Canadian Institute of Actuaries
Institut canadien des actuaires

# Bias and Fairness in Pricing and Underwriting of Property and Casualty (P&C) Risks

Date: April 2023
Document Number: 223056

# Table of contents

## Executive summary

The purpose of this paper is to provide P&C practitioners with an enhanced understanding of bias and fairness, the considerations that go along with these topics, and tools for detecting, evaluating and mitigating potential bias in actuarial risk assessment models. Its scope is to provide guidance for practitioners performing actuarial work related to underwriting property and casualty (P&C) risks, including but not limited to pricing and modelling. Substantial familiarity with P&C modelling and ratemaking approaches is assumed of the reader throughout the paper, and the intended audience is P&C pricing actuaries. However, the content of the paper is also applicable to other areas of actuarial practice.

Historical issues and examples relating to bias and fairness, such as racially biased correctional services risk assessments, contextualize the issues at hand.

Multiple definitions of bias are available. For the purposes of this paper, the concept is defined as a measurable property of predictive models, while fairness relates to the use of these models in practice to make decisions that have consequences for people, for example in terms of financial security and protection from risk. No single definition of fairness exists; like other ethical concepts, it is constantly being debated and adapted in a democracy. As one AI ethics scholar writes in an article on algorithmic fairness and actuarial practice, "fairness is dynamic and social and not a statistical issue."[1] This does not mean, however, that it cannot be a dimension of actuarial model evaluation. As the UK's Financial Conduct Authority notes, fairness is evaluated as it relates to the harm a decision can cause.

Measurement of bias is the first step in addressing any bias or fairness problems related to the P&C underwriting lifecycle. Bias can arise from multiple sources, including improper data collection, improper models and improper assumptions. All these sources should be considered. Many measurement approaches are possible, and these can then be used to test for the presence of bias in outcomes or results.

Two types of fairness are discussed in this report: procedural and distributive. Numerous techniques for evaluating fairness are available, including the Fairness Tree developed by the University of Chicago's Center for Data Science and Public Policy. Selecting a measurement of fairness is dictated by whether the objective of the measurement is individual or group fairness.

It is recommended that the practitioner set up an ethical framework to determine what constitutes fairness and to identify the appropriate steps to take to ensure pricing algorithms and models meet these standards. Once that framework is set up, it is also recommended that the practitioner document the following: the identification of vulnerable groups, the process used, the timing of the assessment, the measurement approach taken, the choice of thresholds, the use of validation procedures, the interventions undertaken or recommended, the methods used to monitor model performance and the implementation of record-keeping measures. Bias assessment could also form part of a broader internal model governance plan.

This paper is not intended to serve as a step-by-step guide outlining exactly how to address every possible fairness issue that may arise in a rating algorithm. The context of an individual situation is important to consider in practice, and there is no one-size-fits-all solution to ensure fairness in rating. Instead, the paper serves as starting point for the practitioner to better understand concepts of bias and fairness without attempting to be exhaustive on these topics. Accordingly, the recommendations throughout are not prescriptive or binding but rather provide suggestions on how to ensure fairness in

[1] Joi Ito, "Supposedly 'Fair' Algorithms Can Perpetuate Discrimination", *Wired*, February 5, 2019.

rating algorithms. They also serve to highlight some important considerations in this domain, with the goal of enhancing the practitioner's awareness of bias and fairness in actuarial work.

The concepts in this paper are closely linked with Section 1400 of the Standards of Practice.

**Acknowledgements**

This document was prepared by the Bias and Ethics Subcommittee of the Committee on Property and Casualty Insurance Pricing: Qi An (co-chair), Carl Lussier (co-chair), Joshua Snow, Erik Christianson, Alexandre Monette-Pagny and Alina Rogozhnikova.

**Feedback**

Comments on the document should be forwarded to practice@cia-ica.ca.

# Section 1: Intent, scope and cross-references

## 1.1 Intent

Much discussion over the past few years has concerned the potential for bias in pricing algorithms. Various papers have been written on bias and fairness in the fields of ratemaking and predictive modelling, especially in methods used by data scientists.

We remind the reader that an insurance premium is calculated based on a rate per unit of exposure.[2] The rate is determined based on several different rating variables, uses of which are governed by federal and provincial laws to ensure transparency and fairness.

As the volume of available data expands and rating algorithms become more complex, the P&C industry increasingly relies on automated processes, models and machine learning techniques to set premiums. Models and automated processes are fit on multiple data sets, which determine the premium based on a set of input characteristics. The fit is then further refined based on the accuracy of the models' predictions. This can lead to the emergence of unwanted bias in the practitioner's work product.

Although data-driven algorithms appear to rely on objective data, they still generally depend heavily on subjective decisions about how that data should be analyzed (e.g., what characteristics should be included, what categories observations should be sorted into, etc.). In addition, the frequency with which different groups or categories appear in the data can have significant unintended effects on the predicted outcomes. The article "Understanding Bias in Algorithmic Design" by ASME Demand highlights these issues with different examples, including one in which LinkedIn's search algorithm seemed to exhibit a preference for male over female names.[3] Though LinkedIn responded that the apparent gender bias was due to the search function's reliance on word frequency – since men are more likely to have common names than women – the result was nevertheless in an inadvertent potential professional advantage for men.

The intent of this paper is to give practitioners a better understanding of bias and fairness and to provide them with tools to evaluate these in the context of actuarial pricing and modelling.

We recognize that a full treatment of fairness is well beyond the scope of this paper, as fairness is a dynamic concept that is relative to the context in which it is being considered. For that reason, the paper

[2] Claudine Modlin, Willis Towers Watson and Geoff Werner, *Basic Ratemaking,* fifth ed., Casualty Actuarial Society, May 2016.

[3] A.R. Lange and Natasha Duarte, "Understanding Bias in Algorithmic Design," *ASME Demand*, Spring 2017.

focuses more heavily on bias than on fairness, limiting its scope to practical applications of the related concepts through quantitative measurements.

### 1.2 Scope

We aim to provide tools, guidance and considerations for actuaries to reflect on when performing various services, including but not limited to the following:

1. Development of risk segmentation or tiers.
2. Measurement of price differentials, discounts and surcharges.
3. Predictive analytics of other types to determine periodic cost level or growth potential.
4. Any other models for which the actuary deems these concepts to be applicable.

**This paper does not apply to the societal determination of notions of fairness.**

The document should be taken as a whole. While every section is important, the concepts referred to throughout should be considered holistically.

### 1.3 Cross-references

Where this paper refers to the provisions of other documents, the references include the documents as they may be amended or restated in the future, as well as any successors to them, whatever they are called. If any amended or restated document differs materially from the originally referenced document, the actuary should consider the guidance in this paper to the extent that it remains applicable and appropriate.

## Section 2: Historical issues and current evolution

Fairness can be defined as a social construct, the distinct manifestations of which can be assessed through various quantitative measurements. What is fair or unfair evolves over time and in various societal contexts, such that practitioners will need to adapt their interpretation of the concept to the circumstances of its application.

In traditional insurance work, issues related to societal and actuarial fairness are present on the spectrum of risk segmentation potential, i.e., from no segmentation linked to the risk to an extremely granular segmentation. This is because the segmentation process in insurance tries to recognize characteristics that are meant to differentiate the level of risk, or probability of outcome, represented by groups of individuals and to then set a price for coverage.

The processes involved in data collection and model development can introduce bias into automated decision-making and potentially perpetuate unfair outcomes. The following example was documented in *The Globe and Mail* (TGAM) as being inherently biased in the application of scoring models to predict inmate outcomes, insofar as it clearly discriminated against certain segments of the population. It is referenced here to illustrate a non-insurance instance of a biased outcome that is inherently unfair.

> In late 2020, TGAM published a series of articles called "Bias Behind Bars: A Globe Investigation Finds a Prison System Stacked Against Black and Indigenous Inmates,"[4] which referenced Correctional Service Canada's use of a risk assessment to evaluate the security classification of inmates in federal prison, assess their potential for reintegration and determine which programs would be offered to them.

[4] Tom Cardoso, "Bias Behind Bars: A Globe Investigation Finds a Prison System Stacked Against Black and Indigenous Inmates," *The Globe and Mail*, October 2020.

This risk assessment was comprised of several elements, purportedly providing an actuarial score that translated to an expected outcome.

TGAM found systemic bias in the derivation of the score, which led to more adverse outcomes for a portion of the inmate population. Indigenous and Black inmates were significantly more likely to receive a “maximum” security rating than their white counterparts, a classification that impacted their access to treatment programs and led to lower reintegration scores, which in turn increased the likelihood of negative parole decisions.

TGAM went on to note that the risk assessments not only led to harsher incarceration terms for racialized people but also appeared to be based on inaccurate data:

> Within the seven-year period for which TGAM obtained data, Indigenous and Black men are less likely than white men to commit a new offence, and our analysis suggests these scores often overestimate the likelihood they'll land back in prison. According to experts, they create a criminal feedback loop, with negative scores ratcheting up the long-term odds an inmate will go on to re-offend, ultimately compounding the gross overrepresentation of Indigenous and Black inmates in federal custody.

The above example clearly illustrates how the loop from data to score to outcome perpetuates systemic discrimination. There was (and, at the time of writing, still is) a reliance on a model that predicts an outcome not supported by the past seven years of data. TGAM's report was in support of an earlier set of findings by the Canadian Senate a few years earlier.[5]

The next example is from the insurance field. If we follow a strictly technical approach, one could argue that poorer communities should be charged significantly more for the provision of insurance, given the nature of the risk as measured from a purely mathematical, loss-cost basis.[6] The practice of redlining, whereby residents of “undesirable” neighbourhoods – often defined by or coinciding with their racial or ethnic makeup – were denied financial services, including insurance coverage, was based on this purely mathematical interpretation of fairness and accuracy. In the same way that offender risk scores in the above example create a self-fulfilling prophesy of increased recidivism, formal and informal redlining in the US and Canada gave rise to a feedback loop in which already disadvantaged communities were thrown into further disrepair and decline, which in turn reinforced the technical justification of the practice.

A second insurance-related example is the use of gender in rating algorithms. While price differentiation by gender is allowed in Canada under the Charter, this is not the case across all provincial legislation, and it was contentious up to the Supreme Court of Canada's (SCC's) judgment on the matter.[7] Even while the SCC ruled that insurers could continue to use gender as a risk-rating factor in certain contexts, this decision was not unanimous, with two justices in dissent. It is worth noting that across the Atlantic, the EU banned the use of gender in insurance pricing in 2012, citing it as inherently unfair.

We find the following section of the SCC's judgment in *Zurich Insurance Co. v. Ontario (Human Rights Commission)* to be contextually relevant even a few decades after it was originally written:

---

[5] Senate of Canada, *Study on the Human Rights of Federally Sentenced Persons*, Interim Report of the Standing Senate Committee on Human Rights, February 2019, https://sencanada.ca/content/sen/committee/421/RIDR/reports/RIDR_Report_Prisioners_e.pdf.
[6] Ito, supra note 1.
[7] Zurich Insurance Co. v. Ontario (Human Rights Commission), [1992] 2 S.C.R. 321. Available at https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/895/index.do.

> Human rights values cannot be overridden by business expediency alone. To allow "statistically supportable" discrimination would undermine the intent of human rights legislation which attempts to protect individuals from collective fault. It would also perpetuate traditional stereotypes with all of their invidious prejudices. Whether there was an alternative, which in all the circumstances was practicable, must be considered.

These examples are provided as illustrations of cases to which a concept of fairness applied and in which unintended consequences sometimes ensued. The last example is especially relevant, as the SCC provided rationale for its judgment and considerations (though with dissent) as to what fairness should look like in an insurance pricing context. As noted earlier, the concept of fairness is by nature a broad and dynamic one that lends itself to different interpretations and viewpoints. As such, it is a concept that evolves in the society in which it is being used.

Recent discussions of bias and fairness in insurance pricing out of the US include:

Study Points to Rate Bias in U.S. Auto Insurance Industry

NAIC – Innovation and Technology (EX) Task Force

NAIC – Race and Insurance

# Section 3: Definitions

This section outlines the key definitions used in this paper, namely "bias," "fairness" and "ethics."

The reader should be conscious that the definitions found below may be considered more normative than is usually found within the literature. Such is used to contrast the various considerations present when discussing these topics.

## 3.1 Bias

It is important to note that there are multiple methods of defining and measuring bias. As such, the practitioner should give due consideration to the question of which definition should be applied. In the interest of clarity, however, the working group considers the definition offered in the new federal privacy Bill C-27: *Artificial Intelligence and Data Act* to be a good basis on which to initiate the discussion:

> *Biased output* means content that is generated, or a decision, recommendation or prediction that is made, by an artificial intelligence system and that adversely differentiates, directly or indirectly and without justification, in relation to an individual on one or more of the prohibited grounds of discrimination set out in section 3 of the *Canadian Human Rights Act*, or on a combination of such prohibited grounds.[8]

While C-27 is intended to apply to those responsible for artificial intelligence systems – with its more extensive obligations saved for "high-impact" AI applications – its definition of bias highlights the broad requirements that any predictive model that classifies people into categories, such as risk classes, must meet. That is, if a model produces differential outcomes between groups, this unequal treatment must be justifiable by reference to some salient difference between them. In accordance with the *Canadian Human Rights Act*, the bill prohibits characteristics such as race or disability status from being used as inputs in such models, except in cases where they are intended to prevent or correct disadvantage based on or related to the protected grounds.

---

[8] Bill C-27, *An Act to Enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make Consequential and Related Amendments to Other Acts*, 1st sess., 44th Parliament, 2022.

For the purposes of the paper, we can thus define bias as follows:

> Bias in P&C pricing is any situation in which the outcomes of ratemaking models are systematically less favourable to individuals within a particular group and where there is no relevant difference between groups that justifies the difference in premiums or rates.

In other words, a biased outcome is one in which people or groups are assigned higher or lower premiums for reasons not justified by differences in the cost of providing insurance. In P&C pricing, this justification will usually come in the form of a statistical correlation between a variable used for risk classification, like a territorial signal, and underlying risk. More rarely, there may be an understanding of the underlying causal mechanism that explains how a characteristic triggers a process leading to heightened risk and cost, but this higher bar is often difficult to achieve in the P&C space. For this reason, using territorial ratemaking as an example: it is usually sufficient to show that a rating variable is predictive of loss or expense, and practitioners will typically correlate multiple sources of information (e.g., census data, road data, crime statistics, traffic density, weather, proximity measures, environmental statistics and industry experience) to obtain a correlated signal of the loss cost to the geographic unit.[9] In this case, while it is difficult to reach an exact understanding of why a particular geographic point is riskier than any other point on a map, the modelled outcome should nevertheless follow the loss experience or rationale be provided as to why a deviation is appropriate. This reasoning can be extended to non-geographic dimensions of ratemaking and underwriting.

Finally, it is important to distinguish between the above definition of bias and the statistical definition, the latter referring to a situation in which the expected value of an estimator differs from the true underlying value. In this document, the term "bias" does not refer to the statistical definition, which should be considered an unrelated concept.

The practitioner may consider any of the above definitions in their work but will likely need to select the one that is most appropriate to the work being undertaken.

## 3.2 Direct and indirect discrimination

By now, it should go without saying that insurers are precluded by human rights legalisation from using certain variables – such as race, disability status, sexual orientation and other characteristics associated with systemic discrimination or marginalization – for risk classification. More problematic, however, is the possibility that superficially neutral data may stand in for membership in a protected group, either intentionally or unintentionally, by capturing that status in a roundabout way. The challenge of indirect or proxy discrimination in ratemaking has become more pressing with the accelerating evolution and adoption of AI, and a growing body of research examines its ethical and policy implications. This paper does not weigh in on those important conversations or make recommendations aimed at regulators, policymakers or companies looking to develop internal policies. Rather, it offers practical tools and guidance to the P&C practitioner who wants to take a considered and actuarially appropriate approach to these issues in their day-to-day work.

To round the above discussion on bias, then, we consider the following concepts of discrimination to be relevant to the practitioner:

- A pricing model avoids direct discrimination if no discriminatory features (i.e., characteristics protected by human rights legislation such as the *Canadian Human Rights Act*) are used as rating factors.

[9] Pinnacle Actuarial Services, *Territory Rating Review Report*, special report for the Financial Services Regulatory Authority of Ontario, October 2022, 10.

- A pricing model avoids indirect discrimination if it avoids direct discrimination and, furthermore, the non-discriminatory features are used in such a way that they do not allow implicit inference of discriminatory features from them.[10]

Of note, it is the observation of the working group that the requirements of implicit inference listed above can be met while still obtaining differential outcomes between groups.

## 3.3 Fairness

No single definition or measure of fairness exists. As noted above, the practitioner should understand that "fairness is dynamic and social and not a statistical issue."[11]

When considering fairness, the practitioner should ask the following questions:[12]

- Who is harmed by the pricing bias or the potential for bias?
- How much are these individuals harmed?
- How significant is the pool of people harmed?
- Is the product/service essential?
- Does society view price discrimination as egregious/socially unfair?

In scenarios where bias could cause significant harm and unfairness based on the above considerations, the practitioner should exercise additional care. This is especially important where there is a risk of perpetuating existing biases or making protected groups more vulnerable to existing societal inequalities.

Fairness can typically be split into two different categories: procedural and distributive. For pricing purposes, practitioners should try to identify and consider these separately. Current insurance regulations tend to focus more on procedural fairness than distributive fairness.

- Procedural fairness refers to how insureds are treated throughout the pricing process. Decisions relating to how to handle missing values in the data or which variables to include in a pricing model would fall into this category.
- Distributive fairness relates to the distribution of pricing outcomes across various insureds. This type of fairness is not linked to the pricing process but rather to the results of the process and its impact on insureds.

## 3.4 Contrasting bias and fairness

It is important to note that bias and fairness are not the same. In the context of P&C pricing, bias does not necessarily imply unfairness; similarly, a lack of fairness does not necessarily imply bias.

Bias in pricing arises due to factors like the data used, the model parameters, the type of model chosen and practitioner assumptions. By contrast, fairness depends both on the outcome of a model and the context in which that outcome is applied. This includes many factors that are external to the model outcome and the interaction of these factors with said outcome.

Bias in this context will be a static concept. A process, model or outcome that is biased today will remain biased over time (unless it is corrected). However, due to the dynamic nature of fairness, an outcome that is fair today may not be seen as fair in the future.

[10] M. Lindholm, R. Richman, A. Tsanakas and M.V. Wüthrich, "Discrimination-Free Insurance Pricing," *ASTIN Bulletin: The Journal of the IAA* 52, v. 1 (2021): 55-89.
[11] Ito, supra note 1.
[12] Adapted from Mary Starks et al., "Price Discrimination in Financial Services," Financial Conduct Authority, July 2018.

## 3.5 Ethics

The Preamble to the CIA Rules of Professional Conduct states that the Rules "identify the professional and ethical standards with which a member must comply and thereby serve the public interest." It goes on to state that "in addition to these Rules, a member is subject to applicable law and rules of professional conduct or ethical standards that have been promulgated by a recognized actuarial organization for the jurisdictions in which the member renders professional services."

**Respect of the law:**

Practitioners have an unequivocal obligation to uphold the law. This includes protected class legislation that sets out characteristics that are not permitted to be used as variables for classification (e.g., the prohibited grounds of discrimination set out in section 15 of the *Canadian Charter of Rights and Freedoms*). The practitioner must consider the legal implications and the various frameworks, ethical or legal, that are in place across multiple jurisdictions to ensure that work outputs are within the applicable frame.

For example, discrimination based on social condition is prohibited in Quebec, age is a protected ground in New Brunswick, and auto insurers are not permitted to use credit information for auto insurance rating or underwriting in Ontario. However, the same piece of information may be legal to use in a different province or for a different line of products.

**The duty to respect the letter and intent of the law is the responsibility of the practitioner.**

We recommend that practitioners familiarize themselves with the CIA Rules of Professional Conduct and all applicable laws. The ethical principles discussed in Section 4 below serve to enhance one's understanding of how to apply these laws and ethical standards and do not substitute the existing legal and professional requirements.

# Section 4: Analysis and measurement of issues

## 4.1 Overview

Bias can be introduced, identified and measured at each step of the pricing process, from data collection and preparation to modelling, predictions and so on. Section 4.2 below outlines the tools the practitioner can use to detect the existence of potential bias and establish whether that bias may be deemed "unfair" according to the relevant legal, ethical and professional standards as discussed above. We recommend checking for bias at each point of controls, using relevant, often multiple, metrics that apply to the type of work being performed.

If the practitioner finds that potential bias (as defined in section 3.1) exists, they will need to consider the trade-off between accuracy and fairness (as defined in section 3.2) when deciding on a course of action to correct it. The focus of the last two sections (Ethical Framework and General Procedures) is to provide tools to help the practitioner evaluate notions of fairness and possible social costs.

## 4.2 Measurements

### 4.2.1 Measurement of bias

There are multiple sources from which bias in pricing can arise. Some examples include:

- Data generation processes, including contextual information about the circumstances or social conditions in which the data originate (i.e., data may reflect historical inequalities).
  - An example, described in more detail in Section 2, would be insurers' reluctance to provide coverage in certain neighbourhoods. This leads to a lack of maintenance to

properties, which then becomes embedded into the technical approach used by actuaries to determine loss cost, in effect perpetuating discrimination.

- Incomplete or unrepresentative data (e.g., data that does not cover the domain space).
  - An example would be collecting data only from those who volunteer to participate.
- Improper models or algorithms.
  - An example would be using linear models to try to explain nonlinear relationship between characteristics and observations.
- Improper assumptions.
  - An example would be an improper assumption about the ordering of categorical variables.

Practitioners should consider all the above sources of bias as well as any others they are aware of.

When evaluating a pricing model for bias, practitioners should examine the set of outputs that are produced by the model to check for anomalous results. The following measures can be considered to determine if bias is present in the pricing model:

- Differences of averages predicted values across classes.
- Ratios of averages predicted values across classes.
- Difference of accuracy, which is the number of corrected predictions relative to total number of predictions, across classes.
- Difference of error rate, which is the number of incorrect predictions relative to the total number of predictions, across classes.
- Disparate impact, which is the ratio of predictive values across classes.

In a 2022 research paper for the Casualty Actuarial Society (CAS), "Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance," Roosevelt Mosley and Radost Wenman classify different fairness measures into three categories: independence, separation and sufficiency.[13] Independence poses the strongest requirement and measures similarities of predictions across classes. Separation poses a weaker requirement by allowing us to condition on the same observed values before comparing predicted values across classes. Finally, sufficiency allows us to condition on the same predictive values before comparing observed values across classes. The paper lists examples of measures in each category and also illustrates them in a pricing context.

### 4.2.2 Purpose of measurement

In determining which methodology to use to measure bias, the practitioner may want to decide the type of fairness they are trying to achieve.

- Individual fairness ensures similar individuals are treated similarly.
- Group fairness ensures that certain statistical measures are equal across groups (i.e., classes).

If the objective is to achieve **individual fairness**, then measurements include various types of differences or ratios of averages between specific samples in different classes.

If the objective is to achieve **group fairness**, then measurements include differences, or ratios, of averages between the groups (i.e., classes) themselves.

As bias is considered in the context of ensuring that we are not reinforcing inequalities for certain groups, practitioners are encouraged to achieve group fairness. Note that the practitioner's view of the true

[13] Roosevelt Mosley and Radost Wenman, "Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance," Casualty Actuarial Society Research Paper Series on Race and Insurance Pricing, May 2022.

relationship between the groups can impact what they deem an appropriate measure of bias. For example, if the practitioner believes that all groups have similar abilities with respect to the task they are evaluated on, even if this cannot be observed, then demographic parity metrics at the group level should be used. If the practitioner believes that what they observe reflects the true ability of each group with respect to the task, then individual-level metrics, such as the average of differences in prediction for risks with the same observed values across groups, should be used.

Many open-source evaluation tools, such as the AI Fairness 360 framework developed by IBM, are available for practitioners and other researchers to share and assess algorithms for bias.

Another important consideration in determining which type of fairness to aim for is whether we are seeking a fair representation or fair outcome. A well-known decision-making tool is the Fairness Tree developed by the University of Chicago's Center for Data Science and Public Policy, which helps practitioners audit model outputs by posing key questions that direct them to different branches of the tree to identify corresponding evaluation metrics.

### 4.2.3 Other relevant considerations

The practitioner can refer to the following sections of the CIA's Standards of Practice (SOP) and the American Academy of Actuaries' Actuarial Standards of Practice (ASOP) for further discussion of bias and fairness:

- SOP 1440 – Sufficiency and reliability of the data
- SOP 1450 – Model appropriateness and limitations
- SOP 1460 – Quality assurance processes
- ASOP 12 – Risk classification
- ASOP 25 – Credibility

## 4.3 Ethical framework

Consider the following example:

> The measurement techniques in Section 4.2 have enabled the practitioner to identify a bias in a pricing model. The premiums produced by the model inadvertently favour individuals of one race over another, with the bias introduced through pricing differentiations based on the postal code rating factor.
>
> The practitioner determines that race was not used as a rating factor when the rates were set, and the geographical differences are due to differences in the frequency and severity of claims – thus accurately capturing the need for higher or lower premiums. However, even when the discrimination is unintended, race-based premiums are unfair based on the definition of fairness discussed in this paper.
>
> The practitioner has multiple courses of action available, from correcting the pricing bias immediately to doing nothing.

In situations like this, we recommend that the practitioner set up an ethical framework to evaluate the relevant notions of fairness and their possible implications in a consistent, verifiable and justifiable manner.

The ethical framework aims to enable the practitioner to gain a better understanding of how to apply ethical principles when determining what constitutes fairness and objectivity in the outcomes produced by pricing models or algorithms. Appendix 2 provides examples of key ethical theories, their corresponding ethical frameworks and how these vary when applied.

We note that the examples in Appendix 2 are not exhaustive, and that practitioners will need to identify the approach that is most suitable to deal with the issues unique to their work.

Below is a process that can be followed to apply an ethical framework.

**1. Define the ethical issue.**

Clearly formulate the ethical issue under consideration. Determine which aspects of the process or the result might be perceived as unfair and which areas require ethical judgment.

**2. Understand the environment in which the decision is being made.**

Be aware of the diversity inherent in the Canadian business and social environment, as well as the history of systemic inequality that still impacts marginalized groups on the basis of their race or ethnicity, gender or gender expression, sexual orientation, disability status, socio-economic class and other social identifiers.

Understand what it means to operate as an actuarial practitioner within Canada, including the requirements of the CIA's Code of Professional Conduct and other professional, regulatory and ethical expectations.

**3. Consider the parties involved.**

Identify who will be impacted by the decision and how. Determine if any groups should be focused on and assessed in more detail.

**4. Gather information.**

Collect all available and relevant information while assessing the reliability of its sources. Use this information to identify any potential gaps in knowledge that make some ethical approaches not feasible.

**5. Formulate actions and consider alternatives.**

Formulate what potential courses of action could be taken and consider the impact (e.g., cost, benefit, time taken, resources required, etc.). Determine if alternative actions could be taken and associated impacts and outcomes.

**6. Review.**

Consider the decision made before actions are taken:

- Consider whether the decision complies with the CIA Code of Professional Conduct.
- Consider your view of the decision.

Consider the decision made after actions are taken:

- Consider the consequences of the actions – intended and unintended.
- Consider your assessment of the outcome given its consequences and whether anything could have been done differently.

## Section 5: Next steps for the practitioner

The practitioner should consider the following next steps once bias has been measured or detected. The action taken will depend on the materiality of the bias in the context of fairness. These considerations can be noted within the internal model governance documentation with reference to the applicable SOP. Specific bias measurement and assessment documentation may be established to track identified biases, remediation action taken and future considerations or recommendations.

There may be no immediate solution to remediate a detected bias that is evaluated to be materially unfair. In this scenario, the practitioner is encouraged to develop a remediation or mitigation plan. This could include collecting additional data elements or data from additional sources. Alternately, it may require more substantial changes that could take time to implement.

As a starting point, the practitioner should consult areas of the SOP that are particularly relevant to the next steps, including 1440 Data, 1450 Models, 1460 Quality Assurance and 1490 Documentation. This list is not exhaustive and other sections of the SOP are also likely to be relevant.

The next section outlines the steps that practitioners can take to measure, document and remediate bias and resulting unfairness in a pricing process.

## 5.1 Bias and Fairness Assessment Documentation

### 5.1.1 Framework

The practitioner can set up and document a bias and fairness assessment to have objective guidance on how bias can be identified, measured and treated. It is recommended that the framework be consistent between various ratemaking models, provided it remains applicable. This will assist in promoting objectivity and comparability in the bias and fairness assessment process. This framework could be included as part of a model assurance framework as per Subsections 1490 Documentation and 1460 Quality Assurance of the SOP. Practitioners are advised to document the framework and validate it through peer review and internal sign-off processes before using it to assess the model bias.

### 5.1.2 Documentation considerations

The bias assessment framework could include the following sections, among others. Practitioners could also develop custom frameworks to deal with areas of concern specific to their pricing area.

**1. Identification of vulnerable groups**

The practitioner could consider which vulnerable or protected groups are likely to be impacted by any potential bias leading to unfairness in their ratemaking process. Vulnerable or protected groups can often be defined based on Canadian legislation. If the definition of the vulnerable group is not conventional – for example, if there are additional vulnerabilities applicable to the business but not specified by Canadian legislation – we recommend documenting the rationale for identifying the group as vulnerable. It should be noted that vulnerable groups for which a pricing outcome is biased may not be obvious or apparent a priori. Therefore, the bias measurement process may cast a wider net than just these identified groups.

It is suggested that the practitioner consult with a varied group of stakeholders to ensure that the approach is multidimensional enough to encompass a variety of points of view.

**2. Identification of process**

The practitioner could consider which areas of the ratemaking or modelling process should be subjected to bias scrutiny, referring to the definition of bias and potential sources of bias detailed above. Consideration of a bias causing material unfairness in a model's outcome could be used to focus this analysis on areas that are deemed the most critical to biases leading to these unfair outcomes.

The use of a transparent methodology and governance framework for machine learning or AI-based pricing algorithms will foster trust, ensuring that detection of any bias is done early to either mitigate or avoid it. It will thus help to reduce or prevent potential negative outcomes and their impact.

### 3. Timing of assessment

The practitioner can consider when the model or process should be tested for bias, given the data, assumptions and methodology chosen. We recommend bias assessments when a new model is built or enhanced, as well as regular reviews to ensure no bias is introduced through changes to the existing model's inputs or calibration.

### 4. Measurement

The practitioner should consider the most applicable method to measure the presence of bias in a ratemaking model or process. Numerous measurement techniques are proposed in section 4.2 above.

The rationale for choosing the measurement methodology should be documented.

### 5. Threshold

To maintain accuracy of the rates, it might not always be prudent or acceptable to eliminate all bias present in the model. The practitioner should consider the threshold at which the bias is considered materially unfair and thus in need of corrective action.

We recommend referring to the ethical framework in the sections above to determine the threshold at which differential treatment or pricing can be considered unfair. As noted, we recommend keeping the ethical framework used to determine this threshold consistent between models and clearly documenting the rationale for choosing it.

Finally, the threshold should be reviewed regularly to ensure that it is still acceptable in the context of current regulation, legislation, CIA standards and societal views on fairness.

### 6. Validation

We recommend that the practitioner subject their findings to peer review. A peer reviewer can evaluate the assessment process, methodology and conclusion as well as any ethical judgments involved.

We recommend seeking additional review or sign-off from a senior leader (CRO).

### 7. Interventions

We recommend establishing which corrective actions are possible within the bias assessment framework, as well as determining at which type and level (high, medium or low) of corrective action is required. For example, if bias is found but is below an internal threshold, it might be considered of low materiality or immaterial, with no corrective action needed. However, in the event that bias exceeds the internal threshold and is unfair, the practitioner will need to consider which actions they can take to reduce or remove it. We recommend determining how material the bias needs to be, and in which area, to trigger various corrective actions. Potential actions, among others, include:

- Correcting bias within one's own business function, if minor.
- Reporting the bias to a more senior business leader.
- Notifying the regulator.
- Notifying the impacted customer group.

To assist with the intervention process, we recommend documenting what action was proposed, what action was implemented, when the intervention took place, and when the bias is expected to be corrected.

**8. Monitoring**

It is recommended that corrective action is monitored to ensure that it has the desired effect. Regular review and reporting of biases will help ensure that interventions can be implemented to keep biases below acceptable thresholds.

**9. Record-keeping checklist**

The bias assessment process should be documented to ensure a record is available in the event that it is requested by interested parties, such as internal senior leadership, regulatory authorities or other actuarial practitioners. The documentation can include reference to the ongoing regular assessment of the process or model, even if bias was not detected.

Please refer to the attached spreadsheet for an example of a regular bias assessment checklist that can be included in the modelling documentation as evidence that bias assessment was conducted.

## 5.2 Reporting within model governance

Bias assessment could form part of broader internal model governance. Although the bias assessment documentation can be kept as a supporting file, it is recommended that the practitioner includes high-level information in the following sections of a model governance report.

- **Model bias**: If bias (as defined in this paper) has been identified in the data, assumptions or outputs, the practitioner should state that bias is present in the bias section of the model governance documentation and indicate how material it is.
- **Weaknesses and limitations**: If bias is present, any distortions in the output of the model can be included in the weaknesses and limitations section of the model governance documentation. If the bias is considered material, the practitioner can include the development plan to reduce or remove the bias within an adequate correction timeline.
- **Appropriateness checks**: The review of the known biases, as well as potential new biases, may form part of the appropriateness checks conducted on a model. The practitioner can include bias assessment documentation as evidence that the check was conducted.
- **Model validation**: Biases should be reviewed as a part of an overall model peer review and independent validation. The practitioner can include the bias assessment documentation as evidence of peer or independent validation review. If the practitioner is an independent validator, we recommend requesting the results of the bias assessment as defined in this paper.
- **Model changes**: If any changes or enhancements to a model are made, the practitioner should conduct a bias assessment to confirm that the changes did not introduce or exacerbate bias within the model.

# Section 6: Further research

Social and economic biases are encountered by actuarial professionals on regular basis, and consideration of the fairness of these biases should be at the forefront of actuarial work. This paper provides an overview of bias, fairness and related ethical concepts and serves as a starting point for identifying and mitigating social bias in P&C pricing. While the paper is intentionally broad, we recognize that the application of bias and fairness considerations is paramount in the machine learning domain. We recommend further research focused on machine learning and predictive modelling, including but not limited to:

- Expanding on sources of bias, particularly in machine learning outcomes.

- Expanding on tools to identify fairness in complex predictive models.
- Identifying approaches to using machine learning audits to identify social biases and unfairness.
- Exploring the application of penalty factors on unfairness in machine learning techniques.
- Examining ethical approaches to fairness and bias mitigation techniques.

Further research can be extended to other areas of actuarial pricing such as life, pensions and health.

# Section 7: Illustrative example

The practitioner wishes to check their property book premium rating algorithms for potential bias in pricing with respect to any vulnerable groups. The goal is to establish whether there is any bias in the pricing algorithm that may be perceived as unfair and if any corrective actions should be taken.

**Identification of vulnerable groups**

The practitioner considers the following categories of vulnerable groups, which may be present within the insured population:

- Elderly people
- People with disabilities
- Historically marginalized racial and ethnic populations, including Indigenous people
- Newcomers to Canada
- Low-income earners

**Identification of bias**

The practitioner considers the property pricing structure and rating factors that could vary by policy risk within the property book of business. While the majority of the rating factors were associated with the presence of risk-mitigating devices, property descriptions or policyholder choices, the practitioner identifies the following rating factors that may result in bias:

- "Age of insured" – impacts policyholders based on their age.
- "Type of property (home, condominium, vacation property)" – may impact newcomers and low-income earners, who might be more likely to live in condominiums.
- "Postal code" – may impact marginalized racial and ethnic populations, who might be more likely to live in certain neighbourhoods.
- "Number of people living in the property" – may impact newcomers and low-income earners, who might be more likely to live with relatives or roommates.
- "Insurance amount (building and contents)" – may impact low-income earners, who are more likely to be underinsured or improperly insured; may also impact people with disabilities, as low building limit to high content limit could indicate insurance of enhancements to accommodate a disability or mobility issue.

The practitioner examines the identified rating factors for potential bias.

**Measuring bias**

The practitioner applies the "differences of averages" bias measure. If an insurance company charges the policyholders premium rates proportionate to the risk they pose, the average loss ratios are expected be constant across risk cohorts. The practitioner calculates the average loss ratio for each risk group for each considered variable and notes that the loss ratios remained moderately uniform (within 0-5%) for all risk variables except "postal code."

For “postal code,” the average loss ratio produced volatile loss ratios fluctuating between 0-15%, depending on the region.

### Materiality threshold

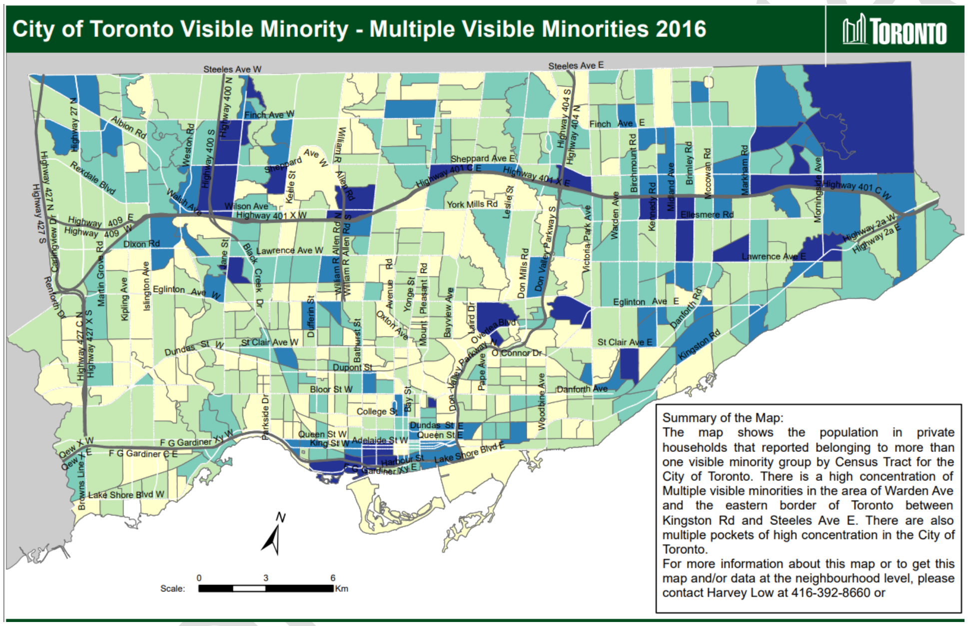
The practitioner consults internal materiality guidelines and determines that an error equivalent to 7% of the gross written premium is deemed to be material. The practitioner decides to use the same materiality threshold to determine whether the loss ratio differences are “material” and worth examining further.

In their examination of the postal code variable, the practitioner notices that for some Toronto and Greater Toronto Area (GTA) postal codes, the premium charged to the policyholders is materially higher than premium suggested by the risk experience in these areas.

The practitioner concludes that material bias is present in the postal code variable.

### Is there a fairness issue?

While “postal code” appears to proxy for racial and ethnic minorities, the practitioner wants to establish whether the vulnerable groups are in fact impacted. To this end, they consider alternative sources of data to determine the distribution of minorities in Toronto and the GTA, finding the map below.




Source: City of Toronto, Statistics Canada 2016 Census.

Overlaying the postal code profiles atop the map, the practitioner establishes that visible minorities are paying unusually high premiums relative to the claims paid. The practitioner concludes that this treatment may be viewed as unfair by the public, regulators or the CIA.

### Corrective actions

The practitioner investigates how the postal code–based rating factors were generated and establishes that they have not been updated since 2010. The claim experience and exposure profiles have changed dramatically over the last 12 years in certain geographical areas.

The practitioner raises this issue to the senior business leaders within the department, prepares a write-up and asks for peer review to validate their findings.

The practitioner identifies following potential corrective actions:

| Action | Advantages | Disadvantages |
|---|---|---|
| Update all postal code factors to be based on more recent claim experience and exposure. | • More accurate and fair premiums will be charged to all policyholders.<br>• In some cases, premiums charged to policyholders were too low. The correction will allow the insurer to collect more adequate premium revenue relative to risk exposure. | • Data collection and evaluation is expensive and time consuming.<br>• Based on average loss ratio analysis, it seems that the "postal code" factors were still appropriate, on average, for most of the policyholders. |
| Update only the postal code factors associated with vulnerable groups. | • More accurate and fair premiums will be charged to vulnerable groups.<br>• More focused data collection and analysis is less expensive and time consuming. | • Postal code factors will remain outdated for the majority of policyholders, resulting in a marginal loss of profitability for the insurer. |
| Do nothing. | • As 95% of policyholders will see little to no change in their premium, most of the customers are being charged fairly.<br>• No costs are incurred to correct the postal code factors. | • The mispricing for vulnerable groups, if known publicly, is likely to result in reputational damage to the insurer.<br>• Not updating the postal code factors soon could result in expansive and expensive pricing disparity over time. |

The practitioner's employer's business mandate is to be customer-centric and to treat all customers with fairness and integrity. The practitioner's pricing department holds that it is ethical to ensure fair premiums are charged to all customers.

After internal discussions within the practitioner's pricing department, it is concluded that the postal code factors associated with vulnerable groups will be corrected manually with immediate effect. The full updated of all postal code factors will be conducted and implemented in due time.

**Monitoring**

After the changes are implemented and the postal code factors updated, the changes to postal code and other rating factors are regularly monitored to ensure the factors used are based on the most recent and credible claims experience and exposure data.

**Record keeping**

The analysis conducted by the practitioner and the corrective actions implemented are documented and kept by the practitioner's pricing department.

# Appendix 1: Various definitions we rely on

**Measuring bias**

Some rating factors, such as age, gender and marital status, are outlawed in some provinces in Canada and cannot be used directly for pricing. Other socioeconomic rating factors, such as ethnicity, race and religion, are banned everywhere, not just in certain provinces.

If we consider the loss ratio for a group of policyholders, and the loss ratio split by "undesirable" rating factor, the loss ratio is expected to be very similar if no bias in the data is present. Alternatively, it could indicate that bias is being introduced by an unintended proxy.

The statistical significance of the loss ratio differences could be tested via p-value.

**Data science ethics**

The UK's Institute and Faculty of Actuaries and the Royal Statistical Society have produced *A Guide for Ethical Data Science*.[14] Its five themes are:

1. Seek to enhance the value of data science for society.
2. Avoid harm.
3. Apply and maintain professional competence.
4. Seek to preserve or increase trustworthiness.
5. Maintain accountability and oversight.

The concepts contained in this paper can also be applied to ethical considerations that actuaries should be aware of when performing pricing work.



[14] "A Guide for Ethical Data Science," Royal Statistical Society and Institute and Faculty of Actuaries, 2019.

# Appendix 2: Examples of ethical frameworks

The main normative ethical theories are outlined below.

**Utilitarianism**

The utilitarian approach is concerned with the ethical consequences of a particular action. The basic principle of utilitarianism is that "the most ethical choice is the one that will produce the greatest good for the greatest number."[15]

The utilitarian approach is the most common one for making ethical decisions with consequences that concern a large number of individuals. The practitioner should weigh all potential actions and pursue the ones that provide the most benefit and least harm to the majority.

**Deontology**

The deontological approach focuses on the principles guiding the individual who is making a decision or performing a particular action. This approach "requires that people follow the rules and do their duty."[16]

Under this approach, the decisions made by the practitioner should conform to the contractual obligations with their employer, CIA's Code of Conduct, regulatory requirements, and laws.

**Virtue ethics**

The virtue ethics approach focuses on the moral standing of the individual rather than the morality of a single action. The virtue approach assumes that "by practicing being honest, brave, just, generous, and so on, a person develops an honourable and moral character."[17]

The approach assumes that actions taken are ethical if these are of a person with high moral standing. The practitioners with a good professional reputation and operating under an employer with a focus on customer-centric values and principles are expected to make a morally correct decision in ethical uncertainty.

**Three ethical frameworks**

The normative theories discussed above can be used to set up a distinct framework for ethical decision-making. Each of the three frameworks are useful for making decisions; however, they each have associated limitations the practitioner needs to be aware of. The frameworks are summarized in the table below:



[15] "Utilitarianism," *Ethics Unwrapped*, McCombs School of Business, University of Texas at Austin.
[16] "Deontology," *Ethics Unwrapped*.
[17] "Virtue Ethics," *Ethics Unwrapped*.

| | Utilitarian framework | Deontological framework | Virtue framework |
|---|---|---|---|
| **Focus** | Future effects of actions<br>Who would be impacted directly and indirectly | Current and future obligations | Character traits that motivate the decision-making process |
| **Action** | Seeks best consequences | Seeks to comply with requirements | Seeks to develop own character |
| **Considerations** | Goals and objectives<br>Cost-benefit analysis | Contractual obligation<br>Legislation and regulation<br>CIA Code of Conduct | Personal values<br>Personal and professional reputation<br>The character of the business |
| **Limitations** | Might not be suitable when the decision relates to minorities.<br>Often does not align with ordinary moral decisions or conform to moral appraisal.<br>Not enough resources/time/information to assess all courses of action. | No consideration given to the results of actions.<br>Multiple actions are not ranked in order of merit. | Assumes decision-maker already is of sound character – it might not be the case in practice<br>Decision is subjective, as it can vary based on the decision-maker's virtue |

Source: Adapted from Sally Haslanger, "Three Moral Theories," (lecture notes), MIT Open Courseware, 2017, and Paul Lewis and Johann le Roux, "The Ethics of Claims Assessment Practices in the South African Life Insurance Industry" (PowerPoint presentation, 2019 Convention of the Actuarial Society of South Africa, October 22-23, 2019).

Returning to the example in the discussion box in Section 4.3, the practitioner can apply the above ethical frameworks in their decision-making process as follows:

1. The practitioner might choose to do nothing.

- **Justification based on the utilitarian approach**: For the majority of insureds, the pricing was accurate. Corrective action would improve the rates of the affected minority but would result in higher rates for most of the other groups. There is also the risk of premiums being too high for better risks (regardless of race). Thus the potential adverse impacts on the business outweigh the potential benefits of doing the correction.
- **Justification based on the deontological approach**: The regulations state that insurers are not allowed to use race as a rating factor. The ratemaking model in this case does not do so; hence the insurer complies with the rule of law and may choose to do nothing.

2. The practitioner might choose to introduce the corrective action immediately.

- **Justification based on the utilitarian approach**: Upon more detailed investigation, it is found that excluding the postal code factor and pricing instead based on region would reduce accuracy but result in better rates for the preferred class of policyholders (who have better driving records regardless of race). The correction would result in an investment of cost and time but would improve profitability over the long run; hence it is fully justifiable.

- **Justification based on the deontological approach**: The regulations clearly state that “race” cannot be used as a rating factor even by proxy. The insurer needs to make the correction to comply with legal and regulatory requirements.
- **Justification based on the virtue approach**: Even if there were no laws prohibiting the use of race in pricing and correction to the ratemaking models would be costly and reduce accuracy, it is morally wrong to differentiate pricing based on race and to perpetuate racial inequality in society.

# Appendix 3: Generic definitions of bias

Generic, basic and detailed definitions of bias are included here for reference. Definitions 1 to 3 are basic definitions, while 4 and 5 are more detailed. Definition 4 is a definition of bias in the context of data science.

1. "A tendency (either known or unknown) to prefer one thing over another that prevents objectivity, that influences understanding or outcomes in some way."[18]

2. "The action of supporting or opposing a particular person or thing in an unfair way, because of allowing personal opinions to influence your judgment."[19]

3. "An inclination or prejudice for or against one person or group, especially in a way considered to be unfair."

4. "Bias has several definitions, and its common usage is decidedly negative. We typically use it to mean systematic favouritism of a group. Generally speaking, 'bias' is derived from the ancient Greek word that describes an oblique line (i.e., a deviation from the horizontal). In Data Science, bias is a deviation from expectation in the data. More fundamentally, bias refers to an error in the data. But, the error is often subtle or goes unnoticed."[20]

5. "… 'bias,' a term that we define broadly as it relates to outcomes that are systematically less favourable to individuals within a particular group and where there is no relevant difference between groups that justifies such harms."[21]



[18] *Open Education Sociological Dictionary*, "bias (*n.*)."
[19] *Cambridge Dictionary*, "bias (*n.*)."
[20] Will Goodrum, "Statistical and Cognitive Biases in Data Science: What Is Bias?" *Elder Research* (blog), October 6, 2017.
[21] Nicol Turner Lee, Paul Resnick and Genie Barton, "Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms," *Brookings*, May 22, 2019.

# Cited References


*Bill C-27: An Act to Enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make Consequential and Related Amendments to Other Acts*, 1st sess., 44th Parliament, 2022.

Cardoso, Tom. "Bias Behind Bars: A Globe Investigation Finds a Prison System Stacked Against Black and Indigenous Inmates." *The Globe and Mail*. October 2020.

"Deontology." *Ethics Unwrapped*. McCombs School of Business, University of Texas at Austin.

Goodrum, Will. "Statistical and Cognitive Biases in Data Science: What Is Bias?" *Elder Research* (blog). October 6, 2017.

Haslanger, Sally. "Three Moral Theories." Lecture notes. MIT Open Courseware. 2017.

Ito, J. "Supposedly 'Fair' Algorithms Can Perpetuate Discrimination." *Wired*. February 5, 2019.

Lange, A.R. and Natasha Duarte. "Understanding Bias in Algorithmic Design." *ASME Demand*. Spring 2017.

Lewis, Paul and Johann le Roux. "The Ethics of Claims Assessment Practices in the South African Life Insurance Industry." PowerPoint presentation. 2019 Convention of the Actuarial Society of South Africa. October 22-23, 2019.

Lindholm, M., R. Richman, A. Tsanakas and M.V. Wüthrich, "Discrimination-Free Insurance Pricing." *ASTIN Bulletin: The Journal of the IAA* 52, v. 1 (2021).

Modlin, Claudine, Willis Towers Watson and Geoff Werner. *Basic Ratemaking,* fifth ed. Casualty Actuarial Society, May 2016.

Mosley, Roosevelt and Radost Wenman. "Methods for Quantifying Discriminatory Effects on Protected Classes in Insurance." Casualty Actuarial Society Research Paper Series on Race and Insurance Pricing. May 2022.

Pinnacle Actuarial Services. *Territory Rating Review Report*. Special report for the Financial Services Regulatory Authority of Ontario. October 2022, 10.

Royal Statistical Society and Institute and Faculty of Actuaries, *A Guide for Ethical Data Science*, 2019.

Senate of Canada, *Study on the Human Rights of Federally Sentenced Persons*, Interim Report of the Standing Senate Committee on Human Rights, February 2019.

Starks, Mary et al., "Price Discrimination in Financial Services," Financial Conduct Authority, July 2018.

Turner Lee, Nicol, Paul Resnick and Genie Barton. "Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms." *Brookings*. May 22, 2019.

"Utilitarianism." *Ethics Unwrapped*. McCombs School of Business, University of Texas at Austin.

"Virtue Ethics," *Ethics Unwrapped*. McCombs School of Business, University of Texas at Austin.

Zurich Insurance Co. v. Ontario (Human Rights Commission). [1992] 2 S.C.R. 321.

Canadian Institute of Actuaries
360 Albert Street, Suite 1740
Ottawa, ON K1R 7X7
613-236-8196
head.office@cia-ica.ca
cia-ica.ca
seeingbeyondrisk.ca