Third Party Moral Hazard and the Problem of Insurance Externalities

Gideon Parchomovsky  
*University of Pennsylvania Carey Law School*

Peter Siegelman  
*University of Connecticut School of Law*

Follow this and additional works at: https://scholarship.law.upenn.edu/faculty_scholarship

Part of the Banking and Finance Law Commons, Behavioral Economics Commons, Cognition and Perception Commons, Finance Commons, Finance and Financial Management Commons, Insurance Commons, Insurance Law Commons, Law and Economics Commons, Law and Psychology Commons, and the Litigation Commons

Repository Citation

https://scholarship.law.upenn.edu/faculty_scholarship/2811

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship at Penn Law by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.
Third Party Moral Hazard and the Problem of Insurance Externalities

Gideon Parchomovsky* and Peter Siegelman**

Revision of Nov 4, 2021

Insurance can lead to loss or claim-creation not just by insureds themselves, but also by uninsured third parties. These externalities—which we term “third party moral hazard”—arise because insurance creates opportunities both to extract rents and to recover for otherwise unrecoverable losses. Using examples from health, automobile, kidnap, and liability insurance, we demonstrate that the phenomenon is widespread and important, and that the downsides of insurance are greater than previously believed. We explain the economic, social and psychological reasons for this phenomenon, and propose policy responses. Contract-based methods that are traditionally used to control first-party moral hazard can be welfare-reducing in the context of its third-party analog, so new approaches are required.

* Robert G. Fuller Jr. Professor of Law, University of Pennsylvania School of Law and Wachtel, Lipton, Rosen and Katz Professor of Corporate Law, Hebrew University.

** Phillip I. Blumberg Professor of Law, University of Connecticut Law School. For excellent comments and suggestions we thank Matt Adler, Emad Atiq, Ronen Avraham, Ian Ayres, Roy Baharad, Tom Baker, Oren Bar Gill, Jonathan Borowsky, Rick Brooks, Giuseppe Dari-Mattiacci, John Donohue, Yoav Dotan, Lee Epstein, Michael Frakes, Brian Galle, Michele Graziaidei, Ehud Guttel, Daniel Hemel, Paul Heaton, Louis Kaplow, Peter Kochenburger, Gary Klein, Alexandra Lahav, Adi Leibovitch, Daryl Levinson, Adi Libson, Daniel Markovits, Pat McCoy, Ittai Paldor, Mitch Polinsky, Ariel Porat, James Quiggle, Arti Rai, Barak Richman, Dan Schwartz, Rick Swedlof, Al Sykes, Kristin Underhill, Abe Wickelgren, Eyal Zamir and participants in seminars at Villanova, Duke, Harvard, Stanford, Yale and the Hebrew University law schools, and the 2019 American Law & Economics Association meetings. The editor and two anonymous referees made extensive suggestions from which we greatly benefitted. Special thanks to Steve Shavell for substantive comments and intellectual support. For outstanding research assistance, we thank Uria Beeri, Paul Cotler, Shira Efron, Nadav Israeli, Daniel Margolin, Roi Ohayon, UConn Law Librarians Adam Mackie and Anne Rajotte, Daniel Margolin and Ben Weitz. Parchomovsky thanks the Israel Science Foundation for research support.

Electronic copy available at: https://ssrn.com/abstract=3554010
The problem of moral hazard has preoccupied insurance economics: it has been the subject of almost 1,600 scholarly articles since Arrow’s (1963) pathbreaking analysis and the term has appeared in more than 850 judicial opinions since the 19th century. But scholars and courts have focused exclusively on how insurance alters the insured policyholder’s incentives for loss-prevention. The possibility that insurance might influence the conduct of those who are not parties to the insurance contract—a kind of insurance externality—has been noticed by a few scholars, but remains undeveloped and uncharacterized as a general phenomenon. We call this effect “third party moral hazard,” defined as the influence of insurance on the loss-creation or claiming behavior of non-parties to the insurance contract. Our main contribution is to demonstrate that third party moral hazard is both widespread and significant, and to offer an assessment of causes and consequences.

Consider several representative examples: Kidnap insurance covers some or all of the ransom demands made by kidnappers. There is evidence that the presence of kidnap insurance leads to more kidnappings, and several countries have banned such insurance, or attempted to do so. Insurance seems to raise the profitability of kidnapping and encourage “entry” by kidnappers (Weil 2014). An identical logic applies to insurance against cyber-ransoms, which has also been linked to an increase in attacks (Dudley 2019).

Third party moral hazard has also permeated the healthcare industry, where over-billing and outright fraud are enabled by the presence of insurance. Liability insurance, too, gives rise to third party moral hazard. For example, Directors and Officers Liability Insurance has apparently led to more litigation against corporate defendants (Donelson et al 2018). Some of this increase may have been strike suits designed to extract settlement offers; but insurance can also lead to an increase in meritorious claims that would not otherwise have been brought (for example, if the defendant is judgment-proof: Shavell 1986; Levin 2019).

What do these examples (and there are many more) have in common? First, each illustrates how the presence of insurance can create an incentive for loss- or claim-creation by those who are not party to the contract of insurance. A second commonality among most—but not all—of these examples is that they represent a form of unproductive rent-seeking behavior (Tullock 1966), in the sense that they are attempts to extract or transfer wealth from one party to another. Some of

---

1 Key insights were further developed by Pauly (1968), Holmstrom (1979) and Shavell (1979).
2 A recent exception is Balafoutas et al (2017, 1), who experimentally study what they term “second-degree moral hazard” in the market for credence goods. This occurs when “the supply side in a market . . . react[s] to anticipated moral hazard on the demand side by increasing the extent or price of the service.” In their experiment, taxi drivers charged roughly 17 percent more to passengers who let it be known that their expenses would be reimbursed than to those who didn’t. Kerschbamer et al (2016) report similar results in a field experiment involving computer repairs: customers who reported that they were insured received substantially higher repair bills. For further discussion, see Huber (2020). For a recent discussion of general equilibrium effects of insurance and its effects on third parties, see Baharad (2021).
these activities constitute insurance fraud, and some of them are crimes (Manes 1944), but the concept of third party moral hazard is both broader and narrower than fraud or criminality. Moreover, third party moral hazard even reaches some activities with positive welfare effects, such as the pursuit of meritorious suits against defendants whose assets, without insurance, would be insufficient to pay for a judgment rendered against them. In that instance, the presence of insurance incentivizes claims that should be brought, but would not have been in its absence.

Our analysis draws from and contributes to two literatures. The first is insurance law and economics (Schwarcz and Siegelman 2017; 2015). Despite its neglect by scholars, our most important point is simply that third party moral hazard matters. Imagine that 75 percent of all policyholders are subject to conventional moral hazard, which leads them to reduce their level of precautions by 10 percent. Assume further that this drop in precautions leads to a 5 percent increase in losses for this group, so that total losses rise by 3.7 percent. Even if only 1 percent of the public is subject to third party moral hazard, the damage done by this group’s deliberate loss-creation could easily be as large as the primary moral hazard effect. The purposeful loss-creation of the few may be more harmful than the modest additional negligence of the many.

Third party moral hazard has other important consequences for insurance markets as well. For one thing, it often cannot be effectively controlled by many of the standard techniques that insurers use to limit its first-party cousin (Ben-Shahar & Logue 2012). Neither deductibles nor loss controls (contractually-mandated precautions) are likely to be effective against third party moral hazard; and some loss controls that reduce traditional moral hazard may have negative consequences when used against its third party variant. Moreover, the costs of third party moral hazard (and of attempts to limit it) will be passed on to insurance customers. This in turn reduces the demand for insurance, limits the efficacy of risk-spreading, and can even induce follow-on selection effects. So in addition to the actual costs of third party moral hazard, there

---

3 For example, under Connecticut law,
[a] person is guilty of insurance fraud when the person, with the intent to injure, defraud or deceive any insurance company: (1) Presents ... to any insurance company, any ... statement ... as part of ... any application for any policy of insurance or a claim for payment or other benefit pursuant to such policy of insurance, knowing that such statement contains any false, incomplete, or misleading information concerning any fact or thing material to such application or claim; or (2) assists, abets, solicits, or conspires with another to prepare or make any ... statement that is intended to be presented to any insurance company in connection with ... any application for any policy of insurance or any claim for payment ... knowing that such statement contains any false, incomplete, or misleading information concerning any fact or thing material to such application or claim for the purposes of defrauding such insurance company.

are the costs of combating it, and spillover effects on insurance consumers and the industry as a whole.

We also draw on criminal law theories of “displacement” (Reppetto 1976; Ayres & Levitt 1998), to argue that precautions against third party moral hazard may have different efficiency consequences from those taken against conventional risks covered by insurance. Human-caused risks such as third party moral hazard have a strategic aspect that naturally-occurring misfortunes lack (Hirshleifer 1953). If an actor’s precautions simply divert harmful activity somewhere else, without reducing overall risk, social and private benefits will diverge. Private efforts to deter many human-engendered strategic risks, even if successful at the individual level, are therefore not as socially beneficial as they appear to be. That insight feeds back into insurance theory. Insurer-mandated loss controls will often divert, rather than prevent, human-caused losses or claims. Once third party moral hazard is taken into account, the welfare consequences of loss controls become more complicated than ordinary accounts (Ben-Shahar & Logue 2012) tend to suggest.

Accordingly, to ameliorate third party moral hazard, it is necessary to adopt a combination of procedural and substantive mechanisms that allow insurers’ to deal with the problem. Our suggested remedial approaches range from precautions mandated by insurance companies, to governmental assistance to the authorization of qui tam suits.

Part I provides a series of case studies of third party moral hazard in several domains of insurance. Part II considers several mechanisms that can explain the phenomenon. Part III suggests some policy responses.

1. THIRD PARTY MORAL HAZARD IN ACTION

In this Part, we draw on evidence from existing studies to discuss the existence and empirical significance of third party moral hazard in several domains of insurance. The cases we discuss span a range of insurance contexts and offer abundant evidence that the phenomenon is worthy of attention.

1.1. Liability Insurance

1.1.1. Directors & Officers (D&O) Liability Insurance

Directors and Officers ("D&O") liability insurance protects corporate officials, as well as the corporation itself, against potential liability arising from negligent actions and omissions that harm the corporation (Baker & Griffith 2007a). D&O liability insurance pervades the corporate world (Baker & Griffith 2007b) and covers virtually all public corporations in the U.S. and Canada.

Champions of D&O liability insurance contend that it lowers “the cost of compensating risk-averse directors and officers and encourages them to take appropriate business risks” (Holderness 1990, 115-16). Detractors maintain “that liability insurance largely nullifies the disciplining potential of litigation, causing directors and officers to be less attentive to their duties to shareholders” (Holderness 1990, 116), a version of conventional moral hazard.
Regardless of these effects, the key question for present purposes is whether the existence of D&O liability insurance creates third party moral hazard, in the form of enhanced incentives for litigation. At first blush, it seems clear that the answer is “yes” (Parsons 2003). Dan Puchniak (2012, 17) listed “the high levels of directors and officers (D&O) liability insurance,” as one of the principal factors that contribute to the high rate of derivative litigation in the U.S. Lin et al (2011, 507) note the possibility that D&O insurance “can attract frivolous shareholder suits.”

Studies from other countries seem to suggest the same conclusion. For example, Mark West (2001) studied the spike in derivative litigation in Japan in the 1990s, and suggested that the presence of insurance was the main driver of the litigation trend. He suggested, however, that the causal mechanism was not necessarily third party moral hazard, but rather standard moral hazard: corporate directors and officers became less circumspect in the fulfillment of their duties once they knew they were insured (West 2001, 376).

A careful examination of the relationship between the presence of D & O insurance and derivative litigation rates reveals a more nuanced picture, however. The high rate of derivative litigation in the U.S., relative to other countries, cannot prove, on its own, that D&O liability insurance generates a third party moral hazard problem. Nor can the fact that companies with high D&O liability insurance coverage get sued more often. There are three possible explanations for the fact that firms with D&O liability insurance are sued more often. The first is standard moral hazard. Directors and managers who have insurance tend to be less diligent in the performance of their obligations. The second is that firms with high coverage are more poorly managed to begin with: adverse selection leads the worst risks to purchase the most insurance. The third explanation is third-party moral hazard: the presence of insurance directly incentivizes litigation.

The three explanations are not mutually exclusive, and there are empirical studies that provide support for each. For example, a recent study reports that firms with high D&O liability insurance premiums are indeed more likely to be involved in litigation. In exploring the root cause for the correlation, the authors arrive at the tentative conclusion that “the association we find is more likely attributable to opportunism or moral hazard in the managers’ actions” (Gillan & Panasian 2015, 805). The article did not consider the possibility of third party moral hazard, however.

A study that used data on Canadian corporations also reported significant correlation between D&O insurance premia and “the quality of firms’ governance” (Core 2000, 476). Core (1997, 84) reports that “firms with greater litigation risk and higher distress probability are more likely to purchase D&O insurance and carry higher limits). In analyzing the reasons for this finding, Core endorses the second explanation, i.e., that higher D&O liability insurance is indicative of poor governance structures. Here, too, we found no reference to the third party moral hazard hypothesis.

Other empirical studies, however, are consistent with third party moral hazard. Tom Baker and Sean Griffith (2007b) examined the elements that affect settlements of class actions, using a qualitative research methodology, i.e., interviews with industry participants responsible for
settling securities cases. While Baker and Griffith do not focus on third party moral hazard or expressly mention the term, the answers to their questions suggest that industry participants view D&O insurance as a critical element that shapes their litigation strategy. As one respondent, a claims head, asserted: “I think it is easier to get money out of an insurance carrier than it is out of an insured. Why? Because it is a third party’s money” (2007b, 806-07). Baker and Griffith further reported that another respondent, a plaintiff’s lawyer, was far more specific about the effect of D&O liability insurance on plaintiffs’ motivation to sue, explaining that plaintiffs’ lawyers “sue for insurance” (2007b, 805).

Finally, Donelson, et al (2018) reported a strong correlation between firms’ levels of D&O insurance and their likelihood of being involved in securities litigation. Their study takes advantage of a unique feature of New York law that mandates disclosure of D&O insurance premia. No such requirement exists in the law of other states. The study then compares the exposure to securities litigation of New York firms and non-New York firms. They note that information about D&O premia have two possible effects on litigation: First, high premia can signal to lawyers a potential for high settlement rates. Second, high premiums may indicate that the insurer thought that the firm represented a high legal risk. Analyzing a large database of class actions brought against firms between 1998 and 2010, they found that “the relation between premiums and litigation is stronger for firms incorporated in New York, compared to firms incorporated elsewhere. As this relation is based on premium levels (rather than disclosure presence), this implies that the disclosure content affects litigation” (Donelson et al 2018, 530). The authors conclude that if D&O premia were known in all states, it would have led to an increase of 12 to 19 percent in securities class-action litigation over the period they studied. The potential costs of non-meritorious lawsuits are not trivial, and include “lawyer fees, nuisance settlements for cases that survive a motion to dismiss, and lost managerial time,” as well as an “undermin[ing] of the credibility of the U.S. securities litigation system” (Donelson et al 2018, 531).

1.1.2. No-Fault Automobile Insurance

No-fault automobile insurance is designed to simplify the payment of lower-value claims and reduce claims-processing costs. Rather than litigating who was at fault in smaller claims, no-fault makes a policyholder’s own insurer responsible for compensating him or her. In New York, for example, medical and other expenses totaling less than $50,000 are covered by each insured’s (mandatory) Personal Injury Policy (PIP). More serious injuries can be litigated in the usual manner, with the injurer’s insurer responsible for compensating the victim.

Even though no-fault operates as first-party insurance, it offers substantial opportunities for third party moral hazard. The extent of the problem is hard to assess. Data are difficult to obtain and—if they are from insurers—almost impossible to verify. But there is abundant anecdotal evidence demonstrating elaborate and well-organized schemes to falsify and exaggerate claims,

4 “Applying the higher dismissal rate of New York firms to the broader Compustat/CRSP sample projects to between approximately 160 to 260 additional non-meritorious suits that might have been filed over the 13-year sample period, which is equivalent to 12% to 19% total securities class actions over this period” (Donelson, et al 2018, 531).
run as sophisticated criminal enterprises in several states (Dornstein 1998).\(^5\) According to testimony by a Michigan doctor who ran an MRI facility outside of Detroit, for example, a prominent “TV lawyer,” Mike Morse, repeatedly leaned on him to exaggerate the severity of injuries detected in MRI scans of patients referred to him by Morse (Wisely & Reindel 2017). A network of chiropractors and physical therapists was also used to overbill for treatments and make and receive referrals (Gunalabam, 2017). The same pattern was documented in a criminal complaint against more than two dozen defendants filed by the United States Attorney for the Southern District of New York in 2015 against numerous medical clinics that were described in the indictment as “medical fraud mills that routinely billed automobile insurance companies under the No Fault Law for medical treatments that were either (i) never provided and/or (ii) unnecessary, because the Patients did not medically need the treatments” (Bharara, 2015).

Although New York law requires that medical clinics be “owned, operated and controlled” by a licensed medical practitioner (a doctor), the defendants apparently hired doctors to serve as fronts for the clinics, while secretly owning and controlling the operations of the clinics themselves. The no-fault clinics also employed “runners” (recruiters) who were paid $2,000-$3,000 per patient brought into the clinic (depending on the quality of the accident report filed with the patient’s injury).

This was not the only such organized scheme. A 2018 article in the New York Times (Rashbaum et al 2018) reported that “countless phony companies were cropping up [in the early 2000’s] to exploit so-called no-fault auto insurance laws in New York.” These companies “staged or exaggerated car accidents” and then set up billing companies “to collect money from insurers...”

Estimates of the magnitude of third party moral hazard in no-fault insurance vary substantially. A widely-cited insurance industry report put no-fault fraud losses in New York at $230-$240 million in 2009 (Insurance Information Institute, n.d.). But others have disputed that estimate, and the details of the calculation never seem to have been made public, so it is difficult to verify.\(^6\) According to the National Association of Insurance Commissioners, the industry’s

\(^5\) Consistent with the exaggeration of losses so as to cross the liability threshold for tort claims, Anderson, et al (2010, 97-111), find evidence in some no fault states of discontinuities or kinks in the distribution of claims: fewer-than-expected claims just below the threshold and more-than-expected claims just above it. They also conclude (2010, 111) that there has been “an increase in prevalence of suspicious claims in no-fault states over time.”

\(^6\) According to one (partisan) source (Israel 2011),

[t]he breadth of fraud, suggested by the industry, is simply exaggerated and untrue. . . Indeed, even a cursory review of the industry statistics lead them to be questioned. For example, one industry-funded source claims that no-fault fraud cost $240 million in 2009. . . However, a simple multiplication of the total number of reported questionable claims to the Insurance Department last year by the average cost of the entire no-fault claim that is being discussed today actually equals to $116 million, literally half of what is being asserted by the industry. And yes, while $116 million in suspicious [sic] fraud sounds like and is a lot, it . . . [amounts to] roughly 1 percent of [the $9.9 billion in] premium dollars collected by NY auto insurers [in 2009].
incurred losses on PIP policies in New York in 2009 amounted to $1.6 billion (National Association of Insurance Commissioners, 2009-10). So the $240 million estimate for no-fault fraud (almost all of which is probably third party moral hazard) amounts to 15 percent of all losses; opponents have suggested that this is too high by a factor of 2, but even if we put the figure at 7.5 percent, it translates to a loss of $120 million.

1.2. Health Insurance

The tangled relationships among insurers, individual policyholders, drug companies, doctors, and hospitals are rife with examples of third parties creating or exaggerating expenses so as to take advantage of what insurance will pay for (Delgado et al 2014). Indeed, the malign effects of third-party moral hazard in health care are the major theme of a recent 550-page book by legal scholars Charles Silver and David Hyman (Silver & Hyman 2018). In the forward to that book, a former Dean of Harvard Medical school opined that the “root cause[ ]” of waste and excessive spending in healthcare is “the . . . overreliance on insurance and other forms of third party payment” to cover medical expenses (Flier 2018, xx). In Silver and Hyman’s view, healthcare is all about third party moral hazard writ large: “health care is expensive because it is insured. . . . Insurance makes healthcare more expensive than it would be if people paid for it themselves” (Silver & Hyman 2018 at 295).

Economists have long recognized the potential for conventional moral hazard in health care, and have suggested that doctors might serve as vital checks on the tendency of insured patients to demand more care than is optimal (Arrow 1968, 538). But since Evans (1974), economists have also understood that “physician-induced demand” can be a source of increased healthcare spending (see, Johnson, 2014 for a survey). Notably, however, the theoretical literature seems not to have focussed on the role of insurance in demand inducement—Johnson’s survey pays no attention to insurance, nor does it list any sources that do.

1.2.1 Drug Rehab Clinics

The presence of insurance has contributed to serious wrongdoing in the provision of rehabilitation services for opioid addicts (Lurie 2019; Wooten 2019). Some states now host a sizeable industry of for-profit rehab centers that offer recruiting bonuses to freelance “patient brokers” to attract addicts whose treatment can be billed to their insurer. Such rehabilitation services are considered “essential benefits” under the Affordable Care Act, and must be offered without annual or lifetime caps, creating a large pool of money that unscrupulous centers can tap into. Their business model is to bring in patients for “treatment” that is completely ineffective, precisely so as to generate a relapse and readmission shortly thereafter. This practice is known in the addiction community as “the Florida shuffle,” “a cycle wherein recovering users are wooed

---

\[7\] See 42 U.S.C. § 18022(b)(1)(E) (2010) (defining Essential Health Benefits to include “Mental health and substance use disorder services, including behavioral health treatment”).
aggressively by [for profit rehabilitation centers] and freelance ‘patient brokers’ in an effort to fill beds and collect insurance money.\(^8\)

Rehab centers can charge private insurers as much as $50,000-$100,000 per month of treatment; and private insurance claims for treatment of opioid users rose from $75 million in 2011 to $650 million in 2014 (lurie 2019). The availability of generous insurance payments has helped create an entire industry of phony rehabilitation programs whose purpose is precisely not to cure addiction, but rather to maintain it as a source of ongoing revenue.\(^9\)

1.2.2. Copay Coupons for Branded Drugs

Health insurers have an obvious interest in incentivizing patients (and their doctors) to choose lower-cost generic drugs instead of branded drugs that are medically identical. One mechanism used to create such incentives has been to set a higher copayment by the patient if she chooses the branded drug instead of the generic alternative (Frakt, 2017; Dafny et al, 2018). Manufacturers of branded drugs have figured out a way to subvert this incentive, however: copay coupons. Suppose the copay on the brand-name drug is set at $65, while the generic version has a copy of $10. The manufacturer gives the patient a coupon that covers $55 of the branded copay, yielding the same out of pocket cost as for the $10 generic. If the seller makes an additional $130 sale (the price paid by the insurer) by diverting the consumer from a generic drug, the $55 copay subsidy generates a profit of $75.

The effects of copay coupons are large and negative, according to a recent study: “[C]oupons increase branded sales by 60+ percent, entirely by reducing the sales of bioequivalent generics” (Frakt 2017, 91). Copay coupons do not simply divert customers from cheaper generic substitutes to branded alternatives, however. They also enable higher prices: the “coupon holds the consumers’ prices fixed at a low level. That allows the manufacturer to raise the overall price without losing sales. This raises spending, too, but for the insurer” (Frakt 2017). According to the best recent evidence, copay coupons cause several billion dollars in additional spending per year (Dafny et al 2018).

Copay coupons are a compelling example of drug companies’ exploitation of insurance to charge higher prices, but there are others. For example, Besanko et al (2016) conclude that the presence

---

\(^8\) https://www.fixthefloridashuffle.com/florida-shuffle. Even worse, reporting by Lurie (2019) uncovered examples in which brokers actually induced recovering addicts to relapse so that they could be referred to a “treatment” center.

\(^9\) Mitch Polinsky (personal communication) pointed out to us that a contractual solution to the problem of drug rehab centers is available if insurers could simply buy and operate them themselves. The idea of such “Integrated Delivery Systems” is attractive, with some influential backers. As Enthoven (2009) put it, “an organized, coordinated, and collaborative network that links various healthcare providers to provide a coordinated, vertical continuum of services to a particular patient population or community” can both reduce costs and improve the quality of healthcare.
of insurance has enabled manufacturers of oral chemotherapy drugs to raise prices beyond the estimates of the additional value the drugs create for consumers.

1.3. Kidnap insurance

The economics of kidnapping has been the subject of considerable analysis, both empirical and theoretical (Vannini et al 2015; Fink and Pingle 2014). There is less on the impact of insurance, however (Shortland 2017 and Shortland 2019 are important exceptions). A key finding of most theoretical models is that the volume of kidnappings will depend on the average amount that victims are willing and able to pay. In other words, there is an upward-sloping supply of kidnaps, so that the higher the expected ransom, the greater the number of kidnappings (Vannini et al 2015; Detotto 2015).

The question of interest to us is whether the presence of insurance enhances victims’ willingness or ability to pay ransoms. If so, the availability of insurance would likely cause more kidnaps. Common sense would seem to suggest an affirmative answer (Clendenin 2006; Kenney 2008; Dutton 2016; Block and Tinsley, 2008), as would at least some theoretical models of the kidnap market. Fink and Pingle (2014, 491) point out that

\[ \text{the provision of kidnap insurance might have an effect on the probability that kidnappings are contemplated. For instance, a potential kidnapper might hear about a successful exchange of a victim for ransom that may have been facilitated with the aid of an insurance company, or a potential kidnapper might conclude from the presence of a kidnap insurance market that his potential victims are on average willing to pay more ransom compared to the situation when no kidnap insurance market exists.} \]

Indeed, in Fink & Pingle’s game theoretic model, “the existence of a competitive insurance market [always] increases the maximum [ransom] demand a family is willing to pay.” The size of the supply-side response to this increase is left open, however. Believing that kidnap insurance aggravates the problem of kidnapping and the costs associated with it, Venezuela, Colombia, and Italy have all banned kidnap insurance at one time or another (Vannini et al 2015, 8). And anecdotal evidence suggests that kidnaps fell in the aftermath of such bans, although the causal link is unclear. (For example, Bohlen (1978) points out that Italian policies that banned kidnap insurance also dramatically discouraged the reporting of kidnappings.)

---

\^10 In the Fink & Pingle (2014, 493) model, the positive effect of insurance on the volume of kidnapping occurs whenever “insurance increases [victims’] maximum ransom offer,” which in turn depends on some technical conditions. Even if insurance increases total kidnappings, it might still reduce fatal kidnappings, since higher ransom payments “push[ ] some victim offers above the net willingness to kill of the marginal kidnapper, reducing the fraction of fatal kidnappings.” The net effect on the total volume of fatal kidnappings depends on the relative size of the kidnap-increasing and fatality-reducing effects.
2. THE MECHANISMS OF THIRD PARTY MORAL HAZARD

In this Part, we investigate the possible causes of third party moral hazard. Since we are dealing with a phenomenon that has received little systematic scholarly attention, we offer several possible mechanisms by which the presence of insurance causes third parties who are not privy to the insurance contract to engage in risk causing or harm causing activities. We consider four explanations: deep pockets/rent-seeking, poorer detection, depersonalization and spillovers from traditional moral hazard.

2.1. Deep Pockets

The most straightforward explanation of third party moral hazard is that insurers are much wealthier than policyholders, making them more attractive targets for ripping-off, simply because they can cover larger losses. In many contexts, an uninsured entity’s ability to pay is significantly constrained by its wealth (Smith and Wright, 1992), but that is virtually never true for an insurer.

For example, consider a staged motorcycle accident that was captured on a recent viral video with more than 4 million views (Viralhog 2017). Shot from the dashboard camera of a car travelling down a residential street in the U.K., it shows a motorcyclist wheeling his bike backwards into the slow-moving car, slamming the cycle into the car, jumping on the hood and smashing his head repeatedly against the windshield. The “victim” ends up on the ground, writhing in pain. Meanwhile, a friend of his films the aftermath of the incident. Absent insurance, the money available to pay for an extracted settlement would likely be very low, reducing the profitability of this kind of scheme. The mandate that British motorists carry very generous third party auto liability insurance\footnote{The Road Traffic Act, 1988, c. 52 §143(2) (Eng.) (as amended) requires coverage for any vehicle of at least £1,200,000 ($1.7 million) in property damage.} thus makes such scams more profitable.

Some rent-seeking schemes are only possible because of the presence of insurance. Organized criminals in Sicily apparently developed a practice of deliberately inflicting significant and gruesome injuries to homeless people and drug addicts, in whose name they had previously taken out insurance policies (D’Emilio 2018).\footnote{Eleven people were arrested, and the amounts of money in question allegedly reached several hundred thousand Euros in this case.} Such behavior would obviously not be profitable if the victim were “uninsured,” since there would simply be no wealth to transfer in the absence of insurance.

We note that the mere existence of deep pockets does not necessarily imply that rent-seeking is always the explanation for third party moral hazard. For example, as Leslie Levin (2019, 113) has noted in the context of legal malpractice litigation,
[m]any solo and very small firm lawyers do not carry LPL [lawyer’s professional liability] insurance. If there is no insurance . . . experienced plaintiffs’ lawyers will almost never take on the malpractice case. Plaintiffs’ lawyers know that even if a case is meritorious, they will not receive their contingent fee because there will be no money to pay the judgment.

Our key point is that the same logic applies both to those creating losses (or feigning harm) and those who experience genuine losses greater than the injurer’s wealth: there is no profit in making a claim—whether real or phony—for a loss that the injurer cannot pay. Thus, by guaranteeing that there will be someone who is able to pay, the presence of insurance promotes claiming, both meritorious and unmeritorious. It is that increase in claiming, whether for good or ill, that constitutes third party moral hazard. Put differently, pockets can sometimes be “too shallow” as well as too deep. When the presence of insurance induces additional claims by relaxing liquidity constraints, third party moral hazard can actually be beneficial.

2.2. Depersonalization

Insurance companies are not flesh and blood individuals. They are psychologically “remote” corporate entities that are understood to be in the business of paying-out to cover losses. That could lead third party loss-causers to feel less inhibited when dealing with an insurer than with an individual who has to come up with the funds from her own bank account. For example, Tennyson (1997, 260) concludes (based on a survey of consumer attitudes) that “tolerant attitudes toward fraud will be expressed more often by individuals who have negative perceptions of insurance institutions.” Tennyson (2008) draws on a wider variety of evidence to reach similar conclusions. Money paid out by an insurer probably has a different moral and psychological valence for loss-causers than money taken from the pocket of a particular individual.

---

13 This phenomenon is similar to John Nyman’ (2004) “access motive” for health insurance coverage. Consider a kidney transplant that costs $300,000 and will be required by 0.1% of the population. Each person would be willing to pay $300 for coverage of a 0.1% risk of needing the operation, but no one has the $300,000 to spend on her or his own. Without insurance, therefore, there would be no transplants; with insurance, the cost of the procedure is no longer a constraint on any individual’s ability to afford it. As Nyman recognized, moral hazard can sometimes be welfare-enhancing. It also bears a close relationship to Pauly and Held (1990)’s analysis of “benign moral hazard” for coverage of preventative care.

14 A more traditional economic explanation concerns the diminishing marginal utility of wealth, which drives demand for insurance. Given that dollars are especially valuable in low-wealth states of the world, uninsured individual tortfeasors should be willing to invest heavily to contest liability since they are spending in high-wealth, no-liability dollars to avoid having to pay in low-wealth (post-liability) dollars. This makes them unattractive targets for third party loss-causers, especially compared to insurers, who “only” have routine profit motives at stake in deciding whether to contest a claim.
Baker (2001) describes the moral economy of “blood money” in tort litigation, and suggests that these kinds of social/psychological concerns are widespread. Baker reports that plaintiffs’ lawyers rarely ask for more in damages than the limits of the defendant’s insurance policy: funds obtained from an insured’s personal wealth (beyond what is covered by insurance) are known as “blood money,” and there is a widespread norm among plaintiff’s lawyers against going after such assets (unless the tortfeasor has deliberately chosen to under-insure). These findings are consistent with the hypothesis that loss-causers might feel differently about taking money from insurers than from identifiable individuals.

The fact that the Federal Rules of Evidence bar juries from hearing about a defendant’s liability insurance (or lack thereof) lends further support to the depersonalization theory. F.R.E. 411 states that

[...] evidence that a person was or was not insured against liability is not admissible to prove whether the person acted negligently or otherwise wrongfully. . . .

This so-called “Insurance Exclusionary rule” originated in 19th century common law, and is apparently based on the idea that “a jury confronted with a sympathetic plaintiff and informed that a well-healed defendant is insured may, despite lack of evidence of wrongdoing, find the defendant liable and impose a large award to provide for the plaintiff” (Greene et al (2008, 196)). The presence of insurance seems to matter to jurors; it probably also matters to claimants (whether bona fide or not).

2.3. Poorer Detection or Worse Bargaining

Beyond enhancing ability to pay, another reason why the presence of insurance may make for more loss-causing by third parties is that it can enhance willingness to pay: insurers may be in a worse position to detect false or fraudulent claims than are the individuals to whom such claims would otherwise be made. Or, insurers might be more generous (pay out more per dollar claimed) than individuals who are paying out of their own pockets.

We don’t have much evidence on these issues, but consider the example of no-fault auto insurance fraud in New York. According to a veteran fraud prosecutor (Ferguson, 2011), fraudulent claimants were able to swamp insurers’ claims-handling and investigation resources by submitting phony or exaggerated claims for reimbursement at a rate that was too high for insurers to handle within statutorily required time limits for payment of claims. Insurers thus ended up paying-out on claims that an individual tortfeasor would almost certainly not have conceded.

More generally, bad faith liability for an insurer’s failure to pay a claim could lead to similar effects (Browne, Pryor & Puelz 2004). If insurers are subject to treble damages for wrongly failing to pay a claim, they will at the margin be induced to pay some questionable claims that an individual paying out of his or her own funds would have been willing to contest (Tennyson & Warfel 2008). In turn, insurers’ reluctance to challenge some claims makes loss-creation more profitable for third parties.
Of course, the reverse possibility also needs to be considered: insurers might well be better at
detecting illegal claims than individuals would be. A high-volume claims-processing operation
could be in position to observe suspicious patterns in claiming that would be invisible to an
individual who only sees a single request for payment. And it is likely that there are economies
of scale in fraud-detection that insurers are better-placed to take advantage of than are individual
payors. As for skill in negotiating, Shortland (2017) reports that kidnap insurers drive
considerably tougher bargains than individuals. These insurers form a kind of loose cartel; and
they are repeat players, not one-shotters, so they internalize the costs that higher ransom payouts
(in today’s kidnapping) cause in the form of more future kidnappings due to deep
pockets/supply side response.

2.4. Spillovers from Ordinary Moral Hazard

Third party moral hazard can also occur as a kind of general equilibrium spillover that flows
from ordinary first party moral hazard. The idea is that when (first) party X obtains insurance,
(third) party Y—whose function is to monitor or protect X—now has less of a reason to do so.
(Recall that there is an insurer who constitutes the “second” party.) The presence of insurance
may thus alter Y’s behavior. We think of this as “communicable” Moral Hazard because it
spreads from insureds to those who interact with them. For instance, insurance may replace or
crowd-out monitoring by others, since the benefits of such monitoring are reduced if one’s
counterparty is insured. Goldberg (2009) argues similarly that corporate demand for liability
insurance is best explained by the desires of suppliers and customers to rid themselves of
various financial monitoring tasks and outsource them to their counterparty’s insurer.

Consider the specialized and generous form of insurance coverage known as jeweler’s block
policies, which cover the theft of precious stones from jewelers. If thieves know that most
jewelers have such coverage, they may be more likely to commit robberies, since they expect that
first party moral hazard on the part of jewelers will make theft easier. More mundanely, there is
evidence that auto or computer repair shops, taxicabs, and other service providers raise prices
for consumers who are not paying out of pocket, and thus face reduced incentives to shop for a
good deal (Balafoutas et al, 2017; Kerschbamer et al 2016). In this scenario, ordinary moral
hazard occurs when insurance reduces policyholders’ incentives to undertake comparison
shopping at a given price level; third party moral hazard is reflected in sellers’ charging higher
equilibrium prices because insurance has made demand less elastic by separating the costs of
search (borne by insureds) from its benefits (which accrue to insurers).

Sometimes, however, the contagion may move directly from the insurer to third parties, even
without any first party moral hazard. Third party moral hazard of this kind might have played a
“[c]onventional moral-hazard explanations of the excessive pre-crisis leverage of the big banks”

15 Judge Posner discusses these policies and their potential for conventional moral hazard in A.M.I.
Diamonds Company v. Hanover Insurance Company, 397 F.3d 528, 530 (7th Cir. 2005).
16 Thanks to Gary Klein for this insight.
and suggests instead that the “insured” financial firms “did not even need to think about the moral hazard of [their own] government bailouts—they merely needed to observe the exceptionally low costs of debt financing offered to them by creditors.” Rather than directly increasing risk-taking by insureds, this is a story in which insurance erodes market discipline: When a bank is understood to be “Too Big to Fail,” its lenders reason that there’s no need to monitor it or charge a penalty for its risky behavior, and it thus faces reduced incentives to be careful.

3. POLICY AND WELFARE IMPLICATIONS

In this Part, we detail the costs of third party moral hazard and the efficacy and efficiency of private sector attempts to limit it. Our bottom line is simple: controlling the loss-causing or claiming behavior of third parties is difficult for insurers, and even when it is possible, it may have counterproductive effects on welfare. We conclude that extra-contractual measures are needed to efficiently limit third party moral hazard, and we suggest several possibilities.

3.1. The Costs (and Benefits?) of Third Party Moral Hazard

In assessing the cost of third party moral hazard, it is important to distinguish between three kinds of costs that such behavior may impose. First, there are “primary” losses. Most economists would argue that payments made to kidnappers or fake accident victims do not themselves constitute social losses. Such payments merely transfer wealth between parties, at least to the extent that they do not use up any resources in the process (Tullock 1967). But kidnappings and other sources of third party moral hazard do generate real losses, including stress and hassle for kidnap victims and their families, the costs of negotiating ransoms, and so on. And even fake accidents can entail some actual harms: for example, if the “victim” in the auto accident had actually sustained an injury in staging the accident, this would constitute a social loss attributable to insurance.

In addition, third party moral hazard generates “secondary” losses in the form of the resources used in creating fake losses (for example, the time and effort to stage them). These expenditures constitute true losses from society’s perspective. So too do additional expenses by insurers to audit for fraud or otherwise prevent, detect, or deter third parties from their rent-seeking behavior.

17 In some cases, victims may suffer real harm, even if their initial injuries were not real. Consider a well-known case involving the deliberate and large-scale mis-diagnosis of silicosis by a ring of doctors and lawyers (In Re Silica Products Liability Litigation 398 F. Supp. 2d 563 (S.D. Tex. 2005)). Most of those diagnosed with the disease were not in fact suffering from it: they were “screened” by fake diagnostic firms that colluded with the plaintiffs’ lawyers. Some of the victims subsequently sued their lawyers, alleging that they “suffered injuries, both financially and mentally, in that they lived their lives [wrongly] believing they had been diagnosed with the incurable disease of silicosis” while their lawyers “reaped the profits of any false diagnosis” Canfield (2013).
Finally, third party moral hazard may also generate through the creation or worsening of adverse selection problems. Notice to begin that even if payments by insurers to “victims” are not true social costs, they are certainly real from the perspective of the insurer paying the claim. As such, they represent financial costs that the insurer has to cover, and in a competitive equilibrium, they will be recovered in the form of higher prices for insurance. Higher prices, in turn, make insurance less attractive to the least risky (marginal) customers. That’s because the marginal customer (the one just willing to buy insurance at the going rate) is typically less risky than the average customer among all those buying. Pushing people out of the market will lower welfare: those who value insurance by more than it would have cost to provide it to them will suffer a loss (Einav & Finkelstein 2011).  

3.2. Self-Help by Insurers

Insurers have developed a well-known arsenal of weapons to combat first party moral hazard, comprising at least the following seven techniques: (a) deductibles and copays (b) exclusions; (c) underwriting to screen high-propensity individuals; (d) experience rating; (e) loss controls; (f) ex post auditing of claims; and (g) ex ante monitoring.

There is a broad consensus among scholars of insurance law and economics that, in combination, these measures are reasonably effective at controlling first party moral hazard in most contexts (Baker 1996; Avraham, 2012; Ben-Shahar & Logue 2012; Shavell 1982). But these mechanisms are inoperative or ineffective when applied to third party moral hazard, and efforts to control deliberate loss-causing behavior have different welfare consequences from attempts to control ordinary moral hazard.

The first four techniques (deductibles, exclusions, underwriting and experience rating) govern risk via the contract between the insurer and the policyholder; they are of little use in constraining the loss-causing behavior of third parties. True, an insured with a substantial deductible will have some reason to prevent loss-causing by third parties, since she or he bears some share of any loss that occurs. But this effect is likely fairly small, and in some cases, may even worsen the problem of third party moral hazard (Sulzle & Wambach 2005). Experience rating also offers an insured a modest reason to limit losses caused by third parties, since even if she is covered for a current loss, her future premiums will increase as a consequence of any claim she makes. Underwriting is presumptively irrelevant to third party moral hazard, since the insured’s character or riskiness is not really at issue in the occurrence of losses caused by third parties. Similarly, exclusions are of limited value in deterring third party behavior, since the primary conduct that is excluded is that of the insured, while the loss is caused by someone else.

That leaves the last three techniques—loss control, ex post auditing, and monitoring—as the primary methods available to insurers to limit third party moral hazard. These may be effective

---

18 Recent theoretical developments also point to a link in the opposite direction, suggesting that the existence of adverse selection can lead to additional fraud (Boyer and Peter 2019).
in some contexts, but they pose significant problems for public policy that are unique to third party moral hazard.

3.2.1. Loss Control

In some settings, insurers can and do provide instruction to policyholders on how to avoid or mitigate risks posed by third parties. Loss control services are frequently bundled with the provision of kidnap insurance, for example (Shortland 2019), and consist mostly of advice on security—how big a wall to build around one’s house, which areas of the city to avoid, and so on.

Third-party generated losses are more difficult to control than those caused by first parties, however. Third parties are active loss-causers, who have agency and strategic capability. Steam boilers do not take countermeasures to increase their ability to explode if they are inspected more frequently. Kidnappers and pirates, however, do respond to efforts to limit their effectiveness. Moreover, some third party activities are simply not amenable to loss control directed at first parties. Drug copays and excessive MRI use cannot easily be mitigated by training policyholders about how to avoid such losses. Indeed, even when consumers are coached to use cheaper MRIs, they don’t do so (Lagasse 2018).

Some kinds of liability insurance—for example, Directors and Officers policies—might seem to be amenable to loss control activities directed against third parties. Insurers could provide companies with advice on implementing best practices for corporate governance so as to reduce the risk of being sued for a violation of corporate law, for example. But in practice, D&O insurers do not seem to make widespread use of loss control, for reasons that are not well understood (Baker & Griffith 2006; we discuss insurers’ creative use of the duty to defend to control third party moral hazard below).

By contrast, Employment Practices Liability (EPL) insurers widely tout their loss-control services as part of their marketing efforts (Talesh 2015). But even when loss control can be used effectively, limiting payouts to third parties has different welfare implications than preventing first party losses. There are at least two important differences between first- and third-party loss controls in liability insurance.

3.2.1.1. Harm Reduction vs Liability Reduction

The first is the distinction between loss prevention and harm prevention. In first party moral hazard, preventing losses is straightforward: the insurer tells the policyholder how to inspect his boiler to minimize the chance it will explode, or how to store chemicals to prevent a leak. But there are actually two kinds of “prevention” at work in the context of liability insurance: forestalling legal liability (conditional on harm) and preventing the occurrence of the harm itself. For example, many EPL insurers advertise their ability to help employers “bullet-proof” their employment policies by establishing proper training procedures, refining language in employee handbooks, and so on (Talesh 2015). Following best practices for Human Resources management
may indeed reduce the likelihood that a plaintiff is successful in a lawsuit against the insured. But preventing liability ex post is not the same as preventing the actual occurrence of losses ex ante, and most employer training probably has little or no effect on actual behavior—that is, on losses resulting from employee misconduct, such as harassment. Grossman (2003, 3) concludes that “[s]exual harassment policies and procedures do not seem to have any reliably negative effect on the incidence of sexual harassment.” Other scholars who have looked at the issue agree (Dobbin & Kalev 2016): loss control in the context of EPL insurance has become a kind of “bureaucratic vaccine against lawsuits for harassment.” (Dobbin and Kelly 2007, 1234 (emphasis added)).

Loss controls that actually reduce the incidence of harms ex ante are different from those that simply forestall liability ex post. Cost-effective harm reduction benefits both the insured and potential victims (who are not injured), with clear positive welfare effects. Forestalling liability (without reducing harm) will presumably deter some third parties from loss-causing behavior: fake victims are less likely to file lawsuits if such suits are less likely to succeed. But if insureds can defeat liability without actually reducing harms, then even victims with legitimate claims are less likely to succeed. Reducing the success rate of plaintiffs (without reducing the incidence of harm) deters deliberate loss-causers, but at a cost of leaving all losses where they fall, even for innocent victims who did nothing to cause their loss. Loss-controls work differently in the context of third party moral hazard.

3.2.1.2. Diversion

A second problem with efforts to control losses from third parties is that they may lead to the strategic diversion of harm-causing activity (Repetto (1976); Mikos (2006) offers a dissenting view). For example, suppose a kidnap insurer advises a policyholder to build a larger wall around his house to deter kidnappers. If the wall does its job, it will reduce the likelihood of third party moral hazard directed at the policyholder (and his insurer). But it will likely do so by deflecting potential kidnappers towards some other victim, with little or no gain to society as a whole. Indeed, shifting the incidence of misdeeds from one party to another will typically produce net social losses—building the wall uses up resources without reducing kidnapping at all, and may even create negative externalities in the form of too many walls.19

Our bottom line is that loss controls can sometimes reduce some kinds of moral hazard caused by third parties. Unlike efforts to limit first-party losses, however, loss control activities directed against third parties are socially suboptimal: insurers will choose the wrong scope and type of loss controls because they do not take sufficient account of the externalities that third-party loss controls engender.

19 A further wrinkle is that precautions that are unobservable to potential wrong-doers may yield social gains. For example, if kidnappers can’t tell who has an alarm system and who doesn’t, then when A installs a new alarm, she raises the average level of precautions among all victims, and hence lowers the expected return to kidnapping generally (Ayres & Levitt 1998).
3.2.2. Monitoring

Another approach to limiting ordinary moral hazard is through monitoring of the insured’s conduct. The conventional wisdom is that if an insured’s actions can be perfectly observed, it becomes possible to write a contract that specifies the degree of riskiness the insured is able to undertake, eliminating the problem of moral hazard altogether (Holmstrom 1979).

Of course, third parties can be difficult to monitor, since they operate in the wild and their identities are not known to the insurer in advance of any loss they cause. But first-party monitoring can be used to deter some kinds of third-party moral hazard, since loss-causers must usually interact with insureds in some fashion in order to create a compensable injury.

For example, automobile dashboard cameras would seem to reduce the possibilities for third party moral hazard in automobile insurance. It is therefore surprising that the technology has not been widely adopted in the US, although it is used in other countries. One explanation is that American auto insurers do not offer discounts for cars with dash cams (dashcameras 2021; Evangelista 2015), as their British counterparts aggressively do, with premium reductions of 10 to 30 percent (Times of London 2018). But that begs the question of why not. We note that mandatory minimum coverage in UK auto insurance is roughly two orders of magnitude higher than in the US. That gives UK insurers a strong incentive to subsidize this kind of monitoring. And since U.S. plaintiffs’ lawyers typically limit their claims to the amount of the defendant’s insurance coverage (Baker 2001), American drivers have reduced incentives to purchase a dash cam to protect their own assets.

3.2.3. Pre-Committing to (first party) Moral Hazard?

Most liability insurance policies give the insurer the duty of covering both the underlying loss (the duty to indemnify) and any legal defense mounted by the insured (the duty to defend) (Silver 2015; Richmond 2016). The logic for doing so is clear: if policyholders paid for their own defense, but insurers covered any damages for which the policyholder was ultimately found liable, the insured would want to spend almost nothing on defending the claim, since any damages would ultimately be paid by the insurer. Usually, therefore, the duty to defend is bundled with the right to control the defense: the party that has to pay (here, the insurer) gets to choose how to organize the defense and how much to spend. Allocating the costs of defense and

\footnote{Roughly one-quarter of British motorists use dash-cameras (Aviva, 2018). They are also very widely used in Russia (Lavrinc 2013).}

\footnote{As noted earlier, English motorists must have at least £1,200,000 ($1.7 million) in property damage coverage for any vehicle. While we have been unable to uncover any official survey of minimum coverage requirements in US auto insurance, our calculations suggest that the mean requirement for liability coverage across US states is only $17,000, with a median of $15,000; two states, Iowa and New Hampshire, have no minimum coverage at all. Findlaw.com (2021).}

\footnote{We owe the key insight in this section to a conversation with Steve Thel.}
the payment of liability with the same party avoids the obvious moral hazard problems that occur when a party has to pay to prevent losses to someone else.

Puzzlingly, however, some liability insurance contracts separate the duty to defend and the right to control the defense. This pattern is observed in D&O insurance (Silver, 2015), as well as areas such as errors and omissions in professional services, and some excess loss policies (Nierengarten 2013, p. 7; Nierengarten 2012; Richmond (2016) provides sample language and citations). Some patent-infringement liability insurance has this feature as well. Technically, this is known as a “duty to reimburse” or as a “defense cost indemnification policy,” rather than a typical duty to defend. Given the logic above, its existence seems puzzling: Why would insurers cede control over the defense and agree to let policyholders freely spend the insurance company’s money to defend against liability that policyholders will not have to pay?

One explanation is that insurers with reimbursement duties do retain some control over defense costs through “co-control” language, which allows them to forestall some moral hazard problems (Nierengarten 2012). But here, first party moral hazard may be a feature, not a bug, because it can be used to control third party moral hazard.

Suppose insureds face the possibility of strike suits that are brought purely for their settlement value. (These are suits that would lose if they were litigated, but for which the costs of mounting a defense are high relative to the costs of bringing a claim (Cooter & Rubinfeld 1989, 1083).) By prepaying (sinking the cost) for an insurance policy that covers generous spending to defend against litigation, insureds are able to make a credible commitment that they will not settle if sued. In turn, that should mean that plaintiffs will be deterred from filing strike suits in the first

23 See, e.g., http://patentinsuranceonline.com/Defense.html, which states that upon compliance with the policy terms, [the insurer] will then authorize the lawsuit filed against you. You have choice (sic) of legal counsel to represent you in the lawsuit; however, [insurer] may suggest reliable and preferred counsel for you to use—but the choice is yours. You also decide the final settlement terms.

24 “A defense-cost-indemnification policy is an insurance policy in which the insurer agrees to pay the costs of defense of a covered legal action and does not undertake the duty to defend. Typically, such policies also cover settlements and judgments.” Restatement of the Law of Liability Insurance § 22 (2019).

25 A public promise by the policyholder herself to vigorously litigate all claims would not be credible if she had to spend her own money to do so: when the time came, the insured would find it in her interests to settle to avoid having to incur the defense costs. But if she had already sunk the costs of insurance coverage, the marginal cost of litigation would be zero and the policyholder would thus want to spend lavishly on defense if she were sued. Leaving the duty to pay for defense costs with the insurer runs into the reverse commitment problem—the insurer would want to settle cheaply, rather than litigate. It is only by separating the decision about how much to spend on litigation from the duty to pay for these costs that the parties can sustain a credible commitment not to settle, and thereby deter litigation. After writing this, we discovered a more formal treatment of these ideas (Lloret & Suarez 2012; for a more elaborate model that situates liability insurance in a mechanism-design context, see Lemus, et al 2020.)
instance, knowing that policyholders will want to freely spend their insurer’s money to hire the best lawyers and use them extensively in defending against any claims. In equilibrium, therefore, defense reimbursement will not actually be necessary, and premiums can be quite low as a result.

If the duty to reimburse is used as a commitment device to deter strike suits in some kinds of liability insurance, an obvious question is why it is not used more broadly. One possibility is that it is employed where strike suits are common and the plaintiff’s bar is highly specialized and strategically sophisticated. That is a plausible characterization of the securities and the patent litigation landscapes. But where the plaintiff’s bar is insufficiently sophisticated, attempts to deter lawsuits by pre-sinking legal costs may not work. For example, most employment litigation is brought by unspecialized and unsophisticated lawyers (Moss (2013)), who might not understand how a pre-paying for defense costs shape defendants’ litigation incentives. Thus, we might expect that Employment Practices Liability Insurance (EPLI) would use a duty to defend, as appears to be the case,26 rather than a duty to reimburse.

3.3. Governmental Assistance

Insurers are constrained by contractual privity in coping with third party moral hazard. The contract between the insurer and policyholder can govern some aspects of the latter’s behavior, but is much less useful in controlling the behavior of third parties.27 Moreover, to the extent that insurers can deal with these problems, they may face the wrong incentives in doing so. These constraints, we believe, create an important role for governmental actors, whose powers are much broader and whose incentives differ.

Some kinds of third party moral hazard are independent crimes. Where that is the case (as, for example, schemes of the type described by Rashbaum et al (2018)), enhanced criminal enforcement is likely to be the best way of deterring large-scale third party moral hazard. As the head of the New York State Trial Lawyers Association put it, “[T]he easiest and the most direct way to [reduce no-fault auto insurance fraud] is better law enforcement efforts.” (Timko 2011, 205). Even where no criminal conduct is at stake, the state can still play a useful role by barring certain kinds of third party moral hazard, as Massachusetts and a few other states have chosen to do in the case of drug copay coupons discussed earlier.

26 We are unaware of any survey of the prevailing treatment of defense costs in EPLI policies. One industry insider has observed that “EPLI policies are of the ‘duty to defend’ variety” (Monteleone 1996, 160). Cf, Mount Vernon Fire Ins. Co. v. Visionaid, Inc., 477 Mass. 343, 345 (2017) (“policy provided that [insurer] had ‘the right and duty to defend any Claim to which this insurance applies’”).

27 Consider a repair shop that seeks to overcharge a customer, knowing she has insurance. First, as noted in Section 2.4, insured customers have no reason to push for lower charges. And while it might be possible to write the insurance contract in a way that required both parties to demonstrate that the repair costs are justified, this would be difficult and costly to implement. And the evidence cited in notes 5, 6, 8, and 14 suggests that in fact, the problem is not actually solved. Detecting overpayment is far from straightforward in most contexts; that is certainly the case in health care, where overbilling has proven very difficult to constrain (Hyman and Silver 2018).
State intervention does risk crowding-out some private enforcement efforts. But because governmental actors do not face the same profit motives that insurers do, public enforcement can complement private efforts. At the margin, insurers have a built-in incentive to deny claims of ambiguous validity. Shifting some of the job of apprehending third party loss-causers to the state therefore eases pressure on insurers to deny claims, and benefits legitimate claimants.

As an alternative to pure public or private enforcement, there are some hybrid models of public-private cooperation that are also worth considering. For example, some insurers have created and funded task forces to work with law enforcement at reducing fraud (National Insurance Crime Bureau, n.d.). These offer the advantages of economies of scale and insider knowledge, combined with the heightened powers of public enforcement.

3.3.1. Technology Standardization and Reporting

An obvious role for collective action in combating third party moral hazard is to regulate the collection and transmission of data that can be used to detect it. For example, US life insurers have established a system of record-keeping for “police departments to check if a homicide victim is named on a life insurance policy. Several perpetrators have been arrested and convicted thanks to this system” (Coalition Against Insurance Fraud 2017, 7). In 2018, English police and dash-camera manufacturers collaborated on a national dash cam safety portal, which “allows owners of any brand of dash cam to submit footage quickly and easily to the relevant authorities” (Muyanja 2018).

3.3.2. Cartelization or Regulation of Industry Structure

As Shortland (2019) persuasively demonstrated in the case of kidnap insurance, some types of third party moral hazard can be managed through the structural design (“governance”) of insurance market contracting. If insurance is provided by a small group of sellers, with close social ties and the ability and incentives to share information, it may be possible to overcome the dynamic coordination problems that would occur when individual victims make separate payments out of their own pockets.

As we explained earlier, one-time victims of kidnapping are more willing to overpay for ransom demands than are repeat players. Given an upward-sloping supply curve for kidnappings, overpaying today will generate additional entry (more kidnappings) in the future. But one-time victims do not care about the effects of their own overpayment on the volume of future

28 Feinman (2010) argues that insurers’ chief strategy for claims-handling boils down to “Delay. Deny. Defend.” To counteract these tendencies, there has emerged an elaborate set of rules, both statutory and judicially-constructed, accompanied by significant penalties, governing what constitutes “bad faith denial of an insurance claim.” For a recent empirical analysis, see Asmat & Tennyson (2014).
kidnappings. By contrast, repeat players (such as kidnap insurers) will take future supply responses into account in assessing how much to pay today. An effective insurance cartel can therefore limit the incentives to overpay that would otherwise accompany individualized ransom decisions. The countervailing effects of monopoly power vs combatting third party moral hazard awaits further analysis.

3.3.3. Qui Tam Suits

In addition to the possibilities just discussed, we propose the recognition of a new private cause of action, an insurance qui tam suit. The idea is to allow private individuals to bring suits on behalf of insurers against persons or entities who file false claims for insurance payments, collecting a fraction of the amount recovered as a bounty. Doing so amplifies disincentives for misconduct; importantly, it accomplishes this goal without altering the delicate balance of power that courts and regulators have sought to create between insurers and their policyholders (Baker, 1997 (liability insurance); Asmat & Tennyson, 2014 (bad faith)).

At its simplest, qui tam litigation expands ordinary notions of standing by offering a bounty to plaintiffs who bring suits to recover losses experienced by someone else, traditionally, the sovereign. For example, in return for surfacing the problem and pursuing litigation against their co-conspirators, a doctor involved in a phony clinic that was set up to treat accident victims in one of the no-fault auto insurance scams in New York (Bharara 2015; Rashbaum et al 2018; Insurance Information Institute, n.d.) would be entitled to a percentage of whatever recovery was obtained for the insurance company.

Our model is based on existing anti-fraud legislation at the federal level, the False Claims Act (31 U.S.C. §3729 (2009)), which was used to recover more than $2.5 billion in Medicare and Medicaid claims in 2016 alone (Nguyen & Perez 2020). While qui tam liability is not without controversy, the evidence in health care and elsewhere suggests that it has been effective. Our proposal would extend this mechanism to private insurers and to other areas of insurance.

3.3.3.1. Why Qui Tam Insurance Suits can be Useful

Litigation theory (Cooter & Rubinfeld 1989) suggests that profit-maximizing private actors should be willing to bring any case with a positive expected value. That would seem to imply that insurance companies already have a good reason to sue any third party who caused them losses (as long as they believe they can recover more than the costs of the litigation). Moreover, insurers can often make use of the law of subrogation, which allows them to pay the claims of their insured and then go after those who caused the loss to recover the amount paid. Given all that, there might seem to be little reason to supplement ordinary liability by bringing in an additional class of plaintiffs (relators).

Engstrom (2007, 1244) discusses nearly a dozen scholarly articles advocating qui tam-like mechanisms in other areas of law.
However, there are at least four reasons why insurers might not choose to pursue third party loss-causers optimally. First, insurers don’t necessarily want to prevent all losses, and may not zealously go after loss-causers. Insurers may prefer higher losses if they increase demand for insurance on the margin (Avraham & Porat, forth.)

Second, theory suggests that insurance fraud is an equilibrium phenomenon (Picard 1996; Picard 2013; Davies 2018, 16-17); under plausible conditions, there will always be some fraud that insurers will optimally choose not to uncover, leaving at least the possibility that others may be able and willing to do so. The logic of this insight is as follows. Suppose initially that 10 percent of all claims are fraudulent. Suppose further that at some cost, insurers could audit each claim and perfectly detect whether it is fraudulent. Even if such audits are profit-maximizing (that is, they cost less than the expected savings they generate) when 10 percent of claims are fraudulent, they will certainly not be profitable once all fraud has been successfully detected and deterred. At that point, of course, there will be no fraud left to detect, so there will be no justification for spending anything to detect it. But if insurers stop auditing altogether, fraudsters can go back to committing fraud, secure in the knowledge that they won’t be caught. Unless insurers can credibly commit to always auditing (even when it is unprofitable), a zero-fraud equilibrium is impossible.

Third, insurers may simply lack information about whom to sue because deliberate loss-causing is invariably covert. Even when the necessary information can be uncovered through audits or other techniques (and that is not always the case), these can be costly and difficult to undertake. But third party “insiders” will often have the information necessary to pursue such litigation—they know how such schemes are organized and implemented, including the techniques used to conceal them. So they will operate with an informational advantage over insurers, and may find it worthwhile to pursue litigation that insurers would not find it cost-effective to uncover.

Finally, the problem is not simply that insurers are uninformed or under-zealous. Indeed, insurers are known to be overzealous—at least at times—for example, in denying claims of loss by their own policyholders (Feinman 2010; Baker 1997; Asmat and Tennyson 2015, 413). Analogously, liability insurers will at times prefer to litigate too aggressively, against the interests of their policyholders (Keeton 1954). In response, the law has restricted insurer power vis-a-vis policyholders, but in ways that may actually advantage third party loss-causers. That is, in structuring relations between insurers and their customers to avoid disadvantaging the latter, the law may at times create opportunities for others to exploit. Our qui tam mechanism is designed to limit such opportunities, but importantly, to do so without strengthening insurers’ positions vis-a-vis their policyholders.

---

30 One area where this can occur is when the plaintiff makes a claim in excess of the limits of the insured’s policy. Under a rule where the insurer pays nothing over the policy limits, it is in a position to gamble (go to trial) with the policyholder’s money: if it wins, it pays only its trial costs; if it loses, the defendant pays the verdict in excess of the policy limits. See Keeton (1954) or Sykes (1994) for analytic details.
For example, most states (and the Restatement of the Law of Liability Insurance) now require that in making a decision to settle or litigate a claim against a policyholder that exceeds the coverage limits of the insured’s liability policy, an insurer must act as “would . . . a reasonable insurer that bears the sole financial responsibility for the full amount of the potential judgment.”  

31 As Allen Sykes (1994) pointed out, the rule can in some cases lead to settlements that are more generous to the plaintiff than would be the case if the insurer had unfettered discretion to litigate.  

32 In turn, that will provide an increased incentive for third party moral hazard. Rather than trying to curb such behavior by repealing the widely-adopted “disregard the limits” rule, qui tam litigation offers a way to reduce the profitability of strike suits without upsetting the carefully-crafted balance of power between insurers and insureds.

The existence of bad-faith liability—for example, for failure to settle, or for delay in processing a claim—serves similar aims as the disregard the limits rule, and has similar consequences. Policymakers have concluded that there is a need to control insurers’ discretion in paying claims in order to protect policyholders against insurers’ dilatory tactics. For example, New York had a requirement that “insurance companies . . . pay . . . claims submitted by no-fault providers within 30 days of their submission.” (Ferguson 2011, 66). But there is some evidence that sophisticated third parties were able to swamp insurers’ claims-handling and investigatory resources by submitting grossly exaggerated claims for reimbursement at a rate that was too high for insurers to handle within statutorily-required time limits for payment (Id.; Bharara 2015). If insurers are subject to damages for wrongly failing to pay a claim or paying too late, they will at the margin be induced to pay some questionable claims (Browne, Pryor & Puelz 2004; Tennyson & Warfel 2008; Asmat & Tennyson 2014). In turn, this makes loss-creation more profitable for third parties.

While some of these problems might be corrected by beefing-up insurers’ discretion over claims disposition, policymakers have concluded that doing so would unduly tilt the balance of power between insurers and policyholders in favor of the former. Our proposed qui tam system therefore incentivizes other actors to supplement the enforcement efforts of insurance companies through a bounty scheme. Such a system boosts deterrence without advantageing insurers in their relations with policyholders.

31 Restatement of the Law of Liability Insurance § 24(2) (2019). “Bears the sole financial responsibility” is another way of saying that the insurer’s settle/litigate decision should be made as if there were no policy limits and the insurer were thus responsible for the entire judgment. “The disregard-the-limits standard was first articulated by Professor Keeton in 1954 . . . and . . has since become the most common test for determining whether an insurer gave ‘equal consideration’ to its insured’s interests in duty-to-settle cases.” Id.

32 The analysis is complicated, and depends on many assumptions. As Alan Sykes points out, if the insured has limited assets, the disregard the limits rule “can increase plaintiff’s expected returns from litigation by reducing the extent to which the insured’s inability to pay large judgments lowers the plaintiff’s expected receipts” (Sykes 1994, 1367). If there are punitive damages for an insurer’s failure to settle, as in some states, that can also raise plaintiff’s settlement demands. “In sum, it is certainly plausible that the disregard-the-limits rule will weaken the bargaining position of the insurer and the insured in important classes of cases.” Id. at 1367.
3.3.3.2. How Insurance Qui Tam Actions Would Work

Qui tam relators can have an important cost or informational advantage over insurers, because they themselves are often insiders, already well-acquainted with the specifics of how a given loss-causing scheme operates. They may not need to investigate, audit, or detect anything, since their jobs are precisely to execute some aspect of the scheme at issue. In other cases, relators may possess detailed knowledge of how fraudulent activity is carried, knowledge that outsiders lack.

Qui tam liability offers potential whistleblowers an incentive to “defect” by turning against their fellow miscreants (Engstrom 2007, 1250). In turn, that should make it more difficult for conspiracies to form in the first place by weakening trust among conspirators, each of whom will have to worry that one of their partners might turn against them to collect a qui tam bounty.

The design of an insurance qui tam action should incorporate some of the same safeguards used in the False Claims Act (FCA) to prevent over-enforcement and agency problems (Engstrom 2007, 1271-72). Just as the Federal government does under the FCA, insurers should play a supplementary role in qui tam litigation. Otherwise, relators could just initiate a lawsuit on their own and then settle the matter at a steep discount, precluding the insurer from litigating similar claims against the defendant. Hence, we would grant insurers the power to oppose settlements that are detrimental to their interests. Only if the insurer declines to litigate the case would relators have the ability to proceed independently.

Finally, and in parallel with the FCA, we would also allow the insurer to intervene and “take control over the litigation, including limiting a relator’s procedural rights” (Engstrom 2007, 1272-73) in appropriate cases in which third party moral hazard is especially damaging. A relator’s award under our proposal would be comparable to those offered under the FCA: 25 to 30 percent if the relator litigates alone and 10 to 15 percent if the insurer takes over.

We believe that the use of qui tam litigation has the potential to deter third party moral hazard, and can do so without strengthening insurers’ hands vis-a-vis their own policyholders.

CONCLUSION

In the process of transferring risk, insurance does not merely reduce the incentives of policyholders to prevent losses. It may also increase the profitability of illicit attempts to transfer wealth, and may facilitate meritorious litigation, when insureds would otherwise lack the wealth to pay for harms they have caused. The key point is that its effects can spill over from the contract between the policyholder and the insurer to influence the behavior of those who interact with them in a variety of complex ways that scholars have only begun to appreciate. Moreover, these insurance externalities are difficult to manage with the contractual tools that insurers have long used to control standard moral hazard problems. We thus need to begin thinking about
insurance in a broader general equilibrium framework that encompasses actors beyond just the insurer and policyholder.

References


Avraham, Ronen & Ariel Porat. forth. Do Insurers Decrease or Increase Risks? Some Thoughts About the Conventional Wisdom. (unpublished ms. on file with the authors).


Clendenin, Meadow. 2006. “No Concessions” With no Teeth: How Kidnap and Ransom Insurers and Insureds are Undermining U.S. Counterterrorism Policy. 56 Emory L. J. 741-. 


Cooter, Robert D. & Daniel L. Rubinfeld. 1989. Economic Analysis of Legal Disputes and Their Resolution, 27 J. Econ. Lit. 1067-.

Core, John E. 1997. On the corporate demand for directors’ and officers’ insurance. 64 J. Risk & Ins. 63-.

Core, John E. 2000. The directors’ and officers’ insurance premium: An outside assessment of the quality of corporate governance. 16 J. L. Econ. & Org. 449-.

Dafny, Leemore, Christopher Ody, & Matt Schmitt. 2017. When Discounts Raise Costs: The Effect of Copay Coupons on Generic Utilization. 9 Amer. Econ. J. Econ. Pol’y 91-.


Delgado, M. Kit et al. 2014. Factors Associated With the Disposition of Severely Injured Patients Initially Seen at Non–Trauma Center Emergency Departments Disparities by Insurance Status. 149 JAMA Surgery 422-.


Detotto, Claudio, Bryan C. McCannon & Marco Vannini. 2015. Evidence Of Marginal Deterrence: Kidnapping and Murder in Italy. 41 Int’l. Rev. L. & Econ. 63-.

Derrig, Richard A. 2002. Insurance Fraud. 69 J. Risk & Ins 271-.


Dobbin, Frank & Erin Kelly. 2007. How to Stop Harassment: The Professional Construction of Legal Compliance in Organizations, 112 Am. J. Soc. 1203-.


Duffie, Darrell. 2018. Prone to Fail: The Pre-Crisis Financial System. 33 J. Econ. Persp. 81-. 

Dutton, Yvonne M. 2016. Funding Terrorism: The Problem of Ransom Payments. 53 San Diego L. Rev. 335-. 


Engstrom, David Freeman. 2007. Harnessing the Private Attorney General: Evidence from Qui Tam Litigation. 112 Colum. L. Rev. 1244-. 


Silver, Charles. 2015. The Basic Economics of the Duty to Defend. In Research Handbook in the Law & Economics of Insurance (Daniel Schwarcz & Peter Siegelman, eds.).


Timko, Nicholas I. 2011. Hearing on No Fault Auto Insurance Fraud re: To Examine Ways To Reduce The Incidents [sic] of No-fault Auto Insurance Fraud In New York Before New York State Senate Standing


