

# The Discriminating (Pricing) Actuary

**Edward W. (Jed) Frees**    *University of Wisconsin - Madison, Australian National University*  
**Fei Huang**    *University of New South Wales*

*Date of this Draft:* March 16, 2021

---

*Abstract:* The insurance industry is built on risk classification, grouping insureds into homogeneous classes. Through actions such as underwriting, pricing and so forth, it differentiates, or discriminates, among insureds. Actuaries have responsibility for pricing insurance risk transfers and are intimately involved in other aspects of company actions and so have a keen interest in whether or not discrimination is appropriate from both company and societal viewpoints. This paper reviews social and economic principles that can be used to assess the appropriateness of insurance discrimination. Discrimination issues vary by the line of insurance business and by the country and legal jurisdiction. This paper examines social and economic principles from the vantage of a specific line of business and jurisdiction; these vantage points provide insights into principles. To sharpen understanding of the social and economic principles, this paper also describes discrimination considerations for prohibitions based on diagnosis of COVID-19, the pandemic that swept the globe in 2020.

Insurance discrimination issues have been an important topic for the insurance industry for decades and are evolving in part due to insurers' extensive use of *Big Data*, that is, the increasing capacity and computational abilities of computers, availability of new and innovative sources of data, and advanced algorithms that can detect patterns in insurance activities that were previously unknown. On the one hand, the fundamental issues of insurance discrimination have not changed with Big Data; one can think of credit-based insurance scoring and price optimization as simply forerunners of this movement. On the other hand, issues regarding privacy and use of algorithmic proxies take on increased importance as insurers' extensive use of data and computational abilities evolve.

*Keywords:* Actuarial fairness; disparate impact; proxy discrimination; unisex classification; credit-based insurance scores; price optimization; genetic testing; big data; COVID-19

---

## 1 Introduction

Discrimination is a topic that touches on the daily lives of almost every person. The word "discrimination" generally has negative connotations; people think of discrimination as occurring when we go into a store and are treated differently because of our hair color, when we apply for a job that we are well qualified except for our age, when we are shunned from a social group because of our ethnicity, heritage, or religious background. From this perspective, discrimination is endemic in our society and a topic that well deserves serious consideration.

This article focuses on discrimination in an insurance context. Insurance is particularly interesting because the entire industry is based on discrimination. Here, we use the word “discrimination” in an entirely neutral way, taking it to mean the act of treating different groups differently – where the groups are distinguished by salient features such as hair color, age, gender, heritage, religion, and so forth – whether such discrimination is justifiable or not. For example, auto insurers often charge younger (presumably riskier) drivers more than older (presumably safer) drivers, but do not make a distinction between brown-haired and red-haired drivers (presumably because the two groups are equally risky). So, discrimination based on age is done routinely, whereas discrimination based on hair color is not. In subsequent sections, we present different arguments about whether insurance discrimination is ethical or is “unfair” and morally indefensible in some sense.

### 1.1 *How Insurers Discriminate*

Insurers collect information on current and potential customers. They collect information about the customers themselves, the entity being insured (whether a person, organization, or physical object such as an auto or home), where the entity is located (that can vary, such as a person or auto), and parameters about the contract desired, among other things. This information, represented as variables or factors, provide the basis that insurers use to form groups and make decisions. And by treating groups differently, they *discriminate* among them. Before describing the process of forming these groups and weighing in on whether or not the use of a specific variable is ethical, we first describe the set of actions that an insurer might take. In short, how do insurers discriminate among customers?

**Issuance, renewal, or cancellation.** The first stage is the decision to insure. We follow the structure of Avraham (2018) who notes that the harshest form of discrimination is the decision on whether or not to issue a policy because of some characteristic, such as the applicant’s religion or ethnicity. This may be at the underwriting stage or even earlier, at the marketing stage. For example, Section 2.3 describes classical issues of so-called *redlining* where insurers simply did not enter geographic districts with high concentrations of African-Americans (who were assumed to be high risks). A similar type of discrimination happens when insurers refuse to renew or when they cancel policies based on some characteristic. Indeed, some jurisdictions have statutes that limit or prohibit the use of a particular characteristic in either issuance, renewal, or cancellation of an insurance contract.

**Coverage.** Another form of discrimination involves restricting coverage in ways that might harm disadvantaged groups. For example, insurance companies might limit disability insurance coverage for people with disabilities that stem from having a human immunodeficiency virus (HIV).

**Pricing.** Even without limitations on issuance and coverage, insurance companies can still distinguish among insureds by simply charging different premiums. For example, an insurer may admit people with various diseases and disabilities into their pool, yet charge these people a higher premium. Is this fair or unfair? Some argue that a specific disease, e.g., cancer, is no fault of the individual and so they should not bear the additional burden of higher premiums. A counter position is that, by admitting high-cost individuals to the pool, this raises the costs for all in the insurer’s pool (even those without the specific disease).

## 1.2 Insurance Prohibitions

The modern-day insurance industry is founded on the ability to differentiate, or discriminate, among risks, known as *risk classification*. Thinking of an insurer as a private company, there are strong economic arguments for permitting insurers to discriminate among risks; Section 3 presents economic reasoning for this permission. There is evidence, consistent across lines of business and international jurisdictions, that insurance policyholders believe that some insurance discrimination is fair, c.f., Schmeiser, Störmer, and Wagner (2014). However, there are also instances where consumers are concerned with “unfair” discrimination. When insurance is mandatory or nearly so, it becomes less of an economic commodity and more of a social good, resulting in different attitudes towards “fairness.” Section 2 describes social justice considerations that underpin notions of fairness. Because of these concerns, regulators and policymakers in many jurisdictions impose restrictions on insurers abilities to discriminate.

Disagreements among stakeholders about what constitutes fairness can arise depending on whether one views insurance as an economic commodity or as a social good. Employees of insurers often think of insurance as an economic quantity and argue that risk classification is morally appropriate, c.f., concepts of *actuarial fairness* in Section 2. Consumer advocates focus on concepts of solidarity and cross-subsidization of risks, thinking of insurance as a social good, or a basic human right. These views depend markedly upon the jurisdiction and line of insurance business, as discussed in Sections 4.2 and 5, respectively.

Insurance discrimination, including the prohibitions from regulators that insurers might face and how these prohibitions affect society, has long been an important topic. Issues of fairness have not been resolved nor are they becoming less relevant; its prominence has lately increased with the increasing importance of “Big Data” where massive data sets and increasing computing power has become commonplace, c.f., Thouvenin et al. (2019). Like other major industries, insurers have at their disposal increasing amounts of information available about current and potential customers. As emphasized in Sections 5.5 and 6.4, more information gives insurers opportunities to differentiate among potential clients at

increasingly granular levels; many are doing so because of real or perceived pressures from competition. However, not all information that could be used should be used - we seek to provide a framework for actuaries and other analysts to think about the use of potentially sensitive information.

## 2 Social Justice Considerations

### 2.1 What is Actuarial Fairness?

To understand what is unfair discrimination and what insurers should be prohibited from doing, let us start with a discussion of “fair” insurance mechanisms. We begin with a historical context, drawing on the work of Frezal and Barry (2019).

#### 2.1.1 Pooling and Solidarity

In the early seventeenth century, mathematicians used games of chance such as rolling dice and drawing lotteries to develop a theory of randomness. For example, the fair price for each participant in a lottery was determined by the sum of wagers divided by the number of participants. Fairness was thus understood within the framework of individual equitable contracts, ones that traded a certain present amount for an uncertain future value. Even though the ex-post results of the game had winners and losers, because the ex-ante probabilities were equal, such games could be accepted as fair.

Much later, in the nineteenth century, a similar mathematical model was used for insurance pricing. From an individual’s perspective, one could again trade a certain present amount, the premium, for an uncertain future value that would provide compensation for an insured loss. For an insurance pool, the sum of future values was no longer certain (unlike the lottery size) but, with the additional regularity in the sum of future value of losses, there were real benefits in pooling of risks. In the modern language of probability, although the amount that an individual might expect to lose for an insured loss remained unchanged, the amount of uncertainty was vastly reduced.

Insurance pooling offered a novel method for coping with the uncertainty of losses. Before this, the only method of coping with potential loss events was individual prudence. With an insurance pool, losses now became the responsibility of the pool. From a moral perspective, the responsibility for the accident could now be thought of as not due to the behavior of the faulty individual but rather attributed to the collective; in this sense, pooling *socializes responsibility*, c.f., Baker (2002).

As further discussed in Lehtonen and Liukko (2011), pooling creates a sense of shared responsibility among a group of people. This, combined with a certain understanding of

equality and justice, creates a type of *insurance solidarity*. It is not the same type of solidarity that one thinks of in political movements which embody a conscious identification with the group, emotional bonds, shared values and beliefs, and so forth. It is solidarity that emphasizes mutual responsibility, reciprocity, and a particular shared understanding of fairness. Insurance provides a mechanism to transfer an individual's uncertainty to a pool; to achieve this agreement, the individual must have faith in the pool.

### 2.1.2 *Responsibility and Actuarial Fairness*

With the shift of responsibility from the individual to the pool, one can imagine that the sense of fairness shifts and depends upon the nature of the pool. If the pool is formed from a small group of like-minded individuals (think of the classical case of a group of farmers forming a collective to restore a member's barn in the event of fire), then there will be little difference between notions of fairness for the individual and for the pool collective. However, modern day insurance is generally sponsored at large levels, either by governments or private corporations, the latter of which can be owned by policyholders (mutual or takaful companies) or by investors (stock companies). Members of the pool may feel a type of (insurance) solidarity but the responsibilities of the pool depend on its nature.

**Stock Insurance Company.** At one end of the spectrum is the case of the pool of contracts issued to individuals by a for-profit stock company. Here, the pool can be thought of as a sum of bilateral contracts that leaves out the collective dimension of insurance. Actuarially fair pricing is based on the expected value of the uncertain event at stake, taken to be the risk transferred from the insured to the insurer. In this context, fairness means that each customer should pay for their own risk and only their own risk. As will be discussed in Section 3, there is ample basis for this position from economic theory.

**Government.** At the other end of the spectrum is the case when the pool is owned by a government entity where such contracts constitute *social insurance*. Subsidies, from one group to another, are common in social insurance. Governments regularly engage in social policy such as the redistribution of risk or income. The use of insurance to subsidize the underprivileged is consistent with what many view as a government's core mission. In social insurance, there can be a large variation in how strictly the principle of actuarial fairness is followed.

**Group Insurance.** Between these two ends of the spectrum, there is substantial variation in principles of fairness depending on who owns the pool and the nature of the contractual arrangement. For example, consider a disability income contract issued to a large group, such as a university. Because the employer (university) pays all or a major portion, premiums rated by risk factors are not a major issue (unlike the individual market). In

general, in group insurance, the amount of socialization is greater than one would find in the corresponding individual market.

As described in Baker (2002), most people do not think that their premiums will go to pay other people's claims. Instead, they think of it as a type of savings account and often expect that over the course of a lifetime the deposits made by each person should roughly equal the withdrawals on that person's insurance account. Thinking of premiums as going to pay others claims emphasizes the social aspect of insurance. That is, the losses, understood as belonging to the collective, are also borne collectively (Lehtonen and Liukko (2011), p. 35).

**Mutual Company.** In a mutual insurance company, the policyholders are both customers and owners. Unlike stock companies, ownership rights of the mutual policyholders are not transferable. By eliminating stockholders with their separate and sometimes conflicting interests, potential conflicts between owners and customers over dividend, financing, and investment policies are internalized. This is the major benefit of the mutual form of organization. As the owners of the pool are the policyholder themselves, this suggests that the amounts of cross-subsidies among groups or socialization would be greater than in an organization with a for-profit motive. However, in point of practice, mutuals compete with stock companies and so many of their practices are indistinguishable from stock companies. A small academic literature examines differences between stocks and mutuals, c.f., Braun, Schmeiser, and Rymaszewski (2015). In part due to the policyholder's owner stake in the company, evidence from this literature suggests that policies offered by stock insurers are overpriced relative to policies of mutuals. Nonetheless, we know of no study that has confirmed nor disproved the conjecture that mutuals discriminate differently than stock insurers.

**Takaful.** As another example, consider modern takaful companies, c.f., Maysami and Kwon (1999). In many senses there are similarities between takaful and mutual companies. Classical Western insurance appears to violate the Islamic prohibition of gambling (as well as the Islamic prohibition of usury). Instead takaful insurance offers, not as a bilateral contract, a transfer of a known risk to a collective enterprise by which Muslims pool resources to aid one other in the event of loss. Responsibility of the loss shifts from the individual to the collective and so aspects of fairness shift.

### *2.1.3 Insurance as a Social Good*

Attitudes towards fairness also depend upon whether an insurance product can be viewed as a social or type of public good. A social good is something that benefits the general public such as clean air, clean water, and literacy. One characteristic of a public good is that is *non-excludable*, that is, it cannot be provided unless others can also enjoy it. For example,

if you erect a dam to stop flooding - you protect everyone in the area (whether or not they contributed to erecting the dam). Thus, viewing an insurance product as public good would argue against excluding members of society.

For example, in many countries health insurance is likely to be seen as a social good where access to a certain level of healthcare is guaranteed for all. This is even true in a country like the U.S. which generally has taken longer to improve access to healthcare than other countries. In contrast, life insurance is more often seen as a private (non-public) economic commodity. Life insurance can enhance the financial security of the family of a policyholder but is voluntary and is not viewed as a necessity. Other insurance lines, such as long-term care and disability insurance arguably fall somewhere in the spectrum between social and economic private goods, c.f., Prince (2019).

If an insurance product is thought to be a social good, a related question is whether members of the public have equal access to the product. Specifically, the issue is whether there is an impact which puts members of a select “protected” group at a disproportionate disadvantage compared with members of a similar group. Issues of such *disparate impact* are difficult for individual insurers to address but are important for public acceptance of the insurance marketplace, c.f., Miller (2009). Regulators have been grappling with the question of whether laws, which prohibit discrimination based on race, religion, or national origin, could or should cover instances of disparate impact on underserved or protected classes of consumers emanating from the use of predictive modeling and analytics.

## 2.2 Characteristics of Sensitive Variables

Grouping, or classifying, insureds into homogeneous categories for the purposes of risk sharing is at the heart of the insurance function. Many variables that insurers use are seemingly innocuous (e.g., blindness for auto insurance) yet others can be viewed as “wrong” (e.g., religious affiliation), “unfair” (e.g., onset of cancer for health insurance), “sensitive” (e.g., marital status), or “mysterious” (e.g., AI produced). When regulators and policymakers decide that it is not permitted to use a variable for classification, it is thought of as creating a **protected** class. By and large, the choice of whether a variable should be used for insurance purposes is a normative one. Although actuaries and other financial analysts determine insurance premiums from the available information, the choice of which variables to use is a societal one in which many actors participate.

Nonetheless, it is helpful to understand what variable attributes influence society’s assessment of whether it is fair for insurance purposes. When identifying whether or not a variable contains sensitive information, we use a structure drawn from Avraham (2018) and Prince and Schwarcz (2020).

- **Control.** If an insured has control over an attribute, it is generally deemed to be an acceptable variable to be used for insurance purposes. For an example from auto insurance, consider a variable that indicates whether the car is high performance, capable of going at fast speeds and expensive to replace in the event of an accident. An insured chooses whether or not to purchase a high performance vehicle and so vehicle type is generally deemed to be an acceptable variable. In contrast, race and sex at birth are examples of characteristics over which insureds have no control. Naturally, questions of degree of choice enter; for example, smoking in life insurance is generally now accepted as a rating factor whereas religious affiliation is generally prohibited.
- **Mutability.** Does the variable change over time (such as age) or stay fixed? It is possible that rating by age is tolerable because we all get the same chance to be on the winning side and the losing side of it over the course of a lifetime.
- **Statistical Discrimination.** A variable should have some predictive value of an underlying risk. If it does not, then it is generally viewed as unacceptable for insurance purposes. As a rule of thumb, the better predictor of risk the characteristic is, the more tolerable such discrimination becomes. However, some pricing variables may not have such predictive abilities; Section 3.3 will sharpen this precept by distinguishing between “risk-based” and “non-risk” price discrimination in insurance.
- **Causality.** It is generally acceptable to use a variable if it is known to cause an insured event. For example, an individual diagnosed with cancer will generally be unable to purchase life insurance. Naturally, establishing causality is a much higher bar than mere correlation, or predictive ability, for a risk. For example, decades were spent building scientific evidence before it was widely established that smoking causes premature deaths, c.f., Peto (1994).
- **Limiting or Reversing the Effects of Past Prejudice.** Does an insurer’s use of a trait perpetuate negative stereotypes or otherwise subordinate disadvantaged groups? The historical use of the characteristic as a method of discrimination is also relevant; that is, whether the characteristic defines a socially salient group that has been disadvantaged in the past. In that sense, discriminating based on skin color is more problematic than based on eye color.
- **Inhibiting Socially Valuable Behavior.** Does an insurer’s use of a trait inhibit or prevent socially desirable activities? Section 5.1 describes how individuals, fearful of being denied life insurance, avoid participating in genetic testing research. As another example, Prince and Schwarcz (2020) cite U.S. laws that prohibit insurers from



discriminating on the basis of intimate partner violence because such reporting could dissuade victims of violence from seeking needed medical care or police intervention.

Whether a characteristic is socially suspect or sensitive is context-dependent, depending on the jurisdiction and the line of insurance business, as described in Sections 4.2 and 5.

### 2.3 Indirect Discrimination

One of the difficult questions in insurance discrimination is the treatment of related variables; use of variables related to a prohibited variable constitutes *indirect discrimination*. These are variables that, although they do not have the usual characteristics of an unfair variable (Section 2.2), have undesirable effects on society. A classic example is *redlining*, a term that refers to the practice of drawing red lines on a map to indicate areas that insurers will not serve, areas typically containing high proportions of minorities.

Specifically, we can define indirect (insurance) discrimination as consisting of three elements: *proxy discrimination*, disparate impact, and whether or not the discrimination could be avoided by other means. We address each in turn.

Proxy discrimination, also known as *indirect statistical discrimination*, occurs when insurers discriminate based on a facially-neutral characteristic, such as the size of the car engine, that correlates with a protected class, such as gender (in many studies, men like to drive cars with big engines). In the case of redlining, insurers discriminate based on geographic area (such as an urban area) that is correlated with race. Specifically, by avoiding certain urban areas, they also avoid large groups of potential minority customers; this amounts to at least partially making insurance decisions based on a protected variable.

It will be helpful to think about two types of proxies: one where an identifiable surrogate such as geographic area serves as a substitute for a protected variable such as race, and one where the proxy is produced by an algorithm that summarizes the effects of many variables. Section 5.5 on big data emphasizes the increasing importance of the second type of proxy discrimination as insurers utilize increasingly sophisticated algorithms and growing sources of data. Empirical aspects of proxy discrimination are discussed further in Section 6.

The second element is disparate impact, i.e., whether there is an impact which puts members of a protected group at a disproportionate disadvantage compared with members of a similar group. Returning to the redlining example, this practice puts minority neighborhoods at an economic disadvantage. Baker (2002) states: “A neighborhood redlined by insurance companies is a more risky place for banks to lend. Without good financing opportunities, fewer people invest in the neighborhood, and without investment the neighborhood becomes an even more risky place for banks, causing further decline.”

These notions of proxy discrimination and disparate impact are drawn from the legal literature where it can be sometimes difficult to infer a precise mathematical formulation. For our purposes, we think of proxy association as the relationship between a protected variable and a set of one or more surrogate variables whereas disparate impact occurs when there is a relationship between a protected variable and an outcome of interest (e.g., insurance purchase). In Section 6.4, we supplement these definitions by drawing from the machine learning literature.

The third element is whether the criterion is motivated by a *legitimate business necessity*. If it is, then discrimination may be legal even in the event of producing a disparate impact. For example, Section 5.3 will describe models that optimize insurer's profitably objectives, a legitimate business motivation, at the expense of disfavoring customers with fewer market options who tend to be low-income and minority consumers.

As another example, the Council of the EU adopted Directive 2004/113/EC Guidelines on the Application of the Gender Directive, "the use of risk factors which might be correlated with gender [...], as long as they are true risk factors in their own right" is still permitted.

### 3 Economic Considerations

One approach to pricing is to think of an insurance contract as a type of financial investment. From this viewpoint, financial investors base risk transfers considerations on a *law of one price* that is dictated by forces of supply and demand in a competitive market. An advantage of this approach is that many issues of insurance discrimination become moot as prices are given by an external marketplace. Many readers will enjoy thinking about pricing of insurance contracts in the context of financial economics asset pricing theory, summarized by Bauer, Phillips, and Zanjani (2013).

However, even in personal lines (where there is much more homogeneity than in commercial lines), there is substantial heterogeneity among insurance products when considering the variety of contract features (deductibles, limits, coinsurance, and so forth), risk factors of the entity insured (e.g., auto or home), and risk factors of the insured (e.g., attitude toward risk) that exist. Because of this heterogeneity, insurance pricing is focused on the underlying *cost of producing the good or service*.

Like any firm, the price that an insurer charges is determined by the quantity where marginal cost equals marginal revenue. However, unlike other industries, determining marginal costs is difficult in insurance. In part this is because the production cost is random. By definition, insurance contracts are based on contingent events whose financial outcomes are uncertain. Further, even when an insured event does occur, the actual cost of an insured

claim may not be known for a long period of time.

As further developed in Section 4.1, insurance prices are based on the expectation of losses, a concept coined as an *actuarially fair* price in Arrow (1963). In a simple model, an actuarially fair price is the result of an assumption of zero profits and ensures that the insured will buy full insurance coverage. It is the foundation of insurance pricing.

### 3.1 *Adverse Selection, Moral Hazard, and Incentives*

Simple models can provide insights but are naturally limited in addressing the numerous features of real contracts. One feature particularly relevant to potential insurance discrimination is unequal access to information, known as *information asymmetry*.

Insurers traditionally face *adverse selection*, a problem that can arise when consumers know more about their own risk characteristics than insurers. Insurers argue that by knowing about risk factors, the entire marketplace is better. Indeed, the entire purpose of risk classification is to mitigate the problem of adverse selection. Extending this line of thought, the more information that insurers have about policyholders, the more effective is risk classification; this in turn results in a better marketplace for all.

Another type of adverse selection can occur when an insurer has less information than other competing insurance companies about the risk levels of its customers, c.f., Cather (2018). This can result in *cream skimming*, since the innovative insurer targets the best risks who, like cream in a container of fresh milk, rise to the top of a pool of policyholders.

Another classic type of information asymmetry is *moral hazard*. Insurers worry about insureds' attitude toward safety; by purchasing insurance, insureds have the incentive to take on more risks (thus, increasing the probability of a risky event). For example, after purchasing of homeowners insurance, the insured may become lax in watching for fires (smoking in bed, not checking for frayed electrical wires). One way to mitigate this risk is through the installation of fire alarms.

Insurers also have to be wary of their own moral hazard. For example, if they acquire a protected variable such as political affiliation, then they have to be careful that this knowledge does not implicitly bias their pricing processes even if they do not use this information actively. One way to mitigate this risk is to simply avoid acquiring such protected information.

An implication of moral hazard is that people tend to increase their risk unless given incentives not to. Conversely, people may also *reduce* their risks when given incentives to do so. Indeed, much of modern risk management is predicated on introducing risk mitigation tools to reduce the impact of insured events. Classic examples include lower premiums for sprinkling systems in fire insurance and no smoker discounts in life insurance (Avraham,

Logue, and Schwarcz (2014)).

Insurers worry about traditional adverse selection and moral hazard because information asymmetries favor policyholders over insurers. In contrast, as emphasized by Schwarcz and Siegelman (2017), much of insurance law is designed to protect policyholders because of information asymmetries that favor insurers over policyholders. For example, policyholders are often insufficiently knowledgeable about an insurance policy terms and conditions, the insurers' financial strength, and the appropriate type of policy for a consumer's needs. Regulators are concerned that insurers may be able to exploit these deficits in policyholder information or sophistication by providing more limited coverage than policyholders believe they are purchasing, or by adopting excessively aggressive claims-handling strategies.

More recently, consumer advocates have been concerned that additional big data information, discussed more in Section 5.5, puts consumers at a disadvantage. For consumer advocates, more data information for insurers means that:

- Insurers can cherry pick at a granular level.
- Insureds do not have equivalent new tools to compare quality of policies and performance of insurance companies.

Consumer advocates argue that mandatory and de facto mandatory purchase of insurance means that free market competition is insufficient to protect policyholders.

### 3.2 *Economic Efficiencies*

Economists largely agree that a competitive market is an efficient one, c.f., Skipper and Klein (2000). Efficiency is achieved because competition forces buyers to pay their maximum demand price and forces sellers to charge their minimum supply price. Competition serves the best interests of consumers in that it provides insurers incentives to attract customers by reducing prices or improving insurance products. Competition policy is about applying rules to make sure companies compete fairly with each other.

One of the barriers to competition is asymmetric information; the insurance industry uses risk classification to cope with this potential problem. What could happen without risk classification? Because of the price differential we might see a reduced pool of insured individuals; this reflects a decrease in the efficiency of the insurance market. Extending this line of thought to multiple periods suggests an exodus of low risks that can lead to a *death spiral* of rising premiums and ends up unraveling the entire market, c.f., Dionne and Rothschild (2014).

However, by and large, insurers are allowed to classify risks. As argued by Tennyson (2007), a large body of academic research supports the conclusion that insurance markets

function in a *workably competitive* manner in the absence of rate regulation. Competition indicators include the number of insurers and their market shares, profitability, and price of their products. At least some markets for some countries (e.g., U.K. auto) generally exhibit the characteristics of a competitive market, FCA (2019) (Annex 3: International Comparisons).

Rate regulation can limit the insurers ability to classify risks and hence threaten competition. Tennyson (2007) describes two types of rate regulation regimes:

- *rate suppression* - reducing average rates for all consumers, and
- *rate compression* - reducing rates for some consumers (usually high-risks) relative to others (usually low-risks).

Both have negative consequences for insurance markets. Rate suppression runs the risk of driving average premiums below competitive levels, reducing insurer returns below a competitive rate of return. Rate suppression will distort insurance supply in the market, reducing competition in the long run. Rate compression can have similar effects by reducing rates for some consumer groups below competitive levels. As an example of evidence of the negative effects of rate regulation, in a classic article Blackmon and Zeckhauser (1991) document the negative effects of rate suppression and compression for the automobile market in Massachusetts.

### 3.3 *Price Discrimination*

The act of charging different prices for identical products is known in economics as *price discrimination*. To apply this to insurance, we need to specify that identical products also means identical production costs. We could, for example, have two auto policies that promise to pay exactly the same benefits for a loss. But, their prices may differ depending on risk factors such as the insured's ability to drive and attitude towards risk, the type of vehicle itself (inexpensive family car versus a pricey sports car), the location where it is being driven (city versus rural), and so forth. So, the expected loss would be different causing the prices to differ. This type of *risk-based* price discrimination is the norm in insurance pricing. In contrast, Thomas (2012) uses the phrase *non-risk price discrimination* for the insurance situation where prices may differ for the same coverage and underlying risk characteristics.

Price discrimination is common in other industries. For example, airlines regularly charge higher prices for flights during the week (e.g., Monday to Friday) because these are typically taken by business travelers. This is an example of *first-degree price discrimination* where the price is based on the buyer's willingness to pay. Second-degree discrimination involves quantity discounts, whereas third-degree discrimination reflects different prices for different

consumer groups, e.g., discounts for senior citizens (known as “honored citizens” in Portland). Price discrimination is not possible in a perfectly competitive market because there are many firms competing for the price, c.f., Lukacs, Neubecker, and Rowan (2016).

In the insurance industry, first-degree price discrimination is common in large commercial insurance where it is assumed that buyers are sophisticated and willing and able to negotiate prices. Third-degree price discrimination is common in personal insurance, at least in some jurisdictions such as within Europe and the United Kingdom. In particular, prices for renewing customers are often distinguished from risk-identical new customers, with different (usually lower) price offers made to new customers; “paying customers to switch.” To illustrate, the work in Adams et al. (2015) on general insurance auto confirms that some consumers pay much higher prices if they stay with the same insurer, particularly for a long period of time. This practice is motivated by so-called *price optimization* models that are described in Section 5.3.

For personal insurance, some jurisdictions allow price discrimination but others take a dim view of it. For example, in the early to mid-1800’s U.S. voluntary associations of insurers were organized in part to enforce uniform rates among the insurers. Uniform rates were desired so that rates were adequate to protect against insolvencies and were not unfairly discriminatory. From Miller (2009),

The primary concern with unfairly discriminatory rates, often stated at the time, was that rich and powerful insureds could unfairly negotiate lower rates than were being charged to less influential insureds, even though their degree of risk and underlying insurance costs did not warrant a lower rate.

Is price discrimination appropriate for insurance? Thomas (2012) summarizes five aspects of insurance that makes it different from other marketplaces.

- (1) **Ability to discriminate.** Insurers can differentiate prices because of the quality of their data, the general confusion surrounding the pricing process, and the consumer’s inability to “re-sell” the product.
- (2) **Price discrimination in insurance does not facilitate new markets.**
- (3) **Price discrimination may undermine utmost good faith.** Laws mandate that customers provide information about their risks truthfully and, in many jurisdictions, insurers are also permitted to share information for the purposes of preventing fraud. If this information is used for other purposes, then over time this could undermine public acceptance of the doctrine of utmost good faith.
- (4) **Price discrimination may undermine justifications for risk-related pricing.**

- (5) **Distributional effects of price discrimination.** Allowing price discrimination may introduce cross-subsidies and have uneven effects on different parts of society.

Insurers may wish to use non-risk related factors to achieve legitimate business goals, such as maximizing profit or increasing customer retention. In these cases, firms may unwittingly discriminate, knowing only that a facially-neutral practice produces desirable outcomes. Section 5.3 describes the specific case of price optimization models where use of non-risk related factors is prohibited in many jurisdictions.

## 4 Actuarial Aspects of Rate Regulation

Section 1.1 provides an overview as to ways in which insurers discriminate; the focus of this section is on the pricing function. This is because actuaries are heavily involved, and hence influential, in pricing. Further, many regulations are geared towards pricing prohibitions, known as *rate regulation*. In addition, one can argue that prices are intimately related to whether or not someone is offered coverage (one could “price someone out of the market”) and the amount of coverage.

### 4.1 Pricing

Like any business, pricing is critical in insurance. As described in Section 3, one aspect in which insurance differs from other industries is that the cost of the good is random and may not be known for many years after the sale of the product. This has led the actuarial profession to think deeply about what this “cost” entails.

Nonetheless, prices are often based on the **cost of insurance**. These are the costs of transferring a risk from the policyholder to the insurer. As described in standard actuarial textbooks such as Friedland (2013) and Werner and Modlin (2016) (see also Chapter 7 of the open source [Loss Data Analytics](#)), insurance costs consist of the (1) losses (compensation provided by the insurer for the insured claim), (2) expenses associated with the policy and claim, and (3) cost of capital (costs of keeping monies necessary to fund the insurance operation). Prices based on insurance costs are sometimes known as *technical prices*.

As described in the online supplement Frees and Huang (2021), many jurisdictions are silent on insurance rate regulations and so market prices are influenced by forces of supply and demand. As with other businesses, the cost of a product is an important but may not be the sole determinant of a price. Additional factors include the market availability of alternatives (e.g., costs of self-insurance, prices offered by competing firms) and marketing considerations such as customer loyalty. However, in jurisdictions where rate regulation

is prominent, prohibitions are in terms of technical prices. Because our focus is on rate regulations, we focus on cost-based technical prices.

Also from Frees and Huang (2021), many jurisdictions exhibit insurance rate regulations in one form or another. At the time of this writing, the U.S. is the country that most actively regulates rates and so we use this to motivate the discussion. From the U.S.-domiciled Casualty Actuarial Society's [Statement of Principles Regarding Property and Casualty Insurance Ratemaking](#) (Principle 4), "a rate is reasonable and not excessive, inadequate, or unfairly discriminatory if it is an actuarially sound estimate of the expected value of all future costs associated with an individual risk transfer."

In the U.S., insurance is regulated at the state level. The National Association of Insurance Commissioners (NAIC) is an organization that provides standards that states may adopt. As described in the model rating law NAIC (2010), the rule is that ... *Rates shall not be excessive, inadequate or unfairly discriminatory.* It further defines an unfair discriminatory rate as ... *Unfair discrimination exists if, after allowing for practical limitations, price differentials fail to reflect equitably the differences in expected losses and expenses.*

As an expectation, cost-based prices are naturally influenced by the choice of rating factors and this is where regulation comes into play. Regulators prohibit the use of certain variables - how do these prohibitions affect rating schemes?

#### 4.2 *Extent of Regulation*

Insurance regulations may consist of *applicable acts, statutes, regulations or any other binding authority (such as accounting standards and any regulatory guidance that is effectively binding)*, as described within the [International Standards of Actuarial Practice \(ISAPs\)](#) of the [International Actuarial Association](#). In most jurisdictions, their enforcement is overseen by an insurance supervisor or regulator, many of whom follow the standards developed by the [International Association of Insurance Supervisors \(IAIS\)](#). According to their website, the IAIS "is the international standard-setting body responsible for developing and assisting in the implementation of principles, standards and other supporting material for the supervision of the insurance sector." These include standards known as insurance core principles (ICPs): on insurer solvency, sales practices, agent licensing and policy forms, for example. Interestingly, there is no discussion within the ICPs on rate regulation.

The extent of insurance rate regulation varies by jurisdiction. At one end of the spectrum, the phrase *active rate regulation* means that the regulator is heavily involved in determining rates. This could mean government mandated rates in which regulators dictate the rates to be charged. Alternatively, regulators may only require approval of rates, either in advance or concurrent with policy offerings. Friedland (2013) (Chapter 27) summarizes the spectrum



of rate regulation. To illustrate, for the U.S., Borselli (2011) provides additional details including types of regulatory system organized by state and line of business. The other end of spectrum is competitive rating or open competition systems. Borselli (2011) compares active rate regulation to open compensation systems. He notes that historically many European countries operated under active rate regulatory environments but now regulators of European members states do not have the right to regulate insurance prices

For more concrete descriptions, in a supporting document Frees and Huang (2021) we describe several major regulatory jurisdictions:

- The U.S. is the largest general insurance marketplace. It is also the most actively regulated jurisdiction with a coordinating body (the National Association of Insurance Commissioners) that develops model laws that *may* be adopted by individual states.
- The European Union is the second largest marketplace. It has a coordinating body (the European Commission) that develops legislative directives that *must* be implemented by member countries.
- China, Japan, and Australia, are the third, fourth and tenth, respectively, largest marketplaces. They illustrate the variety that even single country regulatory environments may exhibit regarding discrimination issues in insurance.

## 5 Prohibitions by Line of Business

In addition to jurisdiction, as emphasized by Avraham (2018), insurance prohibitions vary greatly by line of business. It is notable that the markets in which rate regulation is most common - automobile insurance, health insurance, workers compensation, medical malpractice, and homeowners' insurance – are all markets in which insurance is mandatory or in which universal coverage is thought to be socially desirable, c.f., Tennyson (2007).

In commercial lines of insurance such as general liability and professional liability, the policyholder is a firm. As a general rule, insurance rate regulation tends to be absent from commercial insurance. In part this is because firms typically have more resources than individuals and so do not suffer the same imbalance of information asymmetry as described in Section 3.1. As exceptions to this rule, medical malpractice and workers' compensation continue to be highly regulated. These two lines provide coverages that tend to be mandated by government regulations. Even though highly regulated, because the purchasers of insurance are generally organizations, discrimination issues are not as relevant. In the following, we focus instead on personal insurance.

### 5.1 *Unisex Risk Classification*

The European insurance marketplace was rocked in 2011 when the European Court of Justice concluded that any gender-based insurance discrimination is prohibited, European Union (2012). Prior to this ruling, gender was regularly routinely used for pricing insurance.

Broadly, what characteristics of this trait would lead society to prohibit its use for insurance discrimination? Referring to the Section 2.2 categories, sex at birth is certainly beyond the control of the insured. Insurers have economic motivation for using gender as a predictor because, for many lines of business, it exhibits helpful predictive abilities even though its causal nature can be debated. An important motivation for prohibiting gender as a rating variable is to limit the perpetuation of negative stereotypes, so that men and women would receive equal treatment in the access to and supply of all goods and services. Equality and respect for human dignity and human rights are among the core values of the European Union (EU), reflected in several EU directives targeting discrimination. Initially signed in 1957, [Article 10 of the Treaty on the Functioning of the European Union](#) states:

In defining and implementing its policies and activities, the Union shall aim to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation.

Like other major upheavals, this turning point was precipitated by other events. In the U.S., a series of legal rulings and new laws led to prohibitions based on gender for retirement systems, c.f., McCarthy and Turner (1993). The U.S. Supreme Court issued decisions that prohibit pension plans from using separate mortality tables for men and women to determine contributions and benefits.

- 1) From the case of Los Angeles Department of Water and Manhart (1978), employers cannot require women to make larger contributions to a pension plan in order to receive the same monthly benefits as similarly situated men.
- 2) From the case of Arizona Governing Committee for Tax Deferred Annuity and Norris (1983), women cannot receive lower monthly benefits than men who had made the same contributions (this case was based on a defined contribution plan).

These decisions were based on the 1964 Civil Rights Act that prohibits employment discrimination because of an individual's race, color, religion, sex, or national origin. In 1986, the Equal Employment Opportunity Commission broadened the Supreme Court's prohibitions by forbidding sex-based differences in any employee benefit, even if justified by differences in cost.

Another major event that preceded the 2011 ruling was a European Union Council Directive in 2000 that prohibited discrimination based on racial and ethnic origin, European Union (2000). Prior to this directive, some insurers required an additional premium based on the risk criterion of the applicant being a “foreigner,” c.f., Schmeiser, Störmer, and Wagner (2014). The 2011 ruling was based on European Union Directive 2004/113/EC, European Union (2004) (this earlier directive required equal treatment of men and women but provided an exception for the insurance industry). This ruling applies to all lines of insurance business; in contrast, for example, gender-based pricing of auto insurance is permitted in all but a handful of U.S. states (the exceptions being Hawaii, Massachusetts, Montana, North Carolina, Pennsylvania, and, in 2019, California). As part of the guidelines on the application of the 2011 ruling, the use of risk factors which might be correlated with gender, as long as they are true risk factors in their own right, is still permitted.

The retirement systems line of business is interesting because gender is an important predictor of longevity, c.f., Lemaire (2002). Women, in general, outlive men so they receive pension benefits over a longer period of time. For other lines of business such as personal auto, it is possible to find variables that provide a suitable substitute for gender, c.f., Ayuso, Guillén, and Pérez-Marín (2016). Section 6 explores the concept of proxy discrimination in more detail.

## 5.2 General Insurance and Credit-Based Scoring

An individual’s credit history has long been used in commercial lines of insurance and in life insurance. As described by Brockett and Golden (2007), “Although it has been known since at least 1949 that credit history is related to driving accidents, the advent of high capacity, high-speed computers has made massive personal credit files available, and has made it feasible to routinely use this credit information for predicting insurance losses in personal lines of property and casualty insurance.”

From FTC (2007), general insurers in the U.S. started to use credit history information for automobile pricing in the early 1980’s. It became more routine with the development by Fair Isaac Corporation (FICO) scores in the mid 1990s, first for homeowners and then for auto. According to FICO, in the absence of state prohibitions, 95 percent of automobile insurers and 85 percent of homeowners insurers employ insurance scores in either the underwriting or rating process (NAIC 2012a), see also Morris, Schwarcz, and Teitelbaum (2017).

Credit-based insurance scores, or simply *insurance scores*, are similar to widely known credit scores in that both rely upon an individual’s credit history. This credit history includes prior credit performance (e.g., late payments), current levels of indebtedness (e.g., bankruptcy), length of credit history (e.g., age of oldest account, average age of all ac-

counts), pursuit of new credit (e.g., new accounts, mortgages), and types of credit used (e.g., department, travel, major bank credit cards). However, credit scores predict the risk of credit delinquency and so measures the financial well-being of a consumer. In contrast, an insurance score is designed to predict insurance losses and so assesses how well individuals manages their money.

Credit-based insurance scores summarize an individual's personal financial history; they do not exhibit the characteristics described in Section 2.2 that would lead them to being described as sensitive or *suspect*. However, as emphasized by Morris, Schwarcz, and Teitelbaum (2017), insurance scores are regulated because they potentially correlate with suspect classifications, in particular race and income. For this reason, in the U.S. most states regulate insurers' use of insurance scores in auto and home insurance, and a few states ban their use altogether (Avraham, Logue, and Schwarcz (2014), Avraham (2019)).

Credit-based insurance scores provide a natural example where the ability to predict insurance losses is well established (statistical discrimination) but the causal nature is uncertain. As noted by Brockett and Golden (2007), a poor credit score may not create (cause) an insured loss but it is a measure of underlying biological and psycho-behavioral traits that do affect insured losses. However, from Morris, Schwarcz, and Teitelbaum (2017), this is inconsistent with the fact that two of the major drivers of credit risk are unemployment and health problems, neither of which seems to reflect irresponsible behavior such as reckless driving or lack of fire safety. So, although both sources re-affirm that credit-based insurance scores provide a sound platform for predicting insurance losses, the causal nature remains unclear.

### 5.3 General Insurance and Price Optimization

Technical prices, that are based on expected claims, provide the foundations for most lines of general insurance, at least on the non-commercial, or personal, side. Traditionally, informal judgement has been used to adjust technical prices to become market prices; these adjustments are (i) for consistency among factors, plans, and over time, (ii) for competitors rates, and (iii) for impact on retention, c.f., CAS (2014). *Price optimization* refers to a systematic approach for making adjustments to traditional cost-based technical prices that incorporates customer demand.

Technical prices can be thought of as based on single-period models that focus on costs of insurance including claims and expenses. In contrast, the price optimization approach incorporate models of retention as well as prices of competitors typically by looking over several periods. For example, if an insurer raises prices, then that insurer can expect lower retention; the amounts depend on how sensitive the consumer is to price changes and the

availability of the same coverage from the competition. By looking over several periods, price optimization models can tune prices to achieve an insurer's long-term goals, such as profitability or development of market share.

From an insurer's point of view, the use of price optimization tools is simply a sound business practice that is widely adopted in many industries, including retail and travel. Price optimization moves insurance pricing beyond expected costs to behavior including price sensitivity. At the individual consumer level, it may be that an insurer prices differently two consumers with the same risk profile because their anticipated price sensitivity differs, c.f. Section 3.3. Price sensitivity matters because it affects consumer retention and acquisition expenses for new business are generally higher than expenses for retaining a customer, c.f., CAS and Force (2015).

However, consumers have taken a dim viewpoint of price optimization (see a summary in CAS and Force (2015)). On the one hand, the ability to identify loyal customers suggests that these are the customers who would enjoy lower premiums because of the lower expenses associated with them. On the other hand, these are exactly the customers who are likely to stay (and remain loyal) when faced with price increases. Some critics argue that price sensitivity practices impose an unfair penalty on customer loyalty. Customers more likely to be loyal are less likely to shop for alternatives. Insurers will identify this tendency and so impose price increases on customers, not for their tendency to have high claims but rather for their tendency to be loyal.

Other critics argue that price optimization has been developed to increase insurers' profits by raising premiums on individuals who are less likely to shop around for a better price, and many of these people are low-income consumers. Consumer advocates assert that deviation from cost-based ratemaking through price optimization will disfavor those consumers with fewer market options, less market power and less propensity to shop around in particular, low-income and minority consumers. Thus, although insurers may be optimizing neutral objectives, the result of their actions can result in unintentional proxy discrimination.

Swayed by these arguments, many U.S. insurance state regulators have banned price optimization in personal lines insurance. As another example, price optimization and price discrimination is not illegal in Australia but it does give rise to consumer unease and is being scrutinized by regulators, ESL (2018).

#### 5.4 *Life Insurance and Genetic Testing*

Genetic testing involves a type of medical test that examines chromosomes, genes, or proteins. The results of a genetic test can confirm or rule out a suspected genetic condition or help determine a person's chance of developing or passing on a genetic disorder. There are

many different purposes for testing, including medical (such as diagnosing a genetic disease or predicting disease risk) and non-medical (such as confirming parentage or forensic investigation). Hundreds of genetic tests are currently in use and more are being developed, see, e.g. Born (2019).

Information from genetic tests is potentially sensitive. Following the structure in Section 2.2, the main reasons are because they are not under the control of an individual nor, in most cases, do they change over time. Only in rare instances (such as Huntington's disease) is a genetic condition known to cause an insured event such as death. When thinking about the standards of perpetuating negative stereotypes and historical precedence, the genetic tests themselves are new (and developing), so the historical impact of a specific test is minimal. Nonetheless, as noted by Avraham, Logue, and Schwarcz (2014), "Genetic discrimination in the context of health, life, and disability insurance immediately evokes Nazi Germany and its obsession with promoting the reproduction of more 'genetically desired' people and eliminating 'genetically defective' individuals."

Statistical discrimination is another important attribute. Earlier, Lehtonen and Liukko (2011) wrote "At least for the time being, genetic information is in most cases neither statistically nor economically significant for risk assessment from the insurance companies' point of view. The exceptions are the rare single-gene diseases, such as Huntington's disease, which inevitably or very likely result in death." As of this writing, this remains the case (c.f., Vukcevic and Chen (2018)); today, it is likely that including results from genetic testing will not materially alter an insurer's prices. However, insurers believe that this will change over time; that is, predictive ability of genetic testing will increase over time and will become salient in at least the life, disability, critical illness, and long-term care insurance marketplaces.

Insurers worry about genetic testing information because of information asymmetry concerns. Like the purchase of life insurance, the decision to undergo genetic testing is voluntary. When a potential policyholder has information about his or her health that is not shared with the insurance company, this could lead to anti-selection where poorer risks purchase more insurance and better risks purchase little or no insurance. From an insurer's viewpoint, one solution would be to allow insurers to require genetic testing, just as they are allowed to evaluate other aspects (e.g., weight, hypertension, and so forth) of a person's health. There is some evidence that prohibitions on using genetic testing information may materially affect insurer's claim costs, see e.g. Lombardo (2018).

In policy debates, arguments have been made for the position that genetic information is special and must therefore be treated differently from other types of medical information (sometimes known as *genetic exceptionalism*). One way that genetic testing differs from, e.g., blood pressure, is through the impact that it has on a person's willingness to undergo the

testing for fear of being denied life insurance. As summarized by Prince (2019), “Empirical evidence shows that fear of genetic discrimination has led individuals across the globe to refuse to participate in genetic research projects or to fail to undergo recommended clinical testing.” Nonetheless, this may be simply due to the evolving nature of the science of genetic testing. For an analogy to underscore this point, Born (2019) noted that “. . . over time, other types of medical tests – e.g., tests for cholesterol levels – were first considered controversial when initial evidence showed a wide variation in predicted value.”

The social impact depends on the line of insurance business. In health insurance, the impact of genetic testing is less pronounced because many leading countries in the world offer government-provided health insurance or mandate the purchase of health insurance. Even in the U.S., that does not have government provided health insurance, results of genetic testing are not permitted under the Genetic Information Nondiscrimination Act (GINA). This act prohibits covered health insurers (and employers) from discriminating on the basis of genetic information that includes genetic test results, family medical history, and use of genetic services.

Some international conventions recommend restricting the use of genetic information for insurance purposes. The Council of Europe’s Convention on Human Rights and Biomedicine prohibits the performance of genetic testing as a condition for entering into an insurance contract, (c.f., Lehtonen and Liukko (2011)). The United Nations Educational, Scientific, and Cultural Organization (UNESCO) in 1997 issued a Universal Declaration on the Human Genome and Human Rights. This was followed by a 2003 declaration that argued that genetic data and biological samples should not be accessible by insurance companies, among other actors (c.f. Prince (2019)).

While many countries have specific laws covering genetic testing, most of the regulations are not very detailed. [Table 5.1](#) summarizes genetic testing regulations, from Klein (2017). Joly et al. (2020) provides another perspective on cross-country comparisons. Not only the science, but also the insurance regulation, of genetic testing continues to evolve. Several jurisdictions have passed or are considering legislative changes in the use of genetic testing information in underwriting insurance. For example, Canada passed federal laws in 2017 banning the use of all genetic information for business purposes, see e.g. Lombardo (2018). Within the U.S., Born (2019) documents recent proposals in the state of Florida. Further, in July 2020, Florida has passed [a law](#) that prohibits life insurers & long-term care insurers from canceling, limiting, or denying coverage, or establishing differentials in premium rates based on genetic information. As noted in [Table 5.1](#), Australia has been active in considering genetic testing legislation.

**Table 5.1** : Genetic Testing Prohibitions by Country. *Source*: Klein (2017)

<b>Regulation Category</b>	<b>Country</b>
No regulation	China, Finland, India, Spain, United States
No regulation with written or unwritten codes of conduct from insurance industry groups	Greece, Japan
Prohibitions on insurers requiring applicants to take a genetic test and prohibitions on discrimination if the applicant refuses to take a test	Australia*
Prohibitions or moratoriums on using results from existing tests when policies are below certain limits	Germany, Netherlands, Switzerland, United Kingdom
Prohibitions or moratoriums on using results from existing tests at all, sometimes including use of family history information	Austria, Belgium, Canada, Denmark, France, Ireland, Poland Portugal, Singapore

\* As noted in Vukcevic and Chen (2018), Australia has moved from the third (middle) category to the fifth (bottom) recently.

### 5.5 *Big Data*

As with all institutions, insurers are redefining the way that they do business with the increasing capacity and computational abilities of computers, availability of new and innovative sources of data, and advanced algorithms that can detect patterns in insurance activities that were previously unknown. Sections 6.3 and 6.4 summarize how these advanced algorithms can be used to mitigate discrimination issues.

Conceptually, Big Data does not alter the fundamental issues of insurance discrimination. This point was emphasized in Swedloff (2014), entitled “Risk Classification’s Big Data (R)evolution.” One can think of credit-based scoring and price optimization as simply forerunners of a long-term trend by insurers to gather more and more data about their current and potential customers. One might hope that machine-driven algorithms would eliminate human biases but, as documented by Barocas and Selbst (2016), these algorithms inherit the prejudices of authors of the algorithms and prior decision-makers. As another example, Bartlett et al. (2019) find that the use of algorithmic decision-making of mortgage loans in the U.S. results in significant proxy discrimination of Latin and African-American borrowers although significantly reduces discriminatory practices of face-to-face lenders.

Yet, Big Data is changing the way that insurers do business. With respect to insurance discrimination, Swedloff (2014) argues that the two main aspects of change are *privacy* and *proxy discrimination*. On the privacy front, some of this detailed information is provided



voluntarily by individuals to insurers and suggesting to some that it not be treated as sensitive.. This includes information from global position systems (GPS) that we put in our cars that underpin telematics, comparable devices for our homes (the *Internet of Things*), devices that we wear to improve our health, and so forth.

Still, insurers may also use other information that is not provided directly by individuals. Privacy issues are raised any time a carrier classifies risks on intimate, personal information, like HIV status, marital status, sexual orientation, or genetic information. Although not an insurance case, Swedloff (2014) discusses the highly publicized event where Target, a large U.S. department store, used analytics to predict which of its customers were pregnant. This information was passed on to the marketing arm who sent coupons for maternity clothing, nursery furniture, and the like, to women who were likely to be expecting a child. That is, without asking any customers about their pregnancy status or harvesting that data in particular, Target was able to predict extremely sensitive and personal information about its customers. Consumer advocates fear that similar information, such as where we go, movies that we watch, telephone calls and texts that we make, would be of interest to insurers ostensibly to be used for understanding consumers attitudes towards risk and the likelihood of making insurance claims.

Proxy discrimination, introduced in Section 2.3, occurs when a surrogate, or proxy, is used in place of a prohibited trait. As originally conceived, this proxy is a facially neutral trait, such as the size of an automobile's engine being used as a proxy for gender. In the world of Big Data, an equally important issue is that complex algorithms are being developed using literally thousands of traits (in the simple Target example, the analyst used only 25 traits to develop an effective pregnancy score). Thus, as emphasized by Prince and Schwarzc (2020), proxy discrimination may be unintentional; moreover, the insurer may not even be aware that it is engaging in discriminatory behavior due to the opaqueness of machine-driven algorithms. Proxy discrimination is particularly important for actuaries and so is further discussed in Section 6.

Although consumers may benefit from a marketplace where insurers can more accurately assess risks, there is also a potential loss of *transparency* in insurance pricing. There already exists a low level of consumer understanding and a low level of consumer engagement with insurance purchases. More complex data algorithms will impede efforts at transparency. As discussed by Richman, Rummell, and Wuthrich (2019), machine learning models are often more complex and less transparent than traditional models. Moreover, massive data sets and complex models do not make life easier for regulators. Insurance regulators need to review rating plans that incorporate complex predictive models. Many jurisdictions do not have sufficient in-house actuarial expertise to review such filings.

For a final note on Big Data, some have expressed a concern that highly individualized or personalized rates lose the benefit of risk pooling, e.g., ESL (2018) and Meyers and Van Hoyweghen (2018). Although this could be a problem for some portfolios, most insurance systems are based on diversification of pools of unrelated (independent) risks. This type of diversification does not go away when the risks are different as long as they are priced properly.

## 5.6 COVID-19

A pandemic is a global outbreak of disease and in early 2020 the world has seen the onslaught of a new coronavirus dubbed COVID-19, short-hand with the ‘CO’ for ‘corona,’ ‘VI’ for ‘virus,’ and ‘D’ for disease. The ‘19’ is because it was first identified in 2019 (from the outbreak in Wuhan China). As with other parts of global economy, the disease has rocked the insurance industry. The lines of business most affected on the commercial side include workers’ compensation, business interruption insurance, cyber liability, general liability, and event cancellation as well as health and travel insurance on the personal side. see e.g. Fannin (2020). Interestingly, automobile insurance claims have dramatically reduced (for the first part of 2020 at the time of this writing) due to travel restrictions; people are driving less and getting into fewer accidents.

*Prohibiting Insurance Discrimination Based on COVID-19 Diagnosis.* Insurance legislation is being introduced to prohibit discrimination based on the diagnosis of this disease. For example, the State of Wisconsin passed a [law](#) on 15 April 2020 that includes the following:

This bill prohibits insurers that offer an individual or group health benefit plan, pharmacy benefit managers, or self-insured governmental health plans from doing any of the following based on a current or past diagnosis or suspected diagnosis of COVID-19: establishing rules for the eligibility of any individual, employer, or group to enroll or remain enrolled in a plan or for the renewal of coverage under the plan; cancelling coverage during a contract term; setting rates for coverage; or refusing to grant a grace period for payment of a premium that would generally be granted.

As another example, on 14 April 2020, the [Australian Competition and Consumer Commission](#) has granted interim authorization to the Financial Services Council and its members to ensure front line healthcare workers are not excluded from coverage due to exposure to COVID-19. That means life insurers cannot use the exposure to COVID-19 as a factor for

pricing or applying risk exclusions to any new policy. It is likely that other legal jurisdictions will undertake similar actions. From a narrow actuarial perspective, this legislation is about coverage and rating and so is important for actuaries. Further, in absence of legal restrictions, a competitive market would use diagnosis of COVID-19 like any other part of medical history, potentially predictive of future insurable events.

More broadly, this legislation has several implications. For example, in absence of this legal restriction, rates may well increase for grocery store workers, due to their exposure and increased *suspicion* of a diagnosis of COVID-19. Is this in the best interest of society? To sharpen understanding of the social and economic considerations presented in Sections 2 and 3, we now reflect on these principles in terms of COVID-19.

*Social Justice Considerations of COVID-19 Insurance Prohibitions.* By its very global nature, a pandemic brings out the social responsibility of an insurance pool described in Section 2.1.1. The line of business affected by COVID-19 may be viewed as social/public good (e.g. health insurance) or as an economic commodity (e.g., life insurance) but the responses to COVID-19 are certainly social. Just as when you build a dam it benefits everyone who lives in the areas (not just those who make contributions towards building it), in the same way societal measures for disease prevention (e.g., social distancing) are borne by the entire population.

Section 2.2 introduced variable attributes that influence fairness for insurance purposes. For a COVID-19 diagnosis:

- *Control.* Individuals have few controls as to whether they have disease symptoms due to its widespread impact on society.
- *Mutability.* The variable may change over time but not in a good way.
- *Causality/Statistical Discrimination.* For those that have recently contracted the disease, there is a known pathway to heightened mortality risk and healthcare costs. For those diagnosed with disease but have recovered, there are no known additional risks to mortality nor to morbidity.
- *Limiting or Reversing the Effects of Past Prejudice.* This is not an issue as COVID-19 is a new disease.
- *Inhibiting Socially Valuable Behavior.* If insurers were allowed to rate based on disease symptoms, it is likely that many would refuse testing which would inhibit scientific progress in addressing the disease, similar to genetic testing.

*Proxy Discrimination.* In the U.S., it is known that COVID-19 affects African-Americans more than other ethnic groups and that COVID-19 mortality rates are related to age. Thus,

insurer discrimination based on contracting COVID-19 could be viewed as an indirect path to race and, where pertinent, age discrimination.

### *Economic Considerations of COVID-19 Insurance Prohibitions*

- *Adverse Selection.* It is likely that individuals would know about whether they had symptoms of COVID-19 without going through formal testing - unknown to the insurer. This creates the potential for adverse selection.
- *Competition.* It is unlikely that any insurer will have private knowledge about the nature of the COVID-19 disease suggesting that marketplace competition is not an issue. However, some insurers may elect to pull out of the marketplace (such as with travel insurance), meaning that lack of supply may increase prices.
- *Price Discrimination* is not likely to be an issue with COVID-19.

*Summary of COVID-19 Insurance Prohibitions* For a pandemic, the weight of evidence suggests that **societal concerns dominate** and that a prohibition based on diagnosis, real or suspected, of COVID-19 is warranted. As insurers' lack data about its predictive abilities, it is unlikely that competition will be affected. Prohibitions of this nature increase consumer confidence in the insurance system. At a hopefully not too far date in the future, the COVID-19 will lose its pandemic status and become another disease that we have to deal with. At that time, special legislative actions for COVID-19 will lose their appeal.

## **6 Proxy Discrimination**

Proxy discrimination, when a seemingly innocuous variable is correlated with a protected variable, can be a problem because it produces the same outcomes that would be obtained in the absence of restrictions based on directly predictive traits, c.f., Prince and Schwarcz (2020). This is true whether or not the surrogate is opaque to the insurer and the regulator.

### *6.1 Strategies for Mitigating Proxy Discrimination*

Historically, the focus has been on introducing regulation that prohibits the use of protected variables, such as race, or surrogates thought to be proxies for protected variables, such as credit-based insurance scores. More recently and providing a greater challenge is how to mitigate proxy discrimination when the proxies are produced by opaque machine learning algorithms based on many variables. There are several strategies that policymakers can use to limit this latter type of discrimination, none of which are ideal.

1. **Community Rating.** At one end of the spectrum, proxy discrimination can be completely eliminated by removing the insurer's ability to discriminate entirely. This is the case in community rated plans where all policyholders pay the same price such as common in social insurance schemes.
2. **Approved Variables.** Another option is to specify variables that may be used instead of variables that may not be used. This is the strategy taken in the U.S. individual health insurance market under the Affordable Care Act (ACA). Specifically, insurers may vary rates based on only four factors, (1) whether a plan covers an individual or family, (2) geographic area, (3) age, and (4) smoking status. As described by Prince and Schwarcz (2020), the ACA prohibits discrimination on the basis of prior health history, pre-existing conditions, and sex.
3. **Actuarial Justification.** A third alternative is to restrict the use of protected variables, such as race, religion, and political affiliation, and to further limit the use of rating variables to only those that are actuarially justified, that is, statistically discriminatory. This is the case of the U.S. rules on unfair discrimination where variables induce price differentials that "reflect equitably the differences in expected losses and expenses." There are usually data disclosure requirements for actuarial justification. For example, in the case of insurance based on age discrimination in Australia, "the Commission and the President can require the disclosure of the source of the actuarial or statistical data". And for the case of insurance based on sex discrimination in Australia, there are clients related data disclosure provisions, see the Appendix of Frees and Huang (2021) for further details.
4. **Limited Prohibitions.** A fourth strategy is to only restrict the use of protected variables (including their proxies) such as gender. This is the model followed by the European rules. However, as noted in Frees and Huang (2021), European regulation permits the use of risk factors correlated with gender as long as they are risk factors in their own right.
5. **No Restrictions.** At the other end of the spectrum, an option is to have no prohibitions. This is the case for most lines of commercial insurance.

Within this broad spectrum, there are many variations that regulators could consider particularly for the third and fourth strategies that permit some insurer discretion. One possible solution is to focus on transparency-oriented reforms that require insurers to disclose information on how their algorithms are working and possibly the sources of their data.

## 6.2 Linear Model Strategies

Another possible solution is to require insurers to use only variables that contain no protected information. To see how this might work in practice, consider data that we can represent as  $y$ , an outcome of interest such as an insurance claim and a set of predictor or rating variables. Further decompose the rating variables into components, those that are permitted, or not protected, by the regulators  $X_{NP}$  (non-protected variables) and those that are potentially contentious or protected,  $X_P$ . Following the work of Aseervatham, Lex, and Spindler (2016), we can consider  $y$  to be an auto insurance claim,  $X_P$  an indicator for gender, and  $X_{NP}$  to be a collection of other variables that includes age, type of car, location, and so forth.

In absence of regulatory restrictions, the actuary would consider all variables. For ease of interpretation, consider a “full” linear regression model,  $y = \mathbf{1} \beta_0 + X_P \beta_1 + X_{NP} \beta_2 + \epsilon$ , with predictors of the form

$$\hat{y}_{full} = \mathbf{1} b_0 + X_P b_1 + X_{NP} b_2. \quad (1)$$

In the presence of regulatory restrictions, the actuary could consider a restricted model,  $y = \mathbf{1} \beta_{0,1} + X_{NP} \beta_{2,1} + \epsilon_1$ , with predictors of the form

$$\hat{y}_{restricted} = \mathbf{1} b_{0,1} + X_{NP} b_{2,1}. \quad (2)$$

For a sensitive variable  $X_P$  to be the subject of contention, it often is correlated with an outcome  $y$ . So, one might expect for there to be a drop in the predictive ability when moving from the full information predictors  $\hat{y}_{full}$  to the restricted ones,  $\hat{y}_{restricted}$ . In point of practice, often there is a strong relationship of  $X_P$  with the other predictor variables  $X_{NP}$ . When  $X_P$  is dropped, the other variables serve as proxies for the omitted variable. For example, in their study, Aseervatham, Lex, and Spindler (2016) found that this was largely the case, the interesting exceptions being for younger and older drivers. As another point of practice, if a variable is dropped (such as gender), then it is likely that insurers may seek to incorporate new variables that also serve as proxies (“gender-like”) for the protected variables.

When a variable is dropped, the impact of the other factors changes, as quantified by the regression coefficients moving from  $b_2$  to  $b_{2,1}$ . This is not always desirable and so, to mitigate this drawback, Pope and Sydnor (2011) proposed an alternative predictor

$$\hat{y}_{PS} = \mathbf{1} b_0 + \bar{X}_P b_1 + X_{NP} b_2, \quad (3)$$

where  $\bar{X}_P$  is the average over the protected variables. Here, the coefficients  $b_0$ ,  $b_1$ , and  $b_2$  are from the full model. These predictors are blind to the protected variables in that two

individuals who differ only in their protected characteristics will receive the same predicted value from the model. However, as with  $\hat{y}_{full}$  and  $\hat{y}_{restricted}$ ,  $\hat{y}_{PS}$  can be correlated with the protected variables. In words, the average predicted values will vary across protected groups because of differences in other characteristics across groups. Please note, the Pope-Sydnor model is not restricted to linear models, as shown in Lindholm et al. (2020).

As another option, there may be instances when the actuary would like to have a rating scheme that is totally unrelated to any sensitive or protected variables. For example, one can imagine using only that information that is uncorrelated to a set of protected variables under contention.

To this end, create a set of variables that are uncorrelated to  $X_P$  by defining  $X_1 = (\mathbf{1} \ X_P)$ , the projection matrix  $Q = I - X_1(X_1'X_1)^{-1}X_1'$  and the transformed variables  $X_{NP}^* = QX_{NP}$ . Then, with the new variables, one uses the usual least square procedures to get  $b_2^* = (X_{NP}^*X_{NP}^*)^{-1}X_{NP}^*y$ . Some standard matrix algebra shows that  $b_2^* = b_2$ , that is, the regression coefficients from the transformed variables equals the regression coefficients in the full model. See for example Frees (2009), page 141. From this, the predictors are

$$\hat{y}_{FH} = \mathbf{1} \bar{y} + X_{NP}^* b_2. \quad (4)$$

By construction, these predictors are uncorrelated with the sensitive, protected, variables.

### 6.3 Empirical Example

To see how these strategies might work in a real insurance context, we analyze 4624 claims from Australian automobile insurance drawn from De Jong and Heller (2008). So that our work can be easily replicated, we use the data from the R package `CASdatasets` that slightly differs from the de Jong and Heller book in the coding of the variables. Statistical code, using the freely available software R is in the Appendix of a supporting document Frees and Huang (2021).

For this analysis of claims severity, `ClaimAmount`, we focus on a potential protected variable, `Female`, indicating if the policyholder is female. Other variables relevant to claims severity for these data are `VehValue`, the vehicle value in thousand of Australian dollars, and `DrivAge`, the age and employment status of the policyholder. A preliminary examination of the data (not included here but available, for example, in De Jong and Heller (2008)), show that the distribution of claims is skewed. From this and customary industry practice, we fit a gamma distribution with a logarithmic link. The analysis summarized in Table 1 shows that `Female` is an important predictor of claim amount.

This model could be readily used for predicting claims severity. For illustrative pur-

Table 1: Gamma Regression Model 1 Summary

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	7.984	0.093	85.580	0.000
VehValue	-0.012	0.022	-0.521	0.602
DrivAgeold people	-0.411	0.104	-3.946	0.000
DrivAgeolder work. people	-0.292	0.093	-3.134	0.002
DrivAgeoldest people	-0.353	0.119	-2.967	0.003
DrivAgeworking people	-0.293	0.093	-3.144	0.002
DrivAgeyoung people	-0.199	0.096	-2.073	0.038
Female	-0.179	0.052	-3.450	0.001

poses, the left-hand panel of Figure 1 shows the distribution of fits for the same portfolio of policyholders used to fit the data; the right-hand panel shows the distribution by gender.

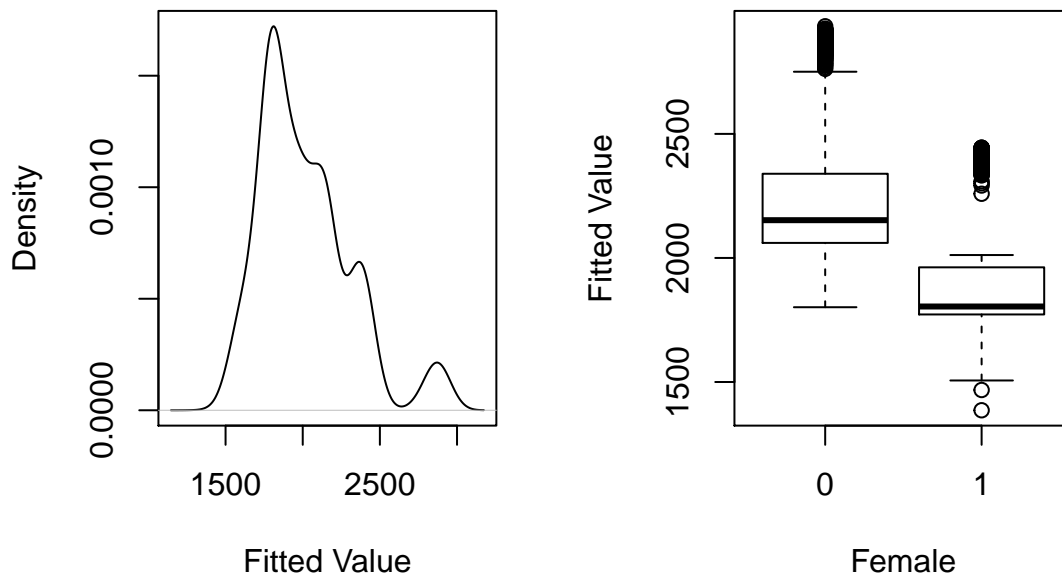


Figure 1: Distribution of Claim Amounts

The right-hand panel of Figure 1 shows significant differences by gender. To address this, the model was re-fit excluding gender and the fitted values are labeled as Model 2. As another alternative, we fit a model using the orthogonalized versions of the VehValue and DriveAge variables, making each variable uncorrelated with Female. This results in Model 3. Then, we developed a proxy for the probability of being female, using automatic variable selection techniques, with VehValue and DriveAge as inputs but also including additional variables in the dataset that were not helpful predictors of ClaimAmount. We added this



Table 2: Gamma Regression Model Summary

	M.1 Coef	M.1 t	M.2 Coef	M.2 t	M.3 Coef	M.3 t	M.4 Coef	M.4 t	M.5 Coef	M.5 t
(Intercept)	7.98	85.58	7.88	89.45	7.61	156.37	8.16	50.19	7.89	54.92
VehValue	-0.01	-0.52	0.00	-0.13	0.00	-0.14	-0.02	-0.88	-0.02	-0.91
DrivAge.old people	-0.41	-3.95	-0.42	-4.01	-0.43	-4.06	-0.45	-4.22	-0.45	-4.26
DrivAge.older work people	-0.29	-3.13	-0.31	-3.24	-0.30	-3.19	-0.30	-3.18	-0.30	-3.14
DrivAge.oldest people	-0.35	-2.97	-0.34	-2.85	-0.35	-2.93	-0.38	-3.13	-0.39	-3.21
DrivAge.working people	-0.29	-3.14	-0.32	-3.39	-0.31	-3.27	-0.30	-3.21	-0.29	-3.11
DrivAge.young people	-0.20	-2.07	-0.21	-2.20	-0.21	-2.16	-0.22	-2.25	-0.21	-2.22
Female	-0.18	-3.45	NA	NA	NA	NA	-0.42	-2.08	-0.43	-2.12
AIC	79365.63	NA	79391.08	NA	79390.49	NA	79382.40	NA	79381.38	NA

predictor to Model 2 and to Model 3, resulting in Models 4 and 5, respectively. Finally, we developed the Pope-Sydnor predictors from Model 1, the results of these fits are labeled as Model 6. This development is detailed in an Appendix of Frees and Huang (2021).

Table 2 summarizes results from the first five models (the repetition of Model 1 is included for comparison purposes). From Model 1, females have significantly lower claims. Similarly, from Models 4 and 5, a higher probability of being female implies a lower expected claim amount. From the AIC goodness of fit statistics, Model 1 is the best fit. Models 2 and 3 are similar and exhibit a markedly worse fit than Model 1. It is interesting that Model 3 does not perform that much worse than Model 2; that is, removing effects of gender from the other predictor variables does not do that much damage to the overall model fit. Models 2 and 3 are significantly improved by including the proxy for being female, as shown in Models 4 and 5.

Table 3 shows the means of the fitted values by gender under each model, including the Pope-Sydnor predictors. These fits are rescaled so that they have the same mean, thus promoting comparability. Figure 2 expands upon this by showing the corresponding distributions. The base Model 1 displays the biggest discrepancy between male and female distributions, suggesting that corresponding prices calculated using this procedure would be the largest among the alternatives considered. Model 2 is the only fitting procedure that does not require knowledge of the protected variable Female. Fits from Model 3 are very similar to those of Model 2 (it turns out that the Spearman correlation between fits is 0.994); further, Model 3 is attractive because it only uses predictors that are uncorrelated with the protected variable Female. The fits from Model 4 uses the same base variables as Model 1 but replaces the protected variable with a proxy. Coefficients of the proxy are determined using the protected variable but, once the coefficients have been determined, the proxy depends only on known covariates, not the protected variable. As with the comparison between Models 2 and 3, Model 5 is similar to Model 4 but uses only covariates uncorrelated to the protected variable (in addition to the proxy variable). Using Model 6, two individuals who differ only in gender will have the same fitted values from the model. However, it does

Table 3: Comparison of Means by Predictors and Gender

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Male	2215.56	2011.80	2013.42	2045.96	2050.74	2007.63
Female	1863.01	2015.06	2013.85	1989.57	1986.00	2018.17

not ensure equal values across the protected groups (two genders) as shown Table 3.

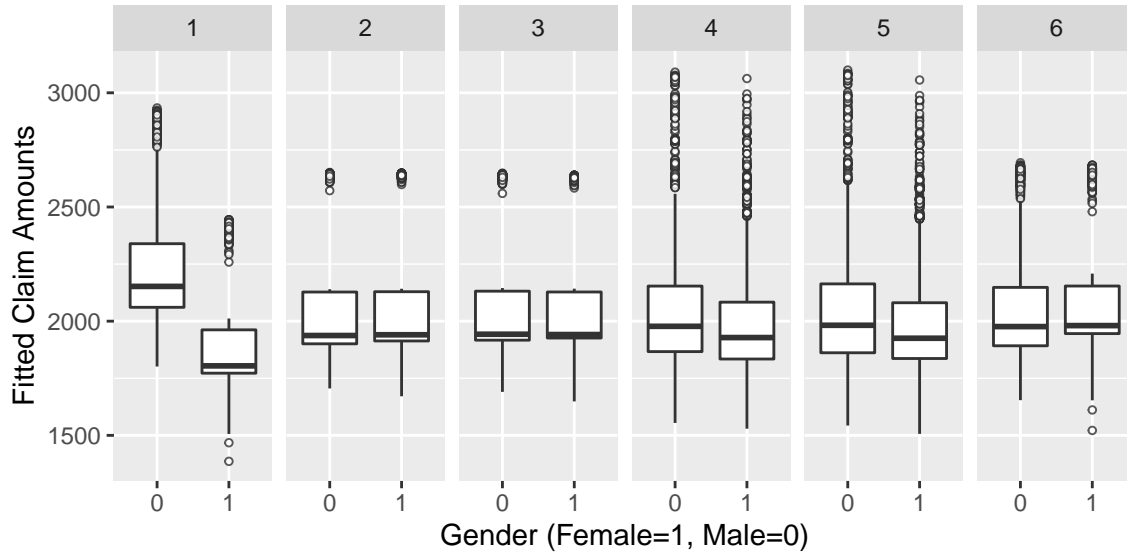


Figure 2: Boxplots of Fitted Claim Amounts by Model and Gender

#### 6.4 Machine Learning Approaches

Recent years have seen an increasing trend in using big data and machine learning techniques in various actuarial practices, as introduced in Section 5.5. They are usually praised for superior out-of-sample forecasting performance, but can also be opaque in insurance discrimination. Mehrabi et al. (2019) identified two potential sources of unfairness in machine learning: biases in data and algorithms. Data (especially big data) can be heterogeneous and creates bias in many different ways, which may lead to unfair results when a model learned from biased data. Algorithms may also lead to unfair decisions even when data are unbiased. To the best of our knowledge, there has been limited research in the actuarial/insurance literature discussing how to measure and manage discrimination using machine learning approaches. For example, Loi and Christen (2019) provide an ethical analysis of private insurance discrimination and fairness in machine learning. They distinguish morally permissible and impermissible forms of statistical discrimination in private insurance and derive some ethical implications for the use of machine learning techniques in the insurance context.

Although the motivating applications are typically not insurance related, the machine learning field has seen an explosion of research on fairness. See, for example, the survey papers Zliobaite (2015), Romei and Ruggieri (2014), Mehrabi et al. (2019) and Chouldechova and Roth (2018). This literature provides discussions and debates on how to define fairness of predictive models and how to measure the performance in terms of discrimination. For example, Kleinberg, Mullainathan, and Raghavan (2016) formalize three core fairness conditions in algorithmic classification that correspond to notions of fairness: calibration within groups, balance for the negative class, and balance for the positive class. They find that except in highly constrained cases, there is no method that can satisfy the three conditions simultaneously. The results suggest thinking about the trade-offs between the notions of fairness.

Following Mehrabi et al. (2019), we summarize three ways for discrimination prevention: pre-processing, in-processing, and post-processing. Data pre-processing removes the discrimination information from the historical data (target or input variables) and then applies regular machine learning approaches for model estimation, for example see Kamiran and Calders (2012) and Calders and Žliobaitė (2013). Most of the linear model strategies introduced in Sections 6.2 belong to this category. In-processing techniques modify the learning algorithms by incorporating changes into the objective function or adding additional constraints to remove discrimination in the model learning phase, for example, see Kamishima et al. (2012). Post-processing modifies a fitted regular model to remove discrimination. For example, Kamiran, Calders, and Pechenizkiy (2010) post-processing decision trees with discrimination-aware pruning and relabeling of tree leaves. The discrimination-free pricing model introduced in Pope and Sydnor (2011) also belongs to this category.

Machine learning approaches generally require knowledge of both protected and non-protected variables. However, for legal or commercial reasons organizations (including insurers) may not hold data of protected variables, such as gender, race and ethnicity, which poses challenges to mitigating discrimination, c.f., Miller (2009). Veale and Binns (2017) introduces and discusses three potential approaches to deal with this problem, including (1) having trusted third parties to store data necessary for incorporating fairness constraints in modelling, (2) building collaborative online platforms to allow diverse organisations to share and access knowledge required to promote algorithmic fairness, and (3) using unsupervised learning and pedagogically interpretable algorithms to incorporate fairness hypotheses for further selective testing and exploration.

Despite the explosion of interest and volume of work that has been produced and published in recent years, the theory and application of discrimination-aware machine learning is still in a nascent state, especially in the context of insurance practice.

## 7 Conclusions

Understanding the insurance prohibitions landscape is important for actuaries and other financial analysts. Actuaries are heavily involved in setting of insurance prices. They are also often influential in determining the scope of insurance contractual coverages as well as whom the company insures, both initially and at renewal. Our position is not that actuaries should dictate whether or not use of information should be restricted or prohibited. Rather, choices regarding insurance prohibitions involve policy choices that should also involve legal and economic scholars, as well as government representatives and advocates for the industry and for consumers. Actuaries can make important contributions to these discussions by quantifying the financial impact of policy alternatives. This article helps actuaries to present financial cost recommendations in a meaningful way by summarizing different perspectives that other participants may entertain when considering insurance prohibitions.

## References

- Adams, Paul, Robert Baker, Stefan Hunt, Darragh Kelly, and Alessandro Nava (2015), "Encouraging consumers to act at renewal: Evidence from field trials in the home and motor insurance markets," *Financial Conduct Authority (FCA) Occasional Paper*, (12). Accessed <https://www.fca.org.uk/publication/occasional-papers/occasional-paper-12.pdf>. Retrieved on Nov 20, 2019.
- Arizona Governing Committee for Tax Deferred Annuity and Deferred Compensation Plans versus Norris (1983), "Supreme court decision: Arizona governing committee for tax deferred annuity and deferred compensation plans versus norris," 463 (No. 82-52), 1073.
- Arrow, Kenneth J (1963), "Uncertainty and the welfare economics of medical care," *The American Economic Review*, 53, 941–73.
- Aseervatham, Vijay, Christoph Lex, and Martin Spindler (2016), "How do unisex rating regulations affect gender differences in insurance premiums?" *The Geneva Papers on Risk and Insurance-Issues and Practice*, 41 (1), 128–60.
- Avraham, Ronen (2018), "Discrimination and insurance," in *The Routledge Handbook Of The Ethics Of Discrimination*, K. Lippert-Rasmussen, ed., Routledge. Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3089946](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3089946). Retrieved on Oct 28, 2019.
- (2019), "Database of state tort law reforms (6.1)," *U of Texas Law, Law and Econ Research Paper*, (e555). Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=902711](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=902711). Retrieved on Nov 9, 2019.
- , Kyle D Logue, and Daniel Schwarcz (2014), "Understanding insurance antidiscrimination law," *Southern California Law Review*, 87, 195–274.
- Ayuso, Mercedes, Montserrat Guillén, and Ana María Pérez-Marín (2016), "Telematics and gender discrimination: Some usage-based evidence on whether men's risk of accidents differs from women's," *Risks*, 4 (2), 10.
- Baker, Tom (2002), "Risk, insurance, and the social construction of responsibility," *Embracing Risk: The Changing Culture of Insurance and Responsibility*, 33–51.
- Barocas, Solon and Andrew D Selbst (2016), "Big data's disparate impact," *California Law Review*, 104, 671–732.
- Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace (2019), *Consumer-lending discrimination in the fintech era*, National Bureau of Economic Research. Accessed [http://www.phd-finance.uzh.ch/dam/jcr:139b84f5-afb8-40ac-b255-83b5386431ff/FS\\_spring18\\_paper\\_Morse.pdf](http://www.phd-finance.uzh.ch/dam/jcr:139b84f5-afb8-40ac-b255-83b5386431ff/FS_spring18_paper_Morse.pdf). Retrieved on Dec 28, 2019.

- Bauer, Daniel, Richard D. Phillips, and George H. Zanjani (2013), “Financial pricing of insurance,” in *Handbook Of Insurance*, Springer, 627–45.
- Blackmon, B Glenn and Richard Zeckhauser (1991), “Mispriced equity: Regulated rates for auto insurance in massachusetts,” *The American Economic Review*, 81 (2), 65–69.
- Born, Patricia (2019), “Genetic testing in underwriting: Implications for life insurance markets.” *Journal of Insurance Regulation*, 38 (5).
- Borselli, Angelo (2011), “Insurance rates regulation in comparison with open competition,” *Connecticut Insurance Law Journal*, 18, 109–68.
- Braun, Alexander, Hato Schmeiser, and Przemysław Rymaszewski (2015), “Stock vs. Mutual insurers: Who should and who does charge more?” *European Journal of Operational Research*, 242 (3), 875–89.
- Brockett, Patrick L and Linda L Golden (2007), “Biological and psychobehavioral correlates of credit scores and automobile insurance losses: Toward an explication of why credit scoring works,” *Journal of Risk and Insurance*, 74 (1), 23–63.
- Calders, Toon and Indrė Žilobaitė (2013), “Why unbiased computational processes can lead to discriminative decision procedures,” *Discrimination and privacy in the information society*, 43–57.
- CAS, Casualty Actuarial and Statistical Task Force (2015), “Price optimization white paper,” National Association of Insurance Commissioners. Accessed [https://www.naic.org/documents/committees\\_c\\_catf\\_related\\_price\\_optimization\\_white\\_paper.pdf](https://www.naic.org/documents/committees_c_catf_related_price_optimization_white_paper.pdf). Retrieved on Nov 9, 2019.
- CAS, Casualty Actuarial Society Committee on Ratemaking (2014), “Price optimization overview,” Casualty Actuarial Society. Accessed <http://www.casact.org/area/rate/price-optimization-overview.pdf>. Retrieved on Nov 11, 2019.
- Cather, David A (2018), “Cream skimming: Innovations in insurance risk classification and adverse selection,” *Risk Management and Insurance Review*, 21 (2), 335–66.
- Chouldechova, Alexandra and Aaron Roth (2018), “The frontiers of fairness in machine learning,” *arXiv preprint arXiv:1810.08810*.
- De Jong, Piet and Gillian Z. Heller (2008), *Generalized Linear Models For Insurance Data*, Cambridge University Press Cambridge.
- Dionne, Georges and Casey Rothschild (2014), “Economic effects of risk classification bans,” *The Geneva Risk and Insurance Review*, 39 (2), 184–221.
- ESL, Emergency Services Levy Big Data Project Team (2018), “Monitoring of insurance premiums in the world of big data,” Emergency Services Levy Insurance Monitor. Accessed [www.eslinsurancemonitor.nsw.gov.au](http://www.eslinsurancemonitor.nsw.gov.au). Retrieved on Nov 9, 2019.
- European Union (2000), “Council directive 2000/43/ec,” *Official Journal of the European Communities*, Publications Office of the European Union. Accessed <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1578080397727&uri=CELEX:32000L0043>. Retrieved on Jan 3, 2020.
- (2004), “Council directive 2004/113/ec of 13 december 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services,” *Official Journal of the European Union*, Publications Office of the European Union. Accessed <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32004L0113>. Retrieved on Jan 3, 2020.
- (2012), “Guidelines on the application of council directive 2004/113/ec to insurance, in the light of the judgment of the court of justice of the european union in case c-236/09 (test-achats),” *Official Journal of the European Union*, Publications Office of the European Union. Accessed <https://op.europa.eu/en/publication-detail/-/publication/f1035369-1189-41d3-937c-aec8aa4f6cfd/language-en/format-HTML/source-112025857>. Retrieved on Jan 3, 2020.
- Fannin, Brian A. (2020), “COVID-19: The property-casualty perspective,” Casualty Actuarial Society. Accessed [https://www.casact.org/research/COVID-19\\_The\\_PC\\_Perspective\\_3-27-2020.pdf](https://www.casact.org/research/COVID-19_The_PC_Perspective_3-27-2020.pdf). Retrieved on April 20, 2020.
- FCA (2019), “General insurance pricing practices: Interim report. Annex 3: International comparisons.” *Financial Conduct Authority (FCA) MS*, (18/1. 2). Accessed <https://www.fca.org.uk/publication/market-studies/ms18-1-2-interim-report.pdf>. Retrieved on Nov 20, 2019.
- Frees, Edward W. (2009), *Regression Modeling With Actuarial And Financial Applications*, Cambridge University Press.
- Frees, Edward W and Fei Huang (2021), “Online supplement to: The discriminating (pricing) actuary.”
- Frezal, Sylvestre and Laurence Barry (2019), “Fairness in uncertainty: Some limits and misinterpretations of actuarial fairness,” *Journal of Business Ethics*, 1–10.
- Friedland, Jacqueline (2013), *Fundamentals Of General Insurance Actuarial Analysis*, Society of Actuaries.

- FTC, Federal Trade Commission and others (2007), “Credit-based insurance scores: Impacts on consumers of automobile insurance,” in *A Report To Congress By The Federal Trade Commission, July 2007*.
- Joly, Yann, Charles Dupras, Miriam Pinkesz, Stacey A Tovino, and Mark A Rothstein (2020), “Looking beyond gina: Policy approaches to address genetic discrimination,” *Annual Review of Genomics and Human Genetics*, 21, 491–507.
- Kamiran, Faisal and Toon Calders (2012), “Data preprocessing techniques for classification without discrimination,” *Knowledge and Information Systems*, 33 (1), 1–33.
- , ———, and Mykola Pechenizkiy (2010), “Discrimination aware decision tree learning,” in *2010 Ieee International Conference On Data Mining*, IEEE, 869–74.
- Kamishima, Toshihiro, Shotaro Akaho, Hideki Asoh, and Jun Sakuma (2012), “Fairness-aware classifier with prejudice remover regularizer,” in *Joint European Conference On Machine Learning And Knowledge Discovery In Databases*, Springer, 35–50.
- Klein, R (2017), “Genetics and life insurance. A view into the microscope of regulation,” *Geneva, Switzerland: The Geneva Association*. Accessed <https://www.genevaassociation.org/research-topics/global-ageing/genetics-and-life-insurance-view-microscope-regulation>. Retrieved on Dec 11, 2019.
- Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan (2016), “Inherent trade-offs in the fair determination of risk scores,” *arXiv preprint arXiv:1609.05807*.
- Lehtonen, Turo-Kimmo and Jyri Liukko (2011), “The forms and limits of insurance solidarity,” *Journal of Business Ethics*, 103 (1), 33–44.
- Lemaire, Jean (2002), “Why do females live longer than males?” *North American Actuarial Journal*, 6 (4), 21–37.
- Lindholm, Mathias, Ronald Richman, Andreas Tsanakas, and Mario V Wuthrich (2020), “Discrimination-free insurance pricing,” *Available at SSRN*.
- Loi, Michele and Markus Christen (2019), “Insurance discrimination and fairness in machine learning: An ethical analysis.” Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3438823](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3438823). Retrieved on Dec 28, 2019.
- Lombardo, Mark (2018), “The impact of genetic testing on life insurance mortality,” Society of Actuaries. Accessed <https://www.soa.org/resources/research-reports/2018/impact-genetic-testing/>.
- Los Angeles Department of Water and Power versus Manhart (1978), “Supreme court decision: Los angeles department of water and power versus manhart,” 435 (No. 76-1810), 702.
- Lukacs, Peter, Leslie Neubecker, and Philip Rowan (2016), “Price discrimination and cross-subsidy in financial services,” *FCA Occasional Paper*, (22). Accessed <https://www.fca.org.uk/publication/occasional-papers/op16-22.pdf>. Retrieved on Nov 13, 2019.
- Maysami, Ramin Cooper and W Jean Kwon (1999), “An analysis of islamic takaful insurance.” *Journal of Insurance Regulation*, 18 (1).
- McCarthy, David D and John A Turner (1993), “Risk classification and sex discrimination in pension plans,” *Journal of Risk and Insurance*, 85–104.
- Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan (2019), “A survey on bias and fairness in machine learning,” *arXiv preprint arXiv:1908.09635*.
- Meyers, Gert and Ine Van Hoyweghen (2018), “Enacting actuarial fairness in insurance: From fair discrimination to behaviour-based fairness,” *Science as Culture*, 27 (4), 413–38.
- Miller, Michael J (2009), “Disparate impact and unfairly discriminatory insurance rates,” in *Casualty Actuarial Society E-Forum, Winter 2009*, 276.
- Morris, Darcy Steeg, Daniel Schwarcz, and Joshua C Teitelbaum (2017), “Do credit-based insurance scores proxy for income in predicting auto claim risk?” *Journal of Empirical Legal Studies*, 14 (2), 397–423.
- NAIC (2010), “Property and casualty model rating law (file and use version),” National Association of Insurance Commissioners. Accessed <https://www.naic.org/store/free/GDL-1775.pdf>. Retrieved on Nov 9, 2019.
- Peto, Richard (1994), “Smoking and death: The past 40 years and the next 40,” *British Medical Journal*, 309 (6959), 937–39.
- Pope, Devin G and Justin R Sydnor (2011), “Implementing anti-discrimination policies in statistical profiling models,” *American Economic Journal: Economic Policy*, 3 (3), 206–31.
- Prince, Anya ER (2019), “Political economy, stakeholder voices, and saliency: Lessons from international policies regulating insurer use of genetic information,” *Journal of Law and the Biosciences*, 5 (3), 461–94.
- Prince, Anya and Daniel Schwarcz (2020), “Proxy discrimination in the age of artificial intelligence and big data,” *Iowa Law Review*. Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3347959](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3347959). Retrieved on Dec 21, 2019.

Richman, Ronald, Nicolai von Rummell, and Mario V. Wuthrich (2019), “Believing the bot - model risk in the era of deep learning.” Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3444833](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3444833). Retrieved on 18 Feb 2021.

Romei, Andrea and Salvatore Ruggieri (2014), “A multidisciplinary survey on discrimination analysis,” *The Knowledge Engineering Review*, 29 (5), 582–638.

Schmeiser, Hato, Tina Störmer, and Joël Wagner (2014), “Unisex insurance pricing: Consumers’ perception and market implications,” *The Geneva Papers*, 39, 322–350.

Schwarcz, Daniel and Peter Siegelman (2017), “Law and economics of insurance,” in *The Oxford Handbook Of Law And Economics: Volume 2: Private And Commercial Law*, Oxford University Press, 481–508.

Skipper, Harold D and Robert W Klein (2000), “Insurance regulation in the public interest: The path towards solvent, competitive markets,” *The Geneva Papers on Risk and Insurance-Issues and Practice*, 25 (4), 482–504.

Swedloff, Rick (2014), “Risk classification’s big data (r) evolution,” *Conn. Ins. LJ*, 21, 339.

Tennyson, Sharon L (2007), “Efficiency consequences of rate regulation in insurance markets,” *Networks Financial Institute Policy Brief*, (2007-PB), 03. Accessed [https://www.indstate.edu/business/sites/business.indstate.edu/files/Docs/2007-PB-03\\_Tennyson.pdf](https://www.indstate.edu/business/sites/business.indstate.edu/files/Docs/2007-PB-03_Tennyson.pdf). Retrieved on Jan 20, 2020.

Thomas, R Guy (2012), “Non-risk price discrimination in insurance: Market outcomes and public policy,” *The Geneva Papers on Risk and Insurance-Issues and Practice*, 37 (1), 27–46.

Thouvenin, Florent, Fabienne Suter, Damian George, and Rolf H. Weber (2019), “Big data in the insurance industry: Leeway and limits for individualising insurance contracts,” *Journal of Intellectual Property, Information Technology and E-Commerce Law*, 10 (2). Accessed <https://www.jipitec.eu/issues/jipitec-10-2-2019/4916>. Retrieved on Nov 5, 2019.

Veale, Michael and Reuben Binns (2017), “Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data,” *Big Data & Society*, 4 (2), 2053951717743530.

Vukcevic, Damjan and Jessica Chen (2018), “Advances in genetics and their impact on life insurance.” Accessed <https://www.actuaries.asn.au/Library/Events/FSF/2018/VukcevicChenPaper.pdf>. Retrieved on Dec 11, 2019.

Werner, Geoff and Claudine Modlin (2016), *Basic Ratemaking, Fifth Edition*, Casualty Actuarial Society. Accessed [https://www.casact.org/library/studynotes/werner\\_modlin\\_ratemaking.pdf](https://www.casact.org/library/studynotes/werner_modlin_ratemaking.pdf).

Zliobaite, Indre (2015), “A survey on measuring indirect discrimination in machine learning,” *arXiv preprint arXiv:1511.00148*.

Online Supplement to:  
**The Discriminating (Pricing) Actuary**

**Edward W. (Jed) Frees** *University of Wisconsin - Madison, Australian National University*

**Fei Huang** *University of New South Wales*

*Date of this Draft: March 16, 2021*

---

*Abstract:* This online supplement provides an extended abstract, a description of regulation by jurisdiction, several appendices, and an annotated bibliography.

*Keywords:* Actuarial fairness; disparate impact; proxy discrimination; unisex classification; credit-based insurance scores; price optimization; genetic testing; big data; COVID-19

---



## Contents

<b>Extended Abstract</b>	<b>2</b>
<b>1 Regulation by Jurisdiction</b>	<b>4</b>
1.1 U.S. Insurance Discrimination Legal Environment . . . . .	4
1.1.1 U.S. Federal Laws and Regulations . . . . .	4
1.1.2 U.S. State Laws and Regulations . . . . .	5
1.2 European Union . . . . .	6
1.2.1 EU Directives . . . . .	6
1.2.2 EU Member Countries Laws and Regulations . . . . .	7
1.3 China . . . . .	7
1.4 Japan . . . . .	9
1.5 Australia . . . . .	10
1.5.1 Australian Federal Laws and Regulations . . . . .	10
1.5.2 Australian State Laws and Regulations . . . . .	11
<b>2 Appendices</b>	<b>13</b>
2.1 Appendix: A Proxy for Being Female . . . . .	13
2.2 Appendix: Australian Regulations . . . . .	14
2.3 Appendix: Chinese Regulations . . . . .	15
2.4 R Code . . . . .	16
<b>Annotated Bibliography</b>	<b>20</b>

## Extended Abstract

The insurance industry is built on risk classification, grouping insureds into homogeneous classes. Through actions such as underwriting, pricing and so forth, it differentiates, or discriminates, among insureds. Actuaries have responsibility for pricing insurance risk transfers and are intimately involved in other aspects of company actions and so have a keen interest in whether or not discrimination is appropriate from both company and societal viewpoints. This paper reviews social and economic principles that can be used to assess the appropriateness of insurance discrimination. Discrimination issues vary by the line of insurance business and by the country and legal jurisdiction. This paper examines social and economic principles from the vantage of a specific line of business and jurisdiction; these vantage points provide insights into principles. To sharpen understanding of the social and economic principles, this paper also describes discrimination considerations for prohibitions based on diagnosis of COVID-19, the pandemic that swept the globe in 2020.

Insurance discrimination issues have been an important topic for the insurance industry for decades and is evolving in part due to insurers' extensive use of *Big Data*, that is, the increasing capacity and computational abilities of computers, availability of new and innovative sources of data, and advanced algorithms that can detect patterns in insurance activities

that were previously unknown. On the one hand, the fundamental issues of insurance discrimination have not changed with Big Data; one can think of credit-based insurance scoring and price optimization as simply forerunners of this movement. On the other hand, issues regarding privacy and use of algorithmic proxies take on increased importance as insurers' extensive use of data and computational abilities evolve.

## 1 Regulation by Jurisdiction

### 1.1 U.S. Insurance Discrimination Legal Environment

The U.S. is the largest general insurance marketplace, to illustrate, it comprised about 37% of the 2017 world market (in terms of gross direct premiums, see Table 7 of GIAJ (2019), SwissRe (2018)). In the U.S., discrimination laws and regulations affecting insurance appear at both the federal (national) and state level.

#### 1.1.1 U.S. Federal Laws and Regulations

There is a host of federal laws directing towards mitigating discrimination in employment. These laws take a broad view of employment and include fringe benefits such as retirement benefits. As noted in Section 5.1 of Frees and Huang (2020), of Frees and Huang (2020), these laws, such as the 1964 Civil Rights Act, are now known to limit the use of gender in retirement benefits. For other limitations, the U.S. [Equal Employment Opportunity Commission](#), charged with enforcing these laws and regulations, decomposes discrimination into the following types: age, disability, equal pay/compensation, genetic information, harassment, national origin, pregnancy, race/color, religion, retaliation, sex, and sexual harassment.

In a separate but related set of laws, the [Genetic Information Nondiscrimination Act of 2008 \(GINA\)](#) protects individuals against employment discrimination on the basis of genetic information. As described in Section 5.3 of Frees and Huang (2020), GINA also prohibits covered health insurers from discriminating on the basis of genetic information that includes genetic test results, family medical history, and use of genetic services.

The U.S. [Department of Health and Human Services Office for Civil Rights](#) enforces federal civil rights laws, including conscience and religious freedom laws. In particular, for individual health insurers, this includes the Patient Protection and Affordable Care Act of 2010 (ACA), the comprehensive health care reform law that addresses health insurance coverage, health care costs, and preventive care. Section 1557 of this act states that individuals shall not be excluded from participation in, be denied the benefits of, or be subjected to discrimination on the grounds of race, color, national origin, sex, age, and disability.

The [Fair Housing Act](#), enacted in 1968, prohibits discrimination in housing-related activities and is now interpreted to include homeowners insurance. Specifically, it is illegal to discriminate because of the race, color, religion, sex, disability, familial status, or national origin of the owner and/or occupants of a dwelling. See, for example, the Department of Housing and Urban Development's [Implementation of the Fair Housing Act's Discriminatory Effects Standard](#).

In a similar vein, the [Equal Credit Opportunity Act](#) is a federal financial regulation

law enacted in 1974. The act prohibits discrimination on the basis of race, color, religion, national origin, sex, marital status, or age in credit transactions.

Despite these laws and the accompanying regulations, insurance discrimination is largely unregulated at the federal level, leaving the states as the primary regulators of insurer discrimination.

### *1.1.2 U.S. State Laws and Regulations*

Enacted in 1945, the McCarran-Ferguson Act clarified that the states regulate and tax the business of insurance. Efforts of the different states are coordinated by the National Association of Insurance Commissioners (NAIC). This organization is governed by the chief insurance regulators from the 50 states, the District of Columbia and five U.S. territories. The NAIC established standards and best practices that each state may choose to adopt, NAIC (2011). To complement this organization, a U.S. [Federal Insurance Office](#) was created in 2010 as an information gatherer to inform the U.S. Congress on insurance matters. The Federal Insurance Office was granted limited authority to enter into covered agreements with other nations on insurance regulatory matters and represents the U.S. with the International Association of Insurance Supervisors.

Despite the structure offered by the NAIC, Avraham, Logue, and Schwarcz (2014) documents a lack of uniformity in state insurance anti-discrimination regulations. Their work demonstrates that affirmative bans of insurer discrimination on the basis of potentially suspect policyholder traits are rare. To illustrate, from their study, “only nine states ban the use of age in auto insurance; only six states ban the use of genetic testing in disability insurance; and only two states ban . . . the use of location or zip code in property/casualty insurance.” This lack of uniformity is also observed by Werner and Modlin (2016), who state:

Some states have statutes prohibiting the use of gender in rating insurance while others permit it as a rating variable. As a result, an insurer writing in multiple states may include gender as a rating variable in those states where it is permitted, but not include it in a state that prohibits its use for rating. Some states may allow the use of a rating variable, but may place restrictions on its use. For example, some states allow credit score to be used for rating personal insurance for new business, but do not allow insurers to raise the rates for renewal risks should the insured’s credit worsen (although they may allow companies to reduce rates if the insured’s credit score improves). Some states also prohibit certain variables from use in the rating algorithm but allow their use in underwriting. Underwriting variables may be used to guide risk selection decisions, but could

also guide risk placement decisions.

## 1.2 *European Union*

The European Union (EU) is the second largest general insurance marketplace. It comprised about 22% of the 2017 world market (in terms of gross direct premiums, see Table 7 of GIAJ (2019), SwissRe (2018)). This does not include the United Kingdom which represented about 4% in 2017.

### 1.2.1 *EU Directives*

The EU is a political and economic union of 27 member countries, not counting the United Kingdom which left in January 2020. The European Commission serves as its executive branch and, among other duties, has the power to propose new laws. For insurance purposes, legislation takes the form of directives, a type of legislative instrument that allows members the freedom to interpret the demands of EU law within their own legislative traditions. For example, Section 5.1 of Frees and Huang (2020), remarked on the 2004 Gender directive with the related court actions. Thus, the requirements set out in the directives have to be interpreted and implemented in each member country by national legislation. In contrast, the U.S. NAIC promulgates model laws and regulations but the states are not required to enact them.

Insurance pricing was essentially deregulated in the EU in 1994 with the introduction of the Third Generation Insurance Directive. Prior to the directive, the European insurance business was mostly embedded in a dense regulatory network. Implementing the 1994 deregulation yielded intensive price competition, primarily lowering prices but in some cases, such as Italy, raising prices, see, for example Hussels, Ward, and Zurbruegg (2005) and FCA (2019).

Equality and respect for human dignity and human rights are among the core values of the EU, reflected in several EU directives targeting discrimination. Initially signed in 1957, [Article 10 of the Treaty on the Functioning of the European Union](#) states:

In defining and implementing its policies and activities, the Union shall aim to combat discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation.

However, as summarized in a report was commissioned by the European Commission, Civic Consulting (2010), there is a patchwork of legislative and regulatory measures across the member countries that deal with discrimination. The majority of countries prohibit any form

of discrimination, with no exceptions, on the grounds of racial/ethnic origin, religion/belief and sexual orientation. Treatment of age and disability is less uniform.

### *1.2.2 EU Member Countries Laws and Regulations*

Some member countries regulate conditions that affect the determination of insurance premiums. One example of this is the automobile insurance bonus-malus system in France. Although auto insurance rate levels are not subject to explicit constraints, the premiums are adjusted by a bonus-malus coefficient (set by law) that considers a driver's past experience. Another set of examples are the restrictions based on non-risk pricing, introduced in Sections 3.3 and 5.3 of Frees and Huang (2020). As described in FCA (2019), Belgium and Italy (as well as non-EU Switzerland) have introduced rules that should limit the use of non-risk based factors in the pricing of insurance.

### *1.3 China*

China is the third largest general insurance marketplace. It comprised about 10% of the 2017 world market (in terms of gross direct premiums, see Table 7 of GIAJ (2019), SwissRe (2018).

As discussed by Chen et al. (2014), there are three levels in the Chinese insurance regulatory system. First, the National People's Congress and its Standing Committee jointly exercise the power to enact laws in China, including the [Insurance Law](#). Second, the State Council, which is the chief administrative authority of China, develops administrative regulations. Third, the primary regulator of the private insurance market is [China Insurance Regulatory Commission \(CIRC\)](#), which is a functional department of the State Council. There are also industry organizations, including the Insurance Association of China (IAC) and "Intermediary Associations" for insurance intermediaries at the provincial and city levels. The IAC, founded in 2001, is a not-for-profit organization authorized by the CIRC and registered by the National Ministry of Civil Affairs. These organizations formulate industry standards and provide professional industry guidance to constrain unfair activities and strengthen self-discipline.

The Insurance Law was first approved and implemented in 1995 and then modified in 2002 and 2009. In the Insurance Law, the item related to insurance discrimination is Item 114, which states that "Insurance companies should obey the insurance regulations and set insurance premium rates fairly and reasonably, which should not harm the legitimate rights and interests of policyholders, insureds and beneficiaries." However, this law does not specify how to judge the "fairness" and "reasonableness" (Zhou (2014)).

There are no insurance discrimination prohibitions in the “[property insurance regulations](#)” and the “[life insurance regulations](#)”. The only discrimination related regulation is in ‘auto insurance rate making procedures and core content (2002)’, which requires insurance rating schedules must be reasonable without any discrimination content. However, it again does not specify nor explain the definition and extent of ‘discrimination’.

The Chinese auto insurance market is comprised of compulsory and voluntary components. The CIRC strictly regulates the pricing of compulsory auto insurance. The only pricing factors are vehicle type and past claim experience. Since 2015, there have been significant regulation changes for voluntary auto insurance. The auto insurance pricing formula depends on a benchmark pure risk premium (determined by the Insurance Association of China), a surcharge rate (management fee, no more than 35%, determined by insurers), and several rate adjustment coefficients, including ones for no claims discount, traffic violation, underwriting, and a “channel.” The benchmark pure risk premium uses only vehicle related information (usage and type) as pricing factors. According to CIRC [price guidelines](#), from 2017 the underwriting coefficient can range within (0.7/0.75/0.8, 1.15/1.25) and the independent channel coefficient within (0.7/0.75, 1.15/1.25) for different provinces. From April 2018, the CIRC approved a pilot experiment in Shaanxi, Guangxi and Qinghai to allow for independent pricing in those provinces. For details of the regulation changes of Chinese auto insurance, please refer to Appendix Section [2.3](#).

Chinese insurers use life tables published by CIRC for life insurance pricing. For example, the current version is [China Life Insurance Mortality Table \(2010-2013\)](#). The mortality rates in the insurance life tables are age, gender, and product specific. In particular, it has three product categories, including annuities/pension, term life insurance, and whole life insurance.

Like many other insurance markets, age is a commonly used pricing factor in the Chinese life insurance market. Chinese auto insurers also set different rates for different age groups in general insurance practice. For example, age groups below 25 and above 60 are usually treated as high-risk groups with 1.05 rate adjustment coefficient, while the age group between 30 and 40 can be treated as low-risk group with 0.95 risk adjustment coefficient. Gender has also been widely used as a pricing factor in the Chinese insurance market. For example, higher rates applies to males purchasing life insurance and females purchasing annuities. In business auto insurance, many insurers set the adjustment coefficients to be 0.95 for females and 1.00 for males with reference to the basic insurance rating table issued by the government. There are no laws or regulations prohibiting the use of disability, race, or religion related factors in insurance pricing.

#### 1.4 Japan

Japan is the fourth largest general insurance marketplace. It comprised about 5% of the 2017 world market (in terms of gross direct premiums, see Table 7 of GIAJ (2019), SwissRe (2018)).

For additional context, the regulator of the Japanese insurance market is the [Financial Services Agency](#). Insurers operating in Japan have formed the [General Insurance Association of Japan](#) whose objective is “to promote sound development and maintain reliability of the general insurance business in Japan, . . .” Among other activities, this organization actively participates in the International Association of Insurance Supervisors (IAIS) in its standard setting activities. Industry statistics and representative rates are gathered by the [General Insurance Rating Organization of Japan](#). This is a Japanese non-profit, private corporation with a membership comprised of general insurance and reinsurance companies.

The Japanese general insurance market is dominated by personal lines, with about 55% from the auto market, 7% from fire dwelling, and 12% from personal accident insurance, (based on 2017 premium statistics from GIAJ (2019)). The auto market is split into a compulsory liability component, known by the acronym CALI, at about 11% of market share, and the larger voluntary component at about 44% of market share. In terms of insurance pricing, CALI premium rates are calculated by the General Insurance Rating Organization of Japan. In this pricing scheme, automobiles are classified by type, size and use, and are designed to be non-profit. As the name suggests, anyone purchasing an automobile must purchase CALI insurance that pays damages for anyone killed or injured by an insured vehicle.

Voluntary insurance includes third-party liability coverage (bodily injury liability and property damage liability), self-incurred personal accident coverage, and own vehicle damage coverage. With respect to liability for bodily injury, voluntary insurance acts as excess cover for compulsory insurance. Prior to 1998, the auto insurance product was similar across firms. Rates and policies were strictly regulated, and underwriting was limited. In 1998, the marketplace was liberalized to a prior approval system where insurers submit rating plans to the Financial Services Agency who examines the rates on the basis of three principles of premium rates, that they be reasonable, adequate and not unfairly discriminatory (see Section 5.1 of Frees and Huang (2020),). Permissible rating factors include driver’s age, gender, driving history, usage, pattern of use, geography, vehicle type, vehicle safety features, and multi-car ownership. In particular, for age, “Any number of categories allowed, but differentials between the highest and lowest rated groups to be within a range of 300%”. As another example, for gender, “Segmentation allowed, but differentials between male and female to be within a range of 150%” (*Source*: GIAJ (2017)).



In Japan, mechanisms for mitigating discrimination appear to be different than in other marketplaces perhaps due to the relatively homogeneous culture. Another possible explanation is that industry organizations work diligently with member insurers to mitigate public concerns of potentially awkward situations prior to legislative action. As one example, the General Insurance Association of Japan publishes *Personal Information Protection Guidelines* to encourage member insurers to protect personal information. This is meant to be in accord with the Japanese Protection of Personal Information Act, brought into effect in 2003 and amended in 2017. To ensure consumer protection, during the advance product approval processes for personal lines, the basic principles of reasonableness, adequacy, and not being unfairly discriminatory are considered, along with regular risk category reviews. Algorithms are also required to be explained within the processes.

As another example, the [Life Insurance Association of Japan](#), another industry organization, has said its member firms do not use genetic information to make decisions about coverage, the subject of Section 5.4 of Frees and Huang (2020). Nonetheless, the association acknowledges that people with genetic diseases and private groups supporting them have voiced concerns about discrimination in the absence of laws prohibiting insurers from gaining such data. See Japan Times News (2019).

## 1.5 *Australia*

Australia is the tenth largest general insurance marketplace. It comprised about 2% of the 2017 world market (in terms of gross direct premiums, see Table 7 of GIAJ (2019), SwissRe (2018)). Australia has anti-discrimination legislation at the federal, state and territory levels. Commonwealth laws and the state/territory laws generally overlap and prohibit the same type of discrimination. Although there are differences in detail, all anti-discrimination legislation reflects the same paradigm for identifying unlawful discrimination Australian Law Reform Commission (2010).<sup>1</sup>

### 1.5.1 *Australian Federal Laws and Regulations*

There are four federal acts containing provisions relevant to discrimination in insurance, the [Age Discrimination Act 2004 \(ADA\)](#), the [Sex Discrimination Act 1984 \(Cth\)\(SDA\)](#), the [Racial Discrimination Act 1975 \(Cth\)\(RDA\)](#) and the [Disability Discrimination Act 1992 \(Cth\)\(DDA\)](#). Prohibitions on the use of genetic information in insurance are most likely to

---

<sup>1</sup>An unlawful discrimination act or omission must be based on one of the grounds or attributes set out in the legislation, such as sex, race or disability; fall within an area of activity set out in the legislation, such as employment or the provision of goods and services; result in some harm or less favorable treatment, whether by direct or indirect discrimination; and not fall within an exception, exemption or defense.

be brought under the DDA. These acts include general provisions applying to the supply of goods and services, including insurance. As an example, Section 24 of the DDA can be found in the Appendix Section 2.2.

Use of age, disability and gender may be allowed in certain circumstances in that ADA, DDA and SDA all contain exceptions relating to the provision of insurance. Insurance exceptions in the acts contain elements similar to those in Section 46 of the DDA, which applies both to the refusal to offer insurance (Section 46(1)) and to the terms or conditions on which it is offered (Section 46(2)). The details of the statement of Section 46(1) can be found in Appendix Section 2.2. There is no exception for discrimination in insurance based on race, even if one can establish the statistical and actuarial relevance of race information.

Australia does not have a stand-alone religious anti-discrimination law, similar to federal laws that protect against discrimination based on age, sex, race, and disability. However, Australia has a legal obligation to protect the right to religious freedom under a treaty known as the International Covenant on Civil and Political Rights. Domestically, there are a variety of laws that protect religious freedom. In August and December 2019, the Federal Attorney-General released the first and second exposure drafts of the Religious Discrimination Bill 2019 (Cth), which will make it unlawful to discriminate on the ground of religious belief or activity in a range of areas of public life. There is no insurance exception in those drafts.

### *1.5.2 Australian State Laws and Regulations*

Each state and territory in Australia has its own anti-discrimination acts with its own insurance exceptions<sup>2</sup>.

In Australian insurance practice, age is commonly used to discriminate, or differentiate, treatment among insureds with few regulatory constraints for the personal lines. For example, auto insurers use the age of the owner/driver of the car as a pricing factor. Life insurers consider the age of the policyholder as a pricing factor. They may also have age constraints to access specific insurance contracts. For example, it is common to require ages 16 - 65 in order to apply for the life cover and to require ages 16 - 55 to apply for critical illness cover or permanently unable to work cover.

Gender and disability are also commonly used pricing factors for the personal lines. For example, auto insurers use the gender and disability of owners/drivers of the car as pricing factors. Life insurers apply the gender and state of health of the policyholder as pricing factors.

---

<sup>2</sup>Discrimination Act 1991 (ACT) s 28; Anti-Discrimination Act 1992 (NT) s 49; Anti-Discrimination Act 1977 (NSW) s 49Q; Anti-Discrimination Act 1991 (Qld) ss 74, 75; Equal Opportunity Act 1984 (SA) s 85; Anti-discrimination Act 1998 (Tas) s 44; Equal Opportunity Act 1995 (Vic) s 43; Equal Opportunity Act 1984 (WA) s 66T.

In the Australian auto market, [Compulsory Third Party \(CTP\) insurance](#) (also called the greenslip) is required in all states and territories of Australia. The CTP insurance provides compensation for people injured or killed when the insured vehicle is involved in an accident. The CTP insurance is highly regulated, as it can be regarded as a social safety net or social good, so social considerations are more important in this case, as introduced in Section 2.1 of of Frees and Huang (2020). The pricing factors allowed to be used in CTP insurance varies across different states. For example, the Australian Capital Territory's CTP insurance scheme is a "community rated" scheme, with all motorists for each vehicle class pay the same amount regardless of their individual risk profile. The Victorian CTP insurance scheme allows the use of three factors to determine the premium, which are the vehicle classification (vehicle type and use), postcode, and eligibility for a pensioner discount. The New South Wales' CTP insurance is the least regulated compared with the other states and territories. It allows the use of more factors for pricing, including for example geographic region, type of vehicle, age of vehicle, distance traveled and Vehicle performance and else.

Voluntary Comprehensive Car Insurance and Third Party Insurance is also offered in the Australian auto market. Comprehensive insurance cover damage to the insured vehicle, damage to other vehicles, damage to or loss of property and theft of the insured vehicle. Third party property insurance only covers damage to other vehicles involved in an accident. Unlike CTP insurance, these voluntary coverages are less regulated, as they are not regarded as a social good and economic considerations dominate more in this case. Insurers can set prices using a range of rating factors for risk classifications. Historically, pricing models used by insurers in Australia have been less advanced. However, most large insurers now have highly sophisticated pricing models, see FCA (2019).

Table 1: Logistic Regression Model of Being Female Summary

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.409	0.162	-2.526	0.012
VehBody.Hatchback	1.707	0.128	13.334	0.000
VehBody.Sedan	1.354	0.119	11.334	0.000
‘VehBody.Station wagon‘	0.950	0.122	7.820	0.000
‘DrivAge.working people‘	0.224	0.077	2.911	0.004
VehBody.Minibus	1.601	0.339	4.721	0.000
‘DrivAge.oldest people‘	-0.478	0.117	-4.098	0.000
VehBody.Hardtop	0.794	0.205	3.876	0.000
VehBody.Coupe	0.863	0.266	3.250	0.001
‘VehAge.oldest cars‘	-0.466	0.102	-4.559	0.000
VehValue	-0.155	0.041	-3.770	0.000
‘DrivAge.old people‘	-0.213	0.094	-2.259	0.024
‘VehAge.old cars‘	-0.186	0.082	-2.262	0.024
VehBody.Convertible	2.240	1.268	1.767	0.077

## 2 Appendices

### 2.1 Appendix: A Proxy for Being Female

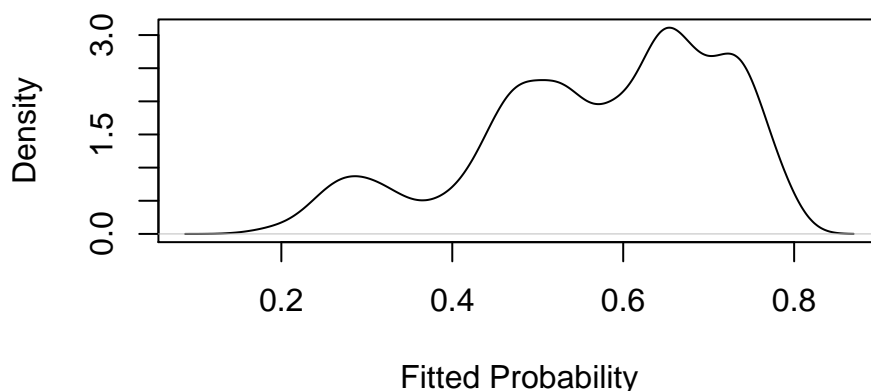
This appendix shows the results of the use of a classical automatic technique, stepwise regression, to develop a desirable proxy for the probability of being female. The *R* code is in Appendix Section 2.4. In addition to the main variables used for predicting claim amounts, `VehValue` and `DrivAge`, it also uses two other variables from the datafile:

- `VehAge` - The vehicle age group, and
- `VehBody` - The vehicle body group.

```
FemLogit1Sum <- summary(FemLogit1)
kable_styling(knitr::kable(coefficients(FemLogit1Sum ),digits=3,
  caption="Logistic Regression Model of Being Female Summary",
  vline = "", linesep = c("", "", "\\addlinespace"), align = "rrrr"), font_size=10)
```

```
plot(density(FemLogit1$fitted.values), xlab = "Fitted Probability",
  main = "Estimated Probability of Being Female")
```

## Estimated Probability of Being Female



### 2.2 Appendix: Australian Regulations

Section 24 of the DDA has the following provisions relating to disability provision:

- (1) It is unlawful for a person who, whether for payment or not, provides goods or services, or makes facilities available, to discriminate against another person on the ground of the other person's disability or a disability of any of that other person's associates:
  - (a) by refusing to provide the other person with those goods or services or to make those facilities available to the other person; or
  - (b) in the terms or conditions on which the first-mentioned person provides the other person with those goods or services or makes those facilities available to the other person; or
  - (c) in the manner in which the first-mentioned person provides the other person with those goods or services or makes those facilities available to the other person.
- (2) This section does not render it unlawful to discriminate against a person on the ground of the person's disability if the provision of the goods or services, or making facilities available, would impose unjustifiable hardship on the person who provides the goods or services or makes the facilities available.

Section 46 (1) of the DDA has the following provisions relating to insurance exemptions:

- (1) This Part does not render it unlawful for a person to discriminate against another person, on the ground of the other person's disability, by refusing to offer the other person:
  - (a) an annuity; or
  - (b) a life insurance policy; or
  - (c) a policy of insurance against accident or any other policy of insurance; or

- (d) membership of a superannuation or provident fund; or
- (e) membership of a superannuation or provident scheme; if:
- (f) the discrimination:
  - (i) is based upon actuarial or statistical data on which it is reasonable for the first mentioned person to rely; and
  - (ii) is reasonable having regard to the matter of the data and other relevant factors; or
- (g) in a case where no such actuarial or statistical data is available and cannot reasonably be obtained—the discrimination is reasonable having regard to any other relevant factors.

Data disclosure requirements for the exception of insurance based on sex discrimination are listed below in SDA:

- if the client gives the insurer a written request for access to the data—either:
- (i) the insurer gives the client a document containing the data; or
  - (ii) the insurer:
    - (A) makes a document containing the data available for inspection by the client at such time or times, and at such place or places, as are reasonable; and
    - (B) if the client inspects the document—allows the client to make a copy of, or take extracts from, the document.

### 2.3 Appendix: Chinese Regulations

We summarize the history of the regulation changes in Chinese voluntary auto insurance market in the table below.

Table 2: The history of regulation changes in Chinese voluntary auto insurance market

Time	Regulation
2001/10	First regulation change after becoming a member of WTO. A pilot experiment in Guangzhou to give the pricing rights to insurers, who use factors related to both the vehicle (type, age, etc.) and driver (age, gender, driving age, etc.) for pricing purposes.
2003	New pricing regulation for the whole country to give the pricing rights to insurers
2006/3	To stop the market chaos, the pricing was again regulated. No more than 30% discount of the benchmark pure risk premium can be applied.
2015-2018	The CIRC makes marketization of the auto insurance market as the objective again and gradually relax the the regulation by increasing the pricing bounds.

## 2.4 R Code

This section provides R Code for the Section 6.3 of Frees and Huang (2020) empirical example and the Appendix Section 2.1.

### Read In the Data

```
#install.packages("CASdatasets", repos = "http://cas.uqam.ca/pub/R/", type="source")
library(CASdatasets)
data(ausprivauto0405)
ausprivauto0405$Female <- 1*(ausprivauto0405$Gender=="Female")
ausprivauto0405$DrivAge = relevel(ausprivauto0405$DrivAge, ref="youngest people")
ClaimsData <- subset(ausprivauto0405, ClaimAmount>0)
```

### Fit Gamma Regression Model

```
AmtGamma1 = glm(ClaimAmount ~ VehValue+DrivAge+Female
, family=Gamma(link=log), data= ClaimsData)
sum.AmtGamma1 <- summary(AmtGamma1)
kable_styling(knitr::kable(coefficients(sum.AmtGamma1),digits=3,
caption="Gamma Regression Model 1 Summary",align = "cccccc|"),
font_size=10)
```

### Plot Distribution of Fits

```
par(mfrow = c(1,2))
plot(density(AmtGamma1$fitted.values, bw=80), xlab = "Fitted Value", main = "")
boxplot(AmtGamma1$fitted.values ~ ClaimsData$Female, ylab="Fitted Value", xlab="Female")
```

### Develop the Proxy for Being Female

```
#Proxy For Female
#install.packages("dummies")
library("dummies")
ClaimsData1 <- subset(ClaimsData, select =
-c(ClaimOcc, ClaimNb, Gender, Exposure, ClaimAmount))
ClaimsData2<- dummy.data.frame(ClaimsData1, sep = ".")
big.mod = glm(Female ~ ., family=binomial(link=logit),data= ClaimsData2)
base.mod <- glm(Female ~ 1 , family=binomial(link=logit),data= ClaimsData2)
# perform step-wise algorithm
stepMod <- step(base.mod, scope = list(lower = base.mod, upper = big.mod),
direction = "both", trace = 0, steps = 1000)
shortlistedVars <- names(unlist(stepMod[[1]]))
shortlistedVars <- shortlistedVars[!shortlistedVars %in% "(Intercept)"]
vars <- gsub("'", "", shortlistedVars)
ClaimsData3<- subset(ClaimsData2, select = c("Female", vars))
FemLogit1 = glm(Female ~ ., family=binomial(link=logit), data= ClaimsData3)
```

## Gamma Regression Summary

```
RemFem <- function(Var){lm(Var ~ Female, data= ClaimsData)$residual}
VehValueFemale <- RemFem(ClaimsData$VehValue)
DrivAge1Female <- RemFem(1*(ClaimsData$DrivAge=="old people"))
DrivAge2Female <- RemFem(1*(ClaimsData$DrivAge=="older work. people"))
DrivAge3Female <- RemFem(1*(ClaimsData$DrivAge=="oldest people"))
DrivAge4Female <- RemFem(1*(ClaimsData$DrivAge=="working people"))
DrivAge5Female <- RemFem(1*(ClaimsData$DrivAge=="young people"))
DrivAge6Female <- RemFem(1*(ClaimsData$DrivAge=="youngest people"))

AmtGamma2 = glm(ClaimAmount ~ VehValue+DrivAge
, family=Gamma(link=log), data= ClaimsData)
AmtGamma1A = glm(ClaimAmount ~ VehValue+ DrivAge1Female+
DrivAge2Female+DrivAge3Female+DrivAge4Female+DrivAge5Female+Female
, family=Gamma(link=log), data= ClaimsData)
AmtGamma3 = glm(ClaimAmount ~ VehValue+ DrivAge1Female+
DrivAge2Female+DrivAge3Female+DrivAge4Female+DrivAge5Female
, family=Gamma(link=log), data= ClaimsData)
AmtGamma4 = glm(ClaimAmount ~ VehValue+ DrivAge+FemLogit1$fitted.values
, family=Gamma(link=log), data= ClaimsData)
AmtGamma5= glm(ClaimAmount ~ VehValue+ DrivAge1Female+
DrivAge2Female+DrivAge3Female+DrivAge4Female+
DrivAge5Female+FemLogit1$fitted.values
, family=Gamma(link=log), data= ClaimsData)

temp1A <- cbind(summary(AmtGamma1)$coefficients[,1],summary(AmtGamma1)$coefficients[,3])
temp1 <- rbind(temp1A,c(summary(AmtGamma1)$aic,NA))
colnames(temp1) <- c("M.1 Coef", "M.1 t")
rownames(temp1) <- c("(Intercept)",
"VehValue", "DrivAge.old people","DrivAge.older work people",
"DrivAge.oldest people","DrivAge.working people",
"DrivAge.young people","Female", "AIC")

temp2A <- cbind(summary(AmtGamma2)$coefficients[,1],summary(AmtGamma2)$coefficients[,3])
temp2A <- rbind(temp2A, c(NA,NA))
temp2 <- rbind(temp2A,c(summary(AmtGamma2)$aic,NA))
colnames(temp2) <- c("M.2 Coef", "M.2 t")
temp3A <- cbind(summary(AmtGamma3)$coefficients[,1],summary(AmtGamma3)$coefficients[,3])
temp3A <- rbind(temp3A, c(NA,NA))
temp3 <- rbind(temp3A,c(summary(AmtGamma3)$aic,NA))
colnames(temp3) <- c("M.3 Coef", "M.3 t")
temp4A <- cbind(summary(AmtGamma4)$coefficients[,1],summary(AmtGamma4)$coefficients[,3])
temp4 <- rbind(temp4A,c(summary(AmtGamma4)$aic,NA))
colnames(temp4) <- c("M.4 Coef", "M.4 t")
temp5A <- cbind(summary(AmtGamma5)$coefficients[,1],summary(AmtGamma5)$coefficients[,3])
temp5 <- rbind(temp5A,c(summary(AmtGamma5)$aic,NA))
colnames(temp5) <- c("M.5 Coef", "M.5 t")
temp <- cbind(temp1,temp2,temp3,temp4,temp5)
kable_styling(knitr::kable(temp,digits=2, caption="Gamma Regression Model Summary"
, align = "cc|cc|cc|cc|cc|cccc"),latex_options="scale_down")
```



## Table by Gender

```
PSFittedF1 <- AmtGamma1$fitted.values*(ClaimsData$Female==1)
PSFittedF2 <- AmtGamma1$fitted.values*(ClaimsData$Female==0)*
  exp(summary(AmtGamma1)$coefficients["Female",1])
PSFittedF <- PSFittedF1+PSFittedF2 #Fitted values if everyone were female
PSFittedM1 <- AmtGamma1$fitted.values*(ClaimsData$Female==0)
PSFittedM2 <- AmtGamma1$fitted.values*(ClaimsData$Female==1)*
  exp(-summary(AmtGamma1)$coefficients["Female",1])
PSFittedM <- PSFittedM1+PSFittedM2 #Fitted values if everyone were male
PSFitted <- (PSFittedF+PSFittedM)/2
Mod1Fits <- AmtGamma1$fitted.values
SumM1Fits <- sum(AmtGamma1$fitted.values)
Mod2Fits <- ( AmtGamma2$fitted.values/sum(AmtGamma2$fitted.values) ) * SumM1Fits
Mod3Fits <- ( AmtGamma3$fitted.values/sum(AmtGamma3$fitted.values) ) * SumM1Fits
Mod4Fits <- ( AmtGamma4$fitted.values/sum(AmtGamma4$fitted.values) ) * SumM1Fits
Mod5Fits <- ( AmtGamma5$fitted.values/sum(AmtGamma5$fitted.values) ) * SumM1Fits
Mod6Fits <- ( PSFitted/sum(PSFitted) ) * SumM1Fits
library(Hmisc)
temp1q<- summarize(Mod1Fits,ClaimsData$Female, mean)[,2]
temp2q<- summarize(Mod2Fits,ClaimsData$Female, mean)[,2]
temp3q<- summarize(Mod3Fits,ClaimsData$Female, mean)[,2]
temp4q<- summarize(Mod4Fits,ClaimsData$Female, mean)[,2]
temp5q<- summarize(Mod5Fits,ClaimsData$Female, mean)[,2]
temp6q<- summarize(Mod6Fits,ClaimsData$Female, mean)[,2]
tempq <- cbind(temp1q,temp2q,temp3q,temp4q,temp5q,temp6q)
colnames(tempq) <- c("Model 1","Model 2","Model 3","Model 4","Model 5","Model 6")
rownames(tempq) <- c("Male", "Female")
kable_styling(knitr::kable(tempq,digits=2,
  caption="Comparison of Means by Predictors and Gender",
  align = "cccccc"), font_size=10)
```

## Plot by Gender

```
# Plot the boxplots with the same y range
library(ggplot2)
d=data.frame(d.type=c(rep(1,4624), rep(2, 4624), rep(3, 4624), rep(4,4624),
  rep(5, 4624), rep(6,4624)), sub.type=rep(ClaimsData$Female,6),
  val=c(AmtGamma1$fitted.values, AmtGamma2$fitted.values,
  AmtGamma3$fitted.values, AmtGamma4$fitted.values,
  AmtGamma5$fitted.values, PSFitted))
p <- ggplot(d, aes(factor(sub.type), val))
p + geom_boxplot(outlier.size=1, outlier.alpha=0.8,
  outlier.shape=21) + facet_grid(. ~ d.type)+
  xlab("Gender (Female=1, Male=0)") + ylab("Fitted Claim Amounts")
```

## Annotated Bibliography

Adams, Paul, Robert Baker, Stefan Hunt, Darragh Kelly, and Alessandro Nava (2015), "Encouraging consumers to act at renewal: Evidence from field trials in the home and motor insurance markets," *Financial Conduct Authority (FCA) Occasional Paper*, (12). Accessed <https://www.fca.org.uk/publication/occasional-papers/occasional-paper-12.pdf>. Retrieved on Nov 20, 2019.

Arizona Governing Committee for Tax Deferred Annuity and Deferred Compensation Plans versus Norris (1983), "Supreme court decision: Arizona governing committee for tax deferred annuity and deferred compensation plans versus norris," 463 (No. 82-52), 1073.

Arrow, Kenneth J (1963), "Uncertainty and the welfare economics of medical care," *The American Economic Review*, 53, 941–73.

Aseervatham, Vijay, Christoph Lex, and Martin Spindler (2016), "How do unisex rating regulations affect gender differences in insurance premiums?" *The Geneva Papers on Risk and Insurance-Issues and Practice*, 41 (1), 128–60.

Australian Law Reform Commission (2010), *Genetic Discrimination In Insurance*, Australian Law Reform Commission.

Avraham, Ronen (2018), "Discrimination and insurance," in *The Routledge Handbook Of The Ethics Of Discrimination*, K. Lippert-Rasmussen, ed., Routledge. Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3089946](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3089946). Retrieved on Oct 28, 2019. This paper puts forth a fair and just theoretical framework for assessing discrimination in insurance. Fairness is judged in terms of six characteristics of each variable, the line of insurance, and the nature of the discriminatory treatment. The trade-off between equality and its consequences is used to assess whether it is just.

——— (2019), "Database of state tort law reforms (6.1)," *U of Texas Law, Law and Econ Research Paper*, (e555). Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=902711](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=902711). Retrieved on Nov 9, 2019.

———, Kyle D Logue, and Daniel Schwarcz (2014), "Understanding insurance antidiscrimination law," *Southern California Law Review*, 87, 195–274 This paper describes fairness and efficiency arguments for introducing insurer prohibitions on underwriting. It uses these arguments to describe U.S. state approaches for nine potentially sensitive characteristics: race, national origin, religion, gender, sexual orientation, age, credit score, zip (postal) code, genetics. This description is for each of five lines of business: auto, disability, health, life, and property/casualty. The paper documents how U.S. state insurance antidiscrimination laws vary a great deal: in substance and in the intensity of regulation, across lines of insurance, across policyholder characteristics, and across states. It also remarks that a surprising number of jurisdictions do not have any laws restricting insurers' ability to discriminate on the basis of race, national origin, or religion, commonly referred to as the "big three" in American law.

Ayuso, Mercedes, Montserrat Guillén, and Ana María Pérez-Marín (2016), "Telematics and gender discrimination: Some usage-based evidence on whether men's risk of accidents differs from women's," *Risks*, 4 (2), 10.

Baker, Tom (2002), "Risk, insurance, and the social construction of responsibility," *Embracing Risk: The Changing Culture of Insurance and Responsibility*, 33–51 Reaching out to insurance institutions for protection cedes responsibility to them. Risk not only creates responsibility, but also socializes that responsibility. Baker argues that insurance is a form of social responsibility and that insurance institutions shape responsibility in five distinct senses: accountability, trustworthiness, causation, freedom, and solidarity.

Barocas, Solon and Andrew D Selbst (2016), "Big data's disparate impact," *California Law Review*, 104, 671–732.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace (2019), *Consumer-lending discrimination in the fintech era*, National Bureau of Economic Research. Accessed [http://www.phd-finance.uzh.ch/dam/jcr:139b84f5-afb8-40ac-b255-83b5386431ff/FS\\_spring18\\_paper\\_Morse.pdf](http://www.phd-finance.uzh.ch/dam/jcr:139b84f5-afb8-40ac-b255-83b5386431ff/FS_spring18_paper_Morse.pdf). Retrieved on Dec 28, 2019.

Blackmon, B Glenn and Richard Zeckhauser (1991), "Mispriced equity: Regulated rates for auto insurance in massachusetts," *The American Economic Review*, 81 (2), 65–69.

Born, Patricia (2019), "Genetic testing in underwriting: Implications for life insurance markets." *Journal of Insurance Regulation*, 38 (5).

Borselli, Angelo (2011), "Insurance rates regulation in comparison with open competition," *Connecticut Insurance Law Journal*, 18, 109–68 The main thesis of this paper is to argue for rate deregulation in the U.S. as a superior alternative to the current regulation model. An extensive historical analysis is provided to support this thesis, as well as some empirical facts about the U.S. market, and a comparison of the European experience

with insurance deregulation. To supplement the analysis, the paper includes a description of rate filing methods being used at the state level.

Brockett, Patrick L and Linda L Golden (2007), “Biological and psychobehavioral correlates of credit scores and automobile insurance losses: Toward an explication of why credit scoring works,” *Journal of Risk and Insurance*, 74 (1), 23–63.

Calders, Toon and Indrè Žliobaitė (2013), “Why unbiased computational processes can lead to discriminative decision procedures,” *Discrimination and privacy in the information society*, 43–57.

CAS, Casualty Actuarial and Statistical Task Force (2015), “Price optimization white paper,” National Association of Insurance Commissioners. Accessed [https://www.naic.org/documents/committees\\_c\\_catf\\_related\\_price\\_optimization\\_white\\_paper.pdf](https://www.naic.org/documents/committees_c_catf_related_price_optimization_white_paper.pdf). Retrieved on Nov 9, 2019.

CAS, Casualty Actuarial Society Committee on Ratemaking (2014), “Price optimization overview,” Casualty Actuarial Society. Accessed <http://www.casact.org/area/rate/price-optimization-overview.pdf>. Retrieved on Nov 11, 2019.

Cather, David A (2018), “Cream skimming: Innovations in insurance risk classification and adverse selection,” *Risk Management and Insurance Review*, 21 (2), 335–66 As stated in the title, this paper documents how advances in classification can lead to cream skimming, a type of adverse selection, against insurers that are slow to adopt such pricing innovations. This is done by providing an overview of the switching literature in economics/marketing and by describing two auto insurance case studies, one on the adoption of credit-based insurance scores and one on the introduction of usage-based insurances programs such as telematics.

Chen, Bingzheng, Sharon Tennyson, Maoqi Wang, and Haizhen Zhou (2014), “The development and regulation of china’s insurance market: History and perspectives,” *Risk Management and Insurance Review*, 17 (2), 241–63.

Chouldechova, Alexandra and Aaron Roth (2018), “The frontiers of fairness in machine learning,” *arXiv preprint arXiv:1810.08810*.

Civic Consulting (2010), “Study on the use of age, disability, sex, religion or belief, racial or ethnic origin and sexual orientation in financial services, in particular in the insurance and banking sectors,” European Commission: DG Employment, Social Affairs; Equal Opportunities. Accessed <https://ec.europa.eu/social/BlobServlet?docId=5599&langId=en&>. Retrieved on March 10, 2020.

De Jong, Piet and Gillian Z. Heller (2008), *Generalized Linear Models For Insurance Data*, Cambridge University Press Cambridge.

Dionne, Georges and Casey Rothschild (2014), “Economic effects of risk classification bans,” *The Geneva Risk and Insurance Review*, 39 (2), 184–221 This paper provides a unified framework for analyzing the economic consequences of insurance prohibitions in environments in which additional information can be learned through potentially costly tests.

ESL, Emergency Services Levy Big Data Project Team (2018), “Monitoring of insurance premiums in the world of big data,” Emergency Services Levy Insurance Monitor. Accessed [www.eslinsurancemonitor.nsw.gov.au](http://www.eslinsurancemonitor.nsw.gov.au). Retrieved on Nov 9, 2019.

European Union (2000), “Council directive 2000/43/ec,” *Official Journal of the European Communities*, Publications Office of the European Union. Accessed <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1578080397727&uri=CELEX:32000L0043>. Retrieved on Jan 3, 2020.

——— (2004), “Council directive 2004/113/ec of 13 december 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services,” *Official Journal of the European Union*, Publications Office of the European Union. Accessed <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32004L0113>. Retrieved on Jan 3, 2020.

——— (2012), “Guidelines on the application of council directive 2004/113/ec to insurance, in the light of the judgment of the court of justice of the european union in case c-236/09 (test-achats),” *Official Journal of the European Union*, Publications Office of the European Union. Accessed <https://op.europa.eu/en/publication-detail/-/publication/f1035369-1189-41d3-937c-aec8aa4f6cfd/language-en/format-HTML/source-112025857>. Retrieved on Jan 3, 2020.

Fannin, Brian A. (2020), “COVID-19: The property-casualty perspective,” Casualty Actuarial Society. Accessed [https://www.casact.org/research/COVID-19\\_The\\_PC\\_Perspective\\_3-27-2020.pdf](https://www.casact.org/research/COVID-19_The_PC_Perspective_3-27-2020.pdf). Retrieved on April 20, 2020.

FCA (2019), “General insurance pricing practices: Interim report. Annex 3: International comparisons.” *Financial Conduct Authority (FCA) MS*, (18/1. 2). Accessed <https://www.fca.org.uk/publication/market-studies/ms18-1-2-interim-report.pdf>. Retrieved on Nov 20, 2019.

Frees, Edward W. (2009), *Regression Modeling With Actuarial And Financial Applications*, Cambridge University Press.

——— and Lisa Gao (2019), “Predictive analytics and medical malpractice,” *North American Actuarial Journal*, 1–17.

Frees, Edward W and Fei Huang (2020), “The discriminating (pricing) actuary.” Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3592475](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3592475). Retrieved on 18 Feb 2021.

Frezal, Sylvestre and Laurence Barry (2019), “Fairness in uncertainty: Some limits and misinterpretations of actuarial fairness,” *Journal of Business Ethics*, 1–10. This article provides a historical overview of the concept of actuarial fairness in an insurance context. It distinguishes two viewpoints of actuarial fairness: (1) as a solidarity among a group of insureds pooling their uncertain results and (2) as a fair contract between an insured and an insurance company.

Friedland, Jacqueline (2013), *Fundamentals Of General Insurance Actuarial Analysis*, Society of Actuaries.

FTC, Federal Trade Commission .and others (2007), “Credit-based insurance scores: Impacts on consumers of automobile insurance,” in *A Report To Congress By The Federal Trade Commission, July 2007*.

GIAJ (2017), *Automobile Insurance In Japan*, General Insurance Rating Organization of Japan. Accessed <https://www.giroj.or.jp/english/pdf/Automobile.pdf>. Retrieved on March 22, 2020.

——— (2019), *General Insurance In Japan Fact Book 2018-2019*, General Insurance Association of Japan. Accessed <https://www.sonpo.or.jp/en/publication/index.html>. Retrieved on March 20, 2020.

Hussels, Stephanie, Damian Ward, and Ralf Zurbruegg (2005), “Stimulating the demand for insurance,” *Risk Management and Insurance Review*, 8 (2), 257–78.

Japan Times News (2019), “Insurance industry to set guidelines to prevent ‘genetic discrimination’.” 7 april 2019,” The Japan Times News. Accessed <https://www.japantimes.co.jp/news/2019/04/07/business/insurance-industry-set-guideline-prevent-genetic-discrimination/#.XnT3jnJ7kc>. Retrieved on March 22, 2020.

Joly, Yann, Charles Dupras, Miriam Pinkesz, Stacey A Tovino, and Mark A Rothstein (2020), “Looking beyond gina: Policy approaches to address genetic discrimination,” *Annual Review of Genomics and Human Genetics*, 21, 491–507.

Kachalia, A and MM Mello (2011), “New directions in medical liability reform.” *The New England Journal of Medicine*, 364 (16), 1564.

Kamiran, Faisal and Toon Calders (2012), “Data preprocessing techniques for classification without discrimination,” *Knowledge and Information Systems*, 33 (1), 1–33.

———, ——, and Mykola Pechenizkiy (2010), “Discrimination aware decision tree learning,” in *2010 Ieee International Conference On Data Mining*, IEEE, 869–74.

Kamishima, Toshihiro, Shotaro Akaho, Hideki Asoh, and Jun Sakuma (2012), “Fairness-aware classifier with prejudice remover regularizer,” in *Joint European Conference On Machine Learning And Knowledge Discovery In Databases*, Springer, 35–50.

Khaitan, Tarunabh (2017), “Indirect discrimination,” in *The Routledge Handbook Of The Ethics Of Discrimination*, K. Lippert-Rasmussen, ed., Routledge. Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3097020](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3097020). Retrieved on Oct 28, 2019. This paper traces the origins of disparate impact and indirect discrimination, beginning with the landmark American case Griggs versus Duke Power Company in 1971 and making its way to the United Kingdom, other Commonwealth countries and the European Union. It explains the nature of the disproportionate impact requirement. The paper also briefly speculates upon the moral bases on which indirect discrimination may be thought to be wrongful.

Klein, R (2017), “Genetics and life insurance. A view into the microscope of regulation,” *Geneva, Switzerland: The Geneva Association*. Accessed <https://www.genevaassociation.org/research-topics/global-ageing/genetics-and-life-insurance-view-microscope-regulation>. Retrieved on Dec 11, 2019.

Kleinberg, Jon, Sendhil Mullainathan, and Manish Raghavan (2016), “Inherent trade-offs in the fair determination of risk scores,” *arXiv preprint arXiv:1609.05807* This paper formalizes three core fairness conditions in algorithmic classification that correspond to notions of fairness: calibration within groups, balance for the negative class, and balance for the positive class. They find that except in highly constrained cases, there is no method that can satisfy the three conditions simultaneously.

Lehtonen, Turo-Kimmo and Jyri Liukko (2011), “The forms and limits of insurance solidarity,” *Journal of Business Ethics*, 103 (1), 33–44 When taking out an insurance, one participates in a risk pool in which each member is reciprocally responsible for others’ risks; this is insurance solidarity. This paper describes relationships among insurance, solidarity, and welfare. It further seeks to clarify the role of solidarity in both private and social insurance.

Lemaire, Jean (2002), “Why do females live longer than males?” *North American Actuarial Journal*, 6 (4), 21–37.

Lindholm, Mathias, Ronald Richman, Andreas Tsanakas, and Mario V Wuthrich (2020), “Discrimination-free insurance pricing,” *Available at SSRN* This paper considers the same discrimination-free pricing formula as that in Pope and (2011). However, they provide a rigorous probabilistic justification for the pricing procedure. They also justify the pricing framework using casual inference. To overcome the bias caused by the discrimination-free pricing formula, they propose bias correction strategies at the portfolio level. Empirical illustrations using both GLM and neural networks demonstrate how to apply the proposed methods for insurance pricing.

Loi, Michele and Markus Christen (2019), “Insurance discrimination and fairness in machine learning: An ethical analysis.” Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3438823](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3438823). Retrieved on Dec 28, 2019. This paper provides an ethical analysis of private insurance discrimination and fairness in machine learning. It distinguishes morally permissible and impermissible forms of statistical discrimination in private insurance and derive some ethical implications for the use of machine learning techniques in the insurance context. In particular, it discusses both social and economic considerations, such as characteristics of sensitive variables, adverse selection, and moral hazard.

Lombardo, Mark (2018), “The impact of genetic testing on life insurance mortality,” Society of Actuaries. Accessed <https://www.soa.org/resources/research-reports/2018/impact-genetic-testing/> [Retrieved on April 11, 2020].

Los Angeles Department of Water and Power versus Manhart (1978), “Supreme court decision: Los angeles department of water and power versus manhart,” 435 (No. 76-1810), 702.

Lukacs, Peter, Leslie Neubecker, and Philip Rowan (2016), “Price discrimination and cross-subsidy in financial services,” *FCA Occasional Paper*, (22). Accessed <https://www.fca.org.uk/publication/occasional-papers/op16-22.pdf>. Retrieved on Nov 13, 2019.

Maysami, Ramin Cooper and W Jean Kwon (1999), “An analysis of islamic takaful insurance.” *Journal of Insurance Regulation*, 18 (1) Takafal means shared responsibility; takaful insurance is a form of insurance that is permitted under Islamic guidelines. This paper describes Islamic socioeconomic principles applied to insurance, including concepts of uncertainty, interest, and investment arrangements. It further describes takaful life insurance, non-life insurance and reinsurance.

McCarthy, David D and John A Turner (1993), “Risk classification and sex discrimination in pension plans,” *Journal of Risk and Insurance*, 85–104.

Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan (2019), “A survey on bias and fairness in machine learning,” *arXiv preprint arXiv:1908.09635* This paper provides a comprehensive overview of types of bias in data, types of discrimination, definitions of fairness, and methods for fair machine learning.

Meyers, Gert and Ine Van Hoyweghen (2018), “Enacting actuarial fairness in insurance: From fair discrimination to behaviour-based fairness,” *Science as Culture*, 27 (4), 413–38.

Miller, Michael J (2009), “Disparate impact and unfairly discriminatory insurance rates,” in *Casualty Actuarial Society E-Forum, Winter 2009*, 276 In the U.S., “unfairly discriminatory insurance rates” are based on the costs of supplying insurance. This paper emphasizes conflicts that arise between cost-based rate practices and other rating practices that accommodate potential disparate impact effects. To motivate the discussion, it provides historical overviews of (U.S.) rate regulation and standards as well as (U.S.) laws and court decisions regarding disparate impact.

Morris, Darcy Steeg, Daniel Schwarcz, and Joshua C Teitelbaum (2017), “Do credit-based insurance scores proxy for income in predicting auto claim risk?” *Journal of Empirical Legal Studies*, 14 (2), 397–423.

NAIC (2010), “Property and casualty model rating law (file and use version),” National Association of Insurance Commissioners. Accessed <https://www.naic.org/store/free/GDL-1775.pdf>. Retrieved on Nov 9, 2019.

——— (2011), “State insurance regulation,” National Association of Insurance Commissioners. Accessed [https://www.naic.org/documents/topics\\_white\\_paper\\_hist\\_ins\\_reg.pdf?59](https://www.naic.org/documents/topics_white_paper_hist_ins_reg.pdf?59). Retrieved on March 6, 2020.

Pope, Devin G and Justin R Sydnor (2011), “Implementing anti-discrimination policies in statistical profiling models,” *American Economic Journal: Economic Policy*, 3 (3), 206–31 This paper proposes an anti-discrimination pricing method, which first estimates the full model to obtain the coefficient estimates, and then averages across the potential protected (SUP) values in the population. This approach can eliminate proxy effects while maintaining predictive accuracy relative to an approach that removing the contentious variables outright.

Prince, Anya ER (2019), "Political economy, stakeholder voices, and saliency: Lessons from international policies regulating insurer use of genetic information," *Journal of Law and the Biosciences*, 5 (3), 461–94.

Prince, Anya and Daniel Schwarcz (2020), "Proxy discrimination in the age of artificial intelligence and big data," *Iowa Law Review*. Accessed [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3347959](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3347959). Retrieved on Dec 21, 2019. As learning algorithms become smarter and big data gets bigger, proxy discrimination represents an increasing challenge to anti-discrimination regimes. This paper describes how these advances can undermine these regime outcomes that include (1) promoting social risk-sharing, (2) preventing the chilling of socially-valuable behavior, (3) limiting the effects of past discrimination, and (4) protecting non-conforming members of groups from being saddled with their group's characteristics.

Romei, Andrea and Salvatore Ruggieri (2014), "A multidisciplinary survey on discrimination analysis," *The Knowledge Engineering Review*, 29 (5), 582–638.

Ruggieri, Salvatore, Dino Pedreschi, and Franco Turini (2010), "Data mining for discrimination discovery," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 4 (2), 1–40.

Saxena, Nripsuta Ani, Karen Huang, Evan DeFilippis, Goran Radanovic, David C Parkes, and Yang Liu (2019), "How do fairness definitions fare? Examining public attitudes towards algorithmic definitions of fairness," in *Proceedings Of The 2019 Aaai/Acm Conference On Ai, Ethics, And Society*, 99–106.

Schmeiser, Hato, Tina Störmer, and Joël Wagner (2014), "Unisex insurance pricing: Consumers' perception and market implications," *The Geneva Papers*, 39, 322–350. Motivated by the European Court of Justice 2011 ruling that forbade sex discrimination in insurance, this paper reports on a survey conducted of customers concerning several price differentiation criteria, including sex. The survey asked about four key insurance products, motor, annuity, life and health, and spanned customers in five European Union countries, the United Kingdom, Germany, France, Italy, and Switzerland. The study showed that consumers accepted differentiation by price criteria, although there were substantial differences by line of business and by country.

Schwarcz, Daniel and Peter Siegelman (2017), "Law and economics of insurance," in *The Oxford Handbook Of Law And Economics: Volume 2: Private And Commercial Law*, Oxford University Press, 481–508.

Skipper, Harold D and Robert W Klein (2000), "Insurance regulation in the public interest: The path towards solvent, competitive markets," *The Geneva Papers on Risk and Insurance-Issues and Practice*, 25 (4), 482–504.

Swedloff, Rick (2014), "Risk classification's big data (r) evolution," *Conn. Ins. LJ*, 21, 339.

SwissRe (2018), "World insurance in 2017," Swiss Re Institute. Accessed <https://www.sonpo.or.jp/en/publication/index.html>. Retrieved on September 3, 2020.

Tennyson, Sharon L (2007), "Efficiency consequences of rate regulation in insurance markets," *Networks Financial Institute Policy Brief*, (2007-PB), 03. Accessed [https://www.indstate.edu/business/sites/business.indstate.edu/files/Docs/2007-PB-03\\_Tennyson.pdf](https://www.indstate.edu/business/sites/business.indstate.edu/files/Docs/2007-PB-03_Tennyson.pdf). Retrieved on Jan 20, 2020. This paper provides an overview of the history, scope, and objectives of U.S. insurance rate regulation. Through economic theory and empirical evidence, it argues that insurance rate regulation entails high costs for society and for insurance consumers, and that alternative policies for meeting regulatory objectives should be considered.

Thomas, R Guy (2012), "Non-risk price discrimination in insurance: Market outcomes and public policy," *The Geneva Papers on Risk and Insurance-Issues and Practice*, 37 (1), 27–46 This paper considers insurance price discrimination. An economic analysis is given of "inertia pricing," where renewal prices are higher than prices for risk-equivalent new customers, also known as "paying customers to switch." Public policy issues relating to insurance price discrimination are also presented.

Thouvenin, Florent, Fabienne Suter, Damian George, and Rolf H. Weber (2019), "Big data in the insurance industry: Leeway and limits for individualising insurance contracts," *Journal of Intellectual Property, Information Technology and E-Commerce Law*, 10 (2). Accessed <https://www.jipitec.eu/issues/jipitec-10-2-2019/4916>. Retrieved on Nov 5, 2019.

Veale, Michael and Reuben Binns (2017), "Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data," *Big Data & Society*, 4 (2), 2053951717743530.

Vukcevic, Damjan and Jessica Chen (2018), "Advances in genetics and their impact on life insurance." Accessed <https://www.actuaries.asn.au/Library/Events/FSF/2018/VukcevicChenPaper.pdf>. Retrieved on Dec 11, 2019.

Werner, Geoff and Claudine Modlin (2016), *Basic Ratemaking, Fifth Edition*, Casualty Actuarial Society. Accessed [https://www.casact.org/library/studynotes/werner\\_modlin\\_ratemaking.pdf](https://www.casact.org/library/studynotes/werner_modlin_ratemaking.pdf) [Retrieved on April 1, 2019].

Zhou, Xuefeng (2014), "On risk classification in insurance law: Reasonable distinction v.s. Discrimination," *Journal of Comparative Law*, 27 (2), 93–108.

Zliobaite, Indre (2015), "A survey on measuring indirect discrimination in machine learning," *arXiv preprint arXiv:1511.00148*.