Reasonable Restrictions on Underwriting

Joseph Heath
University of Toronto
Toronto, ON Canada

Email: joseph.heath@utoronto.ca
Voice: 416-978-8094

Biography

Joseph Heath is Associate Professor in the Department of Philosophy at the University of Toronto. He is the author of three books: *Communicative Action and Rational Choice* (MIT Press, 2001), *The Efficient Society* (Penguin, 2001), and with Andrew Potter, *Nation of Rebels* (Harper-Business, 2004). His recent publications in the field of business ethics include, with Wayne Norman, "Stakeholder Theory, Corporate Governance and Public Management," in the *Journal of Business Ethics* and "Business Ethics Without Stakeholders," *Business Ethics Quarterly* (forthcoming).

Abstract

Few issues in business ethics are as polarizing as the practice of risk classification and underwriting in the insurance industry. Theorists who approach the issue from a background in economics often start from the assumption that policy-holders should be charged a rate that reflects the expected loss that they bring to the insurance scheme. Yet theorists who approach the question from a background in philosophy or civil rights law often begin with a presumption *against* so-called "actuarially fair" premiums and in favor of "community rating," in which everyone is charged the same price. This paper begins by examining and rejecting the three primary arguments that have been given to show that actuarially fair premiums are unjust. It then considers the two primary arguments that have been offered by those who wish to defend the practice of risk classification. These arguments overshoot their target, by requiring a "freedom to underwrite" that is much greater than the level of freedom enjoyed in most other commercial transactions. The paper concludes by presenting a defense of a more limited right to underwrite, one that grants the legitimacy of the central principle of risk classification, but permits specific deviations from that ideal when other important social goods are at stake.

Keywords

insurance, risk classification, underwriting, discrimination, actuarial fairness

Reasonable Restrictions on Underwriting

There are very few issues in business ethics that are as polarizing as the practice of risk classification and underwriting in the insurance industry. Not everyone who seeks indemnification against a particular loss faces the same probability of suffering that loss, or faces a loss of equal magnitude. Thus insurers typically try to ascertain both the magnitude and probability of the loss for which a prospective policy-holder seeks indemnity, in order to determine an appropriate premium level. The ideal is to charge each policy-holder the so-called "actuarially fair" premium, which represents the anticipated cost of compensating that individual for the loss, multiplied by the probability that the loss will occur during the term of the policy. In reality, insurers are often unable to determine the risks that each individual faces. Thus they use a system of more-or-less broad classification, in order to determine which "risk pool" or class an individual falls into. This is used to determine a base premium, which is then "topped up" to cover transaction costs, commissions, and possibly – but not necessarily – profit.

Theorists who approach these insurance practices from a background in economics or business often start from the assumption that the actuarially fair premium represents the most "equitable" (Bossert and Fleurbaey, 2002) or "just" arrangement, such that any deviation from actuarial fairness requires justification.[1] Rates are considered "unfair" if "the insured is overcharged for the loss exposure in comparison with another similar loss exposure" (Outreville, 1998: 149), but there is no question that individuals who present *different* loss exposures should be charged different premiums. On the other hand, theorists who approach the question from a background in philosophy or civil rights law often begin with a presumption *against* actuarially fair premiums and in favor of so-called "community rating," in which everyone, no matter what

their background risk profile, has access to the same insurance policy at the same price (Daniels, 1991; Austin, 1983). Deviations from this baseline are then regarded as standing in need of justification. As a result, the entire practice of underwriting is presented as morally suspect. Tom Baker, for instance, describes the idea of actuarial fairness as "a watered-down form of liberalism that privileges individual interests over the common good and that privileges, above all, the interests of insurance institutions organized on its terms" (2003: 277).

Critics of risk-classification (or more tendentiously, "statistical discrimination") have derived considerable support from a series of Supreme Court decision in the United States that disallowed categorization according to sex in defined benefit pension schemes. The insurance industry also suffered a series of public-relations disasters associated with its underwriting practices, particularly in the United States where the absence of comprehensive public health insurance has made conditions of access to private health insurance a highly charged moral and political issue. This became especially apparent in the early stages of the AIDS epidemic, when insurers began refusing coverage not just to individuals who had contracted the HIV virus, but also those with a record of having been tested for it (on the grounds that only people who engaged in high-risk behavior would elect to test themselves). A similar furor erupted in the 90's when it was discovered that over half of American insurers routinely denied health, life and disability coverage to battered women, on the grounds that victims of domestic abuse had an adverse claims history. These episodes resulted in legislation in several American states imposing restrictions on the "freedom to underwrite" of insurers, preventing insurers from requesting certain types of information from prospective policy-holders, or else directly prohibiting them from charging different premiums to member of different groups (Hellman, 1997; Austin, 1983).

The result has been the development of considerable inconsistency in public policy.

Certain forms of risk classification are prohibited for certain types of insurance, but not others. No general legal principles have been developed to govern the practice either, in part because the ideal of actuarial fairness is rejected by many as inherently discriminatory or unjust. My goal in this paper will be to critically evaluate the latter claim. I begin by examining three different arguments that have been given, purporting to show that the practice of charging actuarially fair premiums is inherently unjust. I will try to show that each of these arguments is informed, in one way or another, by an essential misunderstanding of the mechanism through which insurance serves as a source of cooperative benefit. I go on to consider the two primary arguments that have been offered by those who wish to defend the practice of risk classification. These arguments, I will argue, overshoot their target, by demanding a "freedom to underwrite" that is much greater than the level of freedom enjoyed in most other commercial transactions. Thus I conclude by presenting an outline and defense of a somewhat more limited "right to underwrite," one that grants the legitimacy of the central principle of risk classification, but permits specific deviations from that ideal when other important social goods are at stake. This in turn allows us to develop relatively precise criteria for determining what constitutes a reasonable restriction on underwriting.

I

The world is full of risk. Knowing the probability of various events is extremely useful when it comes to engaging in practical deliberation. Unfortunately, what matters to most of us when we make our plans is not the background probability of an event, but the actual frequency with which it occurs. We know that a fair coin has a 50 per cent probability of landing heads, but we also know that flipping it 10 times is quite unlikely to produce exactly five heads and five tails. As a result, we need to be concerned not just with the mean, but also with the variance –

how far individual outcomes can be expected to deviate from the mean, and how often. However, it is also well known that as the number of tosses increases, the frequency will tend to converge with the probability (a phenomenon often referred to as "the law of large numbers"). In other words, increasing the number of trials induces *statistical stability* (Hacking, 2002: 190-2); it decreases the variance in the distribution. This increase in stability is the central mechanism through which insurance schemes are able to produce welfare benefits for their members.

To see how a group of individuals can benefit from the law of large numbers, it is important to remember that individuals are often risk averse. Consider a farmer who under normal conditions is able to produce 10 tons of grain – enough to feed his entire family well throughout the winter. However, his land is also subject to a highly localized blight, which sometimes wipes out the entire crop. Suppose that the chances of this blight striking his field in a given year are 20 per cent. Although the expected annual output of his field is therefore 8 tons, he would gladly swap a guaranteed revenue of 8 tons for the gamble that he faces between 10 tons or nothing. That way, his family would have a bit less to eat, but they would never risk starvation.

On his own, this is something that he cannot achieve. Suppose, however, that there are 100 small farmers who find themselves in identical circumstances, all facing the danger of this highly localized blight. They might agree to a "risk-pooling" arrangement, under which farmers who lose their crop in a given year are compensated by those who do not. Under this arrangement, the objective risk of blight does not diminish: 20 of the 100 farmers can, on average, expect to lose their crops. However, with the risk-pooling arrangement, each farmer can expect a revenue that will be, with 95 per cent probability, between 7.2 and 8.8 tons (see Moss, 2002: 28-31). Because the farmers are risk averse, this gamble has greater *subjective utility* than the gamble that gives each individual farmer an 80 per cent chance of getting 10 tons and a 20

per cent chance of nothing (even though the two gambles have the same mathematical value).

Thus what insurance offers is a form of superior "risk management," but not necessarily "risk abatement." It does not eliminate the loss, it just redistributes it. Of course, the insurance arrangement has the agreeable consequence of preventing anyone in the community from starving. But it is important to keep in mind that this is not why people buy the insurance. Insurance is not charity. They buy insurance in order to reduce their own uncertainty. If some farmers happened to like the gamble, and thought that it was worth risking starvation in order to get a shot at the full 10 tons of grain, then they would have no incentive to join the insurance scheme.

It is important to note as well that the risk-pooling arrangement is neither a gain from trade nor a straightforward economy of scale, but rather a *sui generis* source of collective benefit. Theorists sometimes mistakenly assimilate the gains that come from *trading* risks with those that come from *pooling* risks (e.g. see Barr, 1998: 111-2; Easterbrook and Fischel 1991, 53). In the former case, two individuals with different levels of risk aversion can generate efficiency gains by exchanging a risk – specifically, the one who is most risk averse can pay the other, in return for a promise of indemnity in the case of an outcome that it too far from the mean. Here, the welfare gain is possible only because one person is less risk averse than the other. In the case of insurance, however, people with the same level of risk aversion and risk exposure can still benefit from the "law of large numbers" mechanism, by agreeing to hold the risk in common. This is the typical arrangement within a mutual society, which was the dominant model of (non-commercial) insurance in the 20th century (Hansmann, 1992). Here there are no investors or stockholders, the company is owned by the policy-holders, and all of the money paid out in claims is simply levied from the policy-holders in the form of premiums.

The mutual society also provides the best model for examining the merits of actuarially

fair premiums. Because there are no investors, the controversial issue of "profit" is taken off the table. Premium levels are determined by "the insurance company," but the company in this case is simply a group of managers appointed to act as agents of existing policy-holders. In this case, it is not difficult to determine how premiums are determined. Take the example of 100 small farmers above. One can start by imagining a meeting at which all the potential policy-holders get together in order to determine the terms of the insurance arrangement. (This is not so fanciful, since many early mutual societies did originate in this way, as witnessed by the fact that publicans were often the founders and record-keepers for early 'friendly societies' [Neave, 1991: 51]). Together the farmers can be confident of producing close to 800 tons of grain, while losing 200 tons to the blight. Each farmer who joins the pool can therefore receive a guarantee of 8 tons of grain, in return for a commitment to contribute all that he is able to grow to the pool (either 10 tons, or 0 tons, depending upon how things work out). This is equivalent to keeping one's own crop, and paying a premium of 2 tons into the insurance pool. (When 80 farmers contribute 2 tons each, it generates the 160 tons needed to indemnify the 20 farmers who lose their crops.) Thus what each farmer pays in the way of a premium is equal to the expected loss that he brings to the pool, viz. a 20 per cent probability of drawing 8 tons while contributing nothing.[2]

Of course, the assumption so far is that each farmer is identically situated – having the same amount of land, the same level of productivity, and experiencing the same probability of suffering from the blight. But what if one of the farmers happened to have a plot of land that was twice the size of anyone else's? If he sought to insure his entire crop (i.e. sought a guarantee of 16 tons), it stands to reason that his premium should be higher. Indeed, the natural thing would be to charge him a premium that was twice as high, in reflection of the fact that the magnitude of the loss that he may impose upon the other members of the insurance pool is twice as large. But similarly, if a farmer had a plot of land that for some reason was twice as likely to be struck by

the blight, then it would also be natural to charge him a higher premium. Indeed, the natural thing would be to charge him a premium that was twice as high. A 40 per cent probability of drawing 8 tons while contributing nothing represents the same expected loss as a 20 per cent probability of drawing 16 tons while contributing nothing.

Thus the reasoning that leads to higher premiums for people seeking indemnity for losses of greater magnitude directly parallels the reasoning that leads to higher premiums for people who present a greater risk of loss. Those who expect to take more out of the pool should be required to put more in. This is the principle that underlies the idea of actuarial fairness, which simply stipulates that the premium paid by an individual should be equal to that individual's expected loss (magnitude multiplied by the probability).

This proposal appears simple, and one can certainly imagine it serving as a basis for agreement in an initial meeting at which individuals get together to form an insurance pool. There are a number of "real world" complications that arise, however, when it comes to determining just what the probability of a given loss is for a given individual. Many philosophers in fact think there is no "fact of the matter" as to whether a particular event, taken all by itself, can be said to occur with some probability. Either it happens or it doesn't. Probabilities belong only to classes of events, as a function of the frequency with which they occur, or as a function of our ability to predict them, based upon the frequency with which they occur (Hacking, 1984). Thus the only way to determine the probability of an individual's loss is to pick out some sort of frequency to which it belongs. With respect to certain events, the individual's own history may provide a sufficient record (so that the insurer is able to use so-called "experience rating" to determine that individual's premium). More often, the individual's own history is inadequate, and so insurers seek to establish "class rate," by finding a larger group to which that individual belongs (Outreville, 1998: 150-151) and seeing what the loss frequency is within that group.

Thus a person who has just learned how to drive has no safety record, and so no basis for estimating his or her chances of having an accident. An insurer may notice, however, that young men have an accident rate that is significantly higher than that of young women, or that single men have a much higher accident rate than married men (Dahlby, 1983). Thus the insurer might respond by classifying individual policy-holders into such groups and using the frequency of losses among members of the group as a way of determining the expected loss that the individual brings to the insurance scheme.

This is where things get controversial. Class rating may seem like a form of guilt by association. Just because *other* young men are terrible drivers doesn't mean that *this* particular young man is going to be one. But of course, this isn't a very helpful observation. If we knew how things were going to turn out in the end, then there would be no need for insurance. All that we have to go on in designing insurance contracts is the *ex ante* perspective. The way to think about the fairness of premium schemes is to imagine all of the policy-holders getting together in an initial meeting, in order to create the insurance pool. In principle, they are not obliged to do business with anyone, and so are not obliged to admit into the pool anyone that they do not want. Furthermore, the terms under which individuals are to be admitted entirely up for negotiation. Under such conditions, if the best available information indicates that young men pose a particularly elevated risk, then these young men are going to have to offer more in the way of premiums in order to secure admittance.

Of course, there are a number of important "second best"[3] problems that arise when it comes to implementing actuarially fair premium schemes. In most cases, insurers will not have all the information that is needed to determine the expected loss that an individual brings to an insurance pool, and therefore cannot actually calculate the actuarially fair premium. They are left having to approximate that premium, using the best information available. Yet one cannot

assume that refining the partition of the insurance pool using new information will necessarily bring the premiums that all individuals pay closer to the actuarially fair level. Thus there is room for significant injustice to arise out of attempts to implement an actuarially fair premium scheme in a world of imperfect (and asymmetric) information (Abraham, 1986: 86; Promislow, 1987: 216). For example, when insurers decided to deny health and disability insurance to battered women, most did not distinguish between those who continued to live with their abusive partner and those who had ended the relationship (Hellman, 1997: 361). Members of the latter group may well have been disadvantaged by the fact that information injurious to their risk rating was easily available (police reports, hospital records, etc.), while potentially exculpatory information (present living arrangements) was either unavailable or unverifiable.

Yet these sorts of "second best" problems are not where critics of actuarial fairness have focused their energies.[4] The most important arguments have all been directed against the principle itself. Critics argue that actuarially fair premiums are inherently unjust, and thus not even what insurers should be aiming for. The discussion that follows will therefore focus upon these sorts of principled objections, setting aside all "second best" considerations. This is not to suggest that problems arising at the level of the "second best" are not serious – in many cases they are enormously so. The discussion will be restricted to the principle of actuarial fairness simply because the arguments about "second best" problems cannot really begin before it is established what a "first best" solution looks like, and this question is still subject to enormous controversy.

II

The most damning criticism of risk classification and "class rating" is the claim that it represents plain old-fashioned discrimination. It penalizes certain individuals, not for their

individual characteristics, but merely because of their membership in a group. This is, in effect, what the United States Supreme Court decided in *City of Los Angeles Department of Water and Power v. Manhart* (435 U. S. 702 [1978]), with respect to the use of sex-segregated actuarial tables in the determination of contribution levels to a defined-benefit pension plan. These types of pensions are essentially insurance products (with the payroll contributions being the premiums). Like life annuities, they generate a stream of fixed payments until the death of the beneficiary, thus providing insurance against the risk of outliving one's retirement savings. (Otherwise put, they represent an arrangement under which multiple individuals pool their retirement savings in order to reduce uncertainty.) Since women on average live longer than men, the expected value of a typical pension of this type is of greater value to women than to men at the time of retirement. Thus the employer in this case created a system of differential contribution levels for its employees, with women paying a higher premium than men. The Supreme Court ruled that this was a violation of Title VII of the Civil Rights Act of 1964, which prohibits discrimination against "any individual because of his race, color, religion, sex, or national origin" by employers.[5]

In order to reconstruct the court's reasoning, it is helpful to introduce a couple of distinctions, implicit in the judgment, but drawn out more explicitly in an influential gloss on that judgment by Lea Brilmayer, Richard Hekeler, Douglas Laycock and Teresa Sullivan (1980). In civil rights law, there is an important distinction between permissible and impermissible grounds for discrimination, and between disparate treatment and disparate impact. It is permissible for employers to treat employees differently, according to some characteristic that they possess, provided that they are able to demonstrate a "business necessity" for so doing. For example, it has been deemed permissible for employers to require that candidates pass a weight-lifting test in order to be considered for certain heavy manufacturing jobs (*Bowe v. Colgate-*

Palmolive Co., 416 F.2d 711 [7th Cir.1969]). This of course will have a disparate *impact* upon women, since women are on average able to lift less than men. But this does not count as disparate *treatment* of women, because the effect is indirect, and is a consequence of a system of discrimination based upon permissible grounds. If, however, the employer were to refuse to consider women for such jobs on the grounds that they are less likely to pass the weight-lifting test, this would constitute impermissible discrimination. The mere fact that some characteristic is statistically correlated with a characteristic that serves as permissible grounds for discrimination does not make it permissible for the employer to use the former as grounds for discrimination as well. In other words, the characteristic "being permissible grounds for discrimination" is not preserved through probabilistic inference. This is as it should be, since an arrangement under which women were excluded from certain manufacturing jobs on the grounds that women in general are less likely to be able to lift heavy weights is clearly unfair to those women who *are* able to lift such weights, and thus do possess the relevant job qualification (however unlikely this may be *ex ante*).

According to the Supreme Court's reasoning in *Manhart*, charging women more for their pensions on the grounds that they are less likely to die young is like excluding women from certain classes of employment on the grounds that they are less likely to be able to lift heavy weights. According to Brilmayer et. al.:

> American women as a group currently live longer than American men as a group, just as they are able to lift less weight as a group. But some women will die earlier than some men, just as some will be able to lift more weight. An employer who pays annuities on the basis of integrated tables in effect distinguishes among his employees on the permissible basis of longevity, for those individuals who live the longest will collect the

> most periodic payments and thus the largest total sum. Of course, the employer's practice may have disparate impact on men, for as a group they may not live to collect as many periodic payments as women. If he tries to avoid this disparate impact by using segregated tables – making larger periodic payments to all men as a group – he distinguishes on the basis of sex. This would be disparate treatment, for individual men and women of equal longevity would be treated differently: both periodic benefits and total benefits will be greater for a man than for a woman of equal longevity (1979: 510-11).

Central to this argument is the idea that "longevity" in this case constitutes the permissible basis for discrimination. Thus equality requires that all employees receive the same *ex post* net benefit from the pension scheme *unless* they differ with respect to longevity (just as all job applications for a position in heavy manufacturing must be considered equally unless they differ with respect to weight-lifting ability). Naturally, using longevity as grounds for discrimination has disparate impact on men, just as a weight-lifting test has disparate impact on women. This is permissible. However, the employer is *not* entitled to use sex as a predictor of longevity, in order to determine pension benefits, because the former is merely statistically correlated with the latter. According to the Court's ruling, even though women are more likely to receive more periodic payments after retirement, "there is no assurance that any individual woman working for the Department will actually fit the generalization on which the Department's policy is based. Many of those individuals will not live as long as the average man. While they were working, those individuals received smaller paychecks because of their sex, but they will receive no compensating advantage when they retire" (*Manhart,* 435 U. S. 702, 708 [1978]). Thus sex-segregated actuarial tables violate equality, by creating a situation in which a

man who lives to the same age as a woman would pay less for the pension benefits received.

This argument is ingenious, and at first glance also seems compelling.[6] And even though it is limited in scope from a legal point of view (to relations between employers and employees, and with respect to only the enumerated categories of discrimination), the moral implications of the argument are much broader. If sound, the argument shows that actuarially fair premiums in general violate the principle of equality. It suggests, for instance, that automobile insurance companies that charge drivers of red cars higher premiums than drivers of beige cars (of the same make and model) violate equality. In this case "having an accident" constitutes permissible grounds for discrimination. Yet risk-rating in accordance with color means that people who drive beige cars and do have accidents pay less for the same benefit received as drivers of red cars who have accidents. Why should we reward people just for driving beige cars?

Yet the analogy that the argument depends upon is clearly strained. First of all, one can see the sense in which individuals are "rewarded" for their weight-lifting ability by being given access to an employment opportunity, but it is odd to think of a defined benefit pension scheme as "rewarding" individuals for longevity (as thus of longevity as a "permissible grounds for discrimination"). What is the point of discriminating on this basis? It is difficult to avoid the impression that the purpose of the pension scheme is being misdescribed. Second, there is the fact that, in the case of weight-lifting, it is the *negative* correlation between being female and the property that is being rewarded that motivates the discrimination against women. Yet in the case of pensions, there is a *positive* correlation between being female and the property that is being rewarded. So if the goal was actually to use sex as a predictor of longevity, and people were being paid *more* for living longer, then that should have led to an arrangement under which women were charged *lower* premiums than men. Here we can see the most serious problem with the framework that the Court used to describe the issue: it is unable to make any sense of the idea

that sex is being used as a *predictor* of longevity. This in turn generates a serious misunderstanding of the purpose of segregated actuarial tables.

The employer in *Manhart* unfortunately muddied the waters by suggesting that the rationale for the sex-segregated tables was fairness to its male employees *as a class*. Instead of having each individual who lives a given number of years receive a net benefit of equal magnitude, they suggested that the goal was to have men as a class receive net benefits that were of equal (average) value to those received by women as a class. But this is clearly a terrible argument. It suggests, as the Court rightly saw, that proponents of sex-segregated actuarial tables did not want to use sex as a predictor of longevity, but rather wanted to *add* sex to longevity on the list of legitimate bases for discrimination. Under such an arrangement, each individual would get the same net benefit as each other individual, unless they differed in longevity *or sex*. But what could possibly motivate adding sex to the list of discriminators, since it obviously results in unequal treatment of individuals? The suggestion that it was being done in order to achieve fairness to classes – so that, in the aggregate, men receive the same average net benefit as women – is a strange rationale. Under such an arrangement, women would be forced to contribute more to the pension scheme, not because they were expected to live longer (this is the rewarded property!), but *merely* because they were women. This is plain old-fashioned discrimination. The Court quite rightly observed that the goal of civil rights legislation is to protect individuals from this sort of treatment, and that the language of the statute explicitly prohibits it.

The problem with the judgment lies in the framing of the question. It starts with the way that the principle of equality is applied (both by the Court, and in the more perspicuous argument of Brilmayer et. al. [1980]). The conflict is not one between equality for individuals versus equality for classes. The relevant contrast is between equality *ex ante* and equality *ex post*. Consider the situation in which one uses a randomizing device, like a coin toss, in order to

allocate an indivisible good between two individuals. A proposed distribution that gives each individual a 50 per cent probability of getting the indivisible good creates a situation that is equal *ex ante*, but of course, the distribution that results from the coin toss (i.e. *ex post*) seems quite unequal, since one person gets the entire good and the other gets nothing. What makes this final distribution acceptable is the fact that the expected value of the lottery *ex ante* was exactly (or roughly, depending upon the *equalisandum*) the same for both individuals. The Supreme Court's approach in *Manhart*, on the other hand, would have us saying that the distribution is actually equal *ex post*, except that "winning a coin toss" represents a permissible ground for discrimination.

Numerous critics of *Manhart* have pointed out that the analogy between weight-lifting (or height) and longevity is faulty, because in the case of weight-lifting one can simply do a test to see how much a person can lift, and so there is no reason to rely upon the statistical correlation between sex and weight-lifting ability. In the case of longevity, on the other hand, there is no way of checking to see when a person will die (Kimball, 1979: 118). Indeed, if it were possible to do so, there would be no reason for insurance in the first place – each person could simply save exactly as much as he or she required for retirement (Kimball, 1979: 133). Yet these critics have failed to articulate the full force of this objection. What the element of uncertainty means, in the case of insurance, is that the principle of equality must be applied *ex ante*. Since there is no way to guarantee that the *ultimate benefit* of entering into an insurance scheme will be the same for all individuals (if there were, people could just save), we must ensure that the *expected benefit* be the same for all.[7] This is precisely what the actuarially fair premium represents – each individual pays a premium sufficient to cover the *expected loss* that her participation in the insurance scheme brings. Thus actuarially fair premiums are not motivated by a commitment to equality for classes rather than for individuals, but rather by a commitment to *ex ante* equality for

individuals.

Thus, contrary to the Court's ruling, the point of risk classification and underwriting is not to ensure that each risk class receives an equal benefit, but rather to ensure that each individual receive an equal expected benefit. The strange idea that "longevity" constitutes permissible grounds for discrimination in pension schemes should be rejected (as should the idea that "having an accident" constitutes permissible grounds for discrimination in automobile insurance). Both ideas are a consequence of the mistaken attempt to apply the principle of equality *ex post* rather than *ex ante*.

Of course, there are a number of complicated "second best" issues that arise with respect to the use of crude partitioning devices, like sex, to estimate the expected loss that an individual brings to an insurance arrangement. These are not at issue here. What is noteworthy about the reasoning of *Manhart* is that it attacks the basic principle of actuarial fairness, claiming that such premium schemes violate equality even under optimal conditions. The discussion here is intended to show that this argument is based upon a misapplication of the principle of equality. There may still be cogent arguments to be made against the use of sex-segregated actuarial tables for defined-benefit pension schemes. It is, however, a mistake to think that the use of such tables is a case of plain old-fashioned discrimination.

III

The second major argument against risk classification and underwriting is based upon the moral intuition that it is unfair to penalize individuals for circumstances that are outside of their control, or for things that are not their fault (Daniels, 1991; Hellman, 1997; Abraham, 1986: 89-92). According to this view, it is acceptable for insurers to penalize a driver with a history of moving violations by charging him a higher premium – after all, he has the option of improving

his driving habits – but it is unacceptable to penalize a young man with higher premiums merely because young men in general have bad driving habits. Similarly, it is thought reasonable to penalize smokers by charging them higher premiums for home insurance, but not people who live in high-crime neighborhoods. And, of course, since people have no control over their sex, race, age, or for the most part, health status, insurers should be prohibited from charging differential premiums on the basis of such characteristics.

Thus Norman Daniels, in an influential article on health insurance, claims that the argument for actuarially fair premiums rests upon a "controversial premise," viz. "that individuals should be free to pursue the economic advantage that derives from any of their individual traits, including their proneness to disease and disability" (Daniels, 1991: 504). (The idea that individuals should not be penalized for their circumstances, in an insurance context, is equivalent to the idea that individuals should not be advantaged by their circumstances, since it is the *relative* premium level that determines what counts as a penalty or an advantage.) Thus Daniels imputes the following argument to proponents of actuarial fairness:

> 1. Individual differences – any individual differences – constitute some of an individual's personal assets.
> 2. People should be free, indeed are entitled, to gain advantages from any of their personal assets.
> 3. Social arrangements will be just only if they respect such liberties and entitlements.
> 4. Specifically, individuals are entitled to have markets, including medical insurance markets, structured in such a way that they can pursue the advantages to be derived from their personal assets (1991: 505).

Daniels goes on to point out that this argument constitutes a direct statement of the basic premises underlying Robert Nozick's libertarianism (1974), which is a highly controversial political philosophy. Many others, including John Rawls (1971) and Ronald Dworkin (2000), believe that the outcome of the "natural lottery" is an effect of brute luck, not desert, and so individuals have no moral entitlement to benefit from their natural endowment. According to G.A. Cohen's influential formulation of this thesis "a large part of the fundamental egalitarian aim is to extinguish the influence of brute luck on distribution," (1989: 931). Anyone who shares this intuition should be unmoved by the argument for actuarial fairness, Daniels claims. Indeed, this is reflected to some degree in current employment legislation, he argues, where "we believe that justice requires us to sever consideration of race, sex or handicaps from deliberations about hiring, firing, and reimbursement for services performed, although in practice we fall far short of what justice demands... Thus we reject, in its most general form, the view that all individual differences can be a moral basis for advantage or disadvantage"(1991: 506).

This is, however, a very odd reading of current anti-discrimination law. Daniels is essentially claiming that actuarially fair premiums are unjust because they conflict with luck-egalitarianism, which is the most widely shared liberal conception of justice. The *prima facie* difficulty with this argument is that it appears to hold insurers to a higher standard of ethical conduct than any other business enterprise. Individuals routinely benefit from their natural endowments (intelligence, beauty, creativity, etc.) or from brute luck (plentiful rain, an early frost, a change in interest rates, etc.) in market transactions, and we think nothing of it. The idea that individuals should only be penalized for their choices and not their circumstances may be part of some luck-egalitarian ideal, but it is not part of what Christopher McMahon has called "implicit morality of the market" (McMahon, 1981).

The reasons for this are not hard to find. The task of carrying out the luck-egalitarian

project of indemnifying individuals against the effects of bad brute luck will in many cases require pure redistributive transfers – i.e. win-lose transformations. Thus any business arrangement (including an insurance scheme) organized along luck-egalitarian principles could leave individuals worse off than if they had never chosen to participate (or bought insurance) at all. Such an arrangement might not even offer them the *prospect* of being better off. Thus in the absence of altruistic preferences, such an arrangement cannot emerge as a result of private contracting. Private contracting is only feasible when there is at least an *ex ante* Pareto-improvement. Why would individuals sign up to pay for someone else's misfortune? This is good samaritanism. There may be a moral case to be made for such behavior, but to propose such principles as a basis for the legal regulation of the marketplace is essentially to argue that there should not be a marketplace.

Thus what Daniels has actually produced is not really an argument for restrictions on underwriting by private insurance companies, but rather a general egalitarian argument for social insurance in the health care sector. He observes that "the design of health-care systems throughout most of the world rests on a rejection of the view that individuals should have the opportunity to gain economic advantage from difference in their heath risks" (1991: 507). But this is precisely why the health care systems that he refers to are operated in the public sector. The fact that there is a strong case to be made for the state to deliver a particular type of service, in accordance with certain principles of distributive justice, does not mean that there is an equally strong case to be made for the state to compel the private sector to deliver that service under the same terms. Most welfare states also offer defined-benefit pension plans that are financed in accordance with principles that impose non-actuarially fair premiums (often aimed at producing a more progressive distribution of retirement income). But this does not mean that the state would be justified in imposing progressive payment schedules on private pension plans.

Daniels acknowledges that his argument may simply militate in favor of public insurance (1991: 518). However, he also wants to suggest that the basic luck-egalitarian principles, which require "community rating" in private health insurance, are not entirely foreign to the marketplace, and that many other enterprises are subject to similar restrictions. This is why he claims an analogy between restrictions on underwriting and anti-discrimination law in other areas of private contracting, and why he argues that anti-discrimination law is based upon a rejection of "the view that all individual differences can be a moral basis for advantage or disadvantage" (1991: 506). But in order to see the problem with this claim, one need look no further than the *Manhart* judgment. The court had no trouble with the idea that height, or weight-lifting ability, or longevity, could count as permissible grounds for discrimination or advantage, even though individuals have very little control over these characteristics. Even sex has been ruled a permissible grounds for discrimination if the employer can show that being of one sex or the other is itself actually necessary for performance of the job. Thus anti-discrimination law does not follow a luck-egalitarian logic. What the law restricts employers from doing is using criteria that are not *relevant* to job performance as a basis for discriminating against individuals. And this is clearly not what insurance companies are doing, when they practice risk classification and underwriting.

Daniels actually sums up the problem with his own view when he writes that the argument for actuarial fairness and the practice of denying coverage to high-risk individuals, "is persuasive only if the important function of health insurance is risk management. Because health insurance has a different social function – protecting equality of opportunity by guaranteeing access to an appropriate array of medical services – then there is a clear mismatch between standard underwriting practices and the social function of health insurance" (1991: 514). This may be true, but one could just as easily argue that current pricing practices in the grocery

industry are acceptable only so long as one thinks that the function of that industry is to sell food to people – however, since the true social function of the grocery industry is to protect equality of opportunity by guaranteeing adequate nutrition for all citizens, there is a clear mismatch between the practices of the industry and its social function. As one can see, this is not an argument for changing pricing practices in the grocery industry, it is an argument for socializing the grocery industry.

When the argument is formulated with respect to groceries, the problems with it become immediately apparent, in a way that they do not when it is formulated with respect to insurance. This is because of a widespread misunderstanding of how the insurance industry works, which leads many people to think that the industry does have a "social function" that extends beyond mere risk management. In particular, it is widely thought that the goal of insurance is not merely socialize *risk*, but rather to socialize the actual *losses* against which individuals seek indemnity. Even very knowledgeable commentators are prone to confusion on this score. Carol A Heimer, for example, writes that "at its most basic, insurance is a social arrangement to reduce the effects of losses by employing the resources of the group to cushion individuals. The key task of insurers is to organize the insurance pools, turn them into groups with a common fate, and act as agents of these groups" (2003: 288). Deborah Hellman argues that legal restrictions on underwriting are desirable, on the grounds that they represent "a first step toward treating the misfortunes of poor health and disability as communal responsibilities" (1997: 359).

But the goal of private insurance is not to pool *losses*. That would require altruism as an economic incentive on the part of a large number of the participants in the insurance pool. Why would one person want to take on someone else's loss? The fact that socialization of losses generates a group benefit does not mean that it generates a benefit for each individual. The reason people sign up for insurance is because of risk aversion, and because they seek to reduce

subjective uncertainty. The reduction in uncertainty is *achieved* through a socialization of losses, but the latter is instrumental to the former, it is not the objective of the arrangement. Thus participants in an insurance scheme will not accept any sort of socialization of losses, only ones that are conducive to the management of certain risks, and on terms under which the welfare benefits stemming from the reduction in subjective uncertainty outweighs the cost of having to indemnify others for their losses.

Of course, there are cases in which there is a powerful argument to be made for socialization of certain losses. An epidemic disease such as AIDS or SARS provides a powerful example. Daniels himself describes the creation of mandatory insurance for high-risk drivers to be a case where "our social interest in guaranteeing a public good... is allowed to overrule otherwise sound (and actuarially fair) underwriting practices" (1991: 510). I believe that health insurance represents a similar case, in which a particular public policy objective trumps the argument for industry practices. It is, however, misguided to transform this into an argument against risk classification in the private insurance industry. Insofar as there is a strong case to be made for socializing losses, rather than just socializing risks, then there is a strong case to be made for the involvement of the public sector.

What then should we say about our intuition that, when people are charged high premiums as a consequence of events or circumstances that are not their fault – such as being abused by their husbands (Hellman, 1997) – that they are being treated unjustly? The first thing to note is that insurance arrangements in general tend to wreak havoc with our intuitions about responsibility and desert. Much of this has to do with a simple tension between our moral reasoning, which is firmly governed by the language of free will, and the perspective that one must adopt when making statistical generalizations. Indeed, François Éwald (1986) has argued that the development of social insurance in the 19th century required a fundamental break with

the central concepts of rights and responsibility that determined the structure of 19th century "liberal" capitalism. Central to this development was the discovery that the rate of industrial accidents was highly predictable, regardless of who was responsible. Since sometimes a worker would be at fault, and sometimes the owner, the most socially efficient arrangement was simply to have both groups set aside a certain amount of money to indemnify the victims under a "no fault" arrangement. Yet even if this is better for everyone involved, it does mean that we can no longer expect the operations of the insurance system to track our intuitions about responsibility. (One can see the same thing with no-fault automobile insurance. Under such arrangements, some people will clearly get benefits that, from a strict view of personal responsibility, they are not entitled to. Does that make the arrangement unacceptable?) Similarly, the way that the insurance industry handles fraud offends the moral intuitions of some people, because it is governed more by a concern over loss-ratio security than by the binary opposition of guilt and innocence. As a result, the insurance industry tolerates a lot of behavior that the criminal justice system would regard as felonious (Ericson, Doyle and Barry, 2003: 340-346).

Second of all, our intuition that it is somehow more legitimate for an insurance company to penalize individuals for the consequences of choices they have made, rather than the circumstances they find themselves in, is perfectly cogent, but it does not count against the principle of actuarial fairness. When dealing with a person who voluntarily runs a risk, or chooses to act in a way that increases the risk of a particular loss, there are strong moral hazard arguments in favor of insurance schemes that penalize or deter such behavior (e.g. increasing premiums). Naturally, in the case of circumstances that are outside the individual's control, there is no moral hazard argument for increasing premiums. But this does not mean that there are no other arguments for charging higher-risk individuals higher-premiums. Most obviously, there is an adverse selection argument (discussed in section VI below), and this argument applies

regardless of whether one is dealing with the individual's choices or circumstances.

Thus the luck-egalitarian argument against actuarial fairness fails. It is based upon the plausible intuition that it is permissible, from the standpoint of justice, to penalize individuals when the high risks that they bring to an insurance scheme are the product of their own voluntary choices – since it at least gives them the option, in cases where they find the insurance too expensive, of changing their own behavior. Yet this does not make it impermissible for insurers to charge individuals higher premiums merely *because* they are high risk.

IV

Finally, it is often suggested in the debates over actuarial fairness that the problem with risk classification and underwriting is that it leaves high-risk individuals unable to afford insurance (Daniels, 1991; Ericson, Doyle and Barry, 2003). This is often felt to be unjust, because it leaves certain individuals excluded from a beneficial social arrangement that the rest of the population is able to enjoy. Yet though this argument occurs with enormous frequency, it is based on a confusion. Naturally, insurance *costs less* for high-risk individuals when they are pooled with a group of low-risk individuals in a "community rating" scheme. That is because they are using the insurance scheme, not just to secure the benefits of reduced uncertainty, but also to externalize some of the costs associated with their losses onto other members of the insurance pool. The problem with this arrangement is that it can easily make insurance unaffordable for the low risks – since the inclusion of high-risk individuals within the insurance pool can drive premiums to a level where low risks are better off "self-insuring" (e.g. putting their own money into a rainy-day fund, or else just tolerating the uncertainty). It is precisely this exclusion of *low-risk individuals* from the insurance market that constitutes the classic efficiency loss associated with adverse selection (Akerloff, 1970).

One might be inclined to think that segregating the insurance scheme, so that higher-risk individuals pay a higher premium than low-risk individuals, simply creates the opposite problem, excluding high-risk individuals from the market. But the two situations are not parallel. With pooling, the reason that low-risk individuals drop out is that the premium level becomes so high that it is no longer worth their while to buy insurance (because the premium significantly exceeds the "actuarially fair" rate). But with a segregated pool, it will still be worthwhile for high risk individuals to buy insurance. Their premiums are high in reflection of the fact that the loss exposure they bring to the insurance scheme is high, but it will still be better for them (assuming risk aversion) to buy insurance than to face the loss without indemnity.

Nevertheless, some commentators do speak as though there were an asymmetry in the position of low-risk and high-risk individuals, such that insurers are less willing to deal with the latter, even when they have the ability to pay. This appears to be based on another misunderstanding of how most insurance markets function. There are in fact two models of insurance (Hacking, 2003: 28). The first is the mutual society described in the first section of this paper. Under such an arrangement, a group of individuals with identical preferences and levels of risk aversion can nevertheless benefit by agreeing to pool their losses and gains. The benefits in this case stem from the reduction of uncertainty thanks to the law of large numbers. The second model, which Hacking refers to as the "Lloyd's of London" model, essentially involves a trade between a risk-averse and a risk-tolerant individual. Lloyd's rich "names" bet on outcomes, much as gamblers bet on horses. The names make money by getting lucky, betting that losses do not occur – collecting the premium but not having to pay out a claim. High-risk individuals, in this case, are like race horses with terrible odds. There needs to be a huge potential payoff in order to persuade anyone to bet on them, and once the odds get bad enough, they can't attract any bettors.

The Lloyd's model of insurance is particular well-suited for dealing with risks for which there is little or no actuarial knowledge (e.g. uncommon, low-risk, events). Thus high-risk individuals may be unable to secure insurance from companies operating on the Lloyd's model. But this is not the case with the mutual society model, which is the one that predominates in the standard categories of health, life, home and automobile insurance. This model is structurally neutral with respect to high and low-risk individuals, since both groups can benefit equally from forming their own insurance pools. Thus there is no reason, in principle, that risk classification should leave high risks any less able to purchase insurance than low risks in the standard run of cases. Generally speaking, if people can afford a loss, then they can afford the insurance to cover that loss. Of course, if they can't afford the loss, then they may not be able to afford the insurance either. But there is no independent issue of whether they can afford the insurance, and thus no special question of justice that arises with respect to the cost of insurance.

Of course, there is considerable evidence to show that high-risk individuals are less likely to purchase insurance when they are put into a high-risk pool and charged a higher premium (e.g. with automobile insurance, see MacAvoy [1977: 38]). The important point is that these people drop out of the insurance market, not because the insurance policy no longer has value to them (as is the case with low-risk individuals who drop out of a community-rated pool), but for some other reason, such as an inability to pay. The way that individuals discount the future is also likely to become a more significant factor as premiums increases, leading many to forego insurance because of a hypertrophied valuation of the present cost of the premium versus the future benefit of the potential indemnification. These, however, are "second-best" problems, which do not speak against the principle of actuarial fairness.

Consider a person who is, by genetic predisposition, almost guaranteed to get a particular form of cancer, and is thus facing the prospect of a significant financial loss for private medical

care. Since there is very little uncertainty in this outcome, what this person needs to do is start saving in order to cover the cost of future cancer treatment. Of course, insofar as there is some uncertainty and this individual is able to find other people who are similarly situated, there is no reason that they cannot get together to pool their savings, and thereby achieve the efficiency gains of an insurance arrangement. Setting aside transaction costs, this insurance is guaranteed to cost less than the cancer treatment, simply because the savings realized by the individuals who happen not to require that treatment are distributed out to all members of the insurance pool. Naturally, some people may not be able to afford the treatment, in which case they may not be able to afford the insurance. But that isn't the fault of the other participants in the insurance scheme, it's a problem of bad brute luck, or perhaps injustice in the distribution of income. The important point is that the anticipated loss is what the high-risk individual is unable to afford, not the insurance to cover that loss.

Thus what many people are articulating, when they worry that risk classification will leave high-risks unable to purchase insurance, is not an objection to the practice of risk classification, but rather a desire to see the losses to which high risk individuals are exposed socialized (and an attempt to use private insurance as a way of achieving this objectives). For example, what they object to is not that some people cannot afford health insurance, but that some people cannot afford cancer treatment. Yet rather than arguing that the latter costs should be directly socialized, by having the state pay for cancer treatment for everyone, they seek to socialize it indirectly (and partially), by externalizing a large segment of the cost onto other policy-holders in the high-risk individuals' insurance plan by imposing restrictions on underwriting (MacAvoy, 1977: 39). Thus the concern that risk classification will leave some individuals unable to afford insurance is often just a misleading way of arguing that individuals should not have to bear the burden of certain losses.

There is nothing intrinsic to the nature of being high-risk that makes a person any less able to participate in an insurance plan than being low-risk. Wherever there is uncertainty, people can benefit from risk-pooling arrangements. Problems arise only when there is an information asymmetry that makes the insurer, or other policy holders, unable to distinguish between low and high risks. Thus it is simply false in many cases to claim that actuarially fair premiums leave high risks unable to purchase insurance. In fact, insurers are sometimes eager to insure high risks, because such accounts generate a larger flow of premiums, and therefore potentially larger investment returns. Ironically, it is often restrictions on underwriting that leave high-risk individuals unable to buy insurance. While insurers can be legally prevented from setting higher rates for certain classes of individuals, they are seldom obliged to sell a policy to anyone who comes along (except in special cases, such as automobile insurance or health insurance in certain jurisdictions). Thus they are often better off not selling policies at all to high-risk individuals (Pauly, 1984). Restrictions on underwriting can therefore motivate "cream-skimming" on the part of insurers, which in turn may leave high-risk individuals unable to buy insurance, even if they are willing and able to pay an actuarially fair premium.

V

Surveying these arguments against actuarial fairness, it is sometimes difficult to avoid the impression that critics have failed to appreciate the full consequences that the rejection of this principle would entail. Most of the cases that have attracted controversy involve members of groups who are already stigmatized or subject to unjust discrimination being denied insurance or charged higher premiums by insurance companies. Furthermore, given that disadvantaged individuals tend to be exposed to higher levels of risk (by virtue of living in high-crime neighborhoods, driving less safe vehicles, eating a poorer diet, working in less safe conditions,

suffering more ill health, etc.), community rating tends to be progressive with respect to income and social class (Abraham, 1986: 76). But if one is to abandon the principle of actuarial fairness, one must do so across the board. And there are many cases in which doing so will not have progressive consequences. With respect to men and women, for instance, while women are the primary beneficiaries of the use of community rating for annuities, men stand to derive an equally large benefit from the use of community rating for life insurance. Similarly, men benefit considerably from the use of community rating in automobile insurance. To select just one example, during a period of intense debate over the use of sex-segregated actuarial tables in the calculation of automobile insurance premiums in Canada, it was calculated that community rating would see the average premium of a single woman between the ages of 21 and 22 rise by over 61.3 per cent, while the average premium of a man in the same category would drop by 11.6 per cent (Dahlby, 1983: 130). The result would be a significant redistribution of wealth from female to male drivers.

While the unfairness of this arrangement seems palpable, one must to be very careful when seeking to articulate the complaint. The natural inclination is to say "Why should women be forced to subsidize the terrible driving habits of young men?" Yet to formulate the argument in this way is to buy into precisely the sort of "fairness to groups" argument that the justices of the United States Supreme Court so effectively dissected in *Manhart*. After all, there is no requirement the women as a group come out the same as men as a group. In order to demonstrate unfairness, it is necessary to show that individual women are treated unfairly by the community rating arrangement. This is the challenge confronting those who would like to show that deviations from actuarial fairness are unjust.

The first thing to note is that, in an idealized "mutual society" insurance arrangement of the sort considered here, premiums are always actuarially fair in the aggregate. This is a simple

function of the fact that total claims paid out are equal to total premiums taken in. Now, as we saw in the first section, the most obvious argument for actuarial fairness in premiums is the principle that those who expect to take more out of the pool should be expected to put more in. In cases where each individual's premium is equal to the expected loss that he brings to the pool, each individual derives a pure welfare benefit from participating in the insurance arrangement – the mathematical value of participation is equal to the mathematical value of the gamble that he faces without insurance, it's just that the former has higher subjective utility because it is less risky. In cases where an individual's premium deviates from the actuarially fair level, it means that this person derives not just a welfare benefit from participating in the insurance arrangement, but also a monetary benefit, since the mathematical value of participating in the insurance pool now exceeds the value of the uninsured gamble (Bossert and Fleurbaey 2002, 114). Furthermore, since premiums are actuarially fair in the aggregate, the fact that one policy-holder derives a monetary benefit from participating in the insurance scheme means that some other policy-holder must suffer a monetary loss. As a result, when premiums deviate from the actuarially fair level there will be cross-subsidization within the insurance pool, or an implicit transfer of wealth between policy holders. Another way of putting it is to say that when an insurance company fails to charge actuarially fair premiums, it allows high-risk policy-holders to externalize some of the costs they face onto other policy holders. Thus the insurance pool, which is intended to be a source of mutual benefit in the form of welfare gains, is used as way of effecting implicit redistributive transfers between individuals (Kimball, 1979: 106).

This sort of cost-externalization is widely regarded as contrary to the basic principles of justice, even by theorists whose work is often appealed to in defense of the principle of community rating. Dworkin, for instance, states that one of the central virtues of his "resource egalitarian" auction is that it forces individuals to take into consideration the full cost that their

choices impose upon others (Dworkin, 2000: 70; see also Gauthier, 1986: 225). When a high-risk individual joins an insurance pool governed by a community rated premium, it generates a negative externality for all the other participants. It is like the person who orders an expensive drink or appetizer when dining in a large party at a restaurant, knowing that the bill is going to be divided up evenly between everyone. When the cost of each diner's meal is shared collectively, it allows those with more expensive tastes to externalize the cost of their actions onto others (which in turn erodes the value of the communal eating arrangement). Thus there is an argument to be made for "internalizing" the externality, by giving each diner an individual bill.[8] The same moral intuition, when applied to the case of insurance, suggests that individuals should be charged an actuarially fair premium.

This argument is, in my view, sufficient to establish a general presumption in favor of actuarially fair premiums, from the standpoint of justice. This is, however, just a presumption. It has not been shown that deviations from actuarial fairness are necessarily unjust. This is because the high-risk individual who joins an insurance scheme does not merely generate a negative externality. She also creates a slight positive externality for all other participants in the insurance pool, via the "large numbers" effect, by virtue of having increased the size of the pool. Thus she is not a pure free-rider. Her inclusion in the pool generates a welfare benefit for all the other policy-holders, even if it also imposes a slight monetary loss upon them. Furthermore, there will be a region in which the welfare benefit generated by expansion of the pool outweighs the welfare loss occasioned by the monetary disadvantage imposed upon the other policy holders. As a result, it can be in the interest of everyone in the pool to accept new members, even when these new members are charged less than the actuarially fair premium. (Just as it can be in the interest of firms to hire more workers, even when doing so depresses average output. It is only when the marginal gains in net output reach zero that the firm should stop hiring.)

There is no question that when premium schemes offer a monetary advantage to some individuals and a disadvantage to others, it results in some people getting a better deal out of their insurance purchase than others. The question is whether this itself is unjust. According to some conceptions of justice, especially strictly egalitarian ones, it may turn out to be so. But there is no question that variations in the level of welfare benefit derived from commercial transactions are tolerated by the "implicit morality of the market." In the absence of price discrimination, for instance, consumers who are further in from the margin derive a larger welfare benefit from their purchases than those who are closer to the margin. A person who would have been willing to pay $20,000 for a flat-screen television may be able to buy it for $5000, because the latter sum represents the most that other more price-sensitive consumers are willing to pay. Even with insurance, highly risk-averse individuals derive more benefit from their policies than less risk-averse individuals, and yet we do not take this to be an affront to justice.

This is not an accidental feature of the market. One of the fundamental features of capitalism is the repudiation of the notion of a "just price" – or of a principled division of the gain from trade – in favor of a competitive determination of price levels. The value of the efficiency gains associated with the establishment of market-clearing prices is taken to outweigh the value of a more egalitarian determination of price levels. The core criterion used to evaluate market transactions is therefore the Pareto-principle, or the requirement that exchanges be mutually beneficial. The desire to achieve market-clearing prices, however, requires a willingness to tolerate transactions in which these mutual benefits are unequally divided. Thus the fact that some people get more out of their insurance purchases than others is not a special injustice, but rather an ordinary feature of commercial transactions in a capitalist economy. The insistence that any deviation from actuarial fairness is unjust, on the other hand, is essentially an insurance-specific version of just price theory.

As a result, the argument for actuarial fairness in premium levels does create a presumption in favor of the justice of such schemes, but it does not show that deviations from actuarial fairness are necessarily unjust. It therefore does not preclude restrictions on underwriting, especially when it can be shown that some *other* important social good is promoted through such restrictions. Governments do this routinely. As long as the insurance scheme remains mutually beneficial for all parties, there is nothing wrong with restrictions on underwriting that for some good reason (i.e. non-capriciously) give some people access to the scheme on preferential terms. The argument for actuarial fairness only precludes restrictions on underwriting that eliminate this benefit for some people (i.e. make it more attractive for low-risks to self-insure). Thus the demand for an unrestricted "freedom to underwrite" is a demand for a degree of freedom for insurers that no other type of business enjoys in a modern market economy.

VI

This brings us finally to the centerpiece of the insurance industry's defense of its underwriting practices: the adverse selection argument. There are both moral and nonmoral versions of this argument. According to the nonmoral version, it does not really matter whether it is just or unjust to charge actuarially fair premiums, it is necessary for insurers to do so if they wish to remain solvent. This is because community rating gives low-risk individuals an incentive to defect from the insurance scheme. Thus if the insurer charges a premium that represents the average loss in a pool that includes both high-risk and low-risk individuals, claims will wind up exceeding total premiums collected, simply because low-risk individuals will drop out of the pool (or high-risk individuals will join in great numbers). This is the adverse selection problem (Akerloff, 1970; Rasmusen, 1989: 230-235). In some cases, these low-risk individuals will defect

to another insurer offering a lower premium and a pool with fewer high-risk individuals. But in other cases, the low risks will drop out of the insurance market altogether. (Moral versions of the adverse selection argument then point to this deadweight loss as a case of injustice perpetrated against these low-risk individuals.)

There is, however, an ambiguity in the adverse selection argument, which can be clarified using the game-theoretic distinction between a Nash equilibrium and the core of a game. A particular insurance arrangement is in equilibrium if no individual participant has an incentive to drop out and "go it alone." The *feasible set*, therefore, represents the set of possible cooperative arrangements that offer individuals expected payoffs higher than those that could be obtained through the non-cooperative strategy of self-insurance. Depending upon how risk-averse individuals are, this set can be fairly large. The *core* of a game, on the other hand, represents an arrangement from which no individual *or coalition of individuals* has an incentive to defect (Ordeshook, 1986). The core will tend to be much smaller (indeed, sometimes it will be non-existent), because not only must the arrangement offer benefits that are superior to what the individual could achieve through self-insurance, it must also be superior to what any subset of insured individuals could achieve by defecting and forming their own insurance pool.

Consider the example of three-person cooperative project, in which the gains from cooperation significantly exceed the returns to individual strategic action. There are various ways in which this "cooperative surplus" could be divided up between the players. In principle, any individual should be willing to accept even a tiny fraction of the cooperative surplus, with the lion's share going to the other two players, so long as the tiny fraction received leaves that individual with a better outcome than she could achieve through defection (i.e. dropping out of the cooperative scheme altogether). Thus highly inequitable divisions of the cooperative surplus will still be within the feasible set, as long as every player gets at least something (Gauthier,

1986: 178). Such divisions may not be in the core, however, because the player who is most disadvantaged may be in a position to make a "divide and conquer" offer to one of the other players, promising to exclude the third player from the cooperative agreement altogether in return for larger payoffs for them both. Obviously this is only possible if the cooperative project is subject to decreasing returns to scale (Gauthier, 1993: 46), but this tends to be the case with insurance. Thus interactions with a risk-pooling structure will tend to have cores, simply because the "large numbers" effect diminishes as the pool grows larger.

An insurance arrangement that deviates significantly from the principle of actuarial fairness in premiums is unlikely to be in the core of the relevant game, simply because the low-risk individuals (assuming they are sufficiently numerous) could all defect and form their own insurance pool. It is sometimes argued, on this basis, that adverse selection problems will begin to show up the moment that premiums take individuals outside the core (Daniels, 1991: 513). This is, however, dubious as an empirical contention. The defection of a *coalition* is a form of collective action and is thus much more difficult to organize than the defection of an individual. It would be difficult to find any workplace, for instance, in which the division of labor and reward was genuinely in the core of the underlying interaction. The most talented group of employees could almost always benefit by defecting from the firm and setting up their own shop. This does happen, but just as often it doesn't. Thus it is not adequate, when considering insurance schemes, to suggest that the *mere* fact that a premium scheme takes policy-holders outside of the core will necessarily generate an adverse selection problem. There are too many other complicating factors. The only thing that can be said with confidence is that a premium scheme that takes some policy-holders outside of the *feasible set* (i.e. makes self-insurance a more attractive strategy) is likely to generate an adverse selection problem, because defection in this case does not involve collective action, individuals simply cancel their policies.

Thus the adverse selection argument does not provide a powerful justification for the principle of actuarial fairness. If premiums had to be in the core, then there would be a strong case to be made on adverse selection grounds for actuarial fairness. But if premiums only need to be within the feasible set, then the adverse selection argument only shows that they must not depart too radically from that principle. There is, however, likely to be considerable zone of tolerance for deviations. In this context, it is worth recalling that many flat-rate "friendly societies" – which essentially eschewed any actuarial basis for the calculation of premium levels – survived well into the 20th century (Emery 1996). Furthermore, it was seldom bottom-up pressure from policy-holders that led to greater risk-classification and differentiated premiums, but rather the aggressive lobbying efforts of actuaries, which date back to the beginning of the 19th century (see Ansell, 1835; Baker, 2003). Although there are certain important exceptions, risk classification has seldom been a defensive response to bottom up adverse selection problems, and has more often been used as a competitive tactic by insurers against one another. In other words, rather than policy holders using private information to purchase insurance at prices below the actuarially fair level, thereby creating losses for insurers, historically it has been more common to see insurers instituting risk classification schemes as a way of creating low-premium pools, which could then be used to entice clients away from rival insurers (Baker, 2003). It is worth recalling that in order for adverse selection to cause a serious problem, absent these competitive tactics, policy-holders must be at least roughly aware of the expected loss that they bring to an insurance pool. Yet individuals are seldom in such a position, simply because they have neither the interest nor the ability to analyze the relevant data. With automobile insurance, for instance, an insurer typically knows a lot more about the accident risk posed by a particular driver than the driver ever will. Thus in practice adverse selection has often turned out not to be the powerful force that economic theory predicts it to be.

As a result, a premium scheme that *merely* takes a group of policy-holders outside the core is unlikely to provoke defection. Under real-world conditions, not only are there significant transaction costs and collective action problems associated with such a defection, but individuals themselves usually lack to information needed to determine whether defection would be advantageous (i.e. whether they are in the core or not). Thus the real danger is not "classical" adverse selection, but rather the prospect that rival insurers, sensing an opportunity, will try to identify and recruit low-risk individuals from the community-rated pool, in part by informing them that they are implicitly subsidizing other participants in their existing insurance pool. In other words, the major problems with community-rated premiums are the various forms of cream-skimming that they encourage.

But because the most important problems arise out of the competitive behavior of insurers, and not strategic behavior on the part of policy-holders, the argument fails to provide a very powerful defense of the idea that risk classification is somehow "necessary" to ensure the financial solvency of insurers. Naturally, if one company is engaging in aggressive risk classification and underwriting, then rival firms may be forced to respond in kind in order to remain solvent. But so long as legal restrictions on underwriting apply equally to all firms in an industry, then this is not a concern. For example, in jurisdictions where sex-segregated automobile insurance policies are not permitted, it is possible in principle for a group of female drivers to defect from existing insurance schemes and form their own pool, with much lower premiums. But since the restrictions on underwriting make it impossible to exclude men from this new pool, or to charge them higher premiums, any such move would be quickly undermined by the number of male drivers who would attracted to the lower premiums as well.

Thus the argument from adverse selection fails to show that an unrestricted freedom to underwrite is a business necessity. There are simply too many ways in which legal restrictions on

underwriting can counteract the tendency, and ensure that insurers who refrain from engaging in a particular type of risk classification are not put at a competitive disadvantage by virtue of that fact. Of course, serious problems can arise from the incentives that are inadvertently created for insurers to find ways around the law. The biggest concern, mentioned above, is that insurers will simply drop entire classes of clients or red-line residential districts in order to avoid attracting high-risk clients (whom they do not have a right to charge extra). There is also the possibility that insurers will avoid certain clients, or charge them extra, in circuitous and indirect ways that are difficult to regulate. For example, although it is generally regarded as discriminatory to refuse individuals insurance coverage on the basis of their sexual orientation, it was suggested at one point that health insurers were refusing to provide coverage to male hairdressers (Aaron and Bosworth, 1994: 269). Needless to say, this sort of behavior not only reproduces the injustice that the original legislation was intended to prohibit, but compounds it in various ways.

Thus restrictions on underwriting must be carefully weighed, and not undertaken lightly, as they do have a strong tendency to generate perverse effects. The argument from adverse selection is important in that it draws attention to the strategic context in which insurance decisions are made. One cannot simply legislate changes in premium levels or underwriting practice without taking into consideration the changes that this will cause in the purchasing decisions made by policy-holders, along with the competitive tactics among insurers. On the other hand, it is far too simplistic to say that any deviation from actuarial fairness is *bound* to generate such perverse consequences. There is far too much friction in the real-world marketplace for that to be the case. Thus the argument from adverse selection adds very little to the basic argument from justice when it comes to supporting the practice of risk classification. Restrictions on underwriting certainly have redistributive effects, but so long as the outcome remains within the feasible set of policy-holders, these redistributions fall within the range that

market institutions normally permit. Furthermore, such restrictions need not create significant adverse selection problems, unless they are so extreme as to drive low-risk individuals from the market entirely.

VII

The argument so far has been focused entirely upon the *principle* of actuarial fairness in premiums. The conclusion has been that actuarial fairness represents a just ideal, but that restrictions on underwriting which move the premium scheme away from actuarial fairness are permissible when doing so is needed in order to achieve some important social good, and when doing so will not create significant Pareto-inefficiencies as a result of low-risk individuals dropping out of the insurance market. Thus I defend a somewhat more limited "right to underwrite" than those who insist upon actuarial fairness, or require that premiums be kept within the core. Neither of these is necessary in order to ensure that the insurance scheme remain advantageous for all parties involved. Furthermore, mutual advantage is all that is required in order for the transaction to be satisfy "the implicit morality of the market." Thus while it remains permissible for firms to charge actuarially fair premiums, it is not necessary, and there is nothing in principle wrong with statutory restrictions on underwriting that take certain policy-holders outside the core. It is wrong, however, for restrictions on underwriting to take any policy-holder outside his or her feasible set.

References

Aaron, H. J. and B. P. Bosworth: 1994, 'Economic Issues in Reform of Health Care Financing', *Brookings Papers on Economic Activity* (Brookings Insititute, Washington, D.C.)

Abraham, K. S.: 1986, *Distributing Risk,* (Yale University Press, New Haven).

Akerlof, G.: 1970, 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism', *Quarterly Journal of Economics,* **84**, 488-500.

Austin, R.: 1983, 'The Insurance Classification Controversy', *University of Pennsylvania Law Review*, **131**, 517-583.

Baker, T.: 2003, 'Adverse Selection and Risk Classification', in Ericson, R.V. and A. Doyle (eds.) *Risk and Morality* (University of Toronto Press, Toronto).

Bossert, W. and M. Fleurbaey: 2002, 'Equitable Insurance Premium Schemes', *Social Choice and Welfare*, **19**, 113-125.

Brilmayer, L., R. Hekeler, D. Laycock and T. Sullivan: 1980, 'Sex Discrimination in Employer-Sponsored Insurance Plans: A Legal and Demographic Analysis', *University of Chicago Law Review*, **47**, 505-560.

Cohen, G. A.: 1989. 'On the Currency of Egalitarian Justice', *Ethics*, **99**, 906-944.

Daniels, N.: 1991, 'Insurability and the HIV Epidemic: Ethical Issues in Underwriting', *The Milbank Quarterly*, **68**, 497-525.

Dworkin, R.: 2000, *Sovereign Virtue (*Harvard University Press. Cambridge, MA).

Easterbrook, F. H. and D. R. Fischel: 1991, *The Economic Structure of Corporate Law* (Harvard University Press, Cambridge, MA).

Emery, J.C.H.: 1996, 'Risky Business? Non-Actuarial Pricing Practices and the Financial

Viability of Fraternal Sickness Insurers', *Explorations in Economic History*, **33**, 195-226.

Ericson, R.V., A. Doyle and D. Barry: 2003, *Insurance as Governance* (University of Toronto Press, Toronto).

Éwald, F.: 1986, *L'État providence* (Grasset, Paris).

Gauthier, D.: 1986, *Morals by Agreement* (Clarendon, Oxford).

Gauthier, D.: 1993, 'Fairness and Cores: A Comment on Laden', *Philosophy and Public Affairs*, **22**, 44-47.

Hacking, I.: 1975, *The Emergence of Probability* (Cambridge University Press, Cambridge).

Hacking, I.: 2002, *An Introduction to Probability and Inductive Logic* (Cambridge University Press, Cambridge).

Hansmann, H.: 1992, *The Ownership of Enterprise (*Harvard University Press, Cambridge, MA).

Heimer, C. A.: 2003, 'Insurers as Moral Actors', in Ericson, R. V. and A.. Doyle (eds) *Risk and Morality* (University of Toronto Press, Toronto).

Hellman, D. S.: 1997, 'Is Actuarially Fair Insurance Pricing Actually Fair? A Case Study in Insuring Battered Women', *Harvard Civil Rights-Civil Liberties Law Review*, **32**, 355-411.

Kimball, S. L.: 1979, 'Reverse Sex Discriminatin: *Manhart*', *American Bar Foundation Research Journal,* **4**, 83-139.

Lipsey, R.G. and K. Lancaster: 1956, "The General Theory of the Second Best," Review of Economic Studies, **24**, 11-32.

MacAvoy, P. W.: 1977, *Federal-State Regulation of the Pricing and Marketing of*

Insurance (American Enterprise Institute for Public Policy Research, Washington, D.C.).

McMahon, C.: 1981, 'Morality and the Invisible Hand', *Philosophy and Public Affairs*, **10**, 247-277.

Moss, D. A.: 2002, *When All Else Fails* (Harvard University Press, Cambridge, MA).

Neave, D.: 1991, *Mutual Aid in the Victorian Countryside* (Hull University Press, Hull).

Ordeshook, P. C.: 1986, *Game Theory and Political Theory* (Cambridge University Press, Cambridge).

Outreville, J. F.: 1998, *Theory and Practice of Insurance* (Kluwer, Dordrecht).

Pauly, M. V.: 1984, 'Is Cream-Skimming a Problem for the Competitive Medical Market?' *Journal of Health Economics,* **3**, 87-95.

Promislow, S. D.: 1987, 'Measurement of Equity', *Proceedings of the Society of Actuaries*, **30**, 215-237.

Rasmusen, E.: 1986, *Games and Information*, 2nd edn. (Blackwell, Cambridge, MA).

Rawls, J.: 1971, *A Theory of Justice* (Harvard University Press, Cambridge, MA).

[1] In their paper "Equitable Insurance Premium Schemes," Bossert and Fleurbaey write: "The basic assumption underlying our analysis is that the most equitable insurance premium scheme is one where the premium paid by each agent is equal to the expected value of the payout of the insurer to this agent," (2002, 114). This assumption is not defended, it is taken as the point of departure.

[2] For a more formal calculation of the "pure premium", see Outreville (1998, 156).

[3] The term comes from the celebrated paper by Lipsey and Lancaster (1956), which showed that if it was impossible to satisfy one of the conditions needed for a competitive market economy to achieve Pareto efficiency, then strict adherence to the remaining conditions would almost certainly make the outcome less, rather than more, efficient.

[4] For a sober discussion of the basic "second best" problems that arise, and some principles to guide insurers, see Abraham (1986: 64-100).

[5] It is worth noting that Title VII discrimination is not the only form of discrimination that underwriting practices might run afoul of. The more obvious suggestion might be that the practice of charging different policy-holders different premiums was a form of price discrimination. This is prohibited in the United States by the Robinson-Patman Price Discrimination Act of 1936, which makes it unlawful "to discriminate in price between different purchasers of commodities of like grade and quality." This is not a powerful argument, however, because it isn't difficult to make the case that insurers are selling a different product to clients in different risk-classes – since the actuarial value of each policy is different, depending upon the risk profile of the individual and the magnitude of the loss. Furthermore, it should be noted that there is considerable tolerance for price discrimination in the market. The Robinson-Patman Act only targets forms of price-discrimination that have anti-competitive consequences. Practices such as price-skimming – when new goods are introduced at inflated prices,

and then dropped over time – are considered quite normal, even though they amount to forms of price discrimination. They are usually regarded as morally unproblematic, simply because those who wind up paying the higher prices are generally those who derive the largest welfare benefit from the purchase. This is also clearly true in case of insurance, since high-risk individuals get more value out of the policies that they purchase.

[6] It is far more persuasive than the arguments of Austin (1983), who merely points to the fact that risk classification will have disparate impact as a way of showing that the practice involves unjust discrimination.

[7] Kimball (1979) overstates the case somewhat, arguing that what employees receive in return for their contributions is simply risk-protection, i.e. the welfare benefits associated with the insurance scheme, and so it doesn't matter what benefits, if any, they ultimately receive. This involves a rather excessive disregard for the benefits. Insurance allows individuals to exchange one gamble for another, less risky one. It is, however, the (at least rough) mathematical equivalence of the two gambles that makes the individual willing to enter into the exchange. Thus the person who acquires an annuity is not just purchasing risk-protection, he is also purchasing an income stream that must be approximately equal to that achievable through savings (i.e. self-insurance).

[8] This analogy is due to Promislow (1987: 217). He uses it to dramatize some of the "second best" problems that may arise with risk-classification under conditions of imperfect information. Consider a situation in which a waiter, while refusing to give individual bills, offers to split the bill into two: one for diners on the north side of the table, one for diners on the south. Assuming that diners should pay their own way, under what conditions would this and would this not be a more just arrangement?