2 The ethics of moral hazard revisited

Rutger Claassen

Historically, the term "moral hazard" came with strong moral connotations, as moral blame attached to those abusing insurance schemes. However, economists have taken moral hazard as a technical term, seeing individuals' risk-adjusting behavior simply as a rational calculated response to insurance and insured situations. But the question still lingers: is morally hazardous behavior – sometimes or always – immoral? This chapter discusses the debate in ethics about this question. It argues that moral hazard is pro tanto morally wrong. The analysis is grounded in the fact that insurance puts people in a fiduciary relationship. They then are under a moral duty to act on behalf of the others in their insurance pool and try to reach the optimum level of social risk. Yet, there are exculpating reasons which diminish moral responsibility. Finally, the policy implications are discussed, illustrated by the moral hazard posed by large banks in the financial crisis.

Moral hazard refers to the tendency for persons to engage in higher levels of risk-taking if they know they are insured against losses. Without response, this may shift the burden of paying for the possible damage onto the shoulders of insurers and/or the other insured parties. They would then pay the price of such "less careful" behavior. A potential moral conflict is born. This phenomenon is ubiquitous in social and economic life. Here are three prominent examples.

The first illustration comes from the financial sector. Large banks, or other private financial institutions, can take high risks because they know that their actions are – implicitly – underwritten by a promise of governments to bail them out should they find themselves in heavy weather. The phenomenon is known as "too big to fail" (Stern and Feldman, 2004). Knowing this, they are likely to engage in riskier actions than they would have, had such an implicit guarantee not been in place. The sale and repackaging of subprime mortgages in the lead-up to the financial crisis of 2008 is one such example of risky actions (Dowd, 2009; Okamoto, 2009; Dow, 2012; Claassen, 2015). The government's motive for bailing out here is that banks are in a position of *systemic* power and importance (Cline, 1984). One bank's bankruptcy may have negative knock-on effects on other banks, given their close systemic

DOI: 10.4324/9781003139249-2

connections, and hence pose a threat to the functioning of the financial system as a whole. Taxpayers – who ultimately pay the bill in the form of higher taxes and/or a larger public debt – perceive the banks as imposing the costs of moral hazard on them.

A completely different context is that of welfare state benefits. Individuals may take fewer precautions than they otherwise would knowing they are insured in case of unemployment, illness, or workplace accidents. For example, they may spend less money and effort in increasing their personal skills to keep themselves employable, they may unreasonably speak up against their employers in cases of conflict being more willing to risk being fired, they may decrease their work efforts, and prioritize non-work activities. Whenever this leads to unemployment, the costs of their "more risky" behavior are shifted – in public welfare systems at least – unto the state, hence, again, the taxpayers. Many – especially conservative – politicians have argued since the 1980s that these effects are large and justify restricting access and generosity of welfare state schemes, so as to avoid overburdening them. The moral hazard argument has become part of the standard arsenal of those who want to roll back the welfare state (Stone, 2002; Heath, 2009).

Finally, a perhaps less familiar example is that of the legal structure of the business corporation (Djelic and Bothello, 2013). In the 19th century, industrialized countries adopted corporate laws which limited the liability of shareholders in a business corporation. Many businesses are now incorporated as Société Anonyme à Responsabilité Limitée (France), Limited Liability Company (U.S.), or Gesellschaft mit beschränkter Haftung (Germany). These arrangements were new at the time, and vehemently opposed by those who saw limited liability as a privilege diminishing responsible business behavior and encouraging high risk-strategies in business life. A businessman, according to these oppositional voices, should remain personally responsible for losses. This is what it means to be an entrepreneur (Ireland, 2010). In the end, the pressures to allow limited liability were too strong, and the practice was widely adopted. The parties onto whom the risk was shifted include a wide range of stakeholders, from employees of the business, creditors, consumers, third parties like local communities and taxpayers; basically, anyone who at some point may suffer losses due to businesses' risk-taking behavior. Regulations (work and safety regulation, environmental regulation, consumer protection regulation) may be seen as attempts to rein in the effect of a moral hazard enshrined into the very legal constitution of business corporations.

As these illustrations show, the phenomenon is wide-ranging. Discussion of moral hazard started in the 19th century, when life insurance and fire insurance were discussed (Baker, 1996). Health insurance became a prominent example when economists started debating the issue after Arrow's (1963) seminal article on the matter. New areas of application are constantly being opened up. For example, the concept has also been used to describe the effects of policies to save desperate refugees on future refugee numbers (Teitelbaum, 2015) and the effects of humanitarian interventions to protect rebellion groups

against genocide on the likelihood of future rebellions (Kuperman, 2008). All these examples vary in their specific characteristics.

Economic analyses of the phenomenon focus on how strong the effect of moral hazard is: how much do individuals in a specific context change their behavior under the influence of insurance? My focus here is different. I assume that there is *some* level of moral hazard going on in examples like the above. It can be a lot, or less so, but for whatever level of moral hazard, we can ask a further, ethical question. This *ethical* question is whether such morally hazardous behavior is *immoral*. Is it merely an economically rational response to conditions or is it an instance of morally blamable behavior? The importance of this ethical question is that any *political* decision about whether or not to insure at all, and/or combat the moral hazards attending insurance, needs to combine economic analysis and ethical judgment. For every case, we need to know how prevalent moral hazard is, and whether it is a moral fault of the insured individual which insurers should be in the business of regulating.

There are only a handful of treatments by ethicists of this question. The leading one is by Benjamin Hale, who argues that moral hazard is *not* immoral. My purpose in this chapter is to argue that moral hazard *is* pro tanto morally wrong. I say pro tanto, because I also outline two important exculpating reasons, which excuse the person posing a moral hazard to others. In Section 1, I start with a brief recap of the debate about the "moral nature" of moral hazard. In Section 2, I introduce the legal category of fiduciary relationships. When people are in such a relationship, they have obligations to act in other-regarding ways. I argue relations under insurance should be seen as fiduciary. In Section 3, I reject Hale's argument, by showing how individuals are under a moral duty to act on behalf of the other people in their insurance pool and try to reach the optimum level of social risk. In Section 4, I discuss two important exculpating reasons, which can lead to diminished moral responsibility on the part of the person posing a moral hazard. In Section 5, I expand the analysis to deal with the special case of public insurance programs, with the bank bailouts as a leading example. Section 6 concludes.

Section 1 Moral hazard as economically rational and morally neutral

I begin with an overview of the debates about the moral nature of moral hazard. Let's start with some history (Baker, 1996; Rowell and Connelly, 2012; Leaver, 2015).

The term "moral hazard" gained traction in the 19th century to designate a certain class of risks for new forms of insurance, such as fire insurance and life insurance. Insurance itself suffered from a moral stigma, since it was associated with gambling and crime (Baker, 1996, pp. 255–259), and its acceptance was contingent, inter alia, on its effectiveness in dealing with moral hazard. 19th-century insurers reacted to moral hazard in two different ways, depending on the two sources of this tendency which they

distinguished: character and circumstance. Some people believed simply to have "bad characters", would use every occasion for fraudulent or careless behavior. These people were to be left without insurance. Other people were in principle believed to be of good character, but nonetheless susceptible to temptation. These people could be insured, but the insurance contract had to be structured in such a way that the temptation to moral hazard was minimized (Baker, 1996, p. 241).

In the 19th century the moral lens was as important as the economic one. The best proof of this was that the standard economic solution for dealing with people of bad character – i.e., charging them higher premiums – was judged unacceptable:

> Not all of life, it seems, was to be ceded to the field of Hazard. What remained to be left outside was the realm of evil crime, fraud, and the suspicious "other". Thus, nineteenth-century life and fire insurers limited the insurance of moral hazards, not because of the complexity or for other technical reasons, but because of ideas about right and wrong, as the term "*moral* hazard" suggests. Insurance was a moral enterprise "deeply interested in the growth of public and private honor", and insurance men had a duty to "[g]uard against moral hazard from without" and "against moral perversion from within". Everyone in the enterprise, both insurer and insured, had an obligation to exclude the immoral. (Baker, 1996, p. 254).

Over the course of time, this moral lens on moral hazard has been lost, and an economic lens has become dominant. How did that happen?

From the economist's perspective, moral hazard is one example of a collective action problem, i.e., a situation in which behavior that is individually rational turns out to be collectively suboptimal. Moral hazard frequently occurs in the context of information asymmetries, as a consequence of what economists call "opportunistic behavior."[1] Individuals weigh costs and benefits, and insurance lowers the costs of a certain loss. This gives them an incentive to be less careful, with the predictable result that the insurance pool is overburdened with claims: "less loss from loss [for the individual] means more loss [for the collective]" (Baker, 1996, p. 270). In response, a fine-tuning of the terms of the insurance contract is needed, so as to minimize the opportunities for careless behavior. In the economist's hands, "moral hazard has become exclusively a property of insurance arrangements and not a property of the individuals who enter those arrangements" (Baker, 1996, p. 271). In the economic analysis, moral hazard is morally neutralized. It no longer denotes "bad character" of immoral persons, but simply refers to instrumentally rational action.

The best-known example of this economic approach is found in Mark Pauly's response to a famous 1963 paper by Kenneth Arrow. Arrow introduced the term with respect to medical insurance, as an example of decision-making under uncertainty (Arrow, 1963). This led Mark Pauly, in reply, to complain about the attitude of

> insurance writers' who 'have tended very strongly to look upon this phenomenon (of demanding more at a zero price than at a positive one) as a moral or ethical problem, using emotive words such as "malingering" and "hypochondria," lumping it together with outright fraud in the collection of benefits.

Instead, Pauly argued:

> the response of seeking more medical care with insurance than in its absence is a result not of moral perfidy, but of rational economic behavior. (Pauly, 1968, p. 535).

Arrow, in turn, replied to this by saying that:

> Mr. Pauly's wording suggests that "rational economic behavior" and "moral perfidy" are mutually exclusive categories. No doubt Judas Iscariot turned a tidy profit from one of his transactions, but the usual judgment of his behavior is not necessarily wrong. (Arrow, 1968, p. 538).

What both sides in this debate agree upon, then, is that moral hazard is an example of rational economic behavior. What they disagree on, is whether this is compatible with a judgment that such behavior is also morally wrong.

Most philosophers would agree with Arrow that conceptually there are two different standpoints. The judgment of immorality simply is made from a perspective of morality, which does not coincide with the perspective of instrumental rationality.[2] Arrow's disentanglement of economic rationality and (im)morality as two separate standpoints however merely opens up the possibility that the two may diverge in the case of moral hazard; it doesn't yet decide that matter. Pauly may still be right that all instances of moral hazard are morally unproblematic. Arrow has not given us a reason why some or all instances of moral hazard are morally problematic. To make such judgments, we need a more fine-grained ethical analysis. Therefore, we now turn to the ethicists' treatment of moral hazard.

There are only a handful of relevant contributions, the most important one being from Benjamin Hale.[3] His treatment vindicates Pauly's position: moral hazard is not immoral, it is morally neutral.

Hale's argument starts from the nature of insurance. The crucial point is that insurance is meant to *raise* people's level of risk-taking. For example, without car insurance, driving a car would be potentially very expensive. The costs of an accident being extremely high, everyone would stop driving or drive so slowly that the speed gains of having car traffic in society would be greatly diminished. With insurance, people can drive cars while assuming a higher risk profile (Hale, 2009, p. 11). Hale is not alone in emphasizing this. A similar argument has been made in the context of social risks. By insuring the accidents of modern industrial society, social insurance helped to make the

development of such a society possible (Stone, 2002, p. 60). From this perspective, moral hazard is not a *moral* problem; on the contrary, higher levels of risk are the sought-after outcome of introducing insurance schemes.

This does not mean that there is no morally problematic behavior under insurance. Such behavior, however, according to Hale is immoral because of *other* reasons than its being the consequence of moral hazard. For example, insurance may give insured persons incentives to lie when filing an insurance claim. However, the wrongness of the moral hazard here is derivative of the wrongness of lying in general. Lying is morally wrong whether or not there is insurance in place (Hale, 2009, p. 9). In terms of the car example, we can perhaps see street drag-racing as an example of this, taking a deliberate chance of causing lethal accidents. This would be morally wrong whether or not there is insurance in place. Hence, we should distinguish three "zones" of risk-taking (my terminology), as summarized in Table 2.1. Insurance aims to get the behavior of those in the insured pool from the first zone ('A') to the second zone ('B'). However, it stays away from a third zone ('C'), which runs alongside A and B (fraudulent behavior is not a matter of taking higher levels of risk, but a qualitatively different kind of behavior, uncorrelated to any level of risk). Note that some of the behavior in zone C can still be *induced* by insurance (without insurance, lying when filing one's insurance claim would not be possible), but it is not morally wrong because of the mere assumption of a higher level of risk under insurance.

This bifurcation between moral and immoral behavior leads to a bifurcated response strategy on the part of the insurer. All behavior under insurance which is classified as moral, is to be legally permitted. All behavior classified as immoral should be ruled out as illegal. Individuals behaving immorally should be sanctioned by expulsion from the insurance pool. But the fact that insured individuals assume a higher risk profile as a matter of insurance is itself to be expected and is actually the very purpose of insurance. This assumes that we can in practice make a reasonable distinction between moral and immoral behaviors under insurance. Street drag-racers (zone C) risking their own and other people's lives can effectively be separated from careful drivers, and their drivers' licenses can be taken from them. By contrast, what we may call moral hazard (in zone B) is in fact morally acceptable behavior, for which an *economic* response is appropriate. This was exactly the conclusion Pauly also reached. At the same time Hale's argument does justice to the earlier "moralizing" approaches of insurers, by recognizing that intrinsically immoral

Table 2.1 Hale's Categorization of Risks (My Reconstruction)

	Zone A: low levels of risk-taking	*Zone B: high levels of risk-taking*	*Zone C: immoral behavior (irrespective of insurance)*
Example of car insurance	Very cautious driving	Normal driving	Street drag-racing
Response	No insurance needed	Insured against loss	Punished, or excluded from insurance

behavior under insurance may occur as much as outside the insurance context, and sometimes be even triggered by insurance (Zone C). Thus, his argument can be seen as an elegant synthesis of both approaches.

In an earlier paper, I accepted Hale's argument (Claassen, 2015, pp. 536–539) for standard private cases of insurance (and then explained that it did not apply in more complex cases, such as the government's relation to banks' position in the financial crisis). I now think Hale's argument was wrong.

Section 2 The moral basis: a fiduciary theory of insurance

My refutation of Hale's argument will be built on the legal theory of fiduciary relationships. Here I first argue that insurance relations are fiduciary relations. In such relations, individuals have moral obligations towards others to act in an other-regarding manner. This provides the groundwork for the argument in the next section.

In private law, many relationships are classified as contractual. In such relationships, private individuals follow their own purposes. In doing so, they generally do not have to take the purposes of others into account, other than as an informational input when trying to realize their own purposes. For example, a seller needs to know the purposes of his clients (to be able to effectively tailor his offer to their wishes, and be successful in business), but he does not have to optimize his clients' purpose-satisfaction himself. Even stronger, he may legitimately try to increase his own gain at the expense of his client, in a bargaining process over the price. Under contractual conditions, individuals act in a self-regarding manner.

Fiduciary relations (even if established through contract) are different. A fiduciary relation arises when one person, the fiduciary, exercises discretionary power over the interests of another, the beneficiary. Beneficiaries are vulnerable, dependent on their fiduciaries' judgments and actions (Frankel, 2011; Miller, 2014). For example, parents are acting as fiduciaries of their children. When doing so, the latter are vulnerable to the actions of the former. As a consequence, in fiduciary relations, the structure of moral obligations and expectations is different from typical contractual obligations. Fiduciaries have an obligation to act *on behalf of* their beneficiaries, in accordance with their best interests. Law expresses this in terms of two duties: a duty of loyalty and a duty of care. The fiduciary is bound to exercise her power in such a way that she does not betray the interests of her beneficiary, and exercise due diligence in the exercise of her mandate (both of these are 'obligations of means', versus the 'obligation of result', characteristic of contractual relations). Trust underlies the relationship as a whole. The beneficiary must have trust in the fiduciary, and the latter must act in a trustworthy manner, for the relationship to get off the ground, and be sustainable. Fiduciaries are expected to act in an *other-regarding* manner. In their actions, they need to adopt the ends of others as their own (Laby, 2008; Smith, 2014).[4]

It may be thought that this kind of relationship is only at home in the family, possibly extended to cover relations with friends, neighbors, and others to whom one is personally close (Frankel, 2011, p. 57). However, fiduciary relations are also an important part of professional and business life. Relations between professionals and clients are routinely described as fiduciary: doctors are fiduciaries for their patients, teachers for their students, lawyers for their clients. The relation between directors and shareholders is also a standard example of a fiduciary relationship, which brings the concept into the heart of economic life. It is interesting that some scholars are also describing the political relation as a fiduciary one: politicians, once elected or appointed to the role of representatives, are fiduciaries for their citizens. After all, they exercise discretionary power (governmental authority) over their citizens, who must obey the laws they make (Fox-Decent, 2011; Criddle et al., 2018). As this overview shows, fiduciary relations are widespread in all major spheres of life: personal, professional, economic, and political. At their origins lies an *incapacity* on the part of the beneficiary, which can be more principled or more pragmatic. Children simply do not have the skills and experience to raise themselves, patients do not have the knowledge to be their own doctors. Shareholders and citizens have neither the time nor the knowledge to govern their corporations and countries. Given such incapacities, it is more efficient or even the only possibility to have a fiduciary perform the actions which otherwise the beneficiary would have to execute. Note that some of these fiduciary relations are described as a *principal-agent* relationship, which is legally classified as one particular type of fiduciary relationship (Frankel, 2011, p. 42). Hence the introduction of fiduciary duties can often also be seen as answers to what economists refer to as "agency problems."

Turn now to moral hazard. Do insurance contracts create fiduciary relations? In law, when the question is raised at all, this is denied. However, I propose to revisit this judgment. There are two main reasons underlying the denial of fiduciary status to insurance relations; I would argue they are both unconvincing.

The first reason for the denial of fiduciary status focuses on the position of the insurer. Insurers, it is said, stand in contractual relations to their clients. This is necessarily so, since their interests are adversarial in the case of a claim. In order to maintain the viability of their business, insurers must critically scrutinize clients' claims, and reject them where appropriate. This adversarial structure impedes the attribution of fiduciary status. Insurers cannot act on behalf of their clients (Richmond, 1999). While this argument is right as far as it goes, it wholly focuses on the insurer-insured relation. However, I would argue that, most fundamentally, insurance is a relation between a set of insured persons, with the insurer serving merely as a hub. Individuals in the insurance pool insure *each other*; they spread risks amongst the members of the group. Insurance through the hub mobilizes the "law of large numbers" in statistics to capture efficiency gains from spreading risks (Heath, 2009; Landes, 2013).

Therefore, I would propose to see every insured person, with respect to her own actions, as being a fiduciary toward the others in the insurance pool, who are her beneficiaries. Vice versa, she is also a beneficiary with respect to each of these as fiduciaries, as far as their actions are concerned. Overall, the insurance pool can be seen as a web of mutual fiduciary-beneficiary relationships. For example, imagine I have bought cell phone insurance at a new insurance company, called Tom & Jerry's (the example comes from Braynen, 2014). I am now under an obligation (a duty of loyalty) to act on behalf of the other insured persons' interests in keeping the costs of coverage as low as possible. To do so, I must avoid actions that bring unnecessary risk to my cell phone (Questions such as "what is 'as low as possible'?" and "what is 'unnecessary' risk?" are addressed below).[5]

The second reason for denying fiduciary status to the insurance relation is that it seems counterintuitive to see everyday actions I undertake for myself as representing the interests of others. When I lie next to the swimming pool and pull my cell phone out of my bag (with some danger of it falling into the pool), I am seemingly performing a self-regarding action. It is hard to see this innocuous action in terms of "representing" other people's interests. Am I not merely after my own interests in such situations? However, this line of thinking suggests a dichotomy that needs to be firmly rejected. Actions can simultaneously be partly self-regarding *and* partly other-regarding. Insurance adds an invisible but real layer of other-regardingness to a whole range of self-regarding actions. Whether we like it or not, the insurance relation with my cell phone insurer is present in the background at every point in which I am performing self-regarding actions with my cell phone. It is my property, but the risk-sharing arrangement has woven a web of mutual dependence with the other cell phone owners in the pool. We now are a community sharing in the risks of cell phone ownership. The liabilities of cell phone ownership are partially socialized. This is not unlike situations where there is interdependence of interests without uncertainty (cooperative schemes that are not insurance schemes). For example, where a boat crew takes part in a sailing race, each is vulnerable to the others in the boat doing their part of the job (i.e., the main risk here is free riding).

Note how the analysis given in this section provides us with a precise handle on Arrow's intuition, that the same action can be economically rational *and* (potentially) immoral. The economically rational part refers to the self-regarding aspect of our actions, the immoral part to the other-regarding aspect. Immorality does not refer to a qualitatively different class of actions (as it did for Hale), but here becomes *a judgment on one aspect of the same action*. The analysis in terms of a fiduciary relationship has prepared the ground for such judgments of immorality, by showing that we are accepting a moral obligation to act on behalf of others when we are part of the same insurance pool. But note that whether we *actually* violate a moral obligation and act wrongfully when we expose ourselves to higher levels of risk still remains to be argued.

Section 3 The moral wrongness of moral hazard

Building on the analysis above, I will now argue that moral hazard is pro tanto immoral. As a first step, I will adjust Hale's conceptual scheme. Then I will argue that there is a pro tanto moral duty to avoid moral hazard.

Remember that Hale treats the issue in terms of three zones (see Table 2.1) where the purpose is to get society from the first zone (no insurance offered) to the second zone. This is socially optimal. I propose to adjust Hale's scheme in several ways. First, I leave out of consideration his zone C – the independently immoral actions, like insurance fraud. This is not because it does not exist, but because it is irrelevant to this analysis. Then I divide the field of his zones A and B into four zones (see Figure 2.1). Finally, I use a curved line instead of the more generic categories of "lower" and "higher" levels of risk; because the justification of moral hazard in terms of "making possible higher levels of risk" obscures three zones (new zones B, C, and D) which all permit higher risk levels than the original zone A without insurance, but are nonetheless very different.

The purpose of insurance, following Hale's argument, is to bring a group of persons from zone A to B. In zone A insurance is not worth having since individuals take very low risk (no demand). The costs outweigh the benefits of insurance. In zone D this is also the case, albeit for opposite reasons. Here individuals take excessive risks, and the costs outweigh the benefits too. Economically rational insurance would be worth having but insurers will not offer it (no supply). In zones B and C, benefits outweigh costs, and insurance will be offered. The shape of the curve is motivated by the idea that there is an optimal level of risk, represented by point S. Below this optimum, social welfare can still be enhanced by taking more risks (which will lead to more opportunities for socially beneficial innovations; but also more "peace of mind" on the part of the insured risk-takers, which is an intrinsic benefit of

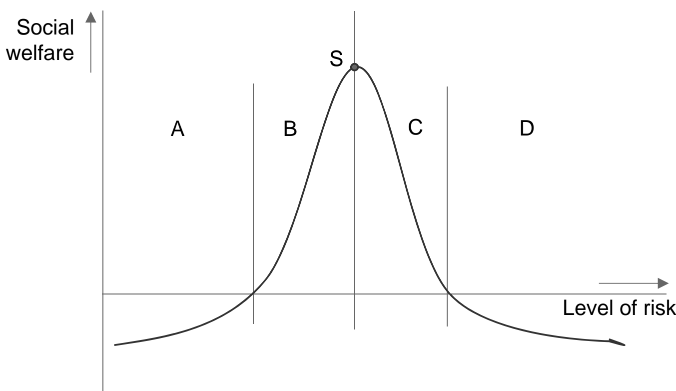

Figure 2.1 Moral Hazard Revisited.

insurance). Beyond the optimum, excessive risk-taking leads to a lower net gain from a social perspective, which at some point turns into a net loss.[6]

From the social point of view, individuals should target the level of risk they take to the social optimum, S. I define moral hazard as referring to the whole of zones B+C. Individuals pose a moral hazard to others *when (i) they increase their exposure to risk under insurance* (as was the intention of the scheme; hence there is a welfare improvement over zone A), *but (ii) they either take too little or too much risk, compared to the socially optimal level S.* They may have good reasons for doing this (i.e., they reach their personal optimum), but by doing so they shift the costs of their behavior to others. If we return to the example of car insurance, imagine that point S reflects the publicly determined speed limit. When all drivers orient their speed to this point, the car insurance system's ratio of social benefits (reduced time of transportation) and social costs (car accidents) is optimal. In zone C, drivers are breaking the speed limit. The speed limit helps individuals to orient their behavior as closely as possible to the optimum. In zone B, they drive too slowly, hence also foregoing some of the gains of the car insurance system. The analysis suggests this may be just as morally problematic, although this rarely is the focus of attention (in the following I will focus on zone C, since moral hazard is mostly thought of where people take too much risk).

All such behavior in zones B and C is *pro tanto* immoral. It may be only slightly immoral, if it is close to S, or very immoral, if further removed from S. I define the moral duty of individuals who are part of an insurance scheme as follows:

MORAL DUTY: Every individual in an insurance pool has a moral duty to target their level of personal risk-taking as closely as possible to point S.

But *why* is this a moral duty? In his article, Hale admits that welfarist doctrines in ethics can see the inefficiencies produced by moral hazard (which he does not question) as immoralities because for these doctrines the principle of efficiency is itself a moral principle. But he dismisses this line of reasoning without much argument (Hale, 2009, 4). In my view, the inefficiency of not reaching the optimum S can be evaluated as immoral (or at least "morally suboptimal", if one wants). Such judgments of immorality come in two different flavors, depending on the situation.

S is a moral requirement in a *weak sense* when the goal of the insurance scheme (i.e., an efficient scheme of cell phone insurance) is a collective goal shared by all the insured, but where this goal does not represent a moral requirement in itself. It is not as if the insured *owe* each other the existence of such an insurance scheme. It is not a moral requirement, but merely a pragmatic convenience for them to have one. In such cases, the moral requirement is generated by the *promise* made between the insured persons when creating their pool. Having opted in, you now have a moral obligation to organize your behavior toward the optimum as much as is possible for you. Nobody denies that breaking one's promises is pro tanto immoral. Here this is the basis of the claim that not targeting one's behavior toward S is immoral. Many

private schemes of insurance fit this description, since the private sphere is characterized by voluntary interactions to reach goals that are themselves not morally mandatory.

S is a moral requirement in a *strong sense* where it is a matter of justice or fundamental rights to implement an insurance scheme in the first place. For example, arguably citizens owe each other insurance against major health risks. Similarly, I argued elsewhere that the government's rescue actions toward big banks, in the end, are at the service of a fundamental right of all citizens to an adequate standard of living (Claassen, 2015, pp. 530–534). Since matters of justice and fundamental rights often become matters of public concern and hence government action, most of the insurance schemes fitting into the strong sense are coercively imposed insurance schemes, located in the public sector (see Section 5 below for specific analyses of such schemes). Here the point is that in such cases, the promises, in the sense of collectively agreed public policies, merely serve to ratify, not to create, this underlying moral requirement. The order of justification hence is reversed, compared to the weak, private cases.

What is true for both weak and strong cases, is that there indeed is a duty of no-waste. The most optimal level (point S) is a moral requirement. Given scarcity of resources, anything less imposes an unnecessary claim on other people's private budgets or on public budgets (thereby representing an opportunity cost). The former should be at the free disposal of private persons themselves, while the latter can be better used to serve other urgent demands of justice. Note that my analysis shows that one does not need to be a welfarist to recognize the morality of efficiency. For private insurance, anyone accepting the basic morality of promises would be committed to the same viewpoint. For public insurance, anyone holding a non-welfarist theory of justice would also be committed to efficiently reaching the aims of that theory of justice.[7]

This general moral requirement is "distributed" over the individuals participating in the insurance pool, who each hold a part of the duty to realize a successful insurance scheme. Here the fiduciary character of the insurance relation kicks in. Each participant is vulnerable to the level of care observed by the others, but each participant also has a discretionary power with respect to their own behavior, to which the others are vulnerable. The duty of loyalty goes beyond duties to act honestly in contractual, adversarial relations, and asks each participant to see themself as a representative of these others, in the relevant respect. The duty does not prescribe particular well-specified courses of action. Rather, in unpredictable circumstances, the right course of action can only be well-judged by the participant themself. They must exercise their own *judgment* to decide what is in the best interests of their beneficiaries (Smith, 2014, p. 148). Fulfilling a fiduciary duty is a matter of having a loyal attitude (Laby, 2008, p. 146). From the perspective of the vulnerable others, this may be frustrating since they cannot control the fiduciary's actions as much as they

perhaps would like to. It also leads to an important problem of uncertainty about what exactly is to be expected, which I discuss further below.

This finishes the explanation and defense of a pro tanto moral duty to avoid exploiting moral hazard at a cost to others. However, this is not the end of the story. As is familiar in ethics, a pro tanto duty is not the same thing as an all-things-considered duty. There may be countervailing considerations that cancel the immorality of moral hazard.

Section 4 Two exculpating reasons

From the definition of the moral duty above ("closely as possible"), it becomes clear that individuals can be excused whenever their level of risk was not under their (full) personal control. This is the basis for two major exculpating reasons. One is general and will be dealt with briefly; the other is particular to insurance relations and will be dealt with at length.

A very general reason, which holds true in and beyond insurance relations, is that one cannot be personally responsible for matters beyond one's control. There is an extensive philosophical literature on this. For example, luck egalitarians have argued that every action is either "a matter of chance" (or circumstance) or "a matter of choice" (or control). This chance/choice dichotomy structures, they argue, our moral landscape. Whenever something bad befalls an individual which was beyond his control ("bad brute luck"), society should compensate them. Whenever something bad happens for which individuals implicitly opted by taking their chances ("bad option luck"), no such compensation is due. Much debate has been waged on which factors belong in which category. For example, Ronald Dworkin (1981) argued that "preferences" (such as a preference for working few hours and shirking on the job) are always under one's control, while "talents" (such as marketable job-related talents) and handicaps are a matter of chance. Hence the untalented may get some compensation at the expense of the talented. Others have argued that the moral landscape should be carved up differently, with both preferences and talents sometimes under one's control, sometimes not (Cohen, 1989).

Regardless of these debates, the relevant point here is a more general one. Individuals who present a moral hazard sometimes cannot prevent themselves from doing so. In these cases, they bear less or even no personal responsibility. Their deviation from moral behavior is excused. This is important, because there is a substantial risk in assessments of moral hazard to "blame the victim." As Iris Young (2011) has convincingly argued in connection to U.S. debates about the welfare state, attributions of personal responsibility often ignore the structural constraints and background injustices under which welfare recipients must act. While I will not go into these matters here, I do think that they are incredibly important in weighing the appeals to moral hazard in policy debates.

The second exculpating reason relates more specifically to the insurance context. This is about information asymmetry, i.e., the difference between the "epistemic states" (Braynen, 2014, p. 43) of the insured and the insurer, and about who is responsible for this difference. One crucial question we have not addressed in the previous section is: how can an individual assess where point S lies? How does he/she know what it means to fulfill the moral duty identified above? Recall the example of the posted speed limit above, in the context of car insurance. Such a limit is an example of a clear, publicly known norm, about which little confusion arises. Hence one can clearly distinguish between those drivers who do, and those who do not target their behavior toward the social optimum. However, moral hazard often is the result of a vague norm, or even absence of norms, about how to behave. Braynen (2014, p. 44) gives the example of an imaginary "Tom and Jerry" company opening a cell phone insurance company, on the basis of flawless calculations indicating 4 in 100 cell phones break by accident. After insurance, the insured report 8 in 100 cell phones break down. There is no fraud, the increase is simply the result of less careful behavior. How can the situation be morally assessed?

The insurance contract is incomplete, to the extent that it doesn't specify acceptable or unacceptable actions to be undertaken with your cell phone. Imagine part of this epistemic uncertainty is disambiguated: Tom and Jerry find out that 25% of the increase is due to water-related accidents (cell phones in swimming pools, etc.), and they subsequently exclude water damage. This reduces reported accidents from 8 to 7 per 100. The insured have now been made morally (and legally) responsible for these 25% of cases. The new insurance contract has reduced the area of uncertainty, i.e., the area for which it is not clear whether the insured are *morally* to blame, because the behavioral expectations have not (yet) been made explicit. Reduced, but not eliminated. For example, is it reckless to do any of the following:

- i put your cell phone in the sun for two hours, causing overheating?
- ii put it on a low table where your two-year-old toddler can grab it?
- iii leave it in your jacket in the wardrobe in a restaurant?

Or are these the kinds of innocent behaviors that often end well, and that typify a more "worry-free" lifestyle, i.e., the kind of life which insurance is meant to make possible in the first place? In other words, are we in zone B, with these behaviors? Some would make a case that as long as a specific kind of behavior, such as (i) to (iii) above (an "action-type", as philosophers would say) has not been specified as problematic by the insurer, then it should be judged morally acceptable. Otherwise, it was up to the insurer to make regulations on the point. The whole of zones B and C should be exculpated in advance, because of this reason. This would make moral hazard morally justifiable after all (and validate Pauly's merely economic lens on the problem).

However, this in my view is flawed. The inability of the other insured persons in the insurance pool to specify all the act-tokens falling under the three action-types described above leaves the situation vulnerable to the judgment of the fiduciary. They have to maintain a standard of "reasonable carefulness" in their actions. The open-ended nature of the fiduciary's discretionary power and responsibility over the cell phone makes any pre-existing assessment by the vulnerable beneficiaries (the others in the pool) practically impossible. Consider that zone C could include each of the following actions:

(i-a)
one leaves one's cell phone in the sun for two hours, mid-spring in Spain (the phone survives);

(i-b)
one leaves one's cell phone in the sun for two hours, on a Caribbean cruise mid-summer (the phone is broken);

(ii-a)
one leaves one's phone on a low table; one's toddler has never violently smashed objects before, although they have by accident sometimes knocked objects over (the phone survives);

(ii-b)
one leave one's phone on a low table; one's toddler has a history of violently smashing objects against the wall (the phone is broken);

(iii-a)
one leaves the phone in a jacket in the wardrobe in a quiet village restaurant, where one knows the place and the owner;

(iii-b)
one leaves the phone in a jacket in the wardrobe in a trendy, crowded Manhattan restaurant that has a sign disclaiming risk to personal belongings.

In each of these three cases, one can argue that the fiduciary has broken a beneficiary's trust by engaging in behavior type-b. To be sure, type-a is also already "beyond the optimum," in that too many broken cell phones come out of type-a behaviors. But these are still to be excused, being reasonably close to S (i.e., in the targeted "zone" of S). Type-b behavior, however, clearly betrays an error of judgment. And in the given context, an error of judgment is a moral failure. There may be a grey zone but, at some point, there is a class of act-tokens about which most of the beneficiaries would say, in hindsight, that the fiduciary acted with insufficient care. This is the class of immoral actions under insurance, which triggers judgments that a person has

exhibited morally hazardous behavior. Hence the discussion here makes a slight modification to the discussion in Section 3. Instead of seeing the optimum as a single point S, it rather is an area around S, a zone in which there is moral consensus.

Accordingly, a closer inspection of the problem of information asymmetry leads us to refine the analysis of Section 3. Moral hazard relates to behavior that falls outside of an area around S; within that area, one is exculpated. The refinement here leaves the analysis of the basic immorality of moral hazard intact while suggesting that there is some scope for behavior that is to be excused.

Section 5 The public context: the imposition of insurance

The analysis so far has explained how and why moral hazard, as defined, is immoral. This was demonstrated using a simple example of private cell phone insurance. This however is important for complex cases in public contexts, such as bank bailouts in financial crises, as well. Such cases build upon – and thus require – the foundational immorality explained so far. Immorality does not kick in *ex nihilo* in such cases; rather it is *aggravated* by these circumstances. To explain whether the public's moral outrage about the 2008–2009 bailouts was justified, we need both.

Public insurance has two special characteristics compared to private insurance. The first one we saw above, when analyzing the reasons for the immorality of moral hazard: public insurance is an answer to a moral requirement itself – in terms of social justice and/or fundamental rights of citizens. Hence the banks' excessive risk-taking generated the kind of moral blame which is always expressed by those in an insurance pool (the beneficiaries) when they find out one of them (the fiduciary) has been taking excessive risks (cf. leaving one's cell phone in the Caribbean sun). The second feature of public insurance is that it is mandatory and coercive: individual citizens are not allowed to opt out. The two features are related. Precisely because, say, health insurance or the stability of the banking system are moral requirements for the public, the costs of maintaining these schemes are imposed on all, through raising mandatory premiums or taxes (or budget cuts on social programs following the bailouts). I will now argue that this aggravates the immorality, because it reduces the response options for the rest of the pool.

To explain this, I first need to introduce some terminology. How to think about the response options of insurers (acting on behalf of all in the pool), in general, when they discover that one or several persons in their pool have been taking excessive risk?

In a perceptive article, Andrew Williams (2009, pp. 501–502) has introduced the idea of a public policy trilemma (I will give my own, idiosyncratic, reconstruction). This trilemma shows how there are three options to respond to moral hazard (see Figure 2.2). Imagine there is a basic requirement of social justice, called "personal health." All people in a society, as a matter of justice, are owed access to medical resources, up to some threshold level,

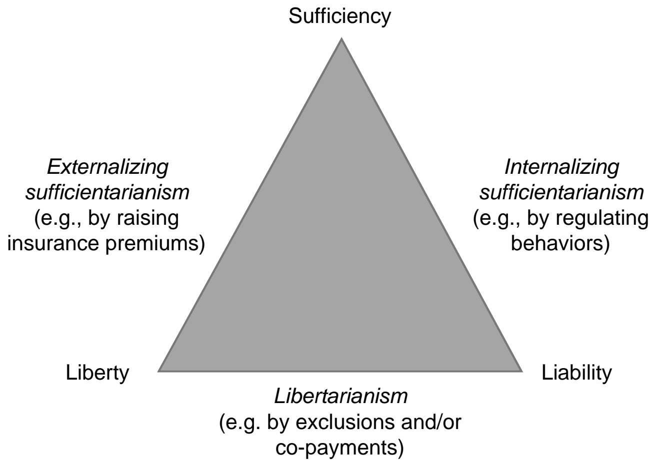

Figure 2.2 Public Policy Trilemma.

which we call "sufficiency." As a consequence, government implements public health insurance. Now a person's excessive risk-taking with one's health (discretionary lifestyle choices) may raise the costs of health care for all (moral hazard). In reply, government can do three things. It can maintain the person's liberty to engage in such risks, but hold him personally liable for this choice, and exclude him from insurance coverage; either altogether or exclusions for coverage of specified medical treatments (A financially based strategy with the same effect is to impose co-payments, co-insurance or deductibles). This will make him fall below the sufficiency threshold. This choice is called "libertarianism" (see Figure 2.2). If applied in the extreme, this abolishes the insurance scheme altogether. In a libertarian society, everyone is free to take the risks they want, but are also liable for the costs themselves. Alternatively, the government can maintain sufficiency for these persons: this is the "sufficientarian" view of justice. However, it can do so in two very different ways. One is by maintaining the person's liberty but canceling his liability. The costs are taken over by the collective. This is called "externalizing sufficientarianism", since the individual's costs are externalized. This will raise insurance premiums for all. The other, third way, is by imposing behavioral regulations and controls on the person, so that they cannot costlessly continue their risky ways. This is called "internalizing sufficientarianism", since they must now internalize the costs of their lifestyle themself. Their liability is maintained, their liberty restricted.[8]

We can learn from this trilemma that no strategy is without its own disadvantages. Out of the three basic moral values (sufficiency, liberty, and liability/responsibility), one can choose a combination of any two, but not all three. The trilemma indicates that an insurer has at least three different strategies at their

disposal to react to moral hazard. The insurer can hence make a moral choice about which values to uphold and which value to give up (mixed strategies are thinkable, since one can position one's strategy anywhere within the areas indicated by the triangle, but any chosen point will involve compromises).

With this conceptual framework at hand, we can now better understand moral condemnations surrounding bank bailouts. Given the nature of the situation, out of the three available strategies, two were unavailable. The unavailability of libertarianism follows from the nature of public insurance itself. *If* one judged that it is a moral requirement to maintain a sufficiency level, then libertarianism is off the table. In the case of the bank bailout, some libertarians did argue that one should let financial institutions go bankrupt, since only this course of action would let liability fall where it should (Miron, 2009). Most people rejected this position, given the effects of such bankruptcies on the economy (hence all citizens), i.e., the systemic nature of the risks involved. But that means the government's only options were to consider internalizing or externalizing sufficientarianism (or a mixed policy of both). The further problem, however, is that internalizing sufficientarianism by its nature is a *forward-looking* strategy. It can only be implemented for future cases. Once some insured persons have engaged in excessive risks, the costs must be borne – either by themselves or by the insurance pool. Hence the only available option to deal with the financial crisis was to bail out distressed banks, thus externalizing sufficientarianism.

Moral condemnations of bank bailouts can be analyzed as a composite of two mutually reinforcing problems. First, moral condemnations of banks' excessive risk-taking have their basis in a violation of the contract to target one's own behavior as closely as possible to the social optimum. Second, the moral anger is aggravated by having no other option to resort to than to bear the costs of this collectively through the bailouts (which will increase public debt, lead to cuts in social programs and arguably inspire future morally hazardous behavior). While the example in this section has been the financial crisis of 2008–2009, one can suspect that a similar analysis applies to other cases of moral hazard which raise public concern. For example, similar concerns should be raised about the bailouts and rescue packages passed during the Covid-19 pandemic (see Block, Chapter 9 of this volume).

Section 6 Conclusion

This chapter has investigated the moral nature of moral hazard. My analysis supports the idea first suggested by Kenneth Arrow that increasing one's exposure to risk in the face of insurance can be immoral. However, it does so by addressing the important argument made in the ethical literature by Ben Hale: insurance schemes are socially beneficial exactly because they allow people to increase their risks. The way out of this conundrum is to recognize that there is a socially optimal level of risk-taking, which is not the maximal level. When

insured persons fail to target their behavior sufficiently closely to this optimal level, they exploit the other members in the insured pool. When doing so, they fail to behave as good fiduciaries toward the others in the pool, thereby disincentivizing reflexive fiduciary behavior from such others.

This logic can become widespread, as illustrated by the financial crisis. If people start to exploit insurance schemes, such behavior may be infectious, i.e., others may be encouraged to do so too. This raises premiums for all and decreases the viability of the insurance pool itself. In the end, the insurance scheme will collapse. What is needed if we want to maintain insurance for all, especially in public policy contexts, is a strategy of what I have called "internalizing sufficientarianism", which regulates the behavior of the individual insured persons with the goal of reducing if not eliminating exploitative risk-taking. This will maintain the viability of the insurance pool, without asking others to pay the price.

Acknowledgment

I would like to thank Txai de Almeida Brito, Michael Bennett, Huub Brouwer, and Yara al Salman for their feedback and discussion. I thank Norbert Gaillard, Xavier Landes, and Maurits de Jongh for their extensive comments on an earlier draft. This work is part of the research program "Private Property and Political Power", with project number 360-20-390, which is financed by the Dutch Research Council (NWO).

Notes

1 For an illustration, see Darbellay and Gaillard, Chapter 4 of this volume.
2 Sometimes the latter standpoint is called "value rationality", following Max Weber's distinction between instrumental and value rationality (this terminology has the merit of making it clear that moral behavior is also rational, in its own way).
3 Besides Hale, only Braynen (2014) and McCaffrey (2017) aim to provide a new ethical argument about the morality of moral hazard. Heath (2009) and Landes (2013) have illuminating discussions of moral hazard, but do not provide a new argument. Claassen (2015) largely follows Hale, with an application to the financial crisis.
4 These two articles link the fiduciary duty to Kant's category of wide, imperfect duties.
5 This shows that we can see the insurer as a fiduciary, but only in the sense that he represents *all the other* insured persons' interests when he is critically examining whether *my* claim is worth accepting. Perhaps a more useful way to put this is that the insurer is a *judge*: he must arbitrate between me (fiduciary) and the others (beneficiaries) in deciding whether or not I have kept my duty of loyalty to them. This clears up another reason for denying fiduciary status to insurance contracts: that there is no concrete beneficiary. Indeed, there is a pool of beneficiaries on behalf of which I must take care. Multi-beneficiary schemes are a familiar figure in fiduciary law (Miller and Gold, 2015).
6 See also Pauly (2007) on the "optimal level of cost-sharing" in health insurance.

7 For my own theory of justice, see Claassen (2018).
8 The prudential rules imposed on financial institutions, especially the systemic ones, exemplify "internalizing sufficientarianism" strategies; see Flores and Gaillard, Chapter 5 of this volume.

References

Arrow, K. (1963) "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, Vol. 53, No. 5, pp. 941–973.

Arrow, K. (1968) "The Economics of Moral Hazard: Further Comment," *American Economic Review*, Vol. 58, No. 3, pp. 537–539.

Baker, T. (1996) "On the Genealogy of Moral Hazard," *Texas Law Review*, Vol. 75, No. 2, pp. 237–292.

Braynen, W. (2014) "Moral Dimensions of Moral Hazards," *Utilitas*, Vol. 26, No. 1, pp. 34–50.

Claassen, R. (2015) "Financial Crisis and the Ethics of Moral Hazard," *Social Theory and Practice*, Vol. 41, No. 3, pp. 527–551.

Claassen, R. (2018) *Capabilities in a Just Society. A Theory of Navigational Agency*. Cambridge: Cambridge University Press.

Cline, W. (1984) *International Debt: Systemic Risk and Policy Response*. Washington, DC: Institute for International Economics.

Cohen, G.A. (1989) "On the Currency of Egalitarian Justice," *Ethics*, Vol. 99, No. 4, pp. 906–944.

Criddle, E., Fox-Decent, E., Gold, A., Kim, S. H. and Miller, P. (Eds) (2018) *Fiduciary Government*. Cambridge: Cambridge University Press.

Djelic, M-L. and Bothello, J. (2013) "Limited Liability and Its Moral Hazard Implications: The Systemic Inscription of Instability in Contemporary Capitalism," *Theory & Society*, Vol. 42, No. 6, pp. 589–615.

Dow, S. (2012) "Economics and Moral Sentiments. The Case of Moral Hazard," in Castro Caldas, J. and Neves, V. (Eds), *Facts, Values and Objectivity in Economics*. London/New York: Routledge.

Dowd, K. (2009) "Moral Hazard and the Financial Crisis," *Cato Journal*, Vol. 29, No. 1, pp. 141–166.

Dworkin, R. (1981) "What Is Equality? Part 2: Equality of Resources," *Philosophy & Public Affairs*, Vol. 10, No. 4, pp. 283–345.

Fox-Decent, E. (2011) *Sovereignty's Promise. The State as Fiduciary*. Oxford: Oxford University Press.

Frankel, T. (2011) *Fiduciary Law*. Oxford: Oxford University Press.

Hale, B. (2009) "What's So Moral About The Moral Hazard?," *Public Affairs Quarterly*, Vol. 23, No. 1, pp. 1–25.

Heath, J. (2009) *Filthy Lucre. Economics for People Who Hate Capitalism*. Toronto, Ontario: HarperCollins.

Ireland, P. (2010) "Limited Liability, Shareholder Rights and the Problem of Corporate Irresponsibility," *Cambridge Journal of Economics*, Vol. 34, No. 5, pp. 837–856.

Kuperman, A. (2008) "The Moral Hazard of Humanitarian Intervention: Lessons from the Balkans," *International Studies Quarterly*, Vol. 52, No. 1, pp. 49–80.

Laby, A. (2008) "The Fiduciary Obligation as the Adoption of Ends," *Buffalo Law Review*, Vol. 56, No. 1, pp. 99–167.

Landes, X. (2013) "The Normative Foundations of (Social) Insurance: Technology, Social Practice and Political Philosophy," *Working Paper-LPF 6*, Torino: Centro Einaudi & Laboratorio di Politica Comparata e Filosofia Pubblica.

Leaver, A. (2015) "Fuzzy Knowledge: An Historical Exploration of Moral Hazard and Its Variability," *Economy & Society*, Vol. 44, No. 1, pp. 91–109.

McCaffrey, M. (2017) "The Morals of Moral Hazard: A Contracts Approach," *Business Ethics: A European Review*, Vol. 26, No. 1, pp. 47–62.

Miller, P. (2014) "The Fiduciary Relationship," in Gold, A. and Miller, P. (Eds), *Philosophical Foundations of Fiduciary Law*. Oxford: Oxford University Press.

Miller, P. and Gold, A. (2015) "Fiduciary Governance," *William & Mary Law Review*, Vol. 57, No. 2, pp. 513–586.

Miron, J. (2009) "Bailout or Bankruptcy? A Libertarian Perspective on the Financial Crisis," *Cato Journal*, Vol. 29, No. 1, pp. 1–17.

Okamoto, K. (2009) "After the Bailout: Regulating Systemic Moral Hazard," *University of California Law Review*, Vol. 57, No. 1, pp. 183–236.

Pauly, M. (1968) "The Economics of Moral Hazard: Comment," *American Economic Review*, Vol. 58, No. 3, pp. 531–537.

Pauly, M. (2007) "The Truth about Moral Hazard and Adverse Selection," *Lourie Lecture Policy Brief No. 36/2007*. Syracuse: Syracuse University.

Richmond, D. (1999) "Trust Me: Insurers Are Not Fiduciaries to Their Insureds," *Kentucky Law Journal*, Vol. 88, No. 1, pp. 1–32.

Rowell, D. and Connelly, L. (2012) "A History of the Term 'Moral Hazard'," *Journal of Risk and Insurance*, Vol. 79, No. 4, pp. 1051–1075.

Smith, L. (2014) "Can We Be Obliged to Be Selfless?," in Gold, A. and Miller, P. (Eds), *Philosophical Foundations of Fiduciary Law*. Oxford: Oxford University Press.

Stern, G. and Feldman, R. (2004) *Too Big To Fail. The Hazards of Bank Bailouts*. Washington, DC: Brookings Institution Press.

Stone, D. (2002) "Beyond Moral Hazard: Insurance as Opportunity," in Baker, T. and Simon, J. (Eds), *Embracing Risk. The Changing Culture of Insurance and Responsibility*. Chicago: University of Chicago Press.

Teitelbaum, M. (2015) "The Truth About the Migrant Crisis. Tragic Choices, Moral Hazards, and Potential Solutions," *Foreign Affairs*. https://s.gifford.ucdavis.edu/uploads/filer_public/b9/28/b9284ac6-06fc-4588-ae30-2b8b38a39ffd/paper_teitelbaum_for-martin-seminar.pdf

Williams, A. (2009) "Liberty, Equality, and Property," in Dryzek, J., Honig, B. and Phillips, A. (Eds), *The Oxford Handbook of Political Theory*. Oxford: Oxford University Press.

Young, I. M. (2011) *Responsibility for Justice*. Oxford: Oxford University Press.