



Official Use

Conferencia Interamericana de Seguridad Social
Inter-American Conference on Social Security

October-2005

English – Or. English

Official Use

Risk in the Workplace: Economic Principles and Policy Alternatives

English - Or. English

For comments or suggestions, please contact Gabriel Martínez
gabriel.martinez@ciss.org.mx or Martha Miranda mmiranda@ciess.org.mx

Risk in the Workplace: Economic Principles and Policy Alternatives

Contents

I. Why the Government Regulates Risk?	3
II. Risk and Efficiency	4
III. Value of a Statistical Life	8
III.1 Using Value of Statistical Life in Litigation.....	11
IV. Assessing Regulatory Performance	12
V. Alternative Policy Strategies	17
VI. Enforcement	20
VII. Social Security Regimen and Workplace Safety Reform	21
VIII. Firms Size and Workplace Safety Reform	27
IX. Overview of the Occupational Risk in Latin America and Caribbean	31
X. Conclusions	34
References	37

Risk in the Workplace: Economic Principles and Policy Alternatives

I. Why the Government Regulates Risk?

Government action in the health and safety arena is justified by the shortcomings in risk information and the externalities present in financial and labor markets. Complete deregulation of health and safety provisions is likely to be far from the ideal economic solution, and governments have used a combination of regulation, social insurance and judicial compensation of damages in search of a balance in the goals of protecting the workers, combating abuse by employers, and providing a cooperative environment for the solution of risk and limited-information conditions. The agencies used by the State to deal with issues of risk at the workplace are regulatory, social insurance and courts. In some countries each of these functions can be performed by a national public agency, while in others the functions can be decentralized and performed by private agents. There can also be overlap in the functions of an agency, which can perform part or all of the functions.

The goal of regulatory agencies that address health and safety risks should be to isolate instances in which misinformation about health risks prevents people from making optimal tradeoffs and to isolate instances where health risks are not internalized in market decisions. It is important that agencies use flexible market-based regulations if their goals are to be achieved at least cost. The existence of a health risk does not necessarily imply the need for regulatory action. In the case of job safety, for example, perceived risks of job hazards lead to compensating differentials in wages for risk. While workers receive wage compensation to make them willing to bear the risk, when the worker does not have access to all the information available to the employer there is a space for abuse; similarly, when both employer and employee lack access to valuable information on risk, there can be a role for a cooperative solution and state regulation. In any case uncertainty about who knows what and when makes it necessary to use a procedure to deal with controversies, uncover information and adjudicate claims. In situations in which the risks are not known to workers, as in the case of dimly understood health hazards or situations in which the labor market is not competitive, market forces might not operate effectively to internalize the risk. Those cases provide an opportunity for constructive, cost-effective government intervention.

Unfortunately, the rationale of correcting the aforementioned problems has not always been the major motivation of regulatory intervention. The simple fact that risks exist has sometimes provided the impetus for the legislative mandates of the health and safety regulatory agencies. Very few regulatory impact analyses explore in a meaningful way the constructive role that the behavior of workers and firms may already play in the regulatory situation that is being considered.

The conventional regulatory approach to health and safety risks is to seek a technological solution either through capital investments in the workplace, changes in the safety devices, or similar kinds of requirements that do not entail any additional care on the part of the individual. The initial faith in the technological approach was so great that proponents envisioned a dramatic improvement in safety from such regulation. The goal of a risk-free society did not recognize the cost tradeoffs involved.

Theoretically, in a fully competitive employment market, market transactions between employers and employees could lead to efficient levels of health hazards and equitable compensations for diseased workers. Workers would not accept jobs posing known risks, unless the positions offered some additional offsetting feature. These risk premiums in turn would establish a safety incentive for the firm. The firm could reduce its wage costs through investment in greater workplace safety. Were this market mechanism fully effective, government intervention

would be unnecessary and should be avoided. However, a complete market approach ignores that for risk premiums to work properly, workers must know the potential risk they face, and they must be able to insure those risks in financial markets, an assumption that is untenable even in rich countries. Additionally, because of the problems of multiple causations, the latency factor, and the difficulty of tracing environmental causes of diseases, such risks are typically not as apparent as safety hazards. In an extreme case, workers may know nothing of the particular risk to which they are being exposed.

An accidental injury or health damage cannot always be observed by employers even after it happens, so that experience will not necessarily give all sides some idea of the risks involved. When health hazards are hidden, workers fail to demand an adequate wage premium for risk, and employers provide too little risk control.

In addition, if workers are warned of a low-probability health risk, they may overreact. Empirical evidence has indicated that individuals systematically over-assess low-probability risk of a magnitude comparable to that of workplace health risks, leading to alarmist market responses. The warning problem is thus double-edged. Unless workers are informed about the potential danger, risk levels will be too great. Employers have little incentive to provide this warning, since doing so would boost the wage rate they must pay and this wage cost will often exceed any reduction in the injury costs to the firm. Government regulations consequently must require that it be provided. On the other hand, any hazard warning program must be designed to convey risk information without inducing undue alarm response.

II. Risk and Efficiency

The potential role of the government is not to eliminate the risk, but rather to address market failures that lead to an inefficient balance between risk reduction and cost. The task of government regulatory agencies is to identify cases in which regulation can generate more benefits to society than the costs that are incurred, and to address the market failures using a cost-effective approach.

In order to achieve those goals, the focus should not simply be based upon rigid technological standards, but on flexible regulatory mechanisms that meet the performance goals. While there has been some movement toward incorporating those lessons from economic theory into regulatory approaches (e.g., the tradable permit system for reducing sulfur dioxide in the US), much regulatory policy ignores both efficiency and cost-effectiveness criteria in their formulation.

The legislative mandates of the health and safety regulatory agencies typically urge the agencies to promulgate standards to promote safety without consideration of the costs. For example it is typical to find in laws phrases such as “adequate margin of safety,” and standards for hazardous materials are to be set at levels that provide an “ample margin of safety.” The goal of the Occupational Safety and Health Act of 1970 in the U.S. is “to assure so far as possible every man and woman in the Nation safe and healthful work conditions.”

Judicial review has attempted to define more precisely those ambiguous goals; however, the courts have ruled consistently that the laws do not require consideration of the costs of the standards. The costs of that unbounded commitment to safety have proven to be considerable. Risk reduction measures should be undertaken only to the extent that their benefits exceed the costs. For example, adding a particular safety device to a machine is desirable only if the benefits of the safety device exceed the costs of modifying the product.

Importantly, benefits are judged not only according to financial consequences but also more broadly on society’s willingness to pay for the reductions in risks, as the appropriate benefit value recognizes the value of the risk reduction that goes beyond the financial effects. Safety is

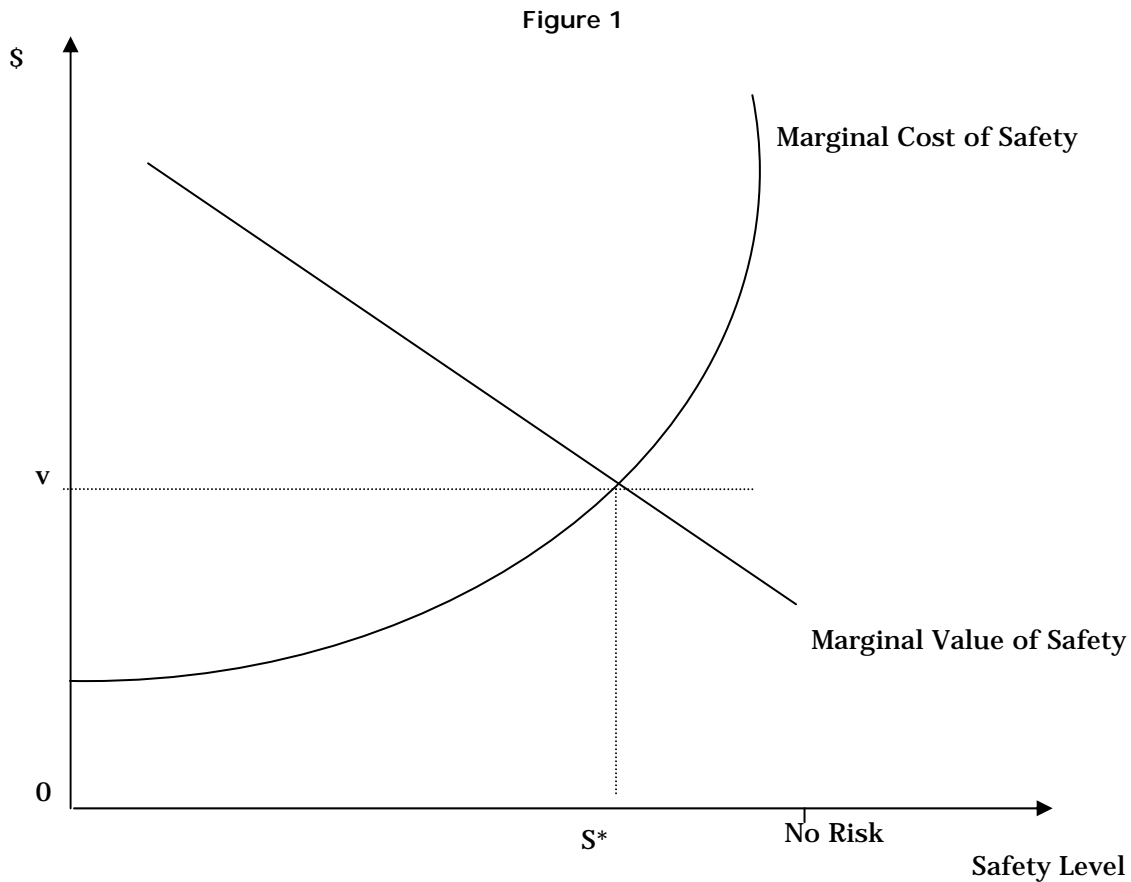
optimized when the marginal benefits equal marginal costs. Often there is a continuum of risk choices, such as the allowable level of exposure to toxic chemicals. As long as the incremental benefits of increased safety exceed the incremental costs, more tightening of the regulation or the imposition of liability on the firm is desirable.

The economic approach to worker safety is the basis for the labor market estimates of the value of life. Wage differentials establish an incentive for firms to promote safety, since doing so will lower their wage costs. In particular, it is primarily the risk-costs trade offs of the workers themselves that will determine the safety decision by the firm.

Figure 1 illustrates how these variables influence the level of safety provided. Suppose a reduction of job-related accidents and that improvements in safety have diminish incremental value to workers. The marginal value of the safety curve is downward-sloping curve, for the initial increments in safety have the greatest value. Workers' marginal value of safety is transmitted to firms through the wage rate the firm must pay to attract and retain workers. The firm can provide greatest levels of safety, but doing so entails additional marginal (or incremental) cost that increase as the level of safety becomes increasingly great, some initial safety improvements can be archived inexpensively. More extensive improvements could require an overhaul of the firm's technology, which would be more expensive. The marginal cost curve is increasing because safety equipment differs in its relative efficacy and the firm will choose to install the most effective equipment per unit cost first.¹

The optimal level of safety will be s^* , this level is short of the no-risk level because promoting safety is costly: at the level of safety provided workers would have been willing to pay v per expect accident in order to avoid such accidents. Additional safety beyond this point is not provided because the cost to the firm for each extra accident avoided exceeds workers valuation of the improvements. The ideal level of harm is not zero. A risk-free society is neither feasible nor desirable because of the inordinate costs of eliminating risk.

¹ Viscusi, Vernon, Harrington, "*Economics of Regulation and Antitrust*", Third Ed. MIT Press (2000).



Cost-effectiveness helps to eliminate some of the most unproductive regulatory alternatives, but it is still not a guarantee that the regulation is in society's interest because it does nothing to guarantee that the net benefits of the regulation are positive (let alone maximized). If market failures present risks to society, the proper economic concern for government policy is to assess the expected number of lives saved by a regulation against the costs of the regulation.

Regulation or litigation is excessively stringent, however, when firms are pushed to enact measures whose incremental costs outweigh incremental benefits. The implications of policies for choices of the level of health and safety are rarely neutral. Ideally, litigation should also help create incentives for efficient levels of safety, but this objective may be compromised when the main focus of the litigation is to provide compensation.

In contrast, government agencies typically focus not on expected outcomes presented by the risk but on worst-case scenarios on the grounds that doing so reflects "conservatism". In other words, when faced with uncertainty about the estimated risk, the regulatory agencies tend to choose assumptions that lead to higher assessments of the risk, thus representing the upper range of risk (rather than the mean risk) suggested by scientific knowledge.

As Nichols and Zeckhauser (1986)² suggest there are two main problems with the regulatory agencies' use of conservative risk assessments. First, overestimating the risk leads to an overestimate of the benefits of the risk-reducing regulation and thus excessive expenditures on risk

² "The Perils of Prudence", Regulation. Spring 1986.

reduction. The second problem is that the practice of conservatism also leads to a misallocation of resources within the risk concerns. Because regulatory agencies use conservative estimates for each parameter in a risk analysis, risks with more uncertain parameters will yield a higher risk estimate than those with fewer uncertain parameters. Also, Zeckhauser and Viscusi (1990) show that regulatory decisions based on cost-effectiveness rankings that use conservative or worst case assumptions are undesirable because they are likely to produce less overall risk reduction than using expected values.³

Thus, even though “risk A” poses a higher risk than “risk B,” if the latter has more uncertain parameters involved in the risk assessment, then the multiplicative effect of conservatism could lead to the agency deeming that the former risk poses a lower threat. That would lead to a misallocation of resources and a missed opportunity for saving lives. Proposed regulatory reforms sometimes call for agencies to base policies on best estimates of the risk rather than on worst-case scenarios. The focus on extremely improbable worst-case outcomes as opposed to realistic assessments of the risk sometimes dominates government policy.

Public or private provision of insurance against workplace accidents exists everywhere in industrial countries and in most of the countries of the Americas. Given the fact that the insurer (be it public or private) has imperfect information on the actions of the insured and on the state of nature, these programs are the source of moral hazard problems. Thus the provision of insurance may affect the incentives of the insured to make prevention efforts or may induce him to report a more severe accident than it is the case. A number of papers provide empirical analyses of the incentive effects of workers' compensation insurance. In particular, they show that an increase in the wage replacement ratio under compensation is associated with an increase in the frequency of reported accidents (e.g., Krueger, 1990)⁴, or in the duration of such claims (e.g. Meyer, Viscusi and Durbin, 1995).⁵

Furthermore, recent studies provide empirical evidence that worker compensation insurance may be used as a substitute for unemployment insurance. Thus workers who are about to be laid-off may have incentives to reduce their prevention efforts or to attempt to prolong their period of absence from work due to an illness or an accident, insofar as the generosity of compensation is greater than that of unemployment insurance. Worker compensation insurance also affects the composition of reported occupational injuries. Indeed, increases in the compensation wage replacement ratio may lead to another form of moral hazard: the reporting of injuries that did not come about off-the-job, but that are difficult to diagnose (e.g., low-back injuries) if the worker compensation is more generous than off-the-job medical insurance.

Risks never exist in isolation. They are part of systems. For that reason, any effort to reduce a single risk will have a range of consequences, some of them likely unintended. If the regulator increases regulation of benzene, a carcinogenic substance, it might lead companies to use an unsafe and perhaps a less safe substitute; it might also decrease wages of affected workers, and decrease the number of jobs in the relevant industry. People who have less money, and who are unemployed, tend to live shorter lives -- and hence occupational regulation might, under certain circumstances, sacrifice more lives than it saves.⁸ Of course the unintended consequences of risk regulation might be desirable rather than undesirable--as, for example, where regulation spurs new pollution control technologies.

³ Zeckhauser and Viscusi, “*Risk within Reason*”, Science 248 (1990).

⁴ Krueger, A.B. (1990), “*Incentive Effects of Workers' Compensation Insurance*”, Journal of Public Economics 41, 73-99.

⁵ Meyer, B.D., W.K. Viscusi and D.L. Durbin (1995), “*Workers' Compensation and Injury Duration: Evidence from a Natural Experiment*”, American Economic Review, 85, 322-40.

III. Value of a Statistical Life

The dominant concern with most health and safety regulations is to prevent fatalities. In order to conduct a benefit-cost analysis of such a regulation, one must assign a value to the benefit of reducing the risk of premature death. The proper benefit measure is the same as in other benefit contexts: society's willingness to pay for the particular benefit. Because government policies reduce risks of death rather than eliminate certain death for identified individuals, the correct benefit value is society's willingness to pay for the reduction in risk. The willingness-to-pay framework for valuing mortality risk reductions yields a measure of the value of a statistical life. In other words, if a regulation would reduce risk by one in one million to everyone in a population of one million, then the regulation would save one statistical life. If the average willingness to pay for that risk reduction is \$6 per person, then the value of a statistical life is \$6 million.

One approach to obtaining estimates of the willingness to pay for risk reduction is by conducting carefully crafted interviews of those exposed to the risk. However, the hypothetical nature of the survey and the possible incentives respondents have to overstate their answers suggest that asking people questions about amounts of money they would pay to reduce hypothetical risks may not be the best measure of how much they actually would be willing to pay to reduce real risks to themselves.

Economists turn to actual market behavior in which people make tradeoffs between risk and money as part of their economic decisions. In the case of jobs, for example, the wages commanded by the job reflect not only various skill aspects but also premiums for unpleasant characteristics of the job such as the job risks that workers must bear. Using detailed data on wages and prices, economists have estimated people's tradeoffs between money and fatality risk, thus establishing a value of statistical lives based on market decisions.

A number of studies began to use extensive labor market data to derive such estimates in the late 1970s. For workers in jobs of average risk, the estimates imply that, in current dollars, workers in the United States receive premiums in the range of \$600 to face an additional annual work-related fatality risk of one chance in 10,000. Put somewhat differently, if there were 10,000 such workers facing an annual fatality of one chance in 10,000, there would be one statistical death. In return for that risk, workers would receive total additional wage compensation of \$6 million. The compensation establishes the value of a statistical life, based on workers' own attitudes toward risks. When economists such as W. Kip Viscusi and Robert Smith first developed estimates in the millions of dollars, some critics suggested that they might be too high. Numerous studies over the past several decades have, however, documented the plausibility of estimates in the \$3 million to \$8 million range.

An interesting implication of those results is that the quite considerable estimates of the value of a statistical life have been derived from actual market decisions involving jobs and products. The estimates suggest that in situations in which there is an awareness of the risk, market forces are powerful and create safety incentives. Thus, we are not operating in a world in which there are no constraints other than regulatory intervention to promote safety. Rather, labor contracts already create incentives for safety that should not be overridden by intrusive regulations, but instead define the overall economic framework in which regulatory interventions can potentially complement the already significant market forces at work.

Regulatory agencies were slow to adopt the emerging value of a statistical life estimate, as they instead clung to more traditional procedures such as the present value of lost earnings (which yield much lower estimates of the benefits of risk reduction) on the grounds that it would be immoral to value human life. The estimate of the value of a statistical life has changed over time and varies substantially across — and even within — regulatory agencies. The value of the statistical life is not a universal constant. Indeed, there is substantial heterogeneity in the value of a

statistical life throughout society as some people are more willing to bear risk than others. Moreover, the quantity of life at risk may be quite different in different regulatory contexts. However, for the most part, such legitimate economic concerns with differences in the value of a statistical life have not accounted for the discrepancies across different regulatory arenas. It can also be pointed out that using a traditional procedure, such as the present value of lost earnings, or a replacement rate (the usual procedure in social insurance), represents also a monetary valuation of human life.

The value estimates adopted by OSHA in the U.S. typically have been in the same vein as those developed through labor market estimates. Thus, OSHA has often used such figures as \$3.5 million to value the risk reduction benefits of its efforts. The EPA used similar numbers in its early efforts, but most recently has been using figures of about \$6 million, which represents the midpoint of market estimates of the value of a statistical life. What is more, EPA, through its Science Advisory Board economics panel and other efforts, has been grappling with frontier issues in the value of a statistical life, such as the differential values that should be placed on the lives of those exposed involuntarily to environmental hazards as opposed to workers who have self selected themselves into relatively high risk jobs.

Countless regulations now attempt to reduce statistical risks. Cost-benefit analysis must generally accompany these regulations, at least if their costs are high, and to undertake that analysis, agencies must turn human lives into monetary equivalents.⁶

What is the source of these numbers? The answer involves real-world markets, producing evidence of compensation levels for actual risks. In the workplace and for consumer goods, additional safety has a price; market evidence is investigated to identify that price. Agency valuations are largely a product of studies of actual workplace risks, attempting to determine how much workers are paid to assume mortality hazards. The relevant risks usually are in the general range of 1/10,000 to 1/100,000. The calculation of the value of a statistical life is a product of simple arithmetic. Suppose that workers must be paid \$600, on average, to eliminate a risk of 1/10,000. If so, the value of a statistical life would be said to be \$6 million.

Additional information comes from contingent studies, asking people how much they are willing to pay to reduce statistical risks of death. Regulators are attempting to produce the optimal level of deterrence by reference to actual market valuation of statistical risks. Nonetheless, serious questions might be raised about the use of these studies by the Environmental Protection Agency (EPA) in the U.S. and other agencies. The underlying studies of market behavior show significant variety in the crucial figures, ranging from \$0.7 million, in 1997 dollars, to \$16.3 million.⁷ The EPA has adopted the \$6.1 million figure on the ground that it represents the median in the relevant studies. But there is a risk of arbitrariness in fastening on that median figure. In fact a more general look at the Value Statistical Life (VLS) data produces further puzzles and wider ranges.

Some studies find no compensating differentials at all, indicating a VSL of zero,⁸ implausibly low, for purposes of policy. Others find that non-unionized workers receive negative compensating differentials for risk, that is, they appear to be paid less because they face mortality risks. Perhaps workers are unaware of the risks that they are running. If so, the labor market studies do not really show how workers are trading off risks for money, and hence they are essentially useless.

⁶ See Stephen Breyer et al., "Administrative Law and Regulatory Policy", 120-35 (5th ed. 2002).

⁷ See EPA, Guidelines for Preparing Economic Analyses 89 (2000). For a detailed outline and discussion, see Richard W. Parker, "Grading the Government", 70 U Chi L Rev 1345, 1485-86 (2003).

⁸ See Peter Dorman and P. Hagstrom, "Wage Compensation for Dangerous Work Revisited", 52 Industrial and Labor Relations Rev. 116 (1998).

This objection cannot be simply dismissed. But current agency practice depends on the judgment that the numerous studies of risk premiums indicate that sufficient numbers of workers are informed to establish a “price” for additional increments of safety. If that judgment is incorrect, then the current numbers need to be rethought, on the ground that labor markets do not show informed tradeoffs between money and statistical risks.

Under the current theory, the regulators’ task would be to use other tools, perhaps contingent valuation studies, to establish those tradeoffs. Do workers voluntarily trade risks for dollars? An obvious objection would be that many of the relevant works have few options, and hence their markets behavior trading off a risk of 1/10,000, say, for an apparently low pay, is in an important sense involuntary. If taken as a normative claim about voluntariness, the claim may be right, when people have few or bad options, their choices might not count as voluntary. But if taken as an objection to VSL studies, the claim is less helpful. Under the theory that we have outlined, the question is how much people are willing to pay to eliminate specified risks.

People are willing to pay less, for risk reduction, if they have less to pay. In any case government does workers no favors by requiring them to “buy” more health protection than they want. Suppose, for example, that workers are generally willing to accept \$600 to run a risk of 1/10,000. If an agency bans that deal, and forbids workers from running that risk at that price, it will not improve workers’ welfare. This is simply a specific example of the general proposition that when people’s circumstances lead them to make harsh deals, they are usually not helped when government blocks those deals. Of course the analysis would be different when those who receive the benefits of regulation do not also pay for it.

Any economy contains a wide range of occupations and industries, and a uniform VSL should not be expected to emerge from each of them.⁹ Indeed, a recent study finds significant differences across both occupations and industries, with blue collar workers showing a higher VSL than others. Both individuals and occupational groups should be expected to show significant differences in their VSL. This is partly because they are risk-averse or risk-inclined, and partly because of differences in terms of both wealth and income.¹⁰ Equally important, risk reduction programs that are aimed at wealthy populations should, under the prevailing theory, produce a higher VSL than similar programs aimed at poor populations. Those who have a great deal of resources will naturally show a higher VSL than those with little. According to the theory that now animates regulatory practice, agencies should use a VSL that corresponds to the actual number for the population at stake.

This point has numerous implications. The workplace studies on which agencies currently rely involve people with income that is below the population-wide median; to that extent, the numbers are too low as applied to a population that is more representative of the nation as a whole. Regulators distinctions between wealthy and poor populations would undoubtedly be quite controversial. Less controversially, recent evidence suggests that older people show a lower VSL than younger people,¹¹ and that distinction might well be incorporated into regulatory policy.

The general conclusion is that use of approximations to the willingness to pay (WTP) to avoid risks or death has a plausible logic to it, that variable numbers make far more sense than uniform ones, and that the real question involves information and administrability. Regulatory programs often affect thousands or even millions of people at the same time, and full individuation is therefore impossible. A rule that calls for minimum levels of air quality cannot provide air quality that is perfectly calibrated to each person’s WTP. Nonetheless, regulators could certainly move in the direction of greater individuation. So long as the WTP figures are accurate, government

⁹ W. Kip Viscusi, “The Value of Life: Estimates With Risks by Occupation and Industry”, 42 *Ec. Inquiry* 29, 39-41 (2004).

¹⁰ See Dora L. Costa & Matthew E. Kahn, “The Rising Price of Nonmarket Goods”, 93 *Am. Econ. Rev. (Papers & Proc.)* 227, 229 tbl.1 (2003).

¹¹ Joseph P. Aldy and W. Kip Viscusi, “Age Variations in Workers’ Value of a Statistical Life” mimeo (2003).

does well to begin with them. Exactly the same is true for courts in wrongful death actions. The VSL should reflect people's willingness to pay to avoid risks. Because the willingness to pay to avoid risks increases with income, poorer people have lower VSLs than wealthier people. Reflecting cross-national wealth differences, VSL is highly variable across nations. 136 Studies find a VSL as low as \$200,000 for Taiwan, \$500,000 for South Korea, and \$1.2 million for India, but \$21.7 million for Canada and \$19 million for Australia.

Table 1: VSL across countries¹²

Nation and Year of Study	VSL (in 2000 US\$)
Japan (1991)	\$9.7 million
South Korea (1993)	\$0.8 million
Canada (1989)	\$3.9-4.7 million
India (1996/97)	\$1.2-1.5 million
Taiwan (1997)	\$0.2-0.9 million
Australia (1997)	\$11.3-19.1 million
Hong Kong (1998)	\$1.7 million
Switzerland (2001)	\$6.3-8.6 million
United Kingdom (2000)	\$19.9 million

III.1 Using Value of Statistical Life in Litigation

Through tort, civil or labor law, courts provide a set of fact-specific awards that attempt to compensate for and to deter wrongful death. The resulting awards are highly variable. For example, in the United States courts have recently given successful plaintiffs as little as a few thousand dollars and as much as tens of millions of dollars. In Latin American countries, labor codes have superseded tort or civil laws, and specialized labor courts are in charge of dealing with litigation related to work risks. In what follows, in the English version, we will use the term tort law in the discussions conducive.

Notwithstanding their overlapping goals, administrative regulations and tort law diverge from one another in dramatic ways. The most obvious difference is that tort law generally disregards the welfare loss to the person who has died; regulatory policy treats that loss as its central and indeed exclusive focus.

While tort law seeks to ensure compensation, especially for family members, regulatory policy is designed to produce optimal levels of risk. Tort law has long focused on the compensation of those still living—a focus that naturally leads to disregard of the deceased, an emphasis on what the plaintiffs have lost, and an interest in a set of highly individuated awards. By contrast, regulatory policy, which assigns monetary values to statistical lives, is concerned above all with producing the right deterrent signal, a concern that might seem to explain the use of a single, uniform number for the valuation of the loss of life.

¹² Posner Eric, A., Sunstein, Cass R., "Dollars and Death", John M. Olin Law & Economics Working Paper No. 222, September 2004.

IV. Assessing Regulatory Performance

Although many agencies use reasonable measures of the value of a statistical life for the purposes of assessing benefits, the cost per life saved for the regulations actually promulgated often far exceeds the estimated benefits. There are, of course, other benefit components that must be taken into account and might increase the benefit estimates of the regulation sufficiently to offset the total costs. However, such influences alone do not account for the fact that the cost expended per statistical life saved often far exceeds the estimated benefits of the risk reduction. Rather, the restrictive nature of agencies' legislative mandates often precludes consideration of costs in the regulatory decision.

A regulatory system based on sound economic principles would reallocate resources from the high to the low-cost regulations. That would result in more lives saved at the same cost to society (or equivalently, shifting resources could result in the same number of lives saved at lower cost to society). The health and safety aspects of any activity will be governed by the combined influence of the technological characteristics of the safety context, the attributes of the individual, and the safety-related behavior that the individual takes. Although buckling the seatbelt potentially could reduce the risk of injury in an automobile crash, buckled-up drivers may change their driving habits such that there would be no overall safety-enhancing effect. Indeed, empirical tests derived by Sam Peltzman failed to indicate any net beneficial effect of seatbelt regulations on occupant safety. While many empirical studies suggest that there is a behavioral response to such regulations, whether the net effect is an increase or decrease in risk is a matter of empirical dispute.

Those and other potential behavioral responses to regulation typically have not been recognized in the design of regulatory policies. There continues to be a substantial gap between research that has documented the potential benefits of safety-enhancing behavior and regulatory approaches that continue to emphasize technological solutions. Failure to recognize and exploit individual behavioral responses leads to regulations that are less successful in promoting safety than would be a more comprehensive approach.

In many policy contexts there is an interaction between regulation and litigation. Many of the economic rationales for government regulation pertain to various forms of market failure. These same forms of market failure often lead to litigation as well, as injured parties seek to obtain damages for the harms that have been inflicted on them because of a lack of appropriate recognition of their economic interests by the party inflicting the harm. The policy task is to coordinate the influences of these two different sets of social institutions, recognizing their different strengths and different functions.

The potential importance of the interaction between regulation and litigation is not a new issue. Traditionally the focus has been on broad conceptual issues, such as the potential for institutional overlap on the creation of economic incentives. The policy concerns arising from these analyses often have focused on fairly narrow policy remedies, such as the provision of a regulatory compliance defense for firms that are in compliance with explicit government standards but are nevertheless subject to litigation.¹³

The different functioning of institutions that litigate and those that regulate is apparent from a look at their roles in promoting health and safety. Consider first the creation of economic incentives. Regulation is generally superior in addressing technical scientific issues because of the importance of expertise in analyzing these regulatory issues. Difficulties arise when these matters are delegated to juries or judges on a case by-case basis. Recent literature has documented the

¹³ Firms could introduce evidence of regulatory compliance as evidence that they were not negligent. In extreme cases, regulatory compliance could be exculpatory.

failings of juries in thinking systematically about risk, as jurors exhibit a wide variety of systematic biases in assessing accident situations, such as hindsight bias in the evaluation of past actions involving risk. Similarly, public opinion often plays a decisive role for public decision making in risk issues.

Government regulations will usually provide a more sound approach to promoting health than litigation does, which by its nature focuses on individual circumstances rather than the functioning of an entire product market. The stringency of government regulations may be excessive from a benefit-cost standpoint, owing to the restrictive nature of regulatory agencies' legislative mandates. When this occurs, regulatory standards for health and safety typically should not require any augmentation through judicial proceedings.

If, however, regulations do not exist, litigation can often help address gaps in the regulatory structure and stimulate regulatory activity. One of the most prominent examples is asbestos risks. Historically, asbestos risks had not been strongly regulated, but the emergence of a wave of asbestos litigation induced the OSHA and the EPA in the United States to set stringent regulations. In this instance, the combination of litigation and subsequent regulation led to inordinately large safety incentives. Litigation may complement regulation when it provides for a transfer of income to injured parties to address the damages incurred.

A general problem with distinct roles for litigation and regulation is that there is no formal or informal mechanism for coordinating the roles of these two institutions. That one institution is imposing economic penalties for a particular type of risk does not prevent the other from also imposing sanctions. That there is a continuing inherent problem in coordinating the roles of regulation and litigation is well documented in the literature.

What is new is that the character of these coordination problems has changed dramatically since the mid-1990s especially in the US. In recent cases no longer was the issue one of litigation creating incentives that overlapped with those resulting from regulation. Rather, litigation was being used as the financial lever to force companies to accept negotiated regulatory policies as part of the litigation. Thus litigation led to regulation, but not regulation that went through the usual rule-making process as a result of a careful analysis by government regulatory agencies subject to their legislative mandates. Rather, the parties in the lawsuit negotiated regulatory changes as part of the package to end the litigation. This has affected other countries, as is the case of the tobacco industry: other countries have adopted regulations on advertisement, tobacco use and taxes following the result of U.S. litigation cases. Asbestos, fertilizers, pharmaceuticals and other industries provide similar examples.

These negotiated solutions have also gone beyond simply specifying regulatory changes. In at least one instance the settlement has led to the imposition of what is effectively an excise tax on products. Rather than imposing a conventional damages award on the defendant, the tobacco settlement imposes charges on customers on a per unit basis in the future. Consider also the case of asbestos, after a long period of regulatory inaction and billions in liability suits the regulators in the U.S. overreacted by trying to cut exposure virtually to zero, OSHA asbestos rules cost on average \$89.3 millions for every life saved and the EPA issued regulations that imposed cost on average of \$104.2 millions for every life saved. Other workplace hazard, such as arsenic exposure in glass manufacturing, is stringently regulated at \$92.5 million per life saved. Unfortunately, there is no coordinating mechanism to ensure the risk level that result from these diverse efforts is appropriate on to keep society from responding excessively to a few targeted hazards while ignoring others.

The increasing cost by the regulator response to litigation cases in fact establishes a tax on the product and paid almost entirely by the consumer rather than a damages payment paid for by the defendant. Citizen interests are not explicitly represented, and, unlike the case of the regulatory changes, there is no mechanism to ensure that these outcomes are in society's best

interests. Moreover, there is typically no procedure for even creating the appearance of the degree of legitimacy usually accorded to governmental policies.

If there is an error in these litigation settlements that impose regulatory and tax changes, the potential adverse consequences could be enormous. The stakes of the tobacco litigation exceeded more than \$200 billion in expected penalties during the next twenty-five years. The regulatory changes also could have significant anticompetitive effects. Although other litigation involves stakes that are typically not as great as those in tobacco, in the effects on particular industries, the influences could be even greater.

In considering the merits of litigation it is useful to assess how it performs from the standpoint of efficient deterrence and efficient insurance. One of the chief functions of a liability system and government regulations is to establish optimal levels of deterrence. Kenneth S. Abraham distinguishes two different types of litigation, each of which has different implications for economic incentives. Litigation that he calls “forward looking” focuses on setting up requirements on firm behavior or a funding mechanism that will directly influence incentives for the future. The settlement of the tobacco litigation was forward looking in that it led to regulatory changes and a damages formula that was largely tantamount to an excise tax on cigarettes. Litigation that Abraham calls “backward looking” is more similar in character to conventional tort litigation. The lawsuits in the lead paint litigation against landlords fall into the backward-looking category. These suits seek to obtain compensation for parties that have been injured.¹⁴

Such compensation will establish payment structures that could potentially alter future incentives because firms will expect to be subject to similar sanctions from future litigation. However, if all such decisions have already been made or if the product is no longer sold, there will be no incentive effect unless these suits impinge on current behavior in some manner. Thus there will be no incentive effect for lead paint manufacturers if lead paint is no longer produced in the country. However, the lead paint suits against landlords potentially could have an incentive effect to the extent that they affect building maintenance, efforts to remove lead paint, and warnings to tenants about lead paint risks. There may also be more general deterrent efforts for landlords beyond lead paint.

A second potential function of social institutions dealing with risk is providing optimal insurance to those who have suffered injuries or illnesses. Regulatory policies usually do not provide any insurance compensation for victims but instead are focused almost exclusively on establishing regulatory standards for health and safety. Insurance functions are typically handled through targeted government programs that focus on the disabled, the poor, or the elderly as the Social Security institutions in several countries.

In contrast, litigation often has as its principal purpose an effort to transfer income to those who have suffered injuries. From the standpoint of optimal insurance this transfer should be sufficient to completely cover the economic loss when people have suffered a financial loss. The desirability of providing this insurance stems from individual risk aversion, which makes insurance of such losses desirable. Governmental entities should be risk neutral except for extremely large losses because they can spread these losses across a large citizenry base. Thus any optimal insurance rationale for transfers to the government must assume that the losses ultimately borne by individual taxpayers will be sufficiently great that risk aversion will come into play.

For individuals sustaining injuries and suffering illnesses, there will be financial losses and effects on individual health. The object of insurance for financial losses is to restore individuals to

¹⁴ Abraham, Kenneth S., “The Insurance Effects of Regulation by Litigation” in “Regulation by Litigation”, W. Kip Viscusi editor, AEI-Brookings Joint Center for Regulatory Studies (2002).

their pre-accident level of utility, but that objective is not pertinent to health effects.¹⁵ Typically, it will not be desirable for an individual to purchase so much insurance (or to receive as much insurance from social security) as to be as well off as he or she would have been had the illness or injury not occurred, because these events reduce people's ability to derive benefits from additional funds.

On the other hand, even enormous transfers of money may not be adequate enough to restore the pre-accident welfare level to one who has become severely disabled. There is also the practical problem of ascertaining a person's psychic losses from such major injuries. But the proper role of the courts will typically fall short of restoring the plaintiff's pre-illness level of utility even when liability for the firm is established.

The cost-benefit default principles: in brief, these principles would be: (1) allow de minimis¹⁶ exceptions to regulatory requirements; (2) authorize agencies to permit "acceptable" risks, departing from a requirement of "absolute" safety; (3) permit agencies to take account of both costs and feasibility; and (4) allow agencies to balance costs against benefits. Taken as a whole, the cost-benefit default principles are making a substantial difference to regulatory policy, both because of their effects in litigated cases and because of their systemic consequences for regulation.

Ultimately, the courts must address the trade off between the competing objectives of deterrence and insurance. De-coupling the insurance and deterrence functions by augmenting a benefit payment with a fine paid by the firm is feasible in job injuries. Moreover, society must also consider the appropriate allocation of institutional responsibility since governmental regulation, if properly designed and enforced, may be more efficient in promoting the deterrence objective. The goal of setting compensation in such a way that it leads to appropriate deterrence must depend, in part, on whether there are demonstrable safety incentives generated by liability cost.

From a theoretical standpoint, any cost positively related to risk levels will foster added safety incentives. Workers receiving more generous benefits are more likely to claim that injuries occurring off the job were actually job related. A growing body of research has found that workers compensation benefits have unfavorable effects on safety. Higher benefits appear to increase both the frequency of work injuries and the number of compensation claims filed. When workers compensation is limited, workers often would rather stay on the job than suffer the loss of being out of work, but often workers compensation combines elements of earnings insurance and welfare assistance. More generous benefits levels also lead to moral hazard in terms of reporting and injury duration, as a result some apparent underinsurance will consequently be optimal for social insurance.¹⁷

Hypothetically higher levels of risk cause workers to demand higher wages, and these wage-risk tradeoffs in turn establish market incentives for safety. However, the introduction of workers compensation establish an additional dimension, higher levels of workers compensation benefits should dampen the wage premium for risk demanded by workers and the funding mechanism for workers compensation should provide incentives for safety. Some economist advocate using an injury tax approach similar to the funding mechanism for workers compensation as a substitute for regulatory policy. In theory this approach could enhance safety levels, however the empirical evidence indicate that the moral hazard effects dominate and that increased benefits therefore lead to greater injury rates.

¹⁵ Optimal insurance satisfies the property that it equates the marginal utility of income when one is healthy to the marginal utility of income when one is ill.

¹⁶ A lower limit below which no action is required.

¹⁷ W. Kip Viscusi, "Product and Occupational Liability", *Journal of Economic Perspectives*, Vol. 5, No. 3 Summer (1991)

Because premiums are not always fully experience rated, particularly for small firms, and because moral hazard problems exist, the extent of the safety incentive effects of workers' compensations is not clear. Empirical evidence in general result that benefits increase the incidence of less severe accidents, the most severe accidents should reflect very little moral hazard effect. Viscusi and Moore's (1989) estimations indicate a substantial safety effect of workers' compensations in reducing fatality rates. Furthermore, if greater safety is promoted by workers' compensation, the required compensating wage differential for risk and the level of injury cost to the firm will both be reduced. Thus, workers' compensation has both direct and indirect effects on workers' wages, and these effects offset both the cost of insurance premiums and the employer's expenditures on safety.

Often the debate in the Social Security arena is about benefit levels not high enough to provide for full income replacement, increases in compensations evoke spiraling premium costs and moral hazard problems, such as false claims and overextended periods of recovery. Although the moral hazard problems are important, workers' compensation also plays a role in reducing workplace fatalities.¹⁸ Statutory benefits are provided to compensate injured or sick workers for lost income, medical and rehabilitation expenses, and, if the injury or illness is serious enough, for a diminished capacity to enjoy the activities of everyday life. Benefit structures vary considerably across schemes. Differences exist in the level and duration of benefits, as well as in limits on the amount of compensation paid.

Any evaluation of workers' compensation benefits should be undertaken in the context of alternative sources of accident compensation and the social security and taxation systems:

- How adequately it compensates injured workers (both in terms of compensation paid and coverage);
- How well it reinforces incentives for employers and employees. In particular, incentives: for safer workplaces, for employees to participate in rehabilitation and return to work, and for employers to facilitate return to work; and
- The degree to which the costs of workplace injury and illness are funded from employer contributions rather than shifted elsewhere.

These criteria involve obvious trade-offs. A 'generous' benefits structure may provide poor incentives for rehabilitation and return to work. Conversely, benefits that impose limits on income replacement (as a means of encouraging return to work) may be regarded as inequitable for workers with serious injuries which respond slowly (if at all) to rehabilitation.

Similarly, if benefits are reduced to provide incentives to employees to participate in rehabilitation and return to work, this may encourage claimants to seek other forms of compensation (which shifts costs away from the workers' compensation scheme). If benefits are increased, cost shifting on to the workers' compensation scheme, by people who have sustained injuries or illnesses outside the workplace, can be encouraged.

There is evidence that changes in workers' compensation benefits are reflected in wages. Using U.S. data, Gruber and Krueger (1990) found that a \$1 increase in expected benefits led to a \$0.86 fall in wages. A more recent study, by Kaestner (1996), which analyzed U.S. workers by age group, found that a \$1 increase in expected benefits reduces wages by more than \$1. This is consistent with the earlier results of Moore and Viscusi and may be evidence that some of the indirect costs of workplace harm are also passed on to workers (1990, pp. 67-68). Kaestner also found that, amongst the youngest age group of workers who were covered by mandated minimum wages, a 1 per cent increase in expected benefits and associated costs led to a 1.5 per cent increase in unemployment.

¹⁸ Moore M., Viscusi W. Kip, "Promoting safety through workers compensation: the efficacy and net wage costs of injury insurance", *RAND Journal of Economics*, Vol. 20, No. 4, winter (1989).

The levels of benefits, conditions of access and the manner in which the benefit is paid (periodic or lump sum) all provide incentives for particular forms of behavior. Taxation and interaction with other income support mechanisms, such as social security, are also important. These incentives are complex and may depend on other scheme features—for example, access to judicial dispute resolution

A range of U.S. studies suggest that a 10 per cent increase in benefit levels leads to an increase in claims of between 4 and 10 per cent (Butler 1983; Butler and Worrall 1983; Worrall and Butler 1989; Johnson and Ondrich 1990; Thomason 1993; Currington 1994; Meyer, Viscusi and Durbin 1995; Hirsch, Macpherson and Dumond 1997). Krueger examined the impact of benefit increases on claim duration in a U.S. scheme. He concluded that:

... the duration of injuries increased by 8 per cent more for the group of workers that experienced a 5 per cent increase in benefits than for the group of workers that had no change in their benefit (1990, p. 1).

However, in the case of death and serious injury, the natural human desire to avoid harm, together with a potential increase in premiums following an increase in benefits, appear to be the dominant influences on behavior. Using U.S. data, Moore and Viscusi (1989) found that increases in workers' compensation benefits resulted in improved health and safety measures by firms and a reduction in the number of fatal accidents.

V. Alternative Policy Strategies

As mentioned earlier, the dominant approach to health and safety regulation has been technology-forcing standards. Unlike the environmental policy arena, there has been no discussion of the potential role of market-based mechanisms for health and safety standards. The two principal areas stressed by economists have been the potential role for performance standards and the desirability of hazard communication strategies.

The impetus for performance-oriented regulations is that it is desirable to obtain regulatory objectives in the most cost-effective way possible. The flexibility of performance-oriented regulations (instead of technology-based regulations) allows the regulated entities to achieve the outcomes at least cost.

Over time, there have been some minor efforts by OSHA and other agencies to adopt performance-oriented regulations that exploit the superior cost-effectiveness of permitting firms to have some discretion. A classical example is the OSHA grain dust standard, which gave firms a variety of performance-oriented alternatives to choose from in order to achieve compliance with the regulation. By giving firms several options, it is possible for them to select the approach that is least costly given the workplace situation but still meets the desired safety objective.

Other kinds of performance-oriented approaches that involve the engagement of protective behavior by workers have not been adopted. For example, one could decrease the control cost of reducing cotton dust exposures by requiring that workers wear light disposable cotton masks or rotating workers out of high exposure areas. Protective equipment such as respirators also may be desirable in other contexts. However, the emphasis of regulatory policies has been on approaches that do not rely on either protective equipment or increased care on the part of workers.

A principal exception has been the emergence of hazard warnings regulations. Unlike technology-forcing regulatory policies that constrain individual choice, hazard warnings potentially can work through the market by providing consumers and workers with needed information. To the extent that the rationale for intervention is inadequate information regarding risks, hazard

warnings can address that shortcoming directly by eliminating the informational gap, thus allowing people to make market tradeoffs that fit their preferences.

At the same time, choices by workers and consumers subject to the receipt of the information would be respected so that market forces would permit people to make choices consistent with their own risk-cost balancing rather than being subject to uniform regulatory standards that almost invariably fail to recognize such differences in individuals' willingness to bear risk.

While hazard warnings are now ubiquitous, that has not always been the case. Beginning in 1927, legislation only required that a dozen of the most dangerous chemicals such as hydrochloric acid and sulfuric acid be labeled "POISON". In 1983, OSHA enacted its communication regulations that for the first time required hazard warnings for dangerous workplace chemicals. Now, both as a result of regulatory interventions and the threat of liability lawsuits, there are extensive warnings for the multiplicity of risks we face. The emerging literature developed by economists and other disciplines has derived many fundamental principles for hazard warnings. Chief among them is that a warning should provide new and accurate information. Properly designed hazard warnings can be among the most beneficial interventions in that they foster improved market performance without imposing any regulatory burden.

It has been suggested that risk-risk analyses should be carried out on safety regulations. Given that efficient regulatory policy seems unlikely, a less stringent evaluative standard would be the use of a "risk-risk" test. Risk-risk analysis weighs the risk-increasing aspects of a regulation against the risk reduction caused by the regulation. Adopting such analysis would encourage the promulgation of regulations that result in a net risk reduction to society. It has been argued that some of the most expensive life-saving regulations may not only lead to negative net benefits, but may also result in an overall increased number of fatalities. Thus the net number of lives saved by regulation will be given by the difference between the lives saved by regulation and the lives lost due to the reduction in employment and wages level. An estimation of the net number of lives saved by a regulation is referred to as a risk-risk analysis (Viscusi, 1994).¹⁹ Risk-risk analysis can be a useful tool for eliminating undesirable policy options in cases where the use of cost-benefit analysis is controversial or not available. However a life-saving regulation that leads to a net reduction in mortality does not imply that it is worth from an economic efficiency viewpoint, since the cost can still exceed the benefits. A net fall in mortality can be viewed as a necessary but not sufficient condition.

The cost of regulation designed to save lives are often expressed in terms of the cost per life saved. Thus to carry out a risk-risk analysis, the cost per life saved can be compared to the income loss that will induce an expected fatality. The key issue in the risk-risk analysis field has therefore been to estimate the income loss that will induce an expected fatality. The income loss that will induce an expected fatality can be inferred from estimates of the WTP for mortality risk reductions.

The first form of risk-risk analysis is straightforward. Although risks to life may be reduced through various regulations, they may also create new risks. If a regulation leads to new manufacturing or construction activity, some workers will be injured or killed as part of such efforts. Thus, extremely ineffective regulations could potentially create more risks than they reduce.

A more thorough form of risk-risk analysis pertains to the economic result first developed in the late 1970s that demonstrates a positive relationship between individual wealth and health. As society has become more affluent, health improved and it is demanded greater levels of safety of all kinds. Regulations impose costs on society and lead to a reallocation of resources that would

¹⁹ W. Kip Viscusi, "Risk-risk Analysis", *Journal of Risk and Uncertainty*, No. 8 (1994).

have been expended on consumption goods—the net effect of which would have been health enhancing. If policies divert health-enhancing resources to extremely ineffective regulatory efforts, the net effect may be to harm individual health.

Any compensation scheme should not only compensate workers but also encourage employers to take additional health precautions, by tying the tax to the riskiness of current operations. In short, occupational-disease or workplace safety policy should take in account the incentive effects of compensations plans on risk levels and coordinate these effects with the incentives created by regulatory programs. A penalty tax (Viscusi 1988) linked to present exposure levels will encourage employers to reduce workplace risks to the point that the marginal cost of additional precautions equals the tax. The latter policy is far more efficient at optimizing risk-reduction investment, encourage employers to clean up their operations rather than just close them down. Viscusi propose a strategy for archiving both compensation and efficient levels of workplace risk combined two types of financing: a targeted tax to help victims of current hazards and a broad based program to help victims of past exposures, who would be compensated through the Social Security system or some other social insurance fund. Victims of occupational disease would be treated the same as victims of diseases caused by hazardous wastes or those of unknown origin. In general, compensation should depend on how a disease affects its victim, not on how it was contracted.²⁰

Workplace conditions should be regulated through a combination of minimum standards and penalty taxes based on the current hazards in the workplace. Thus, the objective of regulatory enforcement should be not to compel compliance with arbitrary technical standards but to induce firms to make whichever safety investment are most feasible to attain the level of safety which is efficient.

²⁰W. Kip Viscusi, "Compensating workplace toxic torts", Proceedings of the Academy of Political Science, Vol. 37 No. 1, New Directions in Liability Law (1988).

VI. Enforcement

Compliance with health and safety standards is expensive, with costs that often run into the billions of dollars. If compliance were discretionary, many firms would choose simply not to comply because noncompliance would be in their financial self-interest. Those relationships have been the focus of economic analyses, which have shown that firms are more likely to comply with the regulation if the costs of compliance are less than the expected penalties associated with non-compliance.

Many standards promulgated by regulatory agencies in safety workplace and environment need an effective enforcement effort is essential to establishing the incentives needed to promote compliance on the part of the affected firms. EPA in the U.S. about water enforcement effort represents perhaps a best-case scenario. All major polluters are required to file regular discharge-monitoring reports and are subject to an annual inspection.

That performance contrasts with OSHA in the US. In any given year, a firm faces a probability of less than one in 100 that it will see an OSHA inspector. Some economists have compared that probability to being less than the chance at seeing Halley's Comet in any given year. Should a firm happen to see an OSHA inspector, he will hand out an average of 2.2 violations per inspection. Most of the violations are for readily monitoreables aspects of the workplace, such as loose handrails and slippery staircases.

The financial incentive for compliance stems from the expected penalties given for violations. If the probability of inspection is low, then the only way to decrease violations is to hand out stiffer penalties. For many years, OSHA penalties tended to be negligible – on the order of \$10 million or less for the entire U.S. economy.

OSHA penalties have been dwarfed by the influence of financial incentives for safety created by the market through compensating differentials for risk. The Clinton administration made a concerted effort to boost the level of penalties, which is beneficial from an economic standpoint to the extent that they establish incentives for compliance using meaningful standards. One of the difficulties with early OSHA enforcement efforts was that the widespread lack of confidence in the soundness of the regulation led inspectors to avoid imposing penalties on the inspected firms. OSHA penalties now total \$82 million per year. Thus, the average penalty per violation is \$1,039. While penalties at that level represent a much-needed improvement in the financial incentives for safety, they are still dwarfed by other financial incentives facing the firm.

Remember that each workplace fatality generates wage premiums in terms of compensating differentials for risk on the order of \$6 million per statistical death. There also are additional safety incentives facing the firm, not the least of which is the influence of workers' compensation benefits. Premiums for workers' compensation are experience-rated, particularly for large firms in the United States, where the function is performed by private insurers, but much of this incentive is lost in countries where experience-rating is limited. From an economic standpoint, one would expect workers' compensation to be a more powerful driver of safety in the workplace than regulations. For much the same reason, compensating differentials for risk, which exceed even the influence of workers' compensation, should be the most powerful force promoting safety.

The role of workers' compensation in providing financial incentives for safety has been considerable. Empirical estimates indicate that, without the financial incentives, worker fatality rates in the U.S. would be one-third greater than their current level. In contrast, the effect of occupational safety and health regulations has been comparatively modest. Early estimates of OSHA's impact on worker safety failed to find any statistically significant effect of the agency. More recent estimates suggest that OSHA may have reduced all worker accidents that involve lost days of work by five to six percent.

The greater influence of workers' compensation stems from two factors. First, that effort creates greater financial incentives. Second, the incentives created by workers' compensation are performance oriented in that changes in the rate of workplace accidents will influence a firm's experience rating. Firms, in turn, can adopt a cost effective way to promote safety in response to the incentives. In contrast, command-and-control regulations give firms less leeway, thus leading to fewer safety gains for any given expenditure on safety.

VII. Social Security Regimen and Workplace Safety Reform

Workers Compensation insurance (WC) in the vast majority of Social Security agencies works as a form of no-fault insurance in case of a workplace accident or illness, where workers give up the right to sue their employer in exchange for a right to compensation. Firms are considered liable for workplace accidents or illnesses and pay insurance premiums as a percentage of their total payroll to a Social Security Institution, which compensates accident victims and pays for their medical expenses related to workplace injuries. In some jurisdictions, insurance premiums are adjusted to reflect the past claim records of firms (experience-rating). WC claims result from work injuries that produce an impairment that can be classified by duration (temporary or permanent) and severity (total or partial).

It is important to consider the implications of the WC system since it operates in a market context. The system may have a variety of effects on employees and employers. Three types of effects can be distinguished. First, WC may influence frequency, duration and nature of claims through a variety of incentive effects. In particular, WC insurance may lead to moral hazard problems which arise when informational asymmetries are used for personal gains. The first is that of ex ante injury hazard. Since insurance covers the financial and medical losses associated with the injury, workers' incentive to exercise care will diminish with increases in coverage. Moreover, because employers fund WC benefits through premiums linked at least in part to their firm's safety record, there is an incentive to increase the investment in health and safety capital when there is an increase in WC insurance coverage. These pressures may result in changes in risk or, more precisely, in the frequency or the duration of injuries.

A second form of moral hazard, termed ex ante causality hazard, arises because it is sometimes difficult to identify which accidents are caused by the job. Therefore, workers may file claims for accidents that have not occurred, or for off-the-job accidents. A third form might be termed ex post duration hazard. With an increase in the insurance coverage, injured workers may be tempted to take action in order to prolong the duration of the period over which benefits are paid out.

A final form of moral hazard, termed insurance substitution hazard, may arise, given that WC is in general more generous than unemployment insurance (UI) or general disability insurance. Workers may be tempted to undertake activities in order to benefit from WC instead of UI, when they are confronted with a lay-off. For instance, they may report false or off-the-job accidents or, given that they have been injured on the job, they may try to increase the duration of their period of recovery compensated by WC.

The decision to file an accident report may also be affected by the level of WC benefits (reporting incentives) since, in some circumstances, an injured worker may have some discretion over whether to ignore an injury and to continue working or to report the injury and to receive WC benefits. Second, the change in risk described above may in turn affect workers' wage through changes in compensating differentials, or simply because social insurance for job injuries will increase the attractiveness of risky employment to workers, thus reducing the required compensating differential. Third, WC benefits may lead to more absenteeism and the loss of firm specific human capital, which in turn may induce productivity effects.

Problems of imperfect information in the “market for workplace accidents” have attracted much attention recently. As noted by Rea (1981), there are at least five possible types of imperfect information that may affect this market: 1) Employers and insurers may not be able to identify workers who are accident-prone; 2) employees and employers may be incorrect in their estimates of occupational risk and of their influence on the level of risk; 3) the employer may not be able to monitor the precautions taken by employees; 4) the insurer may not be able to monitor employers’ and employees’ precautions; 5) the insurer may not be able to monitor the nature of injury. The first type of misinformation leads to what is commonly called adverse selection. The third and fourth types reflect ex ante injury hazard and the fifth type involves ex ante causality hazard or ex post duration hazard.

Are the rates actually in place too high or too low? Pressure groups have their opinions which are probably irreconcilable. Unions want higher replacement rates, while business associations often criticize the cost of workers’ compensation insurance (although the results showing that WC cost is transferred to workers through lower wages put this criticism in a different perspective). Of course, it is difficult to determine theoretically or empirically what is the optimal rate of replacement. Should policy makers be worried by the empirical results showing that increases in WC are associated with increases in the level of workplace risk (duration and frequency of injuries)? Not necessarily, as discussed earlier, these results indicate that employee’s responses to changes in WC are stronger than employer responses. This is a likely outcome if experience rating is not pervasive, which seems to be the case in most jurisdictions, especially for small firms. This could suggest that further increases in WC replacement rate should be accompanied by more intense experience rating if one wants to control total WC costs.

There is also a complex interplay of different types of social security system. These systems include, in varying combinations from one country to another, commercial and social insurance, first-party and third-party covers and an array of public, semi-public and private insurance carriers. In most countries some, if not all, compensation for work injuries is provided through a state social insurance scheme. In some cases, industrial injury compensation is a separate component, but in others integration is almost total, in the sense that the system makes little distinction between work accidents and other sorts of injury. The key example of this approach is found in New Zealand, where the state accident compensation scheme does not distinguish between different sources of injury, except as regards compensation for disease, which is limited to occupational illnesses. At the same time, an employee’s right to sue his or her employer in tort has been abolished in New Zealand. As a result, concepts such as “employers’ liability” and “workers’ compensation” have become almost redundant, because the injured employee has hardly any special rights.

Workers’ compensation models vary a great deal, but they have two key characteristics. First, they provide compensation on a no-fault basis. The claimant is not required to prove negligence or breach of a legal duty on the part of the employer, and fault on his or her own part is usually irrelevant except, perhaps, in the case of willful misconduct or self-inflicted injuries. Second, workers’ compensation systems rarely, if ever, provide “full” compensation for injuries. They aim only to provide reasonable redress for economic losses. Non-economic losses (such as pain and suffering) are rarely compensated, although exceptions are found in the workers’ compensation systems of Switzerland and Sweden. Core benefits include the cost of medical care and rehabilitation, replacement of lost earnings (usually limited to around 70 per cent of income and often subject to a maximum figure) plus funeral costs and benefits for surviving dependants in fatal cases (e.g. 30 to 40 per cent of the deceased’s previous earnings for a surviving spouse in Germany).

Employers' liability models, or tort-based systems, are schemes where the injured employee must establish legal responsibility on the part of the employer if he or she is to secure compensation. The obligation to pay compensation then falls on the latter although, of course, the risk may be transferred to a liability insurer. In most cases tort law requires the employee to prove negligence or fault. However, in some cases the burden of proof is reversed, and in other cases liability may be strict, although strict employers' liability is somewhat anomalous in a system where no-fault workers' compensation benefits are also available. In contrast to workers' compensation models, tort-based systems purport to provide full compensation. The successful claimant is entitled to redress for all losses, both economic and non-economic. This may include, amongst other things, full replacement of lost income, medical costs, and compensation for non-economic losses.

As Parsons (2002) establish, since employers' liability and workers' compensation systems can operate exclusively or in combination, three types of regime are possible:

- A regime where employers' liability is the exclusive remedy for industrial injuries;
- A regime where workers' compensation is the exclusive remedy for industrial injuries;
- A regime which combines employers' liability and workers' compensation.

Regimens (b) and (c) are the most common.

In theory, a State might decide that its injured workers should receive no compensation of any sort in the absence of fault or breach of a legal duty on the part of the employer. There would be an absence of "no-fault" benefits for injured employees and, taking the model to its extreme, no social insurance benefits at all, either general or specific. In fact, no country has gone nearly this far. All have retained some form of workers' compensation scheme and, indeed, there is hardly a country in the world where no such programme exists. In fact, the U.K. government has recently contemplated the abolition of its own no-fault workers' compensation regime but it cannot be supposed that injured employees would be denied all social insurance benefits if the state workers' compensation scheme were to be abolished. They would still be entitled to general disability benefits, albeit at a lower level. In any case, replacement of the state workers' compensation scheme with a privately insured alternative would probably be the most favored option.

Employers' liability systems, which rely upon tort remedies, are often criticized for their inefficiency, with high transaction costs and slow claim settlements. For these reasons many countries, including a number in Europe, have abolished the employee's right to sue in tort, allowing only a claim for defined workers' compensation benefits. For the worker, the loss of tort rights against the employer and of "full" tort compensation is balanced by the right to no-fault benefits that are easier to claim. Of course, when workers' compensation is substituted for the tort remedy in this way the position of the injured employee may appear anomalous, in relation to that of other accident victims. A number of countries have therefore extended the "no-fault/limited compensation" principle to other groups, such as road accident victims. The ultimate result may be a fully integrated no-fault accident compensation scheme such as that of New Zealand. Within Europe, Germany provides a prime example of what is virtually a pure workers' compensation system. Introduced in 1884, the German scheme was the first workers' compensation programme of any nation. It has been imitated worldwide.

Although workers' compensation systems abolish or heavily restrict tort claims against employers, it does not follow that injured workers lose all their rights to bring a tort action. On the contrary, rights against persons other than the employer are often preserved, giving the injured employee alternative targets for legal action. These may include whom responsible for the design, construction or layout of the workplace, safety consultants and, perhaps, directors of the injured employee's firm. However, the most obvious targets are manufacturers or suppliers of defective machinery or equipment used at work and suppliers of hazardous substances used in the workplace. This phenomenon, the displacement or transformation of work injury claims into product liability claims, is clearly evident in the U.S. and is one of the reasons for the extraordinary prominence of the product liability risk in North America. The incentive for legal arbitrage of this sort is a weakness of exclusive remedy workers' compensation systems. A main related issue is that nowadays, ecological, cultural, anti-discrimination and many other regulations can be a source of litigation against an employer, reducing the "value" of the prohibition of civil litigation in cases of work injuries.

In many European countries compensation for industrial injuries is provided partly through an employers' liability system and partly through a workers' compensation scheme. The balance between these two sources of compensation varies, but in most countries employers' liability is of marginal importance, the contribution of tort claims against the employer to the totality of industrial injuries' compensation being small. There are a number of reasons why, in an apparently "mixed" system, employers' liability may be relatively insignificant:

- Workers' compensation benefits may be so generous that few people consider a tort claim worthwhile;
- Tort claims by employers may be limited to cases where there is more than "ordinary" negligence, for example, proof of intent or gross negligence may be necessary;
- Tort claims by employees may be restricted to particular types of accident;
- Claims against employers may be limited to recoveries by workers' compensation insurers, direct claims by employees being restricted;
- Employees may be entitled to extra compensation under collective industrial agreements with their employers.

The U.K., where employers' liability is highly developed, and total employers' liability claim payments actually exceed those under the Industrial Injuries Scheme (the state workers' compensation component of the system), is very much the exception. Thus, in the ranking of European countries according to the degree of penetration of employers' liability within the industrial injuries compensation system, at the top, the U.K. (together with Ireland) at the far end of the scale. At the other extreme would be Germany, Austria and France, where employers' liability is of little or no importance. In between there are countries where workers' compensation is the main source of compensation, but employers' liability plays a rather more significant role, such as Italy and Spain.

Some writers suggest that there is a general trend away from workers' compensation systems towards employers' liability models. The main reason given is the funding problems that demographic trends have produced in welfare and social security systems throughout the world. This has led, in some countries, to cuts in public spending that have reduced the benefits available to injured workers and have increased the incentive to seek tort compensation. The U.K. Government, along with European governments generally, has shown a determination to use private insurance as a means of extending the social security system.

This has further shifted the balance of responsibility for industrial injuries away from the State and towards the employer and his insurers. In the future it is likely that publicly funded hospitals will be given a similar right of recovery against employers' liability insurers, in cases where work accident victims have been treated at public expense.

Also Parsons (2002) indicate that although the trend toward employers' liability is especially strong in the U.K. and Ireland there is evidence of a similar movement in some other European countries, including the Netherlands and Spain. Italy provides a further example. There the public workers' compensation insurer INAIL (Istituto Nazionale per l'Assicurazione contro gli Infortuni sui Lavoro) provides compensation on a no-fault basis for financial losses attributable to work accidents. Compensation for financial loss in excess of that available from INAIL and for pain and suffering is recoverable from the employer, but only when the latter (or a fellow employee) has committed a grave criminal offence and there has been a breach of regulations on safety at work. In this case the employer deducts benefits paid by INAIL from the claim and refunds INAIL.

Because it is often of marginal importance only, employers' liability insurance is rarely written as a separate line of insurance business in Europe. In most European countries the risk, where it exists at all, is insured under public (general) liability policies, either tacitly or expressly. Separate employers' liability policies are found in only a few European countries, such as the U.K., Ireland and Cyprus. The U.K. is one of very few European countries where employers' liability insurance is compulsory for virtually all employees by law. Carriers of the employers' liability risk are invariably private rather than state insurers. In Europe employers' liability insurance is never combined with workers' compensation insurance, although the practice is common in some U.S. states, Australia and Singapore.

Insurance is largely compulsory. In some cases it is provided by the state as part of a fully integrated social insurance scheme, as in the Netherlands. Alternatively, it may be a distinct component within a social insurance programme, such as the U.K. Industrial Injuries Scheme or the workers' compensation programme that forms part of the French national social security system (Sécurité Sociale), and as in Mexico. Again, it can be provided by recognized private insurers, as in the case of the "accident" element of the Belgian, Portuguese and Finnish schemes, which are discussed in the next section. Between the two extremes of state and private provision there is a variety of public and semi-public risk carriers. Examples include INAIL, the statutory public agency, mentioned above, which provides workers' compensation cover in Italy and the German Industrial Injuries Insurance Institutes – non-profit and largely autonomous corporations offering cover for member employers in particular industrial sectors (e.g. mining, gas and water, food, hotel and catering).

Employers (other than self-insured employers) pay insurers a premium based on their total wages or remuneration bill. The premium rate (or percentage rate of the wages bill) charged to any one employer depends on a range of factors such as the size of the employer, the industry in which the employer operates (industry class rating), individual claims experience (experience rating), the financial position of the insurer, and the stage of the insurance market cycle. Small to medium-sized employers are subject to industry class rating, whereas large employers are subject to experience rating.

Premiums are efficient if they are set so as to cover employers' expected scheme costs of work-related injury and illness. An employer's expected scheme cost depends on the likelihood of work-related injury or illness and includes the:

- Medical and income payments made to injured or ill workers;
- The cost of rehabilitating and facilitating the return to work of injured or ill workers;
- Compensation for the diminished quality of life of injured and ill workers; and
- The administration costs for insurers of managing premium pools.

If insurers do not set premiums efficiently or cost-effectively (so as to achieve scheme objectives at least cost to the society) a number of potentially adverse outcomes could arise:

- Unfunded liabilities, where a scheme's liabilities are not covered by its assets;
- Cross-subsidization between employers of different sizes, within a particular industry, in different industries, and over different generations;
- Insufficient levels of compensation;

- Distorted incentives on the part of employers to improve workplace health and safety.

Many policy makers in the workplace safety institutions have expressed concerns about premium setting in both publicly and privately underwritten schemes:

- The politically sensitive nature of the premium rates introduces a risk of rates being depressed for political purposes and this is clearly a contributing factor to the unfunded liabilities and insurance losses seen in all schemes at some point in their history.
- Schemes with unfunded liabilities as a result of 'inefficient and cumbersome' arrangements are under financial pressure to increase premiums to employers regardless of their workplace safety.
- Premiums are increasing despite safer workplaces and falls in the number of work related fatality, injury and illness.
- Employers with a good health and safety record are not being rewarded through lower premiums, while those who do not place the same emphasis on health and safety are not penalized.
- Small to medium-sized employers are 'penalized' by premiums despite good safety practices and claims records; they have no bargaining power in negotiating premiums and premiums offer them limited incentive to mitigate and improve the management of claims.

There are two main approaches to the funding of workers' compensation schemes:

- Pay-as-you-go funding meets the immediate cash requirements of the scheme. Immediate obligations are met such as management expenses and entitlements to weekly compensation, medical and hospital costs, and common law settlements. No assets are accumulated to meet future compensation entitlements or management expenses; this approach applies in Europe and New Zealand.
- Full funding is where sufficient assets are accumulated in the scheme to meet all expected entitlements to compensation, regardless of when they may be paid, and all costs associated with managing claims that have occurred. It is expected that investment income earned on the funds set aside to meet future claims will also be available to meet emerging costs. These earnings, and changes in them, can have a significant impact on the level and stability of premiums.

An important objective of premium setting is to create incentives for employers to improve workplace safety as well as to fund the cost of claims. Clearly, premiums are not the only way of achieving this. The regulatory measures embodied in occupational health and safety legislation is also important to this end.

For premiums to send clear signals to employers they should reflect workplace risks (or the expected scheme costs of work-related fatality, injury or illness). If risks are high, this should feed through into premiums, which in turn should signal to employers the need to invest in workplace safety and rehabilitation. Where there are improvements in safety and rehabilitation, and workplace risks are accordingly lowered, this should be reflected in reduced premiums.

VIII. Firms Size and Workplace Safety Reform

As workers' compensation schemes oblige employers (other than self-insurers) to purchase a compulsory insurance policy to cover their liability for work-related fatality, injury and illness, 'affordable' premiums are keenly desired. If premiums are too high, employers' competitiveness and financial viability would be affected. They might also encourage premium avoidance. Even where premiums are set to reflect the workplace risks facing employers; they may not necessarily be affordable for all employers. Cross-subsidization in premium setting may be introduced to ensure affordability across all employers. It should be noted that cross-subsidization amongst employers with different workplace risks is quite distinct from risk pooling which applies to employers with the same risk (see Box). With cross-subsidization, for example, the premium does not fully reflect an employer's workplace risk and the premiums of other employers facing different risks must increase or decrease to offset this. In practice, however, it is difficult to detect the precise extent of cross-subsidization.

However, cross-subsidization through premium setting dulls the incentive for both high and low risk employers to reduce work-related fatality, injury and illness. Removing cross-subsidization does not mean that an employer must always pay premiums exactly equal to the cost of claims it generates in a year. This would undermine the benefits of risk pooling. There is a difference between the pooling of risk and the cross-subsidization of a service through uniform pricing. If systemic cross-subsidization of premiums is to occur, it should be transparent, publicly justified and kept at levels which would not unacceptably distort incentives to employers to reduce workplace risks. Independent regulatory monitoring of premiums would be a way of achieving these conditions set. Employer affordability of premiums should be directly and transparently dealt with through explicit subsidies, such as given to implement workplace safety and rehabilitation programs, and not through premiums.

The difference between risk pooling and cross-subsidization ²¹

Often workers' compensation insurance is related with the cross-subsidization amongst employers and, indeed, equates risk pooling — which is the basis of insurance — with cross-subsidization. From an economic efficiency perspective, however, there is a difference between the two concepts. This can be illustrated using the following stylized example.

Risk pooling

Suppose there are 1000 employers each facing an identical 1 per cent probability of a claim for work-related injury or illness involving \$10 000 loss a year. On average, there would be 10 such claims amounting to \$100 000 a year. Each employer thus faces an expected loss of \$100 (0.1 times \$10 000) a year. However, in any give year, most employers face no claims, and around 10 employers would be exposed to claims amounting to \$10 000. Without insurance, employers are in a lottery for significant payouts which could threaten their continued viability.

Each employer could diversify or pool the risk by banding together to insure one another. In that case, each employer would pay a certain premium of \$100 a year (given no transaction costs) into a common fund and an individual employer facing a claim would finance its costs from the pooled premiums.

Cross-subsidization

Suppose that there are now two groups of employers with different risk profiles:

- Group A consists of 500 employers facing a 10 per cent probability of a \$10 000 claim a year. On average, there would be 50 such claims for the group amounting to \$500 000 a year; and
- Group B consists of 500 employers facing a 1 per cent probability of a \$10 000 claim a year. On average there would be 5 claims amounting to \$50 000 a year. Effective risk pooling would ensure that the premiums charged to the two groups would reflect their different risk profiles.

Hence, group A employers would pay a premium of \$1000 a year and the group B employers would pay \$100. If, however, no account is taken of the different risk profiles of the two groups in setting premiums, cross-subsidization would ensue. Suppose premiums are set to ensure that the expected total losses for all 1000 employers of \$550 000 (\$500 000 for group A and \$50 000 for group B) were distributed equally. Each employer would then need to pay a premium of \$550 a year. Group A would be paying far less than was actuarially fair given its risk profile (namely, \$1000 a year), and group B would be paying much more than was actuarially fair for its risk profile (namely, \$100 a year). This would significantly reduce the financial incentives for employers in group A to reduce workplace risks while adding an unnecessary cost burden on employers in group B.

An objective of premium setting desired by a number of participants is to ensure stability—or to reduce 'volatility'—in premium changes that result from employers' claims experience, particularly from atypical claims and random variations in claims. Also influencing premium stability has been the volatility of investment returns, especially those returns earned by private insurers. Premium stability can assist business planning and investment for future growth. Premium stability can also ensure the viability of employers. For example, a small to medium-sized employer could have a relatively good record but then encounters a relatively expensive claim. If the cost of this expensive claim has to be met immediately through a substantial increase in premium, it could have implications for the viability of the employer.

²¹ Example taken from "National Workers Compensation and Occupational Health and Safety Frameworks", Australian Government, Productivity Commission Inquiry Report, No. 27, 16 March 2004.

Most workers' compensation schemes impose limits on the amount of premium volatility that employers — particularly small to medium-sized employers — may face. Volatility is suppressed through various forms of premium controls, either directly (for example, through caps on premium increases) or indirectly (for example, by placing limits within experience rating formulae on the extent to which an employer's experience is reflected in premiums). Suppressing volatility in these ways means that employers bear the costs of claims for employers over a long period rather than closer to the time the costs of claims are incurred.

A degree of premium volatility is necessary to transmit incentives to employers about workplace safety and rehabilitation. Suppressing volatility through caps and other premium controls could mute these incentives, it is desirable to have premium volatility in which employers bear a greater proportion of the costs of claims closer to the time they are incurred, rather than have these costs spread over a longer period.

The risk properties of small and large employers differ and, thus, insurers apply different risk rating approaches to each. A small to medium-sized employer suffers from a lack of 'credibility' of their claims experience. Analysis of claims statistics show that, as a group, small to medium-sized employers are expected to have a low number of claims with a smaller proportion of large claims.

These ratios are relatively stable. However, an individual employer faces a far more changeable claims experience. It is difficult for an insurer to interpret an individual employer's claims experience.

Accordingly, it is often necessary for insurers to pool small to medium-sized employer risks such that the premium reflects a category based on some common element such as size or industry. Small to medium-sized employers are thus often charged the industry class rate. Moreover, as noted earlier in relation to industry class rating, the industry classes may be too broad to capture the statistically average risk for small to medium-sized employers engaged in a common activity.

There may be scope for insurers to pool more directly the experience of small to medium sized employers in order to gain the advantages of experience rating. There are proposal to small and medium firms as that a firms group could work together on occupational health and safety and returns to work, and could share resources, such as a safety manager. Groups could be based on existing industry, employer, district, service provider or union organizations. Care would be needed to avoid manipulation; however, such as if the group expels employers on the basis of one or a small number of claims.

Premiums charged to small to medium-sized employers could also include additional financial incentives for reducing workplace risks such as bonuses and penalties for claims performance, and explicit financial incentives for workplace safety and rehabilitation. The credibility problem associated with a small to medium-sized employer does not apply to the same extent for a large employer. Being large, its risk is in effect internally pooled and more predictable over time. Thus, for this group of employers, experience rating can be applied.

If large employers are subject to experience rating, then their exit from a scheme (say to become self-insurers) should not unduly affect the financial position of insurers who administer the premium pool. They would be paying their own way, including paying appropriate amounts towards the fixed administration costs of managing the pool. This of course might not be the case if large employers were cross-subsidizing others in the pool and/or are not contributing appropriately towards the fixed costs.

For premium setting, large employers should be subject to experience rating. Some of the deficiencies of experience rating could be dealt with explicitly to ensure it works better. For example, consideration could be given to simplifying formulae. For small to medium-sized (and new) employers, for whom claims experience is not a good proxy of workplace risk, industry class

rating should apply, accompanied by well designed explicit financial incentives for achieving workplace safety and rehabilitation.

In the publicly underwritten schemes, controls are applied to premiums directly (through premium caps) and/or indirectly (through caps on claims costs that are included in premium calculations). One objective for premium controls is to seek premium stability; another is to ensure affordability for employers.

As noted earlier, premium controls have real costs. They can mute signals and incentives to improve workplace safety and rehabilitation; create cross-subsidies; and lead to the under funding of schemes among other problems as follows:

- Adverse selection, as state workers' compensation schemes are subject to price controls, only "bad risks" have an incentive to seek insurance (rather than self insurance) through government providers. This adverse selection leads to unfunded liabilities in state schemes.
- Moral hazard arising from under pricing, where price caps and price floors exist there does not exist either an incentive for poor performers to improve their workers' compensation outcomes, or rewards for those who have exemplary records. The economic effect of such practices is an increase in workers' compensation claims.
- Cross-subsidies — where price controls are in place, it is inevitably the case that the poor risks are subsidized by the good risks. Again, this practice distorts and retards the economic incentives which would exist in the private market.
- Under-reserving — the effect of a non-market based pricing mechanism is that the liabilities exceed the revenues or assets. By failing to properly price policies initially, the flow on effect is that liabilities are unfunded.
- Cost-shifting — as a state scheme fails to fund itself, the costs are borne by other programs, such as public health, or by future policy holders, so that past losses become a burden for new businesses and employers.

The argument for the independent regulatory monitoring of premiums is that, if left to themselves, insurers (both public and private sector) might charge premiums based on factors not directly related to risks.

- Cross-subsidizing between employers with different risk profiles and between their less profitable workers' compensation business and their more profitable lines of insurance business, when they have market power to do so; and
- Offering incentives for large employers, but not for small to medium-sized employers, because the latter have less bargaining power and there is less margin to be traded in their premiums.
- In publicly underwritten schemes, premiums can also be based on factors other than risks:
- Cross-subsidizing between different employers with different workplace risks;
- A lack of competition leading to complacency; and
- Political pressure on public insurers to suppress premium increases when workplace risks increase to meet objectives such as premium affordability and stability or when elections are imminent.

Insurers, both public and private, might not closely relate premiums to risk where the transaction costs of doing so are too large (particularly in respect of small to medium-sized employers). There is scope for some type of independent regulatory premium monitoring of both private and public insurers. However, this should be light handed. Regulatory monitoring should seek to ensure that workplace risks are reflected in premiums and to make transparent the basis

for setting premiums, including exposing any cross-subsidies. It is to be distinguished from the monitoring that should occur under prudential regulation. The objectives of the latter include ensuring that long-term financial commitments can be met.

IX. Overview of the Occupational Risk in Latin America and Caribbean

The Latin American and Caribbean labour market is one of the fastest growing in the world. In 1980 there were 112 million workers; by 2001 the workforce almost doubled, reaching approximately 219 million (ILO 2005). The occupational safety situation in the region is far from adequate, as show in the Table 1, and characterized by three factors. First, there is a general lack of awareness regarding the importance of a safe and healthy working environment. Secondly, data on occupational accidents, illnesses and deaths tend to underestimate the magnitude of the problem. Finally the region lacks the institutional capacity and infrastructure needed to develop and sustain a safe and healthy working environment. Thus, failure to implement or enforce appropriate safety laws translates into lost production, lost wages, medical expenses, disabilities and deaths.

Workers in Latin America and the Caribbean pay a higher toll of deaths and injuries than in other regions of the world. Murray and Lopez (1996) calculated that work-related fatalities in Latin America and the Caribbean represented 3.2% of total deaths in 1990. This percentage is higher than in any other region in the world. For example, the same study calculated that in China work-related fatalities represented 2.8% of total deaths, in Asia 2.7% and in the developed market economies 2.2%.

Job insecurity and precarious contractual conditions characterize employment in the informal sector. A recent survey of more than 93 studies (Quinlan et al. 2001) provides evidence that precarious workers-temporary staff, workers subject to job insecurity, personnel employed with outsourcing arrangement, part-time workers – are associated with worse conditions. There are various reasons why the informal sector is linked to adverse outcomes. First, firms in the informal sector have higher turnover, thus less incentive to invest in training and to improve conditions. Secondly, regulation and compliance programmes usually focus on permanent employees in large workplaces. On the other hand, the informal sector, by definition, operates outside formal legal standards and regulation.

Workers in small businesses of less than six workers represented around 16% of the total non-agricultural workforce in 1999 (ILO 2000). Occupational safety conditions tend to be worse in small businesses because of the presence of fixed costs and economies of scale in investment and regulation compliance aimed to reduce occupational hazards. Consequently, along with other factors, it is argued that the smaller the industry is, the higher the rate of workplace injury and illness. Therefore, workers in the region may be more concerned with holding on to their jobs than with the possible health effects of hazardous working conditions.

In healthy people, exposure to occupational health hazards may occur up to a certain level without apparent effects because the human body has the capacity to deal with such challenges. However, some individuals can be more vulnerable because of their physical condition, age or gender. For example, toxicological evidence suggests that the health effects of exposure to hazardous chemicals are increased by malnutrition, and low protein diets increase the susceptibility of exposed individuals to the toxic effects of pesticides. The population in Latin America and the Caribbean contains a large proportion of people who are less healthy and therefore more vulnerable to occupational exposure to toxic chemicals or biological agents.

Women's participation in the labour force has been beneficial to their social and economic wellbeing, and has also improved the education and health of their children. However, women are exposed to greater health risks than their male co-workers, and their increased role in the

workplace has not generally been met by adjusting work conditions to gender differences. For example, women of fertile age are more susceptible to occupational hazards that affect reproductive functions. When pregnant, occupational hazards pose risks to the growing foetus, which may lead to congenital defects and miscarriage, as well as long-term impairments to the child's health and development. Women may suffer from muscle-skeletal disorders when the tasks or equipment used are designed for the 'average man' rather than adjusted to their different builds physiology. Moreover, populations that are particularly vulnerable to occupational hazards (children, women and the elderly) are concentrated in informal activities, which are more exposed to occupational hazards. Thus, the negative consequences for the health of these population groups would cumulate.

Occupational safety coverage applies only to workers who are affiliated with Social Security institutions. Thus a large portion of the workforce is excluded. Occupational health research is also likely to be under-funded since estimates show that only about 5% of occupational health research in the world takes place in developing countries. Overall the region has fewer experts, less safety equipment, less monitoring equipment, fewer inspectors and worse enforcement than developed nations, and the situation is particularly bad in the poorer countries.

Occupational accidents in informal-sector employments are usually not included in the figures reported. However, even for the workforce that is covered by reporting systems, the poor identification of work-related accidents, and the effects of some legal and bureaucratic features of the systems, does not guarantee the validity and accuracy of the estimates. Reporting systems are particularly weak in the diagnosis of occupational diseases because of the difficulty of relating the cause of the illness to the working environment. Thus, there is likely to be a high level of misattributions of occupational diseases to other sources.

In the region the only solid information that is available is social security payments for health care and indemnification for work-related disabilities and deaths. Table 2 summarizes the information available from several Latin America countries. The situation is quite different among countries. In Costa Rica, where the National Insurance Institute covers 68.4% of the country's workforce, direct costs of occupational injuries and diseases amounted to US\$47.9 million in 1995, representing nearly US\$70 per insured worker. In other countries, spending was significantly lower. -For instance, in Chile in 1996 expenditures were US\$33.80 per insured person. Of the countries listed here. Mexico spent the least - just US\$21.26 per insured worker. Considering data from all the countries the average social security expenditure per insured person amounted to US\$30.62 per year.

This direct spending by social security institutions is only one part, and a relatively small part, of the social costs of occupational injuries and diseases. Costs that are not directly compensated for by the social security systems are borne by workers, their families, employers, the public health system and taxpayers. For example, in Australia the Industry Commission estimated that injured workers and their families bore around 30% of total costs of occupational illness, employers bore another 30%, and taxpayers accounted for about 40%. A similar distribution of these hidden costs was found in Norway (Quinlan 1999). The fact that employers bear only a portion of these costs has implications for their incentives to create safe workplaces, as will be discussed in the following section. However, we expect that in the region, as well as in other developing countries, the distribution of costs is likely to be different because of the much smaller role of social security and public health systems. In particular, the share of costs borne by workers and their families is likely to be larger because of the large share of workers who do not have formal insurance coverage for occupational injuries and diseases.

Presumably, Latin America and Caribbean countries have many opportunities to improve occupational health and safety in ways that are cost-effective, given the generalized lack of safety provisions and the excessive rates of fatalities and non-fatal injuries documented above. There are many simple measures, such as adequate ventilation and unobstructed work areas that would go a long way toward reducing occupational risks in the region.

Table 1. Occupational Accidents and Fatality Rates, 2002

Country	Fatality rate per '000 workers	Accidents causing 3 days absence per '000 workers	Economically Active Population (%)
Argentina	0.1461	111.4897	36.7
Bolivia	0.1986	166.8949	46.6
Brazil	0.1661	126.7888	48.2
Chile	0.1466	111.8179	39.3
Colombia	0.1466	138.7000	47.4
Costa Rica	0.1584	121.1284	41.2
Ecuador	0.1822	138.8947	45.8
El Salvador	0.1796	136.9104	49.7
Guatemala	0.2243	171.2527	31.5
México	0.1592	121.5253	41.3
Nicaragua	0.2067	157.5156	34.1
Perú	0.1899	144.9008	45.7
Uruguay	0.1486	113.3705	47
Venezuela	0.1433	109.3071	41.7
Latin America and Caribbean	0.2582	14.8258	
U.S.	0.0519	39.5942	
Developed Market economies	0.0424	20.0402	

Source: Own calculations with ILO data (2005).

Table 2. Social Security Expenditure for Occupational Injuries and Diseases

Country (Year)	Cost of the Social Security (US\$ millions) ^a	Workers insured ^a	Proportion of the workforce insured (%) ^b	Cost per person insured (in US\$)	Cost per person insured/GDP per capita (%) ^c
Chile (1996)	122.5	3,624,129	68.40	33.80	0.90
Costa Rica (1995)	47.9	687,114	58.52	69.71	3.29
México (1996)	196.7	9,251,639	26.26	21.26	0.67
Perú (1996)	12.7	509,234	8.31	24.94	1.19
Venezuela (1995)	118.2	2,087,225	27.21	56.63	1.75

Source: Giuffrida, Iunes, Savedoff (2002).

VIII. Conclusions

Government action in the health and safety arena is justified when there are shortcomings in risk information. The goal of regulatory agencies that address health and safety risks should be to isolate instances in which misinformation about health risks prevents people from making optimal tradeoffs and to isolate instances in which health risks are not internalized in market decisions. The effectiveness of work hazard regulations hinges critically on the economic incentives created. Firms' investments in work quality will increase if such allocations will diminish the expected penalties associated with noncompliance with standards. Although the provision of safer work environment will be offset in part by diminished safety-enhancing actions by workers and compensations incentives from social security provisions. The weak financial incentives associated with enforcement effort combine with a ill-conceived nature of enforcement strategy lead to an environment with high cost and rent seeking associated with ineffective regulation.

The existence of a health risk does not necessarily imply the need for regulatory action. For example, as long as workers understand the risks they face in various occupations, they will receive wage compensation through normal market forces sufficient to make them willing to bear the risk; the health risk is internalized into the market decision. In situations in which the risks are not known to workers, as in the case of dimly understood health hazards or situations in which the labor market is not competitive, market forces might not operate effectively to internalize the risk. Those cases provide an opportunity for constructive, cost-effective government intervention.

Unfortunately, the rationale of correcting market failures has never been a major motivation of regulatory intervention. The simple fact that risks exist has provided the impetus for the legislative mandates of the health and safety regulatory agencies. To this day, very few regulatory impact analyses explore in any meaningful way the role of potential market failure in the particular context and the constructive role that market forces may already play in that context.

The conventional regulatory approach to health and safety risks is to seek a technological solution either through capital investments in the workplace, changes in the safety devices, or similar kinds of requirements that do not entail any additional care on the part of the individual. Stated

simply, the conventional view is that the existence of risks is undesirable and, with appropriate technological interventions. That perspective does not recognize the cost tradeoffs involved; the fact that a no-risk society would be so costly as to make it infeasible does not arise as a policy concern of consequence.

The potential role of the government is not to eliminate the risk but rather to address market failures that lead to an inefficient balance between risk reduction and cost. The task of government regulatory agencies is to identify cases in which regulation can generate benefits to society that are worth more than the costs that are incurred and to address market failures using a cost-effective approach.

To achieve those goals, the focus should not simply be on rigid technological standards but on flexible regulatory mechanisms that meet the performance goals. Almost from its inception, health and safety regulation has been the target of proposed reform. Some policy improvements have occurred, such as elimination of some of the nitpicking of safety standards, the increased use of informational approaches to regulation, and enhanced enforcement efforts. However, health and safety regulations have fallen short of any reasonable standard of performance.

The underlying difficulty can be traced to the legislative mandates of the regulatory agencies. Instead of focusing regulations on instances of market failure, the emphasis is on reductions of risk irrespective of cost. The regulatory approach has also been characterized by an overly narrow conceptualization of the potential modes of intervention. The emphasis has been on command-and-control regulations rather than performance oriented standards. More generally, various forms of injury taxes that would parallel the financial incentives created by workers' compensation programs could establish incentives for safety while at the same time offering firms leeway to select the most cost-effective means of risk reduction. A glaring omission from the regulatory strategy has been adequate attention devoted to the role of consumer and worker behavior and the potential for exploiting the benefits that can derive from promoting safety-enhancing actions by individuals rather than relying simply on technological controls.

The firms size are heterogeneous in size, measure in assets, employment or sales, the investment in workers safety is a fixed cost therefore the size determine the average compliance cost to each firm, small firms are in disadvantage if the regulation is inflexible. The participation of employment in small and medium size firms is very important, but there are plenty of private participation coalition opportunities to overcome the size obstacle to comply with regulation. Moreover, the level of stringency of the regulation is also of consequence, if a regulation is very tight many firms will choose not to comply at all, so that a very tight and uniform regulation may actually produce less of a beneficial safety effect than a more flexible regulation for which compliance is feasible by any firm size.

The use of performance standards rather than narrow tight specification standards could enable firms of any size to select the cheapest means of achieving safety objectives. Such flexibility would reduce compliance cost and increase the incentives of firms to develop innovative technologies, individually or by joint organizations, to foster health and safety.

Current regulatory approach have long seized the moral high ground by claiming that their uncompromising efforts protect individual health; less consequential concerns such as cost should not interfere with that higher enterprise. The fallacy of such thinking is that high-cost, low-benefit safety regulations divert society's resources from a mix of expenditures that would be more health enhancing than the allocations dictated by the health and safety regulations. Agencies that make an unbounded financial commitment to safety frequently are sacrificing individual lives in their symbolic quest for a zero-risk society.

There is a need to reform the regulation approach in design safety objectives or standards and enforcement. The enforcement efforts for regulatory standards have largely been inadequate. The

experience thus far has been characterized by safety standards of dubious relevance coupled with lax enforcement. A preferable approach, if the standards strategy is to be continued, is to couple sound economic in regulations design with a stricter enforcement that provides real financial incentives for safety.

Health and safety regulations that have the current inordinate imbalance between costs incurred and risk reductions achieved divert society's resources from a mix of expenditures that would be more health enhancing. Agencies that make an unbounded financial commitment to safety frequently are sacrificing individual lives in their symbolic quest for a zero-risk society.

As Hahn (2005) establish economic analysis have made important contributions to the study of social regulation, including inside workplace regulations. These contributions include the showing the regulation is inefficient in the sense that is possible to get more for less and showing that significant fraction of regulation design in safety are likely to fail a cost-benefit test. The quantitative tools as cost-benefit analysis will support some regulation as they are, in others cases they will suggest the need to change or eliminate regulation. Changing the decision making process in workplace regulation across countries will provide better treatment for uncertainty, resources to asses the effects of regulation on different firms and workers socioeconomic groups, increase regulatory transparency and new institutional approach for improving regulation and social security.

References

- “Breaking the Vicious Circle: Toward Effective Risk Regulation”*, by Stephen Breyer. Cambridge, Mass.: Harvard University Press, 1993.
- “Do Federal Regulations Reduce Mortality?”* by Robert W. Hahn, Randall W. Lutter, and W. Kip Viscusi. Washington, D.C.: AEI – Brookings Joint Center for Regulatory Studies, 2000.
- “Economics of Antitrust and Regulation”*, by W. Kip Viscusi, John M. Vernon and Joseph Harrington, Jr. Cambridge, Massachusetts, MIT Press, 2000.
- “The Perils of Prudence: How Conservative Risk Assessments Distort Regulation,”* by A.L. Nichols and R.J. Zeckhauser. *Regulation*, Vol. 10-2, 1986.
- “Rational Risk Policy”*, by W. Kip Viscusi. Oxford, U.K.: Oxford University Press, 1998.
- “Regulation by Litigation”*, by W. Kip Viscusi, editor. Washington, D.C. AEI – Brookings Joint Center for Regulatory Studies, 2002.
- “Risk and Reason”*, by Cass Sunstein. Cambridge, U.K.: Cambridge University Press, 2002.
- “Incentive Effects of Public Insurance Programs on the Occurrence and the Composition of Workplace Injuries”*, by Denis Bolduc, Bernard Fortin, France Labrecque and Paul Lanoie, Cirano Working Paper 97s-24, May 1997.
- “Effects of Workers Compensation: A Survey”* by Bernard Fortin and Paul Lanoie, Cirano Working Paper 98s-04, February 1998.
- “Incentive Effects of Workers Compensation Insurance”* by Alan B. Kruger, NBER Working Paper 3089, August 1989.
- “The Incidence of Mandated Employer-Provided Insurance: Lessons from Workers Compensation Insurance”*, by Jonathan Gruber and Alan B. Kruger, NBER Working Paper 3557, December 1990.
- “Liability Rules, Compensation Systems and Safety at Work in Europe”*, by Chris Parsons. The Geneva Papers on Risk and Insurance, vol. 27-3, July 2002.
- “World Day for Safety and Health at Work 2005: A Background Paper”*, by International Labor Organization (ILO), Geneva, 2005.
- “The Global Burden of Disease”*, by Murray C.J.L. and Lopez A.D., Harvard University Press, Cambridge MA, 1996.
- “The Implications of Labor Market Restructuring in Industrialized Societies for Occupational Health and Safety”* by Quinlan M., *Economic and Industrial Democracy* 20, 1999.
- “The Global Expansion of Precarious Employment, Work Disorganization, and Consequences for Occupational Health: A Review of Recent Research”*, *International Journal of Health Services* 31, 2001.
- “Occupational Risks in Latin America and the Caribbean: Economic and Health Dimensions”* by Giuffrida, A., Iunes, R.F., Savedoff, W.D., *Health Policy and Planning*, September 2002.

“In Defense of the Economic Analysis of Regulation” by Robert W. Hahn. American Enterprise Institute, 2005.