Economic Incentives and Workplace Safety

Terry Thomason
University of Rhode Island

Chapter 2 (pp. 9-36) in:
Workplace Injuries and Diseases: Prevention and Compensation: Essays in Honor of Terry Thomason
Karen Roberts, John f. Burton, Jr., and Matthew M. Bodah, eds.
Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2005
DOI: 10.17848/9781429454919.ch2

Copyright ©2005. W.E. Upjohn Institute for Employment Research. All rights reserved.
2 Economic Incentives and Workplace Safety

Terry Thomason

The problem of work injuries is a substantial one. Recent estimates put the cost of workers’ compensation benefits paid to injured Canadian workers and their families at more than $6 billion annually, or nearly 1 percent of gross domestic product. In the United States, workers’ compensation benefit payments amount to over $40 billion annually. And workers’ compensation benefit payments represent only a small portion of the economic costs of work injuries. Work injuries also entail losses due to lost production, damage to plant and equipment, and the uncompensated losses suffered by injured workers that are estimated to be as much as four times the cost of benefits (Heinrich et al. 1980).

The remainder of this chapter proceeds as follows. The next section discusses the economic theory of work injuries and illnesses. Specifically, this section examines employer and worker incentives for safety in the absence of government regulation. The chapter then discusses safety incentives created by different types of government regulation. Conclusions are drawn in the final section.

ECONOMIC THEORY OF WORK INJURIES AND ILLNESSES

Work injuries are an unwelcome by-product of economic activity. In part, they are random events, but they are also, to some extent, under the control of workers and employers. Employers can reduce the number of workplace injuries and illnesses by investing in safer technology, providing workers with personal protective equipment (such as hard hats and safety glasses), training workers and their supervisors, etc.;
workers can avoid accidents by following safe work practices and by taking greater care on the job.

Both parties incur costs when an accident occurs. Workers’ costs include potential loss of income and medical expenses associated with treatment and rehabilitation as well as intangibles, such as pain and suffering and disability that reduces the ability to enjoy leisure activities. Employers’ costs include interruptions in production and damage to capital equipment and physical plant.

Since accident prevention also entails costs to employers and employees, public policy should encourage employers and employees to minimize the combined costs of accidents and accident prevention that are incurred by both workers and employers. It is possible to spend both too much and too little on accident prevention. Investment in accident prevention is *socially efficient* when total costs are minimized, that is, when an additional dollar spent on prevention reduces accident costs by exactly one dollar.

As indicated, both employers and workers affect workplace health and safety. We can expect that—if they are rational—both actors will make accident prevention decisions that are *privately efficient*. That is, we may expect that each will make decisions that minimize their own accident and accident costs individually; however, their decision making process may not consider costs that are incurred by the other party.

However, under some conditions, it is at least arguable that employers do consider the workers’ accident costs when making investments in workplace health and safety and thus make socially efficient decisions as well. To understand this argument, let us consider a world where there are two types of employers, those with safe workplaces and those with hazardous ones. Assume that workers employed by safe firms do not risk having an accident or illness while at work—i.e., the probability of injury or illness is zero—while one of every ten workers employed by hazardous firms will have an occupational accident each year. Let us further assume that workers are aware of the probability of accidents at both types of firms and that they are free to choose the type of firm for whom they will work.

Under these assumptions, we may expect that if everything else were equal—i.e., the compensation package and other terms and conditions of employment—all workers would prefer employment at the safe firms. In order to attract workers, hazardous firms will be forced
to increase wages above the level paid by safe firms. In other words, we would expect to find that hazardous firms pay a compensating differential and that the magnitude of this differential will be related to the workers’ expected accident costs, including the cost of lost income, medical expenses, pain and suffering, etc.

So, for example, let us assume that the average cost of accidents for workers is $10,000 and that the average annual salary of workers in safe firms is $40,000. Since the probability of an accident at a hazardous workplace is 0.1, then expected accident costs at that workplace are $1,000 (= 0.1 × $10,000). This means that hazardous employers must pay their employees an annual salary of $41,000 for employment at a hazardous firm to be equally attractive as employment at a safe firm. Thus, the employer’s accident costs include the expected accident costs borne by workers. Importantly, employers will be able to reduce the compensating differential and, consequently, their accident costs, by reducing the incidence of workplace accidents and illnesses.

The economic model presented in the preceding paragraphs rests on a number of key assumptions, which many have questioned. In particular, the model requires that workers have complete and accurate information with respect to the risk of injury or death and an absence of barriers to worker mobility, i.e., that workers are free to move in and out of the labor market or between employers at relatively low cost. However, critics point out that it is likely that either workers do not have access to good information about injury risks or barriers to mobility prevent workers from moving to safer jobs. As a result, wage differentials due to the risk of injury either do not arise or they are inadequate, i.e., they do not fully compensate workers for the risk of injury.

Do employers, in fact, pay a compensating differential to workers exposed to greater risks of injury or illness? To answer this question the researcher must address a number of methodological issues that are not easy or simple to resolve, and existing statistical evidence is decidedly mixed. By and large, research investigating the relationship between the risk of fatal injury and wages has found a risk premium, while studies examining the relationship between wages and non-fatal risks have not (see Viscusi 1993, for a recent review of this literature). However, Dorman and Hagstrom (1998) demonstrate that even fatal-risk differentials are extremely sensitive to the regression specification.
Importantly, there is no evidence to demonstrate that the risk differential is fully compensating, even for fatal injuries. In addition, research suggests that, after controlling for the risk of injury and a variety of other factors affecting wages, the wage differential is substantially larger for unionized workers than for nonunion workers (Olson 1981; Dickens 1984; Fairris 1992; Siebert and Wei 1994; and Sandy and Elliott 1996). This result, which indicates that union workers get a greater premium for the same level of risk, is difficult to reconcile with the hypothesis that wage differentials compensate workers for the expected cost of accidents. Finally, psychological research suggests that people overestimate the likelihood of a low probability event and underestimate the likelihood of a high probability event (Viscusi 1993). This systematic bias implies that workers will generally demand a risk premium that is less than fully compensating.

**Workers’ Compensation**

Workers’ compensation provides cash benefits to workers who are unable to work as the result of an occupational injury or illness as well as medical benefits and rehabilitative services to all who are injured as the result of a workplace accident.

These benefits have the effect of reducing accident costs for workers and, consequently, the risk premium paid by hazardous employers. As a result, we may expect that the worker’s incentive for avoiding workplace injuries will have been reduced because their accident costs have been reduced by the medical and cash benefits provided by the workers’ compensation program, a problem known as risk-bearing moral hazard in the insurance literature. We might also expect that workers’ compensation benefits would increase the workers’ willingness to expose themselves to greater risks on the job, but that these benefits would also increase the likelihood that workers would report an injury that would have otherwise gone unreported or even falsely report a nonwork-related injury as occupational. This latter problem is known as reporting moral hazard. In either event, because workers’ compensation reduces the cost of workplace accidents for workers, we would expect it would also reduce the compensating wage differential. In fact, there is some statistical evidence indicating that as compensation becomes more generous, the risk premium for hazardous work is reduced.
Do workers’ compensation benefits affect employers’ incentives to prevent workplace accidents? The answer depends on the way in which compensation benefits are funded. If the employer is liable for workers’ compensation benefits paid to his or her firm’s injured workers, then the employers’ incentive structure will be unchanged by the introduction of workers’ compensation. However, if there is no relationship between employer costs and worker benefits, then the employer’s incentive to prevent accidents is reduced by workers’ compensation benefits.

In Canada and the United States, workers’ compensation benefits are funded through a payroll tax paid by employers. A two-step process determines tax (or assessment) rates in most provinces. In the first step, industrial classifications are used to group firms who share similar risks of workplace injury or illness, so that banks are grouped with other financial institutions, for example, food stores are grouped with similar retail establishments, etc. The recent historical accident record of each of these classifications, known as rate groups, is used to determine the base assessment rate for each group. The assessment rate is set so as to provide sufficient income to fund all workers’ compensation benefits paid to workers and any expenses associated with workers’ compensation program administration.

In the second step of the rate-making process, known as experience rating, the base assessment rate for some firms is adjusted to account for the firm’s individual safety record. In other words, the assessment rate for firms with better than average safety records (lower injury rates) is reduced, and the rates of firms with worse than average safety records (higher injury rates) are increased.

Both steps of the rate-making process should reduce the injury rate relative to a regime where all employers are charged an identical assessment rate unrelated to the risk of injury. Variation in the base assessment rate means firms in hazardous industries pay a higher base assessment rate than firms in relatively safe ones, so that the cost of goods and services produced by firms in hazardous industries increases relative to a regime in which a flat assessment rate is charged to all employers. In turn, this reduces consumption of goods and services in hazardous industries relative to safe ones and subsequently employment; as the proportion of employment in safe industries rises, the overall accident rate will drop.
However, the base assessment rate is only marginally related to the firm’s accident experience. If the firm is not experience-rated, the employer does not consider workers’ compensation assessments to be part of the cost of accidents, since it cannot affect costs by preventing accidents. However, if the firm is experience-rated, then a reduction in the accident rate directly reduces its subsequent accident costs. Thus, if the firm is experience-rated, the employers’ investment in workplace safety will remain unchanged following the introduction of workers’ compensation insurance; however, if the firm is not experience-rated, the employer’s safety investment will decline after workers’ compensation is introduced.

Thus, workers’ compensation unambiguously reduces workers’ safety incentives and increases workers’ incentives to report compensable claims. Furthermore, since not all employers are experience-rated, the overall impact of workers’ compensation is to also reduce, on average, health and safety investments by employers.

Workers’ Compensation and Occupational Injuries: The Evidence

Since the introduction of workers’ compensation pre-dates the collection of injury rate data, there are only a handful of studies that have attempted to directly examine this issue and those that do have produced contradictory results. Chelius (1976) found that the introduction of workers’ compensation programs led to a reduction in fatal accident rates relative to the tort regime that preceded them. However, Fishback (1987) reached the opposite conclusion, using a different (and arguably better) data set.

On the other hand, several studies have attempted to determine whether there is a relationship between the generosity of workers’ compensation benefits and the work injury rate. As indicated, economic theory suggests that, where workers’ compensation insurance is less than perfectly experience-rated, the accident rate should be positively related to workers’ compensation benefit generosity.

A large number of studies using different methodologies and data sources have found the expected positive relationship between benefit levels and injury (or workers’ compensation claim) rates. Studies of the U.S. workers’ compensation include Butler and Worrall (1983) and Chelius (1982) who examined state-level claims and injury data and
Hirsch et al. (1997) who used longitudinal survey data to estimate the impact of benefit generosity on the probability that a worker would file a workers’ compensation claim. Canadian studies include Thomason and Hyatt (1997), who examined provincial injury rates and Thomason and Pozzebon (1995), who used data on individual workers to estimate claim probability. Uniformly, these studies have found that higher levels of workers’ compensation benefits are associated with higher injury or claims rates or a higher probability that a worker would initiate a compensation claim.

GOVERNMENT REGULATION AND WORKPLACE SAFETY

There are at least three approaches to the regulation of occupational health and safety, all of which have been adopted by policy makers in one form or another at one time or another. The first—and the one most commonly identified as occupational health and safety regulation—involves the promulgation of rules prescribing or proscribing specific policies and practices by employers, which are enforced through on-site inspections and monetary penalties for infractions. The second approach comprehends systems of general safety incentives that reward or punish employers on the basis of safety and health outcomes rather than behaviors that are thought to affect those outcomes. This second approach is embodied in the experience rating of workers’ compensation assessments, whereby employers’ compensation costs are tied to their accident experience. The third approach, termed internal responsibility, pervasive in Canada, is designed to improve safety and health conditions through workers’ empowerment and involves three principal elements: 1) the worker’s right to refuse to perform unsafe work; 2) the worker’s right to information on the nature of workplace hazards; and 3) joint labor-management safety and health committees, which are given a mandate to oversee safety and health conditions in the workplace.

The Economics of Regulation

Occupational safety and health regulation seeks to change behavior of the employer by changing the cost-benefit calculus described in the previous section, through imposition of monetary penalties or other
sanctions. Specifically, regulatory sanctions lower accident prevention costs by the expected value of the sanction. In other words, firms considering adoption of a particular safety practice must now weigh expected costs of the sanction that will be imposed if they fail to do so. Sanction costs are characterized as “expected” because, under some—if not all—regulatory regimes, penalties are not imposed unless a violation is detected.

There are two costs that must be considered by efficient regulators: the administrative costs of regulation (the cost of staff involved in enforcement and adjudication), and the cost of regulatory effort (the imposition of sanctions whose expected costs are either too great or too small). Sanctions are too small (large) if the costs of accident prevention, including the expected savings from the avoidance of sanctions, are less (greater) than associated accident costs. The cost of error is equal to the difference between accident costs and the cost of accident prevention if the regulation in question is adopted and enforced. The goal of efficient regulation is to minimize the sum of these costs.

**Direct Regulation of Workplace Hazards**

As indicated, direct regulation attempts to change employer behavior by promulgating regulations that prescribe or prohibit specific employer or worker practices. Regulations are enforced through workplace inspections and penalties for noncompliance. Critics argue that direct regulation fails to recognize important variation across firms with respect to technology and other characteristics. In other words, a safety practice that is efficient for one employer may not be efficient for another, so that there are potentially substantial error costs. In addition, a system of direct regulation in which the regulator agency responsible for promulgating rules is one-step removed from the workplace and is, therefore, slow to respond to technological change. Once again, this could result in substantial error costs.

Furthermore, as Dorman (1996, p. 197) notes, “Most occupational risks are transitory . . . Safety features mandated by law may be unavailable or malfunctioning from time to time, but inspectors are not likely to know this.” In other words, the probability of detecting noncompliance is low so that the regulators must substantially increase the magnitude of the sanction imposed. Finally, direct regulation is costly to
administer. It requires an extensive bureaucracy to develop standards, inspect workplaces, and to resolve disputes with employers concerning the appropriateness of penalties.

Most research examining direct regulation is confined to an examination of the effects of the U.S. Occupational Safety and Health Act (OSHA) of 1970. At best, this research has produced mixed results with respect to OSHA’s effectiveness, although more recent research—and, in some ways, methodologically superior—tends to find results that support the hypothesis that direct regulation reduces injury rates. (This question is also taken up by Mendeloff [1979, Chapter 11], with some comparative discussion on the United States and Canada.)

Much of this early research involved a time-series analysis or cross-sectional pre- and post-OSHA comparisons of aggregate injury rate data. By and large, these studies were unable to find the expected reduction in the incidence of workplace injuries (Smith 1973; Mendeloff 1979; Currington 1986). However, Smith (1992, p. 566) notes data problems render such comparisons problematic: “Because the Occupational Safety and Health Act fully covers the private sector, and because before-and-after comparisons are generally infeasible, a convincing study of the overall effects of the Act has not been—and may never be—done.”

Another group of studies has evaluated the impact of OSHA enforcement activity—that is, the effect of inspections and fines—on the incidence and severity of workplace injuries. Following Smith (1992), these studies may be classified into two categories: those using aggregate industry injury rate data and those using plant level data.

Enforcement variables used in research examining industry aggregate accident rates include lagged measures of the probability of inspection and the expected penalty for an OSHA violation. In general, these studies found little or no effect for OSHA enforcement activity. For example, Viscusi (1979) was unable to detect a statistically significant relationship between injury rates and either inspection probability or the expected penalty. In a later study, Viscusi (1986) found that OSHA enforcement reduced the lost workday incidence rate by a modest 1.5 to 3.6 percent, although Smith (1992) argues that this result may have been a statistical artifact—the product of changes in employer reporting behavior resulting from a change in OSHA inspection strategies.

Arguing that these lagged penalty data were as much a measure of employer noncompliance as a proxy for a deterrent effect, Bartel and
Thomas (1985) estimated a system of structural equations in which the probability of inspection and penalties per inspection were treated as endogenous. They found that while OSHA significantly reduced employer noncompliance, there was little relationship between noncompliance and the lost-time injury rate. However, these authors conclude that OSHA indirectly reduced accident rates by placing a greater regulatory burden—in the form of increased inspection probability—on firms with higher injury rates.

A study of industry aggregate injury rates in Quebec by Lanoie (1992) found a statistically significant negative relationship between inspection probability and the lost-time injury rate. However, the likelihood of a workplace health and safety inspection by an officer of the Quebec government was positively associated with injury severity, measured as average number of workdays lost per injury. In addition, Lanoie failed to detect a statistically significant relationship between probability of penalty and either frequency or severity of work injuries.

Research using plant level data have generally reached more optimistic conclusions about OSHA’s effectiveness, although these studies have also produced mixed results. Two types of studies have been conducted. Earlier research compared firms that had been inspected early in the year with firms that had been inspected late in the year, hypothesizing that inspection effects should be more evident for the former group of firms than for the latter group (Smith 1979; McCaffrey 1983). Using data from 1973 and 1974, Smith found that 1973 inspections reduced injury rates by about 16 percent while 1974 inspections induced a 5 percent reduction, although the latter relationship was not statistically different from zero at conventional levels. McCaffrey failed to find a statistically significant effect using data from 1976–77. As Scholz and Gray (1990, p. 299) note, taken together, these results suggest that “the easily accomplished reductions in risk that OSHA inspections could impose may have already been implemented in 1976, leaving more complex issues of risk reduction, less amenable to quick fixes.”

As Smith (1992, p. 569) points out, because these early studies lacked data on citations and fines resulting from inspections, they were only able to measure the abatement of injuries following an inspection; as a result, these studies were unable to measure OSHA’s “deterrent” effect. Replicating this research, Ruser and Smith (1991) used a measure
of inspection probability based on the average inspection frequency for similar firms to estimate the deterrent effect. On the basis of this analysis, they concluded that: “there is virtually no evidence of a deterrence effect” (p. 231).

Interestingly, recent plant-level studies, which use explicit before-and-after comparisons of the same firms, provide evidence for a more sanguine assessment of direct regulation. Using a unique data set that allowed tracking of inspections and penalties for a large sample of individual firms over a seven year period, Scholz and Gray (1990) estimated both the deterrence and abatement effects of OSHA enforcement activity. They found that a 10 percent increase in enforcement resulted in a 1 percent reduction in the accident rate, a much larger effect than detected in prior research, although one that the authors describe as “modest” (p. 302). This reduction was primarily due to a “deterrence” effect and, specifically, an increase in the probability of inspection, as opposed to an increase in the average penalty.9

Importantly, Scholz and Gray argue that their results indicate that economic models of occupational safety and health regulation, which assume that firms optimize when making safety and health choices, fail to account for the limited information processing capacity of managers.10 Due to their limited capacity, managers do not optimize, but often engage in “fire-fighting,” responding to problems as they become more significant relative to other issues. As evidence, Scholz and Gray find that an unexpected increase in the accident rate in one year will lead to a reduction in injuries in the next, and vice versa. In addition, they find a lag between OSHA enforcement activity and a change in firm health and safety—a result that they claim is evidence of an organizational learning curve. Ruser (1985) obtained similar results.

Nonetheless, overall the extant evidence suggests that OSHA has, at best, resulted in a modest improvement in workplace health and safety in the U.S. However, advocates of direct regulation argue that these disappointing results are primarily due to dillident administration and a lack of funding than to a fundamental flaw in this type of regulatory regime. In particular, they point to two problems. First, the process of adopting permanent health and safety standards under OSHA is slow and cumbersome. Governed by the Federal Administrative Procedures Act, the law requires a Notice of Intended Rulemaking and a subsequent proposal, both of which must be published in the Federal Register. This
is followed by a Public Hearing and comment period where all interested parties are invited to submit comments, which the agency must consider before promulgating a standard. After they are issued, standards are subject to judicial review, and the Supreme Court has ruled that the agency must provide substantial evidence that the standard is based on a “significant” risk. Second, agencies responsible for administering the Act are substantially underfunded, a problem exacerbated during the Reagan-Bush administrations. Dorman (1996, p. 193) notes that there are more fish and game wardens in the U.S. than occupational safety and health inspectors.

Critics have expressed greater apprehension over the impact of direct regulation on economic productivity (Burton and Chelius 1997). There is a public perception, shared by some economists, that the proliferation of industrial regulation in the 1960s, particularly with respect to occupational health and safety and environmental protection was responsible for anemic productivity growth since that time. Research is sparse, however. One study, estimating annual total factor productivity for 450 U.S. industries between 1958 and 1978, found that OSHA accounted for around 19 percent of the productivity slowdown of the 1970s (Gray 1987). Viscusi (1996) has estimated the cost and benefits of five OSHA regulations and found that for four of these, the costs of the regulation exceed the benefits in terms of lives saved. However, Stone (1997) challenged Viscusi’s estimates, claiming he ignored other benefits, such as the reduction in injuries and illnesses. His reanalysis of one of these regulations showed it was in fact efficient, when these other benefits were considered.

Nonetheless, if one assumes that the direct regulation of workplace safety is inefficient public policy, then it is possible that direct regulation could actually result in the deterioration of worker health. Keeney (1994) has argued that a reduction in disposable income due to these regulatory costs can lead to changes in spending on safety and healthcare more generally, greater stress due to job loss, and risky behavior such as increased alcohol and tobacco consumption.

**Internal Responsibility System**

A principal criticism of direct regulation is that it fails to recognize firm heterogeneity, so that standards appropriate for one firm are likely
Economic Incentives and Workplace Safety  21

to be inappropriate for another. In contrast, the internal responsibility system is highly adaptable to the particular circumstances of the firm and is flexible so that it can respond relatively quickly to technological change. The design of safety “standards” is in the hands of the parties themselves—labor and management—who are intimately familiar with plant operation and who are therefore well placed to implement regulatory standards that are effective and efficient. In addition, administrative costs, which are principally borne by employers, at least initially, are relatively low. Enforcement is in the hands of the firm’s workforce so that the probability of detecting a violation will be high.

On the other hand, the success of the internal responsibility system is critically dependent on employee bargaining power. It is likely that internal responsibility is less effective in nonunion workplaces than in union ones. In addition, unions are political organizations that necessarily respond to the preferences of their memberships. And safety and health are often given a relatively low priority by a rank and file that sometimes appears to be more interested in wages and job security. Furthermore, the internal responsibility system can be used by employees to shirk legitimate work assignments or by labor unions as leverage in collective negotiations with employers. Finally, there are concerns that labor members may lack the expertise, particularly in the realm of occupational health, to either design effective standards or monitor firm compliance.

Unlike either direct regulation or general financial incentives, there is little direct evidence on the efficacy of the internal responsibility system. Most of this research has examined joint health and safety committees (JHSCs) and much of it uses data on Canadian workplaces. By and large, however, the Canadian studies either examine process issues or factors determining the relative effectiveness of JHSCs rather than the question of whether or not they reduce injury rates or otherwise improve worker health compared to workplaces without such committees. In addition, these studies often rely on subjective reports by the participants rather than objective evidence. Nevertheless, some useful information relevant to the question of the effectiveness of the internal responsibility system may be gleaned from this research.

For example, Shannon et al. (1992) find that lower accident rates are found in firms where the JHSC includes a senior manager; where labor members had access to professional expertise; and where the JHSC
had a broad mandate rather than a narrow one. Furthermore, Tuohy and Simard (1993) find that JHSCs were more effective in reducing accident rates when the committee had an equal number of labor and management members and where there are well-established operating procedures. In other words, both studies indicate that JHSCs are more effective when employers give them greater resources and support.

Three studies directly examine the issue of whether JHSCs ameliorate workplace safety. Cooke and Gautschi (1981) combined OSHA administrative data with the results of a survey of 113 manufacturing firms in Maine to investigate, among other things, whether joint labor-management safety programs affected firm injury rates. They obtained mixed results, which depended on firm size. Large firms with joint safety programs had lower injury rates than large firms that did not have a joint program. However, this result was statistically significant only for firms with more than 300 employees and only at the 0.10 confidence level. For small firms, the opposite result was found; firms with joint programs had higher injury rates. Boden et al. (1984) surveyed 290 large (more than 500 employees) Massachusetts firms but failed to find a relationship between the presence of a joint safety committee and workplace injury or illness rates. Importantly, both of these studies—Cooke and Gautschi (1981) and Boden et al. (1984)—use cross-sectional, rather than longitudinal data, and are, therefore, limited in their ability to address the question of whether there is a causal relationship between JHSCs and workplace safety.

On the other hand, the most careful examination of internal responsibility found that the internal responsibility system generally and JHSCs in particular were associated with lower levels of workplace injuries and illness (Lewchuk et al. 1996). This study used administrative data from the Ontario Workers’ Compensation Board supplemented with data from two surveys. The authors find that both enactment of internal responsibility legislation and the introduction of JHSCs were negatively and significantly related to the workplace injury rate. Specifically, they find that JHSCs may reduce lost-time claims by as much as 18 percent relative to similarly situated firms without JHSCs. Importantly, they also find that joint committees were more effective at reducing injury rates in unionized firms than in nonunion firms.

The latter results suggest that unions play an important role determining the effectiveness of JHSCs. Similarly, Weil (1991, 1992) has
argued that unions improve the effectiveness of direct regulation. Using 1985 OSHA data from the United States, he has shown that unions increase inspection probability; inspection intensity, as measured by the duration of inspections per employee; and the scope of the inspection, i.e., whether or not the inspection resulted in a physical examination of the workplace. Weil also found that unions increased the number of citations as well as the severity of the penalties. These results suggest that there may be a synergy between direct regulation and internal responsibility, at least for unionized workplaces.

On the other hand, critics cite anecdotal evidence that shows that unions use regulatory agencies and, in particular, occupational safety and health agencies to enhance their power in organizing campaigns and in collective bargaining (Northrup 1997). JHSCs would seem to offer similar opportunities for unions to enhance their organizing and collective bargaining outcomes. However, Schurman et al. (1998) note that complaint-based inspections in unionized firms result in a higher percentage of violations than similar inspections in nonunion firms and argue that this contradicts an interpretation that unions use safety regulation to gain organizing and bargaining advantage.

Hebdon and Hyatt (1998) present conflicting evidence with respect to this issue. They use Ontario data to examine factors influencing the probability of a refusal to do unsafe work or the probability of a health and safety complaint. In general, they found that while the probability of both events was higher where there is a contentious industrial relations environment, they found no evidence of concerted harassment of employers during collective negotiations.

More generally, we might expect that unionization could lead to more optimal health and safety conditions. Workplace health and safety has characteristics of a public good in that consumption is neither rival nor excludable. In addition, free rider problems may prevent unorganized workers from negotiating the optimal provision of safety conditions by the employer. That is, workers will be individually reluctant to reveal preferences because they fear that they will pay the full cost of safety. Employers must therefore rely on information gleaned from the labor market. However, such information necessarily reflects the preferences of workers who are very different than the average worker; these marginal workers are younger and are less likely to have family responsibilities. Among other things, marginal workers are likely to be
less concerned about workplace hazards and should be less willing to trade off wages for increased safety.

On the other hand, unions, which are democratic political organizations, are more likely to reflect the preferences of the average workers. In fact there is some evidence that unions in fact respond to the safety objectives of more senior workers while management is more likely to be influenced by the preferences of marginal workers (Kahn 1987; 1990).

**General Financial Incentives**

Both direct regulation and internal responsibility attempt to regulate the safety process, imposing sanctions on employer behaviors thought to affect the accident rate. In contrast, a regime using general financial incentives regulates safety outcomes, imposing sanctions based on employer performance with respect to results-based workplace safety measures. One proposal for general financial incentives is the injury tax, whereby the government imposes a monetary penalty for each work-related injury or illness (Smith 1974). A more prosaic form of general financial incentives is experience-rated workers’ compensation insurance, as discussed in the previous section, whereby the firm’s compensation assessment is based, wholly or partially on its accident experience.

Like the internal responsibility system, a system of general financial incentives imposes no specific requirement vis-à-vis firm health and safety practices, allowing firms to select the most appropriate means for attaining its safety goals. Furthermore, under a system of general financial incentives, administrative costs will be lower than those incurred under either direct regulation or the internal responsibility system. However, because experience-rating adjustments to workers’ compensation assessments are based on the firm’s claim experience rather than its accident experience, experience rating provides employers with incentives to engage in claims management as well as accident prevention. Claims management includes a number of less than desirable practices, including retaliation against workers who initiate compensation claims and legal challenges to legitimate claims by injured workers. In addition, for actuarial reasons, true experience rating is not feasible for small firms.
There is substantial research investigating the impact of experience rating on the frequency and severity of work accidents. In general, these studies have found considerable evidence that experience rating is associated with lower injury rates, although there are a few exceptions (Hyatt and Thomason 1998). However, research that fails to find the expected effect is, in general, methodologically weaker than studies that do (Hyatt and Thomason 1998). Studies investigating injury severity have generally produced mixed results. There are two possible explanations for the disappointing results with respect to injury severity: either employers have less ability to affect severity than the incidence of injuries or the effects of experience rating on incidence overwhelm the severity effect. That is, on the margin, experience rating induces employers to reduce the frequency of less severe injuries. In either event, severity studies are generally less informative and will not be reviewed here. However, a brief review of injury rate research follows.

Research examining the impact of experience rating on workplace safety, most of which uses U.S. data, falls into one of three categories. The earliest studies exploited the fact that U.S. experience-rating formulae are different for large and small firms, so that large firms are more likely to be experience rated and are more extensively experience rated than small firms. Since a difference in injury rates between large and small firms could be ascribed to firm size effects unrelated to experience rating—such as, scale economies in accident prevention efforts—these studies examined the relationship between benefit generosity and accident rate. As indicated previously, empirical research conclusively demonstrated work injuries are positively related to benefit levels. However, if experience-rating induces firms to improve workplace safety, then this relationship should be attenuated in large firms relative to small ones. That is, as benefit levels become more generous, experience rated firms will increase their safety investment, partially offsetting the increased level of injuries resulting from worker moral hazard. Several studies found this hypothesized relationship (Ruser 1985; Butler and Worrall 1988; Ruser 1991); only one failed to do so (Chelius and Smith 1983).

As indicated, the positive relationship between benefit levels and the work injury rate is primarily attributable to a reporting effect; workers are more likely to report an injury when benefit levels are high than when they are low. It is unlikely that fatal claims are subject to this
reporting phenomenon, so that the relationship between benefit levels and fatal injury probability should more accurately reflect the impact of benefits on employer behavior. Four studies have examined this relationship, and three found that the incidence of fatal injuries was negatively associated with higher benefit levels, as expected if experience rating has safety-enhancing effects (Moore and Viscusi 1989; Ruser 1991; and Durbin and Butler 1998). Only Butler (1983) failed to find the hypothesized negative relationship.

Several studies have taken advantages of “natural experiments” to compare injury rates before and after the implementation of an experience rating program. Chelius and Kavanaugh (1988) examined injury rates of two New Jersey colleges before and after they elected to self-insure and ceased to be covered by private compensation insurance. Chelius and Smith (1993) compared occupational injury rates for small firms in Washington, which gives experience-rated discounts to these firms, with injury rates for small firms in states that do not offer these workers’ compensation claim rates in Ontario and British Columbia, respectively, before and after the introduction of experience rating in those provinces. Shields et al. (1997) explored the effect of the implementation of “large-deductible” compensation insurance policies—where insured firms are responsible for the first several thousand dollars of compensation costs—in Texas. Finally, Durbin and Butler (1998) used state-level U.S. data to investigate the effects of both large and small deductible policies as well as a rule change that lowered eligibility criteria for experience rating. With the sole exception of Chelius and Smith (1993), these experiments found that experience rating was associated with lower injury rates.

Out of 14 studies reviewed here, 11 found evidence that experience rating results in an amelioration of workplace health and safety. This evidence was produced by research that is remarkably mixed with respect to both data sources and methodology. And, as indicated, a careful examination reveals that studies failing to detect this relationship were methodologically weaker than those that did. Taken as a whole the evidence is quite compelling: experience rating works.

However, as Hyatt and Thomason (1998) point out, the leap from the observation that experience rating is associated with lower injury or claims rates to the conclusion that experience rate enhances firm safety is short, but perilous. Experience rating may lead to increased claims.
management by employers, who file claims, as well as pro-active staffing practices designed to screen job applicants likely to file a workers’ compensation claim. This has the effect of reducing injury reporting, while leaving workplace hazards undisturbed. Two studies show experience rating increases employers claims management activity.

Examining a large set of administrative records from Ontario, Hyatt and Kralj (1995) found that experience-rated employers were significantly more likely to appeal claims than non-experience rated employers, and that the likelihood of an appeal for experience-rated employers increased as a function of the financial incentives that they faced. Kralj (1994) analyzed a small survey of Ontario employers in which managers were asked to report their impressions of the effects of experience rating on their behavior, i.e., changes in accident prevention and claims management practices resulting from experience rating. He found that while both prevention and claims management behaviors increased, experience rating had a greater impact on accident prevention efforts. Thus, while it is clear that experience rating leads to more intensive claims management efforts, this is not the only effect. Furthermore, claims management is not an unalloyed evil. The denial of fraudulent claims is both equitable and efficient, and there is evidence indicating that a prompt return to work leads to more successful rehabilitation.

Using a survey data set consisting of over 450 Quebec manufacturers, Thomason and Pozzebon (2002) examined the estimated relationship between experience rating and a wide range of firm health and safety and claims management practices. These practices included, for example, the amount of health and safety training provided to workers, the extent to which the firm disputed workers’ compensation claims, the number of in-house personnel devoted to claims management or accident prevention activities, and firm expenditures on personal protective equipment. They found that experience-rated firms were both more likely to engage in more aggressive claims management and to make greater effort to increase workplace health and safety. Interestingly, however, the evidence also suggested that high wage firms are more likely to reduce workers’ compensation claim costs by increasing their accident prevention efforts (relative to their claims management efforts) than low wage firms. This result implies that there may be a “high road” and a “low road” response to experience rating.
CONCLUSIONS

The past 20 years have seen a substantial accumulation of knowledge concerning the effects of various policy options, although much is left to be learned. It is by no means certain that policy makers have fully taken advantage of this knowledge or that they have developed a coherent policy with respect to occupational health and safety problems. Rather, policy has developed in a piecemeal fashion as jurisdictions have experimented with various approaches to these problems.

Until recently, these approaches tended to emphasize direct regulation and, more recently, internal responsibility; general financial incentives are little used. Workers’ compensation programs have only recently introduced experience rating to the assessment process, and in most provinces in Canada, there are restrictions on its application, which substantially limit its effectiveness. For example, in British Columbia the experience rating adjustment is limited to 30 percent of the base assessment rate.

However, considerable evidence indicates that general financial incentives are effective in reducing accident rates. Moreover, experience-rating does not share many of the problems associated with the other two approaches. In addition, both the costs of direct regulation and its apparent limited effectiveness call into question whether a broad application of direct regulation is appropriate.

Nonetheless, general financial incentives, particularly in the form of experience-rated compensation assessments, are not a panacea. Two problems may be identified. First, because accidents are, by definition, random events, general financial incentives are not easily applied to small firms—the small firm’s experience is not necessarily indicative of its underlying safety. Second, due to the long latency of many occupational diseases, it is difficult to assign responsibility to a particular employer. Finally, direct financial incentives assume that firms engage in an optimizing cost-benefit calculus, but the evidence suggests that limited information processing capacity may lead managers to satisfice. Under these circumstances, direct regulation could provide a needed shock to focus managerial attention on safety and health problems.

While this implies a continued role for direct regulation, it also suggests a more limited and targeted approach. More specifically, due to the high costs of direct regulation, the resources required by this
option should be directed at high-risk industries. This would include, in particular those in which there are numerous small firms, such as construction and those in which there is a relatively high probability of catastrophe—that is, an accident in which there is significant loss of life—such as underground mining. In addition, these resources should also be directed to the problem of long latency occupational disease, where it is unlikely that general financial incentives will be effective. This includes funding research that would investigate the relationship between occupational exposures and subsequent disease development as well as funding for monitoring workplace exposure.

**Key Messages**

- Much has been learned in the last two decades regarding effective policies to reduce disabling injury at work.
- Both direct regulation and internal responsibility have been widely used in Canada, whereas general financial incentives recently have become more pervasive as they are in the United States.
- Financial incentives do appear effective in reaching injury rates whereas the limited effectiveness of direct regulation raises questions about its value except where it may be targeted at high-risk individuals and longer-latency occupational disease exposures.
- General financial incentives are limited in their value for small firms.
- A stronger role for workplace exposure surveillance is necessary.

**Notes**

This chapter reprinted by permission. See Thomason (2003).

1. Accident prevention costs are manifested in higher production costs and lost productivity, which means that there are fewer goods and services including, for example, medical and rehabilitation services for those claimants who are injured or become ill due to a workplace accident or exposure.
2. The $10,000 figure for accident costs subsumes an evaluation of the monetary value for intangibles such as pain and suffering.
3. This example assumes that workers are risk neutral, i.e., they are indifferent be-
between income that will be paid with certainty (such as the wage income available from safe employers, where there is no risk of injury) and income that is uncertain (such as the wage income paid by hazardous employers, where there is a 10 percent chance that the worker will lose wage income due to a work accident). If the worker is risk averse, as is generally thought to be the case, then the worker would demand a salary higher than $41,000 to compensate him or her for the additional risk incurred by working for a hazardous employer.

4. It is important to note that workers will continue to have an incentive to avoid workplace accidents and illnesses even if the wage is fully compensating, i.e., if it compensates workers for all of the expected cost of injury. This is because the worker will continue to incur costs if an accident occurs, unless he or she is able to purchase insurance that covers those costs.

5. Interestingly, one of the few studies that failed to find this relationship used a Canadian data set (Martinello and Meng 1992).

6. It is not feasible to experience rate small firms, i.e., firms with only a few employees. Because work accidents are random events and because their employment base is small, the number of accidents does not provide a reliable estimate of the underlying risk of injury.

7. It is also possible that Occupational Health and Safety regulations could attempt to influence the behavior of employees, although none of the existing regulatory models contemplates this.

8. Alternatively, sanctions raise the costs of not engaging in accident prevention.

9. Scholz and Gray also used their data set to replicate prior research in order to determine reasons for the discrepancy between their results and the results of these earlier analyses. They concluded that Smith (1979) and McCaffrey (1983) failed to find significant abatement effects because they had not accounted for long-term enforcement effects. Smaller deterrence effects found by Viscusi (1986) were attributed to sample differences. Specifically, the Scholz and Gray sample contained plants that were larger, more dangerous, and more heavily inspected than the average manufacturing plant examined by Viscusi. Scholz and Gray hypothesized that the plants in their sample were more amenable to the ameliorative effects of OSHA enforcement than the average plant.

10. The phrase “limited information processing capacity” is not meant to apply only to the abilities (or limitations) of managers. Rather it refers to limitations that constrain us all (including university professors).

11. For example, Meisenhelter (1991) notes that a period of six years elapsed between OSHA’s initial work on a Hazards Communication standard — similar to WHMIS — before it was finally issued in November 1983. Over 200 written comments were submitted totaling over 12,000 pages. There were 19 days of hearings, which produced 4,250 pages of transcripts.

12. A literature search uncovered only two studies examining the effect of an aspect of internal responsibility other than JHSCs on workplace safety. Lanoie (1992) estimated the impact of refusals to do unsafe work in Quebec, using industry aggregate data. He failed to find a relationship between the number of refusals per employee and the lost-time injury rate. However, Lanoie’s data show that
refusals are negatively related to injury severity, although this relationship is only marginally significant in two of his four specifications and statistically not different to zero in the other two. On the other hand, Cousineau et al. (1995), who also used Quebec data, found that refusals were positively related to one type of injury (“struck by or striking against”) thought to be particularly susceptible to safety regulation, while negatively related to two other types (“caught in or between” and “falls or slips”). However, the latter two relationships were not statistically significant.

13. A rival good is one that may be consumed by one and only one person. If it is possible to prevent the consumption of a good, it is excludable. A candy bar is a good that is both rival and excludable, while, clean air is both nonrival and non-excludable. These distinctions are important because it is generally thought that a private market is perfectly capable of efficiently providing rival and excludable goods, but not goods that are nonrival and nonexcludable.

14. Firms that self insure, an option available in most U.S. states for firms that meet certain fiscal requirements, are, by definition, perfectly experience-rated.

15. In most U.S. states, there is a minimum payroll requirement that a firm must satisfy in order to become experience rated.

References


Thomason, T. 2003. “Economic Incentives and Workplace Safety.” In *Prevent-
ing and Managing Disabling Injury at Work, Terrence Sullivan and John Frank, eds. London; New York: Taylor and Francis, pp. 183–204.
Workplace Injuries and Diseases
Prevention and Compensation

Essays in Honor of Terry Thomason

Karen Roberts
John F. Burton Jr.
Matthew M. Bodah
Editors

2005

W.E. Upjohn Institute for Employment Research
Kalamazoo, Michigan
Library of Congress Cataloging-in-Publication Data


Papers from a conference entitled “Workers’ Compensation: Current and Emerging Issues” held at the University of Rhode Island on March 27, 2004 to honor the former director of URI’s Schmidt Labor Research Center.

Includes bibliographical references and index.

2005009746

© 2005
W.E. Upjohn Institute for Employment Research
300 S. Westnedge Avenue
Kalamazoo, Michigan 49007-4686

The facts presented in this study and the observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of the W.E. Upjohn Institute for Employment Research.

Cover design by Alcorn Publication Design.
Index prepared by Diane Worden.
Printed in the United States of America.
Printed on recycled paper.