

# Assurance, discrimination et équité, quel rôle pour les actuaires

**Arthur Charpentier<sup>1</sup> & Laurence Barry<sup>2</sup>**

<sup>1</sup> Université du Québec à Montréal    <sup>2</sup> Chaire Pari

Atelier pluridisciplinaire sur la justice sociale, l'équité  
et les discriminations dans les systèmes algorithmiques, 2021

# Fables, and models

- ▶ “in economic theory, as in Harry Potter, the Emperor's New Clothes or the tales of King Solomon, we amuse ourselves in imaginary worlds. Economic theory spins tales and calls them models. An economic model is also somewhere between fantasy and reality ( . . . ) the word model sounds more scientific than the word fable or tale, but I think we are talking about the same thing”,

Rubinstein (2012)

- ▶ see also Bébéal (1986)

## CLARTE ET VERITE DES COMPTES D'UNE SOCIETE D'ASSURANCE SUR LA VIE

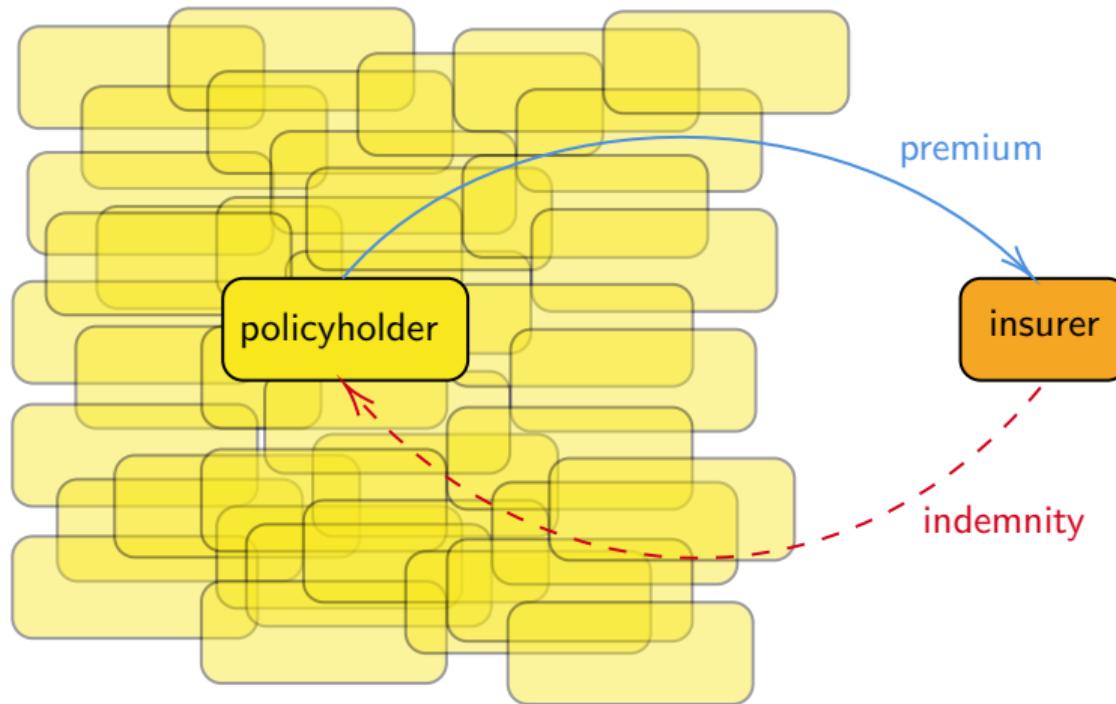
Exposé présenté par M. C. BÉBÉAR  
Membre agrégé

« Il était une fois un assureur sur la Vie fort intelligent, éminent et respecté. Il était considéré par tous comme un grand technicien, voire un savant, qui n'avait pas son pareil pour estimer la prime à demander aux assurés et la valeur des engagements pris par l'assureur. Mais, parfois, un doute le taraudait : sa science était-elle vraiment estimée ? Le vulgaire ne risquait-il pas de s'en saisir, montrer qu'il n'y avait là rien d'admirable et détruire une réputation bâtie à force de patience ? Vite, il fallait se protéger. Et c'est ainsi que notre assureur éminent, respecté et fort intelligent, inventa un langage ésotérique qu'il appela « actuariat ». Et, grâce à lui, des générations et des générations d'actuaires heureux vécurent et vivent encore à l'abri des critiques, adulés de tous. Chaque année, aidés par des comptables pleins de respect, ils mettent au point un rapport annuel fort savant qu'assurés, actionnaires et contrôleurs en tous genres contemplent sans trop comprendre et sans oser poser de questions de peur de passer pour des sots. Le système est parfait. »

(Conte persan)

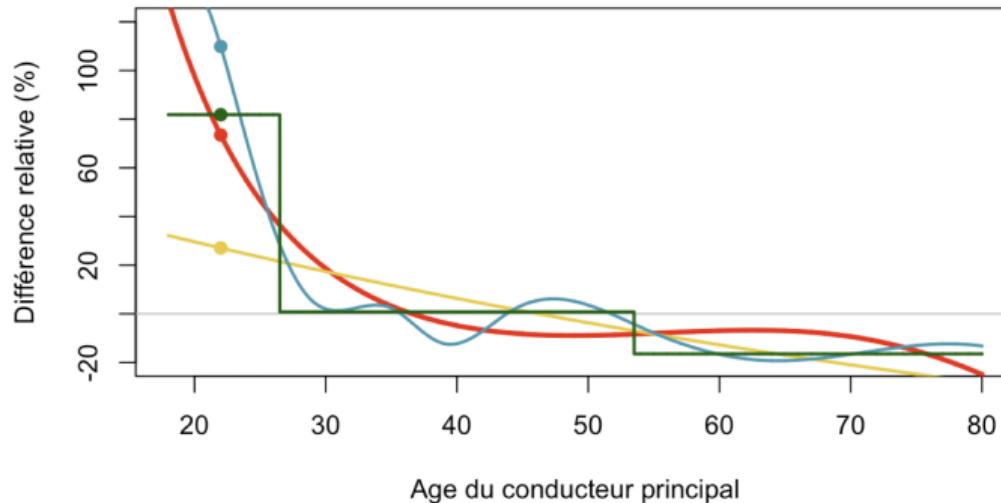
# Insurance and risk pooling & solidarity

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



# Risk Heterogeneity I

- ▶ Motor claim frequency as a function of the **age** the driver (Charpentier (2014))



- ▶ “*actuaries smoothed because smoothing was a ‘mathematical and ethical’ good. McClintock defended his practices to Hughes on ‘moral’ grounds*”, Bouk (2015)
- ▶ possible issues about the narrative part (explanability/interpretability)

## Risk Heterogeneity II

- ▶ Life insurance, survival probability as a function of the **age** and the **gender**
- ▶ Life tables, men vs. women, in 1720 ([Struyck \(1912\)](#), page 231)

men				women				
0	1000	29.0%	45	371	16.6%	0	1000	28.9%
5	710	5.6%	50	313	19.2%	5	711	5.2%
10	670	4.2%	55	253	22.9%	10	674	3.3%
15	642	5.5%	60	195	27.2%	15	652	4.3%
20	607	6.6%	65	142	31.7%	20	624	5.8%
25	567	7.9%	70	97	37.1%	25	588	6.8%
30	522	9.2%	75	61	45.9%	30	548	7.3%
35	474	10.5%	80	33	51.5%	35	508	7.9%
40	424	12.5%	85	16		40	468	9.6%

## Risk Heterogeneity III

- ▶ Life insurance, survival probability as a function of the **age** and the **gender**
- ▶ More recent life tables (TV/TD and INED)

TD 73-77		TV 73-77		TD 88-90		TV 88-90		INED (H)		INED (F)	
0	100000	0	100000	0	100000	0	100000	0	100000	0	100000
10	97961	10	98447	10	98835	10	99129	10	99486	10	99578
20	97105	20	98055	20	98277	20	98869	20	99281	20	99471
30	95559	30	97439	30	96759	30	98371	30	98656	30	99247
40	93516	40	96419	40	94746	40	97534	40	97661	40	98810
50	88380	50	94056	50	90778	50	95752	50	95497	50	97645
60	77772	60	89106	60	81884	60	92050	60	90104	60	94777
70	57981	70	78659	70	65649	70	84440	70	78947	70	89145
80	28364	80	52974	80	39041	80	65043	80	59879	80	77161
90	4986	90	14743	90	9389	90	24739	90	25123	90	44236
100	103	100	531	100	263	100	1479	100	1412	100	4874
110	0	110	0	110	0	110	2				

## Risk Heterogeneity IV

- ▶ Life insurance, residual life expectancy (in years) as a function of the **age**, the **gender** and being a **smoker** (or not), (data [Benjamin and Michaelson \(1988\)](#) 1970-1975, US)
- ▶ [Hoffman \(1931\)](#), [Johnston \(1945\)](#) “*it is clear that smoking is an important cause of mortality*”, [Miller and Gerstein \(1983\)](#)

men		women		
	non-smoker	smoker	non-smoker	smoker
25	48.4	42.8	25	52.8
35	38.7	33.3	35	43.0
45	29.2	24.2	45	33.5
55	20.3	16.5	55	24.5
65	12.8	10.4	65	16.2
				49.8
				40.1
				31.0
				22.6
				15.1

## Risk Heterogeneity V

- ▶ Life insurance, residual life expectancy (in years) as a function of the **age**, the **gender** and being a **weight** (BMI) (data [Steensma et al. \(2013\)](#) US)  
regular [ $18.5; 25\text{kg}/\text{m}^2$ ), overweighted [ $25; 30\text{kg}/\text{m}^2$ ), obesity I [ $30; 35\text{kg}/\text{m}^2$ ),  
obesity II [ $35, 100\text{kg}/\text{m}^2$ )
- ▶ [Crossley \(2005\)](#), [Czerniawski \(2007\)](#) or [Kelly and Markowitz \(2009\)](#)

		men				women			
		regular	over.	obesity I	obesity II	regular	over.	obesity I	obesity II
20		57.2	61.0	59.1	53.5	20	62.8	66.5	64.6
30		47.6	51.4	49.4	44.1	30	53.0	56.7	54.8
40		38.1	41.7	39.9	34.7	40	43.3	46.9	45.0
50		28.9	32.4	30.6	25.8	50	33.8	37.3	35.5
60		20.4	23.6	21.9	17.6	60	24.9	28.1	26.4
70		13.2	15.8	14.4	10.9	70	16.8	19.7	18.2

## Risk Heterogeneity VI

- ▶ handicap and genetic testing
- ▶ “*the insurance industry has generally regarded handicapped persons as undesirable risks*” Baker and Karol (1977)
- ▶ “*the denial of insurance coverage to an individual whose (noninherited) cancer had been long cured would not constitute genetic discrimination, while the denial of insurance to that individual's relatives because of the (erroneous) belief that that type of cancer is heritable would be genetic discrimination*” Natowicz et al. (1992)
- ▶ Schatz (1986), Clifford and Iculano (1987) (HIV experience), Jacobs and Sommers (2015) (inference from drugs)

# Insurance and premium “individualization” |

- ▶ “*il convient en effet de distinguer deux choses lorsque l'on parle d'assurance. La première, l'opération d'assurance, relève de la technique et a une dimension collective, la seconde, le contrat d'assurance, relève du droit et a une dimension individuelle*”, Bigot and Cayol (2020) (also Thiery and Van Schoubroeck (2006), Lehtonen and Liukko (2015))
- ▶ **Individualistic approach**
  - ▶ The individualistic approach to equality analyses fundamental rights, such as the right to equal treatment, in terms of individuals.
  - ▶ An individual cannot be treated differently because of his or her membership in such a group, particularly in a group to which he or she has not chosen to belong.
- ▶ **Group approach**
  - ▶ The insurance tradition, on the other hand, analyses risks, premiums and benefit schedules in terms of groups
  - ▶ Unlike the individualistic approach, insurance classification schemes rely on the assumption that individuals answer to the average (stereotypical) characteristics of a group to which they belong.

## Insurance and premium “individualization” II

- ▶ “at the core of insurance business lies discrimination between risky and non-risky insureds”, Avraham (2017)
- ▶ perfect segmentation with observable latent risk factor  $\Theta$

	policyholder	insurer
loss	$\mathbb{E}[Y \Theta]$	$Y - \mathbb{E}[Y \Theta]$
expected loss	$\mathbb{E}[Y]$	0
variance	$\text{Var}[\mathbb{E}[Y \Theta]]$	$\text{Var}[Y - \mathbb{E}[Y \Theta]]$

$$\text{Var}[Y] = \underbrace{\mathbb{E}[\text{Var}[Y|\Theta]]}_{\rightarrow \text{insurer}} + \underbrace{\text{Var}[\mathbb{E}[Y|\Theta]]}_{\rightarrow \text{policyholder}}.$$

## Insurance and premium “individualization” III

- ▶ statistical segmentation with observable features  $\mathbf{X} = (X_1, \dots, X_k)$  []  
“categorization based on immutable characteristics”, Crocker and Snow (2013)

	policyholder	insurer
loss	$\mathbb{E}[Y \mathbf{X}]$	$S - \mathbb{E}[Y \mathbf{X}]$
expected loss	$\mathbb{E}[Y]$	0
variance	$\text{Var}[\mathbb{E}[Y \mathbf{X}]]$	$\mathbb{E}[\text{Var}[Y \mathbf{X}]]$

$$\begin{aligned}\mathbb{E}[\text{Var}[Y|\mathbf{X}]] &= \mathbb{E}\left[\mathbb{E}\left[\text{Var}[Y|\Theta] \middle| \mathbf{X}\right]\right] + \mathbb{E}\left[\text{Var}\left[\mathbb{E}[Y|\Theta] \middle| \mathbf{X}\right]\right] \\ &= \underbrace{\mathbb{E}\left[\text{Var}[Y|\Theta]\right]}_{\text{perfect segmentation}} + \underbrace{\mathbb{E}\left\{\text{Var}\left[\mathbb{E}[Y|\Theta] \middle| \mathbf{X}\right]\right\}}_{\text{misfit}}.\end{aligned}$$

- ▶ “kanssolidariteit” vs “subsidierende solidariteit”, De Pril and Dhaene (1996)

# Insurance(s) & solidarity

- ▶ **health insurance**
- ▶ **collective insurance**
- ▶ **natural catastrophes**

*“La Nation proclame la solidarité et l'égalité de tous les Français devant les charges qui résultent des calamités nationales”, Constitution du 27 octobre 1946*

*“la solidarité en assurance, c'est décider de ne pas segmenter le marché du risque correspondant sur une base des caractéristiques observables des risques des individus”, Gollier (2002).*

- ▶ **general insurance**

*“Tout ce qui n'est pas défendu par la Loi ne peut être empêché”, Déclaration des Droits de l'Homme et du Citoyen, 1789, art. 5*

# Legal aspects of regulation I

- ▶ “accéder à l’assurance s’entend non seulement de la possibilité même de souscrire un contrat en vue d’une couverture, mais peut-être, également, à un coût économique raisonnable, non prohibitif partant non dissuasif” ([Noguéro \(2010\)](#))
- ▶ le **sex**e ou le **gen**re (art. A. 111-6 du Code des assurances, Commission européenne (Arr. 18 déc. 2012, NOR : EFIT1238658A, relatif à l’égalité entre les hommes et les femmes en assurance, JO 20 déc., mod. par Arr. 3 févr. 2014, NOR : EFIT1400411A, JO 11 févr.)
- ▶ distinction fondée sur l’**âg**e (C. pén., art. 225-1 et 225-2), (“*belonging to a particular race or sex is akin to joining one specific ‘club at the moment of conception, whereas age...*”, [Macnicol \(2006\)](#))
- ▶ la **situation de famille** ou sur l’ **orientation sexuelle** (C. pén., art. 225-1 et 225-2)
- ▶ en raison du **lieu de résidence** d’une personne constitue une discrimination au sens pénal (C. pén., art. 225-1)
- ▶ “*Nul ne peut faire l’objet de discriminations en raison de ses caractéristiques génétiques*” (C. C., art. 16-13)

## Legal aspects of regulation II

	CA	HI	GA	NC	NY	MA	PA	FL	TX
Genre	X	X	•	X	•	X	X	•	•
Âge	X	X	•	X*	•	X	•	•	•
Expérience de conduite	•	X	•	•	•	•	•	•	•
Antécédents de crédit	X	X	•	•	•	X	•*	•	•
Éducation	X	X	X	X	X	X	•	•	•
Profession	X	X	X	•	X	X	•	•	•
Situation d'emploi	X	X	X	•	X	X	•	•	•
Situation de famille	•	X	•	•	•	X	•	•	•
Situation résidentielle	X	X	•	•	•	X	•	•	•
Adresse/code postal	•	•	•	•	•	•	•	•	•
Antécédents d'assurance	•	•	•	•	•	•	•	•	•

# Discrimination by proxies (?) I

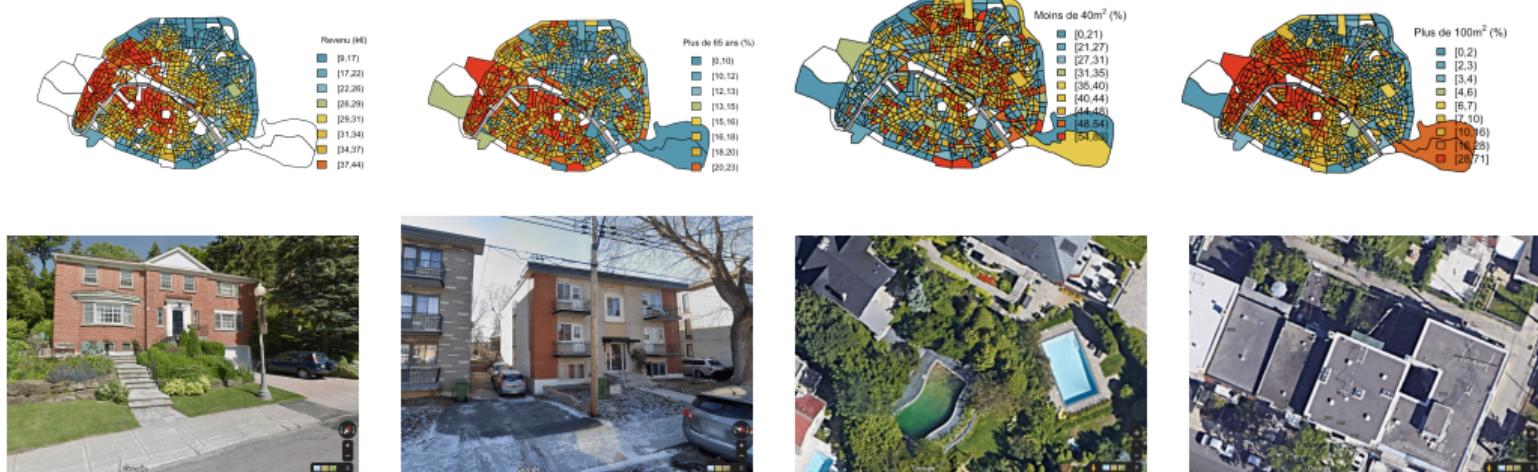
- ▶ **faces**, recently, some insurers considered the idea of using facial recognition to predict some diseases, [Shikhare \(2021\)](#)
- ▶ see also phrenology [Lombroso \(1876\)](#) and [Bertillon and Chervin \(1909\)](#)



# Discrimination by proxies (?) II

- ▶ location (home address of policyholder)

Jean et al. (2016), Seresinhe et al. (2017), Gebru et al. (2017), Law et al. (2019), Illic et al. (2019), Kita and Kidziński (2019)



# Discrimination by proxies (?) III

- ▶ credit scoring, very important in North America

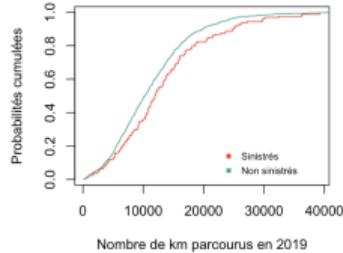
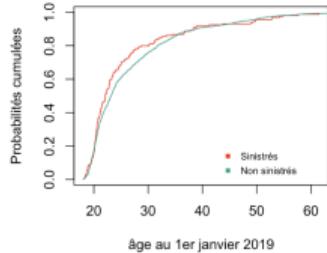
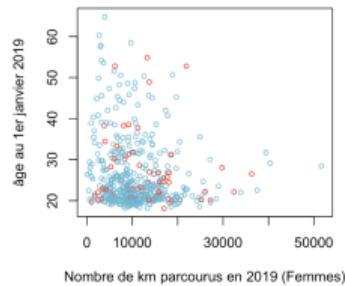
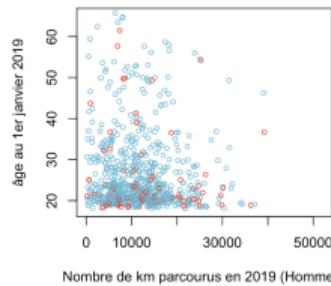
Kabler (2004), Arya et al. (2013), Miller et al. (2003) Bartik and Nelson (2016), O'neil (2016), Lauer (2017), Morris et al. (2017), Kiviat (2019)



source <https://www.incharge.org/debt-relief/credit-counseling/>

# Discrimination by proxies (?) IV

## ► telematic data, behavioral features



E.g. distance driven and gender are correlated... [Verbelen et al. \(2018\)](#)

## Wrap-up

- ▶ “les États membres peuvent décider (...) d'autoriser des différences proportionnelles dans les primes et les prestations des particuliers lorsque l'utilisation du sexe est un facteur déterminant dans l'évaluation du risque, sur la base de données actuarielles et statistiques pertinentes et précises”  
**causal** models ?
- ▶ “the myth of the actuary, a powerful rhetorical situation in which decisions appear to be based on objectively determined criteria when they are also largely based on subjective ones”, **Glenn (2000)**  
“virtually every aspect of the insurance industry is predicated on stories first and then numbers”, **Glenn (2003)**  
importance of the **narrative** (“all models are wrong but some models are useful”, **Box et al. (2011)**).

## References I

- Arya, S., Eckel, C., and Wichman, C. (2013). Anatomy of the credit score. *Journal of Economic Behavior & Organization*, 95:175–185.
- Avraham, R. (2017). Discrimination and insurance. In Lippert-Rasmussen, K., editor, *Handbook of the Ethics of Discrimination*, pages 335–347. Routledge.
- Baker, L. D. and Karol, C. (1977). Employee insurance benefit plans and discrimination on the basis of handicap. *DePaul L. Rev.*, 27:1013.
- Bartik, A. and Nelson, S. (2016). Deleting a signal: Evidence from pre-employment credit checks. *SSRN*, 2759560.
- Benjamin, B. and Michaelson, R. (1988). Mortality differences between smokers and non-smokers. *Journal of the Institute of Actuaries*, 115(3):519525.
- Bertillon, A. and Chervin, A. (1909). *Anthropologie métrique: conseils pratiques aux missionnaires scientifiques sur la manière de mesurer, de photographier et de décrire des sujets vivants et des pièces anatomiques*. Imprimerie nationale.
- Bigot, R. and Cayol, A. (2020). *Le droit des assurances en tableaux*. Ellipses.
- Bouk, D. (2015). *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. The University of Chicago Press.

## References II

- Box, G. E., Luceño, A., and del Carmen Paniagua-Quinones, M. (2011). *Statistical control by monitoring and adjustment*, volume 700. John Wiley & Sons.
- Bébéar, C. (1986). Clarté et vérité sur les comptes d'une société d'assurance sur la vie. *Bulletin de l'Institut des Actuaires*.
- Charpentier, A. (2014). *Computational Actuarial Science*. The R series. CRC Press.
- Clifford, K. A. and Iculano, R. P. (1987). AIDS and insurance: the rationale for AIDS-related testing. *Harvard law review*, 100(7):1806–1825.
- Crocker, K. J. and Snow, A. (2013). The theory of risk classification. In Loubergé, H. and Dionne, G., editors, *Handbook of insurance*, pages 281–313. Springer.
- Crossley, M. (2005). Discrimination against the unhealthy in health insurance. *University of Kansas Law Review*, 54:73.
- Czerniawski, A. M. (2007). From average to ideal: The evolution of the height and weight table in the United States, 1836–1943. *Social Science History*, 31(2):273–296.
- De Pril, N. and Dhaene, J. (1996). Segmentering in verzekeringen. *DTEW Research Report 9648*, pages 1–56.

## References III

- Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., and Fei-Fei, L. (2017). Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences*, 114(50):13108–13113.
- Glenn, B. J. (2000). The shifting rhetoric of insurance denial. *Law and Society Review*, pages 779–808.
- Glenn, B. J. (2003). Postmodernism: the basis of insurance. *Risk Management and Insurance Review*, 6(2):131–143.
- Gollier, C. (2002). La solidarite sous langle economique. *Revue Générale du Droit des Assurances*, pages 824–830.
- Hoffman, F. L. (1931). Cancer and smoking habits. *Annals of surgery*, 93(1):50.
- Ilic, L., Sawada, M., and Zarzelli, A. (2019). Deep mapping gentrification in a large canadian city using deep learning and google street view. *PloS one*, 14(3):e0212814.
- Jacobs, D. B. and Sommers, B. D. (2015). Using drugs to discriminateadverse selection in the insurance marketplace. *New England Journal of Medicine*.
- Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, 353(6301):790–794.
- Johnston, L. (1945). Effects of tobacco smoking on health. *British Medical Journal*, 2(4411):98.

## References IV

- Kabler, B. (2004). Insurance-based credit scores: Impact on minority and low income populations in missouri. *State of Missouri Departement of Insurance*.
- Kelly, I. R. and Markowitz, S. (2009). Incentives in obesity and health insurance. *Inquiry*, 46(4):418–432.
- Kita, K. and Kidziński, Ł. (2019). Google street view image of a house predicts car accident risk of its resident. *arXiv*, 1904.05270.
- Kiviat, B. (2019). The moral limits of predictive practices: The case of credit-based insurance scores. *American Sociological Review*, 84(6):1134–1158.
- Lauer, J. (2017). *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*. Columbia University Press.
- Law, S., Paige, B., and Russell, C. (2019). Take a look around: Using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology*, 10(5).
- Lehtonen, T.-K. and Liukko, J. (2015). Producing solidarity, inequality and exclusion through insurance. *Res Publica*, 21(2):155–169.
- Lombroso, C. (1876). *L'uomo delinquente*. Hoepli.
- Macnicol, J. (2006). *Age discrimination: An historical and contemporary analysis*. Cambridge University Press.

## References V

- Miller, G. and Gerstein, D. R. (1983). The life expectancy of nonsmoking men and women. *Public Health Reports*, 98(4):343.
- Miller, M. J., Smith, R. A., and Southwood, K. N. (2003). The relationship of credit-based insurance scores to private passenger automobile insurance loss propensity. *Actuarial Study, Epic Actuaries*.
- Morris, D. S., Schwarcz, D., and Teitelbaum, J. C. (2017). Do credit-based insurance scores proxy for income in predicting auto claim risk? *Journal of Empirical Legal Studies*, 14(2):397–423.
- Natowicz, M. R., Alper, J. K., and Alper, J. S. (1992). Genetic discrimination and the law. *American Journal of Human Genetics*, 50(3):465.
- Noguéro, D. (2010). Sélection des risques. discrimination, assurance et protection des personnes vulnérables. *Revue générale du droit des assurances*, 3:633–663.
- O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Rubinstein, A. (2012). *Economic fables*. Open book publishers.
- Schatz, B. (1986). The aids insurance crisis: Underwriting or overreaching. *Harvard Law Review*, 100:1782.
- Seresinhe, C. I., Preis, T., and Moat, H. S. (2017). Using deep learning to quantify the beauty of outdoor places. *Royal Society open science*, 4(7):170170.

## References VI

- Shikhare, S. (2021). Next generation ltc - life insurance underwriting using facial score model. In *Insurance Data Science conference*.
- Steensma, C., Loukine, L., Orpana, H., Lo, E., Choi, B., Waters, C., and Martel, S. (2013). Comparing life expectancy and health-adjusted life expectancy by body mass index category in adult canadians: a descriptive study. *Population health metrics*, 11(1):1–12.
- Struyck, N. (1912). *Les oeuvres de Nicolas Struyck (1687-1769): qui se rapportent au calcul des chances, à la statistique général, z la statistique des décès et aux rentes viagèter*. Société générale néerlandaise d'assurances sur la vie et de rentes viagères.
- Thiery, Y. and Van Schoubroeck, C. (2006). Fairness and equality in insurance classification. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 31(2):190–211.
- Verbelen, R., Antonio, K., and Claeskens, G. (2018). Unravelling the predictive power of telematics data in car insurance pricing. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(5):1275–1304.