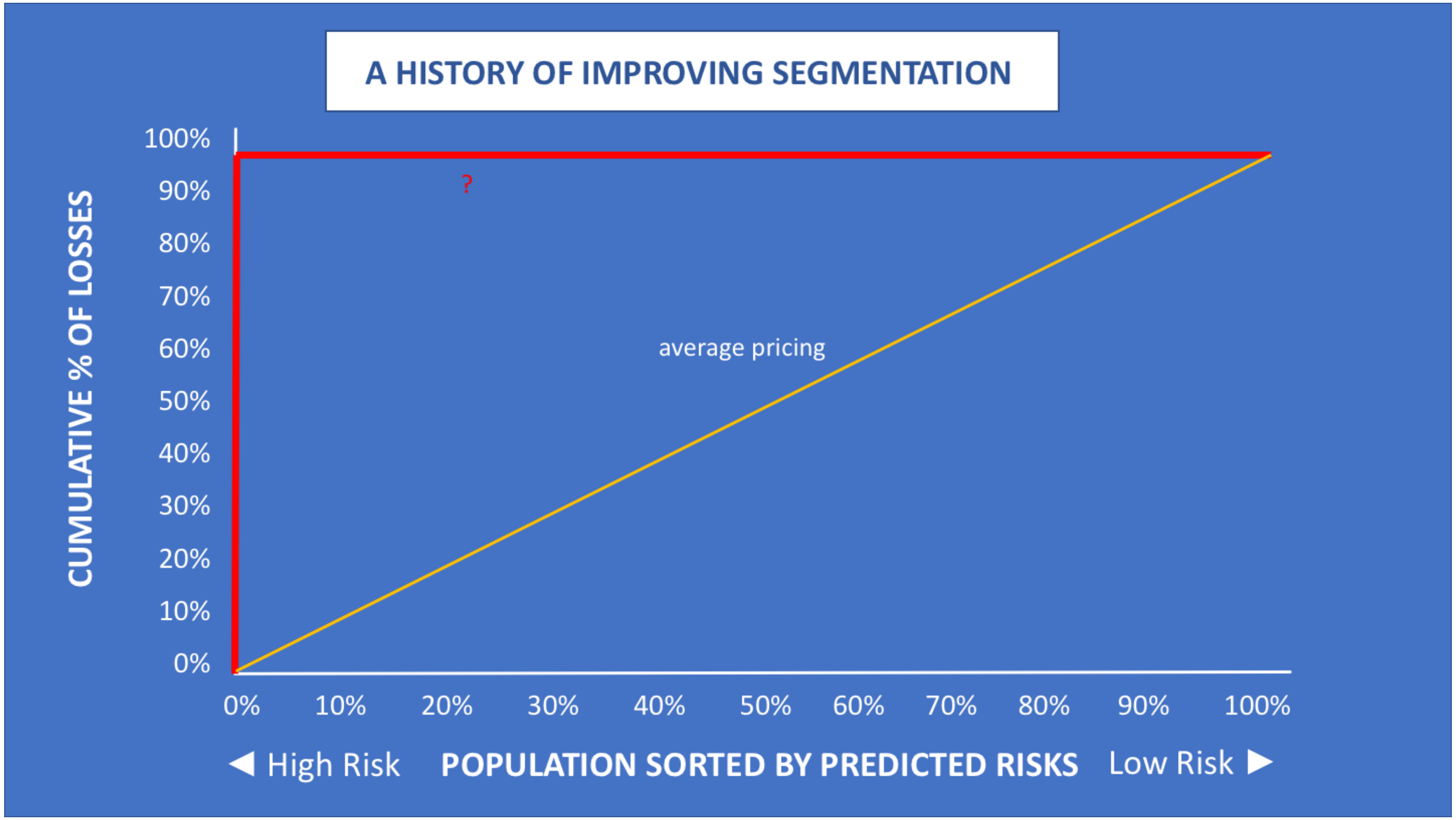
What is this “*average pricing*” ?

Can it be worst than the "average" model ?



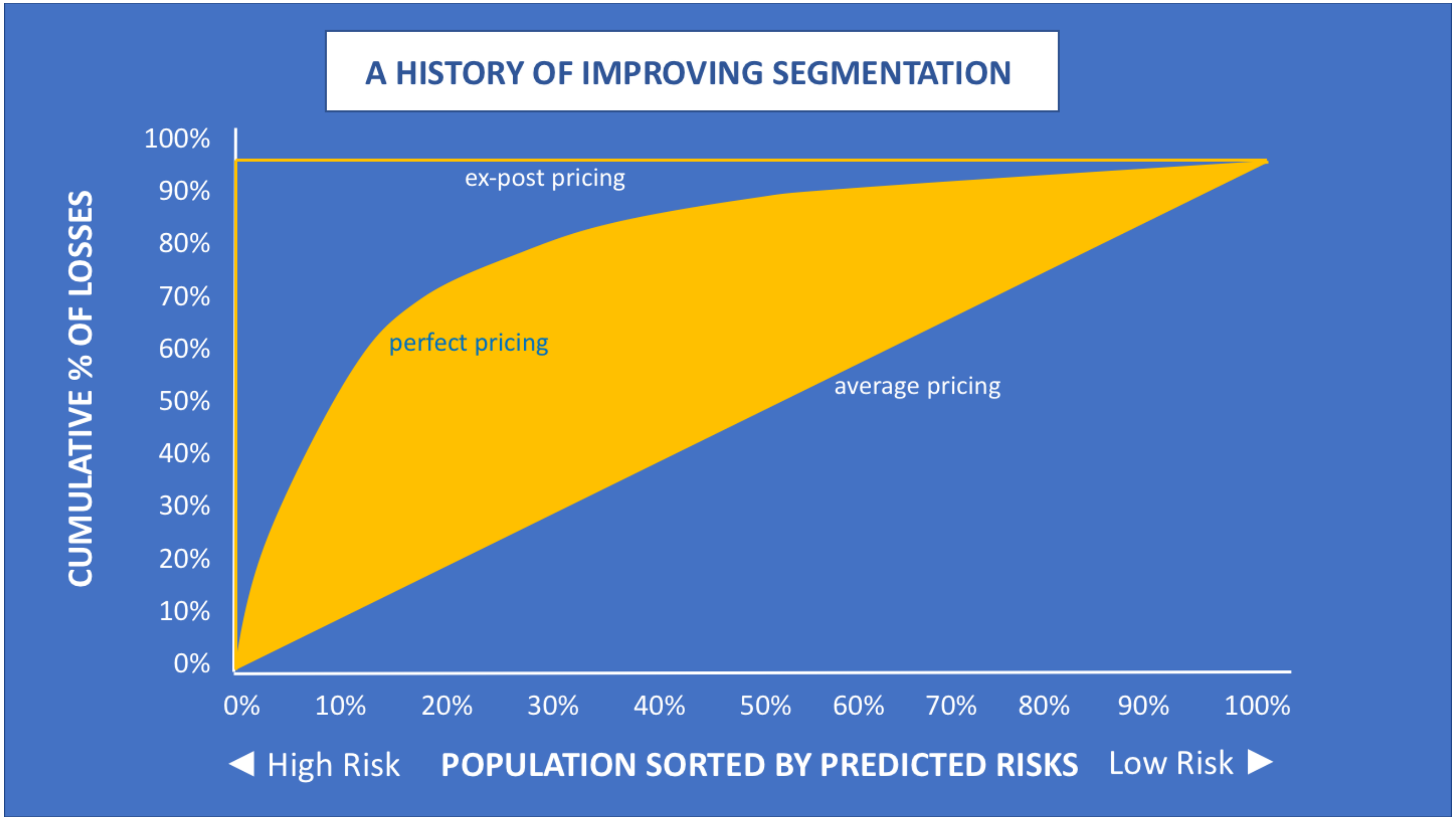
Is it a lower bond ? Is it possible to be below that curve ?

What is in the upper corner ?



What is the upper bond ? Ex-post pricing...

How to understand this (*pseudo*)-Lorenz curve ?

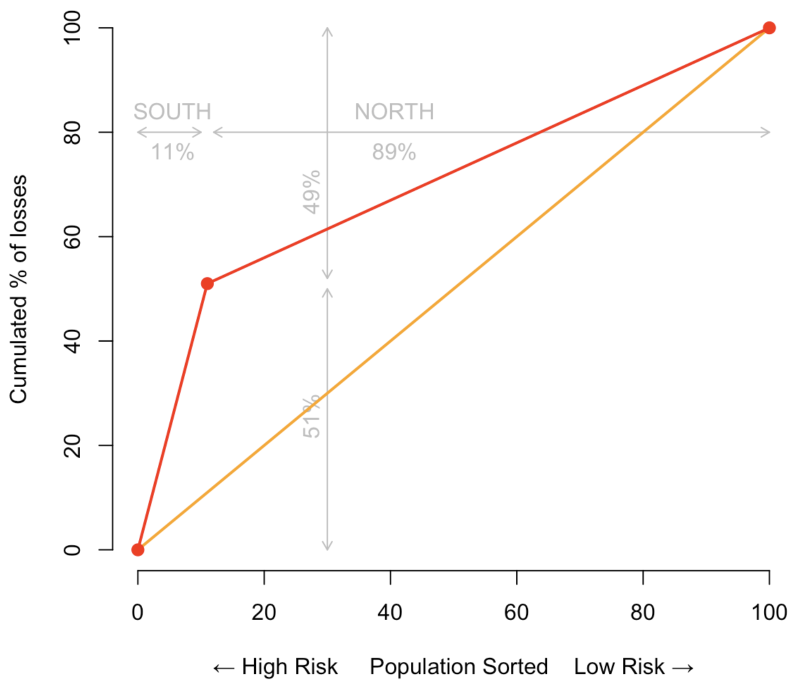


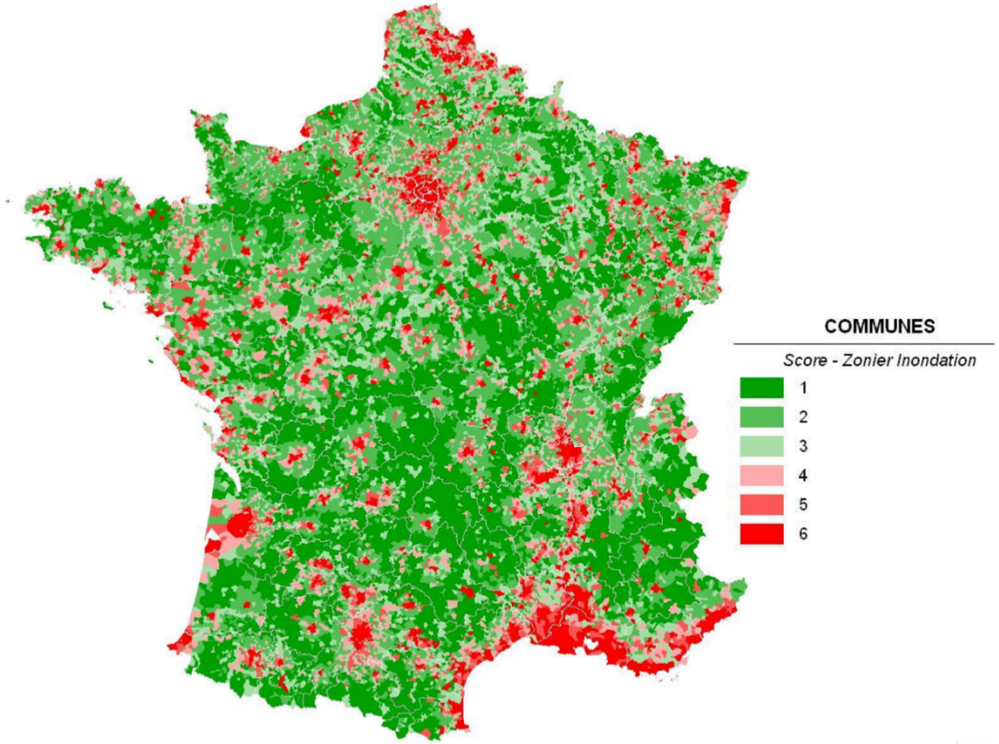
Is there a continuity between mutualization and hyper-segmentation ?

Insurance, Risk Pooling and Solidarity

Consider flood risk

One can look at the "Lorenz curve"



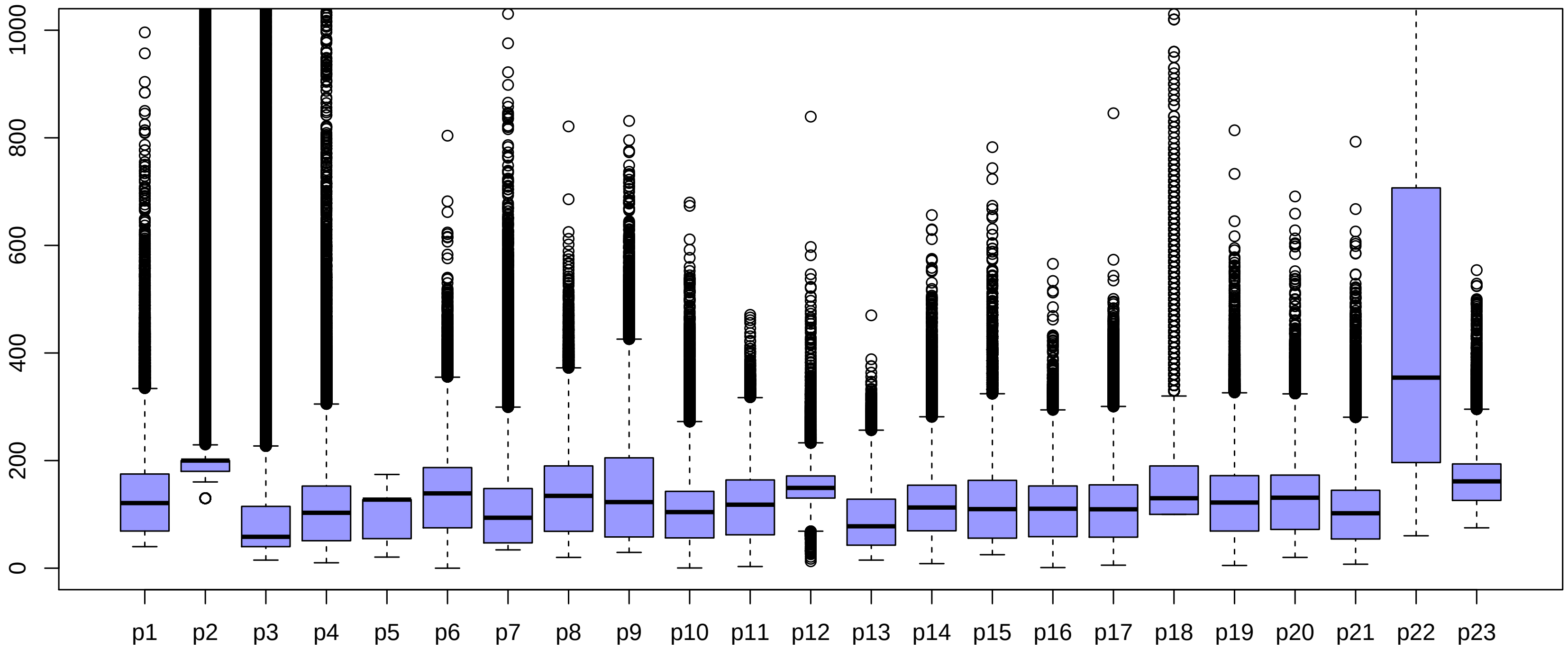


Field experiment: the actuarial pricing games

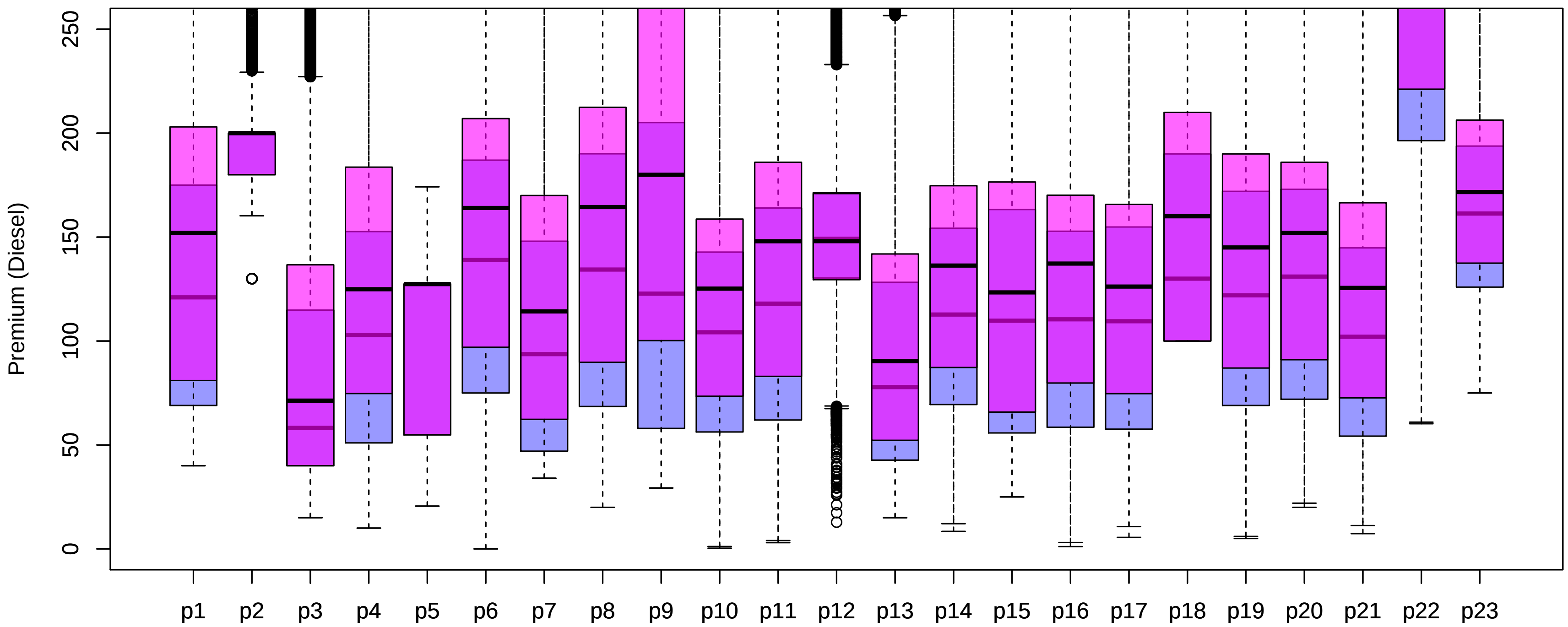
Actuarial pricing is data based, and model based

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

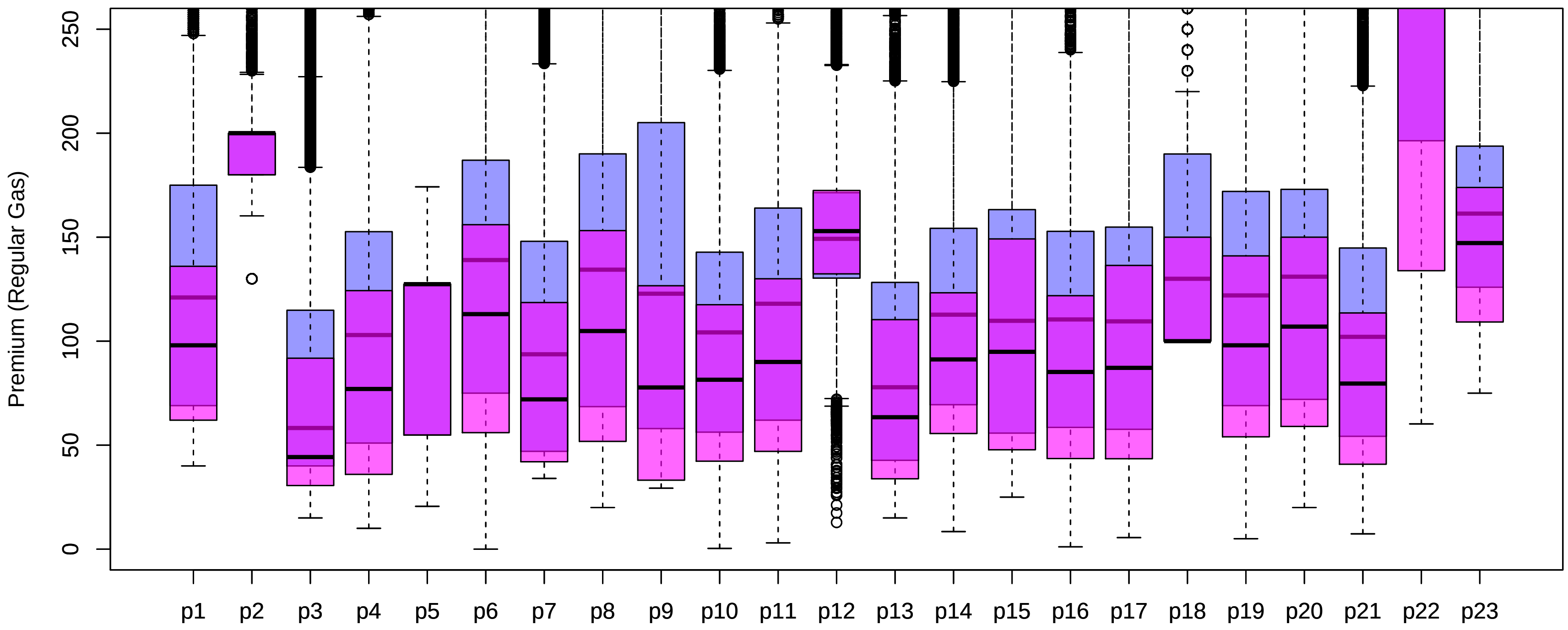
Insurance Ratemaking Before Competition



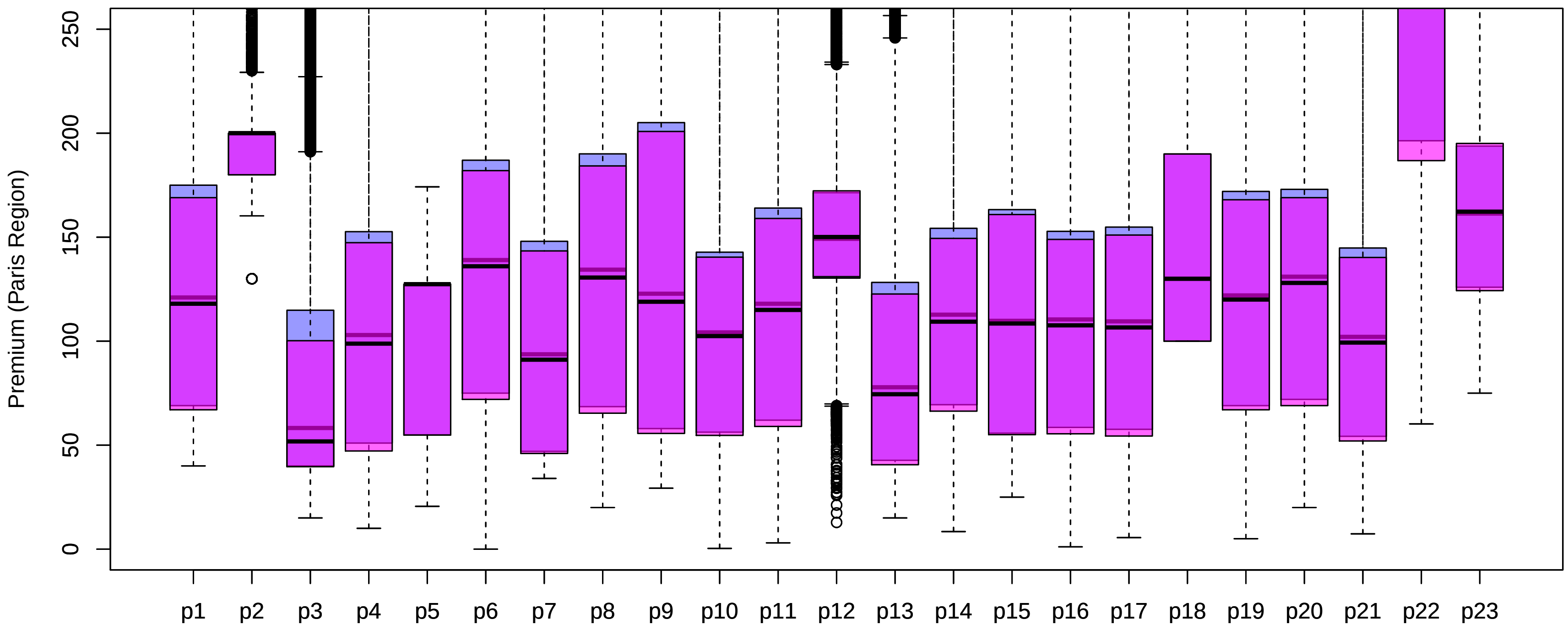
Insurance Ratemaking Before Competition Gas Type Diesel



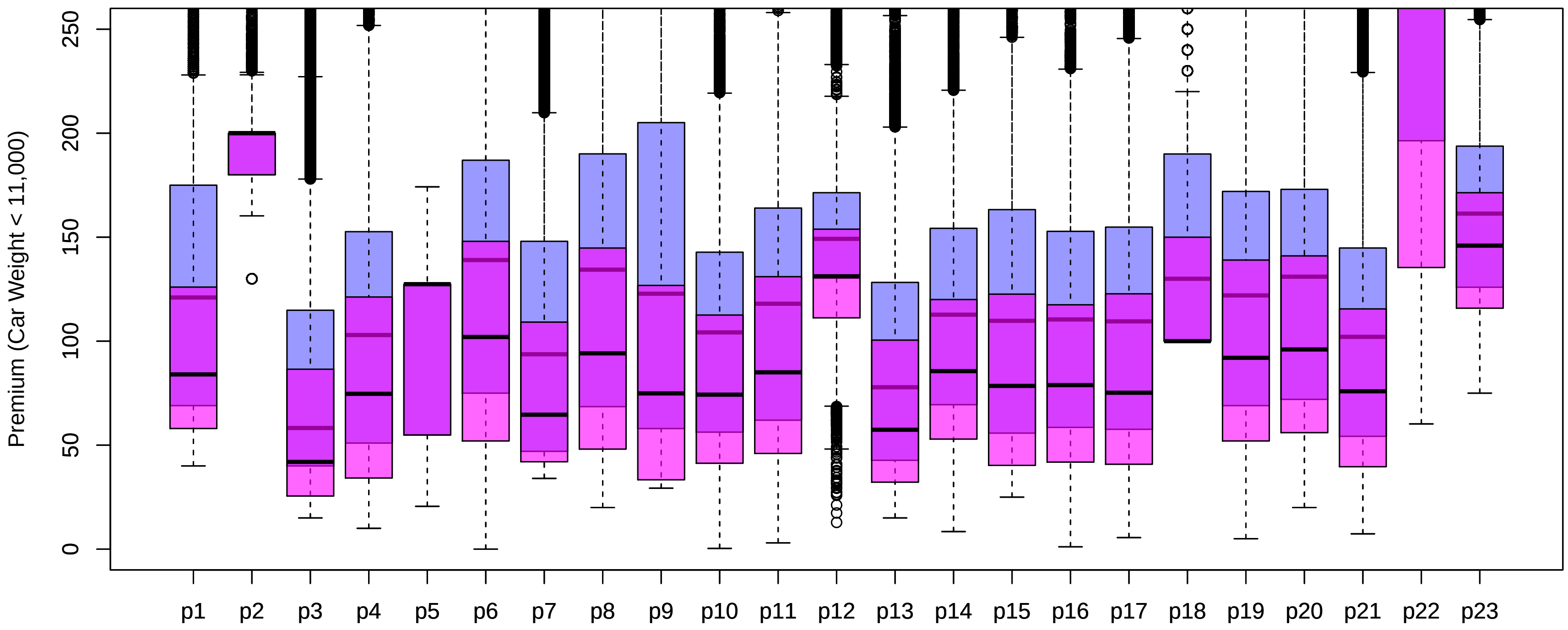
Insurance Ratemaking Before Competition Gas Type Regular



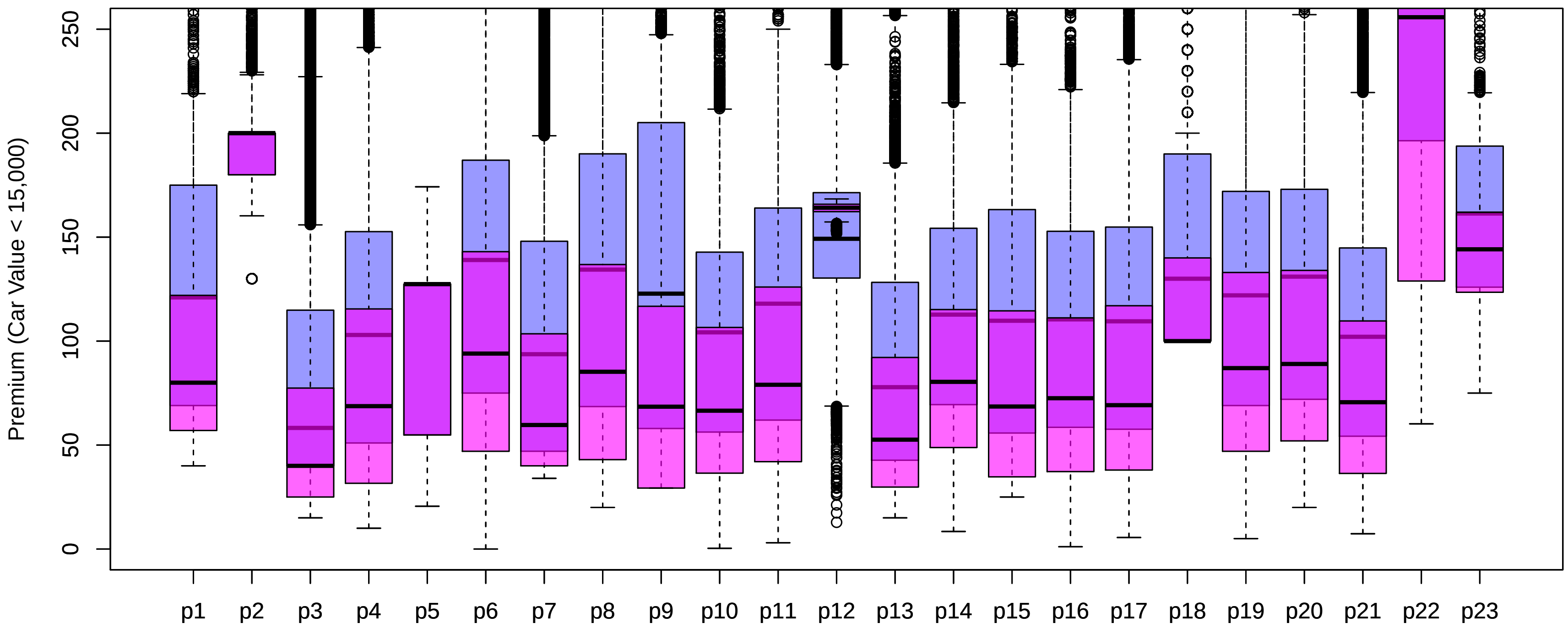
Insurance Ratemaking Before Competition Paris Region



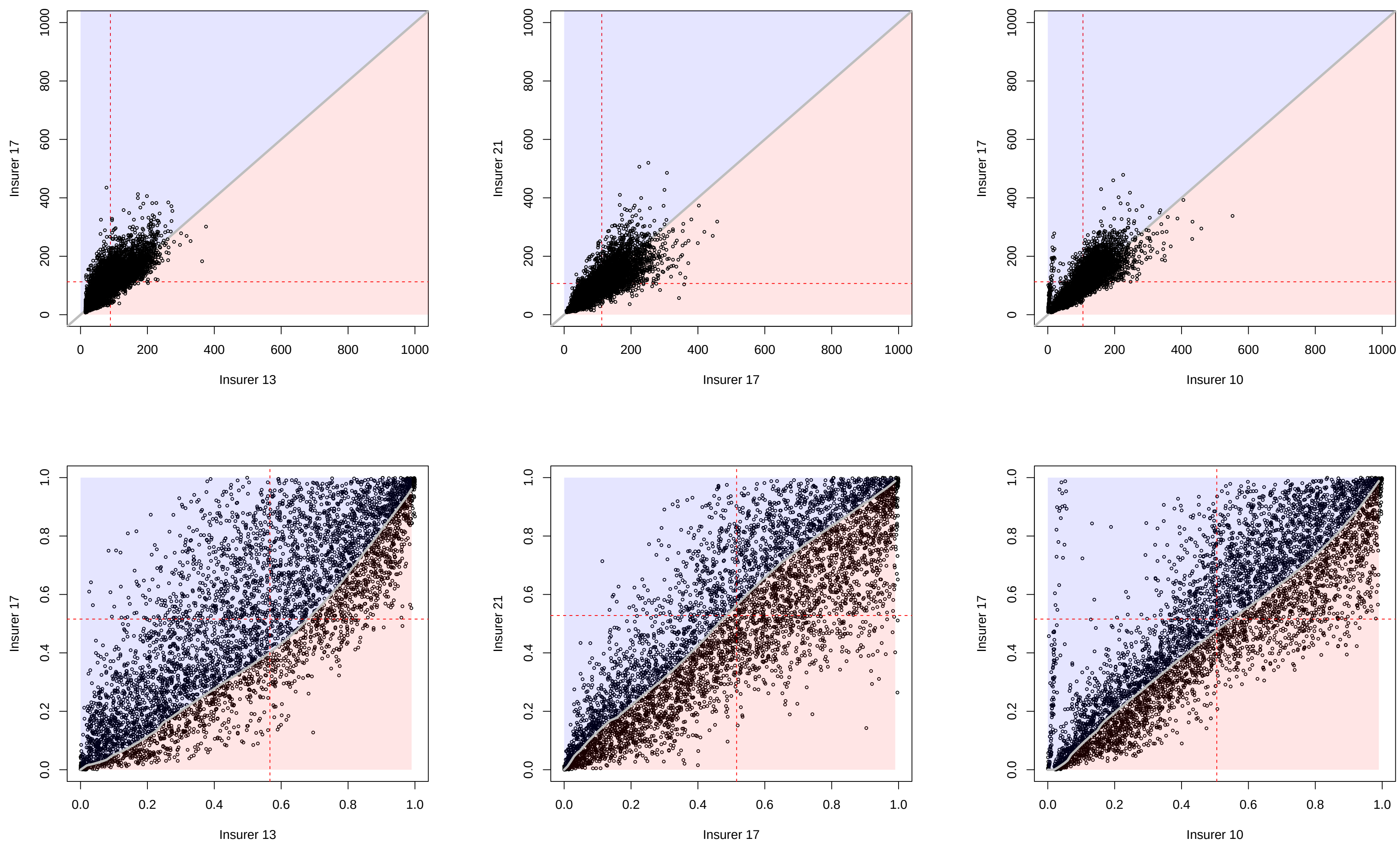
Insurance Ratemaking Before Competition Car Weight



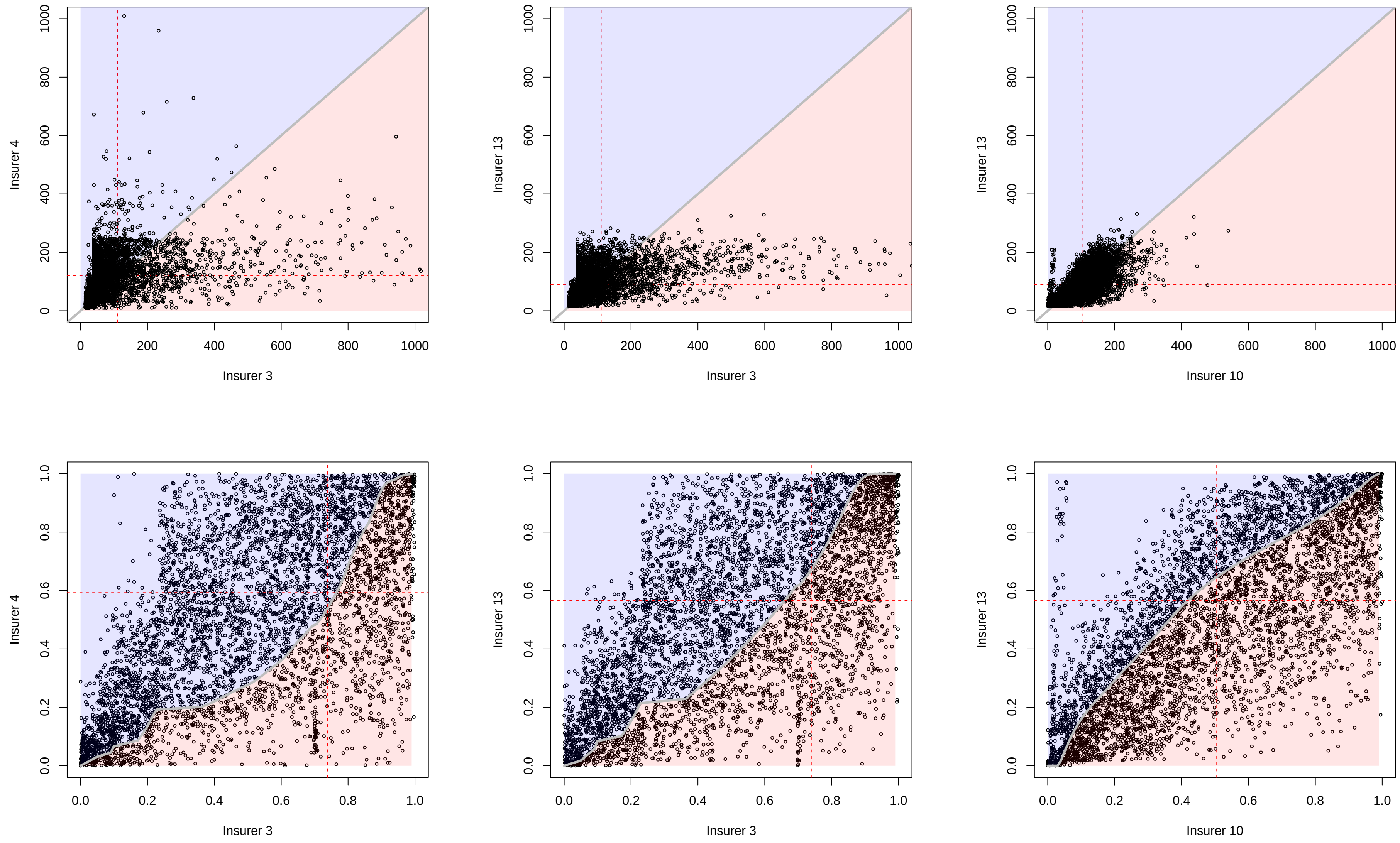
Insurance Ratemaking Before Competition Car Value



Insurance Ratemaking Competition : High Correlation ?



Insurance Ratemaking Competition : High Correlation ?



Insurance Ratemaking Competition

We need a **Decision Rule** to select premium chosen by insured i

	Ins1	Ins2	Ins3	Ins4	Ins5	Ins6
	787.93	706.97	1032.62	907.64	822.58	603.83
	170.04	197.81	285.99	212.71	177.87	265.13
	473.15	447.58	343.64	410.76	414.23	425.23
	337.98	336.20	468.45	339.33	383.55	672.91

Insurance Ratemaking Competition

Basic ‘**rational rule**’ $\pi_i = \min\{\widehat{\pi}_1(\boldsymbol{x}_i), \cdots, \widehat{\pi}_d(\boldsymbol{x}_i)\} = \widehat{\pi}_{1:d}(\boldsymbol{x}_i)$

	Ins1	Ins2	Ins3	Ins4	Ins5	Ins6
	787.93	706.97	1032.62	907.64	822.58	603.83
	170.04	197.81	285.99	212.71	177.87	265.13
	473.15	447.58	343.64	410.76	414.23	425.23
	337.98	336.20	468.45	339.33	383.55	672.91

Insurance Ratemaking Competition

A more **realistic rule** $\pi_i \in \{\widehat{\pi}_{1:d}(\boldsymbol{x}_i), \widehat{\pi}_{2:d}(\boldsymbol{x}_i), \widehat{\pi}_{3:d}(\boldsymbol{x}_i)\}$

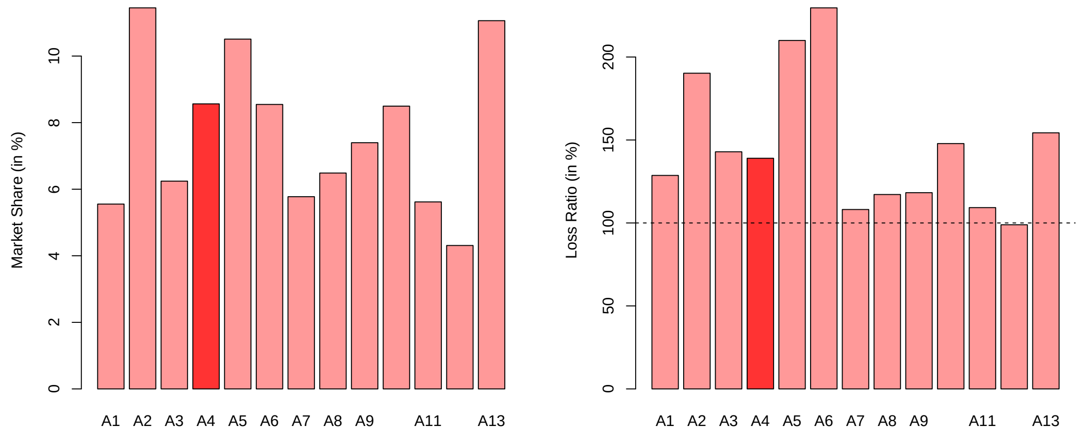
	Ins1	Ins2	Ins3	Ins4	Ins5	Ins6
	787.93	706.97	1032.62	907.64	822.58	603.83
	170.04	197.81	285.99	212.71	177.87	265.13
	473.15	447.58	343.64	410.76	414.23	425.23
	337.98	336.20	468.45	339.33	383.55	672.91

Pricing Game in 2015

Insurer 4

GLM for frequency and standard cost (large claimes were removed, above 15k), Interaction Age and Gender

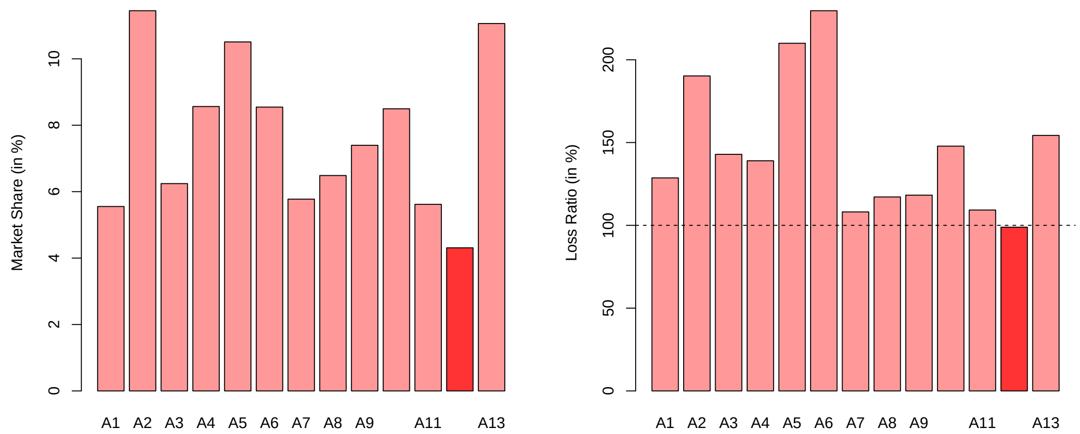
Actuary working for a *mutuelle* company



Insurer 11

Use of two XGBoost models (bodily injury and material), with correction for negative premiums

Actuary working for a private insurance company



Additional References (to go further)

- L'intelligence artificielle dilue-t-elle la responsabilité (2018)
- Les modèles prédictifs peuvent-ils être loyaux et justes (2017)
- L'éthique de la modélisation dans un monde où la normalité n'existe plus (2017)
- Les dérives du principe de précaution (2016)
- Segmentation et mutualisation, les deux faces d'une même pièce (2015)
- La tarification par genre en assurance, corrélation ou causalité ? (2016)
- Big data : passer d'une analyse de corrélation à une interprétation causale (2015)

and much more soon on freakonometrics.hypotheses.org