

SCOR-UQAM RESEARCH REPORT #1 (2023-2024)

Arthur Charpentier (UQAM; charpentier.arthur@uqam.ca)



Fairness and insurance, Project financed by SCOR Foundation for Science

PI: Arthur Charpentier, Professeur UQAM (Université du Québec à Montréal)

- publication of [Insurance, Biases, Discrimination and Fairness](#), Springer,
- 6 internship grants, 4 PhD students involved (2 defended, 2 started), 3 postdoctoral fellowship (1 ended, 2 stated), 3 invitations
- 10 published articles in peer-review journals and conferences, 10 working papers submitted, 5 dissemination articles
- presentations in 29 conferences and workshops, 16 in seminars, involved in the organization of 3 workshops

Main results

- algorithms with theoretical guarantees to **mitigate discrimination** based on Wasserstein Distance and barycenters: application to multiple sensitive attributes with python package
- algorithms to measure **counterfactual fairness** based on optimal transport in probabilistic graphical models, with a sequential approach,
- statistical tools to investigate **calibration** of predictive models, with computational and philosophical perspectives (interpretation of predictive scores)
- creation of (theoretical) **competitive markets** to understand possible **algorithmic collusion** with reinforcement learning techniques
- investigation of **imbalanced regression** and **generative processes**

Context

The integration of predictive models in insurance markets has revolutionized the way insurers assess risk, set premiums, and manage claims. However, the deployment of these models raises critical concerns regarding fairness, particularly when algorithms inadvertently perpetuate or exacerbate existing social inequalities. Ensuring fairness in predictive modeling is essential to maintain public trust, comply with regulatory standards, and uphold ethical principles within the insurance industry.

Predictive models often rely on vast amounts of historical data, which may contain embedded biases related to race, gender, socioeconomic status, and other sensitive attributes.

See <https://freakonometrics.hypotheses.org/>

When such biases are not adequately addressed, the models can produce discriminatory outcomes, such as disproportionately higher premiums or denial of coverage for certain groups. This not only leads to unequal treatment of policyholders but also poses legal risks for insurers under anti-discrimination laws and regulations that govern fair lending and insurance practices.

To promote fairness in predictive modeling within insurance markets, several strategies can be employed. These include implementing bias detection and mitigation techniques during model development, such as fairness-aware machine learning algorithms that adjust for identified disparities. Additionally, transparency in model decision-making processes and ongoing monitoring for unfair outcomes are crucial. Regulatory frameworks also play a vital role by setting standards and guidelines that ensure predictive models are used

responsibly and equitably. Collaborative efforts between data scientists, insurers, regulators, and stakeholders are necessary to create models that not only enhance efficiency and profitability but also uphold the principles of fairness and social justice.

This research project, named (in short) “[fairness and insurance project](#)” is funded for 3 years by the SCOR Research Foundation for Science.

Results

• Overview on insurance and fairness

[Charpentier \(2024\)](#) has proposed a state of the art on issues of discrimination and equity, as they relate to insurance. This reference work has been used in doctoral courses (in Montreal and at ENSAE-Institut Polytechnique in Paris), as well as at the Warsaw Actuarial Summer School. A case study on real data in automobile insurance was published in [Moriah et al. \(2024\)](#), and [Barry et al. \(2023\)](#) revisits ethical and philosophical issues related to machine learning and premium personalization in insurance.

- Charpentier, A. (2024) [Insurance, biases, discrimination and fairness](#). Springer. ISBN: 978-3-031-49782-7.
- Moriah, M. Vermet, F. & Charpentier, A. (2024) [Measuring and Mitigating Biases in Motor Insurance Pricing](#). *European Actuarial Journal*, doi:10.1007/s13385-024-00390.
- Barry, L. & Charpentier, A. (2023) [Melting contestation: insurance fairness and machine learning](#). *Ethics and Information Technology*. doi:10.1007/s10676-023-09720-y

• Mitigating discrimination

We have investigated the use of optimal transport techniques to reduce the risk of predictors, and scores, with fair predictions against a protected criterion, and in particular barycenter approaches (and in particular the Wasserstein barycenter). Discrimination is defined here using a strong demographic parity criterion. In the case of a binary criterion (man or woman in the case of sexism, black or not in the case of racism, etc.), [Charpentier et al. \(2023\)](#) took up the general idea,

which consists in weighting the prediction obtained in the two groups, at an unchanged relative risk level. [Hu et al. \(2024\)](#) discussed the extension to several sensitive variables, and using associativity properties, we showed that we had a simple and interpretable method for returning the fair prediction. While the order doesn't affect the final forecast, it can raise concerns in terms of sequential (or cascade) interpretation. Last but not least, [Hu et al. \(2023\)](#) studied the cost of imposing a parametric model constraint on the fair forecast. This work has been presented at international conferences (AAAI, ECML, the Workshop on Trustworthy Artificial Intelligence, IDSC'24 (Insurance Data Science Conference), or the 17th Financial Risks International Forum) and seminars (København Universiteit, Paris Sorbonne).

- Charpentier, A., Hu, F., & Ratz, P. (2023). [Mitigating discrimination in insurance with Wasserstein barycenters](#). *3rd Workshop on Bias and Fairness in AI, International Workshop of ECML PKDD*.
- Hu, F., Ratz, P., & Charpentier, A. (2024). [A sequentially fair mechanism for multiple sensitive attributes](#). In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 11, pp. 12502-12510).
- Hu, F., Ratz, P. & Charpentier, A. (2023) [Parametric Fairness with Statistical Guarantees](#). *arXiv:2310.20508*.

• Counterfactual fairness

While the barycentric attenuation approach has interesting theoretical properties, it comes at the end of the chain, and is content to correct the models to make them non-discriminatory (post-processing). But to better understand the source of discrimination (certain differentiations may be perceived as legitimate), it is essential to have a causal approach to discrimination. [Côté et al. \(2024\)](#) returns to the use of causal graphs to better define several definitions of equity that may seem contradictory, and proposes an exhaustive framework of the different definitions of discrimination, and in particular indirect discrimination. [Fernandes Machado et al. \(2024\)](#) uses optimal transport approaches to create counterfactuals, and to quantify a ‘conditional average treatment effect’ (using causal anal-

ysis terminology) *mutatis mutandis*. In particular, the proposed approach makes it possible to answer the question ‘what would the decision have been if the person had been white instead of black’. This work was presented at the 2nd Workshop on Fairness and Discrimination in Insurance, JEDA’24, as well as various conferences and seminars (UNSW, SSC’24, Workshop on Optimal Transport).

- Côté, O., Côté, M.P. & Charpentier, A. (2024) [A Fair price to pay: exploiting causal graphs for fairness in insurance](#). *SSRN:4709243*.
- Fernandes Machado, A., Charpentier, A., Flachaire, E., Galllic, E. & Hu, F. (2024) [Sequential Conditional Transport on Probabilistic Graphs for Interpretable Counterfactual Fairness](#). *arXiv:2408.03425*.

• Calibration of predictive models

Calibration is a dimension often neglected in machine learning models, but fundamental to actuarial science, and raises the question of whether a score, returned by an algorithm, can have a probabilistic interpretation. From an epistemological point of view, the question “what does it mean to have a 20% chance of having an accident in a year?” is very close to the questions historically asked in meteorology “what does it mean to have a 20% chance of having rain?” [Fernandes Machado et al. \(2024a, 2024b\)](#) returns to the properties of scores predicted by ensemblistic approaches, to show firstly that accuracy in no way guarantees a good calibration, and secondly that a good calibration can be independent of a good distribution of scores. We show that the usual recalibration approaches often provide no guarantee on these two points. This work has been presented at various conferences (Mathematical Foundations of AI Day, EAJC’24 (European Actuarial Journal conference), WIM’24 (Workshop on Insurance Mathematics), UQAM).

- Fernandes Machado, A., Charpentier, A., Flachaire, E., Galllic, E. & Hu, F. (2024a) [From Uncertainty to Precision: Enhancing Binary Classifier Performance through Calibration](#). *arXiv:2402.07790*.
- Fernandes Machado, A., Charpentier, A., Flachaire, E., Galllic, E. & Hu, F. (2024b)

[Probabilistic Scores of Classifiers, Calibration is not Enough](#). *arXiv:2408.03421*.

• Generative models

We have studied problems very closely related to transfer learning, corresponding to the case where the distribution of the training data does not necessarily correspond to the target distribution (for example, if the population that has purchased a contract is not the one that will be the target when the contract is officially marketed). In particular, some regions are under-represented in the learning base, and the solution is to over-weight these populations. We have studied how to generate dummy data to compensate for the lack of representativeness of certain groups. This work has been presented in peer-reviewed conferences, International Conference on Neural Information Processing (ICONIP’24) or International Joint Conference on Neural Networks (IJCNN’24), and International Conference on Data Science and Advanced Analytics (IEEE-DSAA’24).

- Stocksieker, S., Pommeret, D. & Charpentier, A. (2024). [Generalized Oversampling for Learning from Imbalanced datasets and Associated Theory](#). *Transactions on Machine Learning Research (TMLR)*
- Stocksieker, S., Pommeret, D. & Charpentier, A. (2024). [Data Augmentation with Variational Autoencoder for Imbalanced Dataset](#). *Proceedings of the International Conference on Neural Information Processing (ICONIP)*

Perspectives

For next year, we’ll be focusing our attention on a number of issues

- dealing with fairness with unobserved sensitive attribute, in connection with the “Bayesian Improved First Name Surname Geocoding” (BIFSG) model,
- discussing interpretability and explainability of black box models, when explanatory variables are correlated, and analysis when probabilistic graphical models are given,
- investigating generative models in connection with Wasserstein flows and diffusion schrodinger Bridge.

Appendix

- [Newsletter #1](#) ↗, October 2023-March 2024
- [Newsletter #2](#) ↗, April-September 2024

People involved

Principal investigator (PI)

Arthur Charpentier [in](#)
 Professeur at UQAM
 (Université du Québec à Montréal)



Below, Olivier Côté [in](#) & Agathe Fernandes Machadeo [in](#) who started their PhD thesis last Fall.

Postdoctoral Fellows



François Hu [in](#), Arsene Brice Zotsa Ngoufack, [in](#) & Marouane Il-Idrissi [in](#)

Students (Interns)



Noé Bosc-Haddad [in](#), Florent Cruzet [in](#),
 Suzie Grondin [in](#), Julien Siharath [in](#), Cassan-
 dra Mussard [in](#), Ana Patrón Piñerez [in](#)

Invitations



Laurence Barry [in](#), Fei Huang [in](#), Ewen Gallic
[in](#) Stéphane Loisel [in](#), Mario Ghossoub [in](#) &
 Marie-Pier Côté [in](#),

PhD Students



On top, Philipp Ratz [in](#) & Samuel Stocksieker
[in](#), who defended their PhD thesis, in June.

Publications and recent work

Book

1. Charpentier, A. (2024) [Insurance, biases, discrimination and fairness](#). Springer. [in](#), ISBN: 978-3-031-49782-7.

Published in peer-review journals

2. Stocksieker, S., Pommeret, D. & Charpentier, A. (2024). [Boarding for ISS: Imbalanced Self-Supervised Discovery of a Scaled Autoencoder for Mixed Tabular](#)

Datasets. *Proceedings of the International Joint Conference on Neural Networks (IJCNN).*

3. Stocksieker, S., Pommeret, D. & Charpentier, A. (2024). [Data Augmentation with Variational Autoencoder for Imbalanced Dataset.](#) *Proceedings of the International Conference on Neural Information Processing (ICONIP)*
4. Charpentier, A. (2024) [Quantifying fairness and discrimination in predictive models.](#) *Machine Learning for Econometrics and Related Topics*, Songsak Sriboonchitta, Vladik Kreinovich, Woraphon Yamaka Eds, Springer.
5. Moriah, M. Vermet, F. & Charpentier, A. (2024) [Measuring and Mitigating Biases in Motor Insurance Pricing.](#) *European Actuarial Journal*, doi:10.1007/s13385-024-00390-8.
6. Stocksieker, S., Pommeret, D. & Charpentier, A. (2024). [Generalized Oversampling for Learning from Imbalanced datasets and Associated Theory.](#) *Transactions on Machine Learning Research (TMLR)*
7. Stocksieker, S., Pommeret, D. & Charpentier, A. (2024). [Delving into Deep Smoothed Bootstrap: Application in Imbalanced Regression.](#) *Proceedings of the International Conference on Data Science and Advanced Analytics (IEEE-DSAA'24)*
8. Hu, F., Ratz, P., & Charpentier, A. (2024). [A sequentially fair mechanism for multiple sensitive attributes.](#) *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 11, pp. 12502-12510).
9. Charpentier, A., Hu, F., & Ratz, P. (2023). [Mitigating discrimination in insurance with Wasserstein barycenters.](#) *3rd Workshop on Bias and Fairness in AI, International Workshop of ECML PKDD.*
10. Flachaire, E. & Charpentier, A. (2024) [Oaxaca-Blinder decomposition of changes in means and inequality: A simultaneous approach.](#) *Economics Bulletin.*
11. Barry, L. & Charpentier, A. (2023) [Melting contestation: insurance fairness and machine learning.](#) *Ethics and Information Technology.* doi:10.1007/s10676-023-09720-y
12. Bigot, R. & Charpentier, A. (2024) [Regards croisés dans le champ des assurances et de la responsabilité civile,](#) in *Genre, algorithmes et droit*, E. Bonifay and S. Sereno, Eds.

Chapters

Working papers

13. Niakh, F., Ratz, P., Charpentier, A. & Hillairet, C. (2024) [Risk sharing and Taxation.](#) *soon.*
14. Fernandes Machado, A., Charpentier, A., Flachaire, E., Gallic, E. & Hu, F. (2024) [Probabilistic Scores of Classifiers, Calibration is not Enough.](#) *arXiv:2408.03421.*
15. Hu, H., Charpentier, A., Ghossoub, M. & Schied, A. (2023) [The Multi-Armed Bandit Problem Under the Mean-Variance Setting.](#) *arXiv:2212.09192.*
16. Côté, O., Côté, M.P. & Charpentier, A. (2024) [A Fair price to pay: exploiting causal graphs for fairness in insurance.](#) *SSRN:4709243.*
17. Fernandes Machado, A., Charpentier, A., Flachaire, E., Gallic, E. & Hu, F. (2024) [From Uncertainty to Precision: Enhancing Binary Classifier Performance through Calibration.](#) *arXiv:2402.07790.*
18. Fernandes Machado, A., Charpentier, A., Flachaire, E., Gallic, E. & Hu, F. (2024) [Geospatial Disparities: A Case Study on Real Estate Prices in Paris.](#) *arXiv:2401.16197.*
19. Fernandes Machado, A., Charpentier, A., Flachaire, E., Gallic, E. & Hu, F. (2024) [Sequential Conditional Transport on Probabilistic Graphs for Interpretable Counterfactual Fairness.](#) *arXiv:2408.03425.*
20. Vamparys, X. & Charpentier, A. (2023) [Artificial Intelligence and Personalization of Insurance: Failure or Delayed Ignition?.](#) *Chaire PARI, WP 32.*
21. Hu, F., Ratz, P. & Charpentier, A. (2023) [Addressing Fairness and Explainability in Image Classification Using Optimal Transport.](#) *arXiv:2308.11090.*

22. Hu, F., Ratz, P. & Charpentier, A. (2023) Parametric Fairness with Statistical Guarantees. *arXiv:2310.20508*.

Dissemination

23. Charpentier, A. (2024) *Peut-on diversifier des risques extrêmes ? Risques*.
24. Charpentier, A. & Gallic, E. (2024) *Croissance, décroissance, de quoi parle-t-on ? Risques*.
25. Barry, L. & Charpentier, A. (2024) *Partage de données, à qui profite le crime?. Risques*.
26. Barry, L. & Charpentier, A. (2024) *Le Rapport Langrenoy : lutter contre le désengagement des assureurs dans la couverture des risques climatiques. Dalloz Actualités*.
27. Charpentier, A. (2023) *Est-il nécessaire (et utile) d'être en guerre contre tout ? Risques*.

Interviews

- (2024) *indispensable e controverso uso dell'intelligenza artificiale, Il Sole 24 Ore*.
- (2024) *pour une intelligence artificielle équitable en assurance, ULaval Actualités*.
- (2024) *Arthur Charpentier : IA, biais et éthique en assurance, Expertises Croisées, Nexialog*,
- (2024) *Les événements climatiques ont pesé sur les résultats des assureurs en 2023 Portail de l'Assurance*
- (2024) *Assurances : des collectivités désarmées face aux effets du dérèglement climatique Politis*
- (2024) *Arthur Charpentier pris au pied de La Lettre pour parler d'intelligence artificielle La Lettre de l'Assurance*

Datasets and codes

- [doi:10.57745/P0KHAG](https://doi.org/10.57745/P0KHAG) R package

Presentations

The work done in this research project has been presented in more than a two dozen conferences, and more than a dozen seminars, this first year. We also organized to workshop (each one on one day).

Conferences and Colloquiums

1. Workshop on Climate Change and Insurance (CCI 2024) in Vienna, Austria, September 2024
2. Conference on New Developments in Probability (CNDP), at CRM, Centre de Recherche Mathématiques de Montréal Canada, September 2024
3. *Mathematical Foundations of AI day*, DATAIA Institute and SCAI (Sorbonne Center for AI), Paris, September 2024,
4. Actuarial summer school, *Szkoła Nauk Aktuarialnych*, Warsaw, September 2024
5. 6th edition of the European Actuarial Journal Conference EAJC'24, in Lisboa, Portugal, September 2024,
6. *L'assurance au défi des ruptures, Cerisy Colloquium*(France), September 2024
7. Canadian Economic Association, *CEA*, in Toronto, Canada, June 2024,
8. *23e Congrès des Actuaires*, Paris, June 2024
9. *KU Leuven*, in Belgium, at the ACP (actuarial contact program), June 2024,
10. Centre sur l'intelligence² en gestion de systèmes complexes *CRI2GS*, in Montréal, June 2024,
11. Statistical Society of Canada Annual Meeting (*SSC'24*), May 2024,
12. *Workshop on Trustworthy Artificial Intelligence*, Montréal, May 2024
13. Scientific Discovery in the light of Artificial Intelligence *SDAI'24*, in El Ghazala, Tunisia, May 2024
14. Rencontres des Jeunes Statisticiens *RJS*, in Bordeaux
15. Journées de la Société Canadienne de Sciences Économiques *SCSE* , in Montréal, Canada, May 2024

16. [Workshop on Optimal Transport](#), Cargèse (IES), Corsica, in May 2024
17. Journées de Statistiques [JDS](#), Bordeaux, May 2024,
18. [Optimization Days](#), at HEC Montréal, May 2024
19. [Journée de la Recherche](#), UQAM, Montréal, February 2024
20. Workshop on decentralized insurance and risk sharing, Chicago, July 2024
21. Insurance Data Science Conference, [IDSC'24](#), in Stockholm, Sweden June 2024
22. Workshop on Insurance Mathematics, [WIM'24](#), March 2024,
23. 17th [Financial Risks International Forum](#) "Big Data & Algorithmic Finance", in Paris, March 2024
24. International Joint Conference on Neural Networks, [IJCNN'24](#), IEEE World Congress on Computational Intelligence (WCCI), in 横浜市 (Yokohama) Japan, July 2024
25. 38th Annual AAAI Conference on Artificial Intelligence, [AAAI'24](#), January 2024
26. Journée d'étude sur le blanchiment et la fraude Université de Nîmes, France, March 2024
27. Data Talk Generali, Paris, France, February 2024
28. Colloquium 'Enjeux actuels et futurs des sécheresses', ENS Ker Lann, France, December 2023
29. 3e Colloque International de l'Actuariat Francophone, Paris, France, October 2023
30. Actuarial seminar, [København Universiteit](#), Denmark,
31. [SCOR Foundation Webinar](#), online
32. Chaire Thélem/ ILB, [Université d'Orléans](#) seminar, France
33. Séminaire modélisation financière [Centre d'Économie de la Sorbonne](#), Paris, France
34. Actuarial Seminar [University of Melbourne](#), Australia,
35. Working Group in Risk, [ESSEC](#), Paris, France
36. Seminar [Intact Insurance](#), Montréal, Canada
37. ARC (Actuariat et Risques Contemporains) seminar, [Sorbonne Université](#), Paris, France
38. [Autorité des Marchés Financiers](#), Québec, Canada
39. Pricing Seminar, [Akur8](#), Paris, France
40. Sèminaire [StatQAM](#), UQAM, Montréal, Canada
41. Pricing Seminar, [TD Insurance](#), Montréal, Canada
42. [Data Science Institute](#), Montréal, Canada

Seminars

30. Probability Seminar, [CRM](#), Centre de Recherche Mathématiques de Montréal Canada
31. CIREQ Seminar 'Marcel-Dagenais', [Université de Montréal](#), Canada
32. Seminario de Matemáticas Aplicadas of [Quantil Research Group](#), in Bogotá, Columbia.
33. Summer Seminar in Statistics and Actuarial Science at [UQAM](#)
47. 2nd Workshop on [Fairness and Discrimination in Insurance](#), JEDA'24, in May 2024, at Université Laval, with Fei Huang (UNSW Sydney), David Schraub (Chicago Actuarial Association), Emmanuel Hamel (Autorité des marchés financiers), Laurence Barry (Chaire PARI), Agathe Fernandes Machado (UQAM), Mallika Bender (Casualty Actuarial Society), Christopher Cooney (TD Insurance) and Olivier Côté (Université Laval). Organized by Arthur Charpentier, with Marie-Pier Côté (Laval).
48. Workshop [Quantact Day](#), in May 2024, with Sébastien Jessup (Concordia), Nahid Sadr (Université de Sherbrooke), Benjamin Côté (Université Laval), Marie Michaelides (UQAM), Fabrizio Sabelli (Université de Montréal), Sébastien Legros (HEC). Organized by Agathe Fernandes Machado, with Maciej Augustyniak (Université de Montréal), Sébastien Jessup (Concordia), Fabrizio Sabelli (Université de Montréal),

Organization

André Orelieu Chuisseu Tchuisseu (Université Laval).

49. Workshop [Networks, Games and Risk](#), in December 2023, at UQAM, with Renaud Bourles (Centrale Marseille, Aix-Marseille School of Economics), Vincent Boucher (Université Laval), Federico Bobbio (Université de Montréal), Leonie Baumann (McGill University), Fallou Niakh (CREST-ENSAE, Institut Polytechnique de Paris) and Philipp Ratz (UQAM). Organized by Arthur Charpentier, with Mario Ghossoub (Waterloo).